Optimization
L1-Norm SVM

N Chakradhar[1] Havish[2]

[1]EE16BTECH11022

[2]EE16BTECH11023

March 4, 2019

Introduction

The standard 2-norm SVM is known for its good performance in two- class classication. In this paper, we consider the 1-norm SVM. We argue that the 1-norm SVM may have some advantage over the standard 2-norm SVM, especially when there are redundant noise features.

Note on SVM

An SVM model is a representation of the dataset as points in space, mapped so that the data points of the separate categories are divided by a clear gap that is as wide as possible. So, our objective is to find support vectors(lines) such that it is at maximum distance to its corresponding class.

The problem formulation

The SVM(p-norm) formulation is of the form

$$min||w||_p \tag{1}$$

subject to the following constraints:

$$y_i(w^T x_i + w_0) \geq 1 \tag{2}$$

Using lagrange multipliers, we can convert this to unconstrained optimization problem as follows

$$min||\beta||_p + \Sigma_{i=1}^{i=n} \alpha_i (1 - y_i(w^T x_i + w_0) \tag{3}$$

Clearly, the constrained problem is convex because we know that $||\beta||_p$ when $p \geq 1$ and the constraints are also convex.

Implementation

We would like to frame both the optimization problems using cvxpy and compare their performance on a standard ML dataset.