CSE353 Assignment 7 Linear Hard-margin Support Vector Machine

Due Dec 16 2021 5:00PM

Main TA for this assignment:
Lihan Huang
Email: lihan.huang@stonybrook.edu
Office hours: Thursday 10AM-noon
Zoom: https://us02web.zoom.us/j/3165392097?pwd=WVJpeFVnQUhXSGFRSWROMEJ5S29FUT09

One linear separable data set ('X_LinearSeparable.txt' and 'Y_LinearSeparable.txt') is provided, where dataset X (20 x 2) contains 20 training samples. Each column in X is a sample with feature dimension of 2. Y (20 x 1) contains the ground truth binary labels of all samples.

(a). Implement and apply the linear support vector machine onto the given dataset to obtain the $(\boldsymbol{w}, b)$. Note: use the quadprog() function in Matlab/Python to solve the Quadratic Programming optimization problem.

(b). Identify which training samples in the dataset are support vectors.

(c). Compute the largest margin you achieved from your SVM.

(d). Visualize the training data, highlight the support vectors, plot the decision boundary line and the two lines along the support vectors.

Upload your codes with enough comments and a brief report to Blackboard by the due date & time, including

a) Introduction. Brief summary of what you think the assignment is about,

b) Method. Brief outline of your (algorithmic) approach,

c) Experiments. Tables and/or pictures of intermediate and final results that convince us that the program does what you think it does.

d) Discussions and Conclusions. Any design decisions you had to make and your experimental observations. What do you observe about the behavior of your program when you run it? Does it seem to work the way you think it should? Play around a little with different setting to see what happens. Note, your open-ended exploration is highly valued.