Luke Jaffe
CS6140 Machine Learning
12/19/15

HW6 – SVN and Kernels

Problem 1. SVM Library

A) Spambase
python p1a.py -k {p,l,r}

Polynomial kernel:
Training Accuracy = 80.7246% (3342/4140) (classification)
Testing Accuracy = 81.7787% (377/461) (classification)

Linear kernel:
Training Accuracy = 93.2609% (3861/4140) (classification)
Testing Accuracy = 94.577% (436/461) (classification)

RBF kernel:
Training Accuracy = 94.8068% (3925/4140) (classification)
Testing Accuracy = 93.7093% (432/461) (classification)

B) Digits
python p1b.py

Training Accuracy = 97.02% (9702/10000) (classification)
Testing Accuracy = 92.2% (9220/10000) (classification)

Problem 2. SVM SMO for Spambase

Usage: python p2a.py

Training accuracy: 0.891062801932
Testing accuracy: 0.915401301518

Problem 3. SVM SMO for Digits

Usage: python p3a.py

Simple accuracy: 0.7436
Complex accuracy: 0.7436

Problem 4.

Explain why $0 <= \alpha <= C/m$ is a constraint in the dual optimization with slack variables:

The parameter C is a scaling factor to the slack variable sum term. It is chosen by the user, where a larger C corresponds to assigning a higher penalty to errors. The three conditions listed below are the KKT conditions, used to check for convergence to the optimal point. Any $\{\alpha_i\}$ that satisfy these conditions for all *i* will be an optimal solution to the SVM optimization problem.

a) $\alpha_i = 0 \rightarrow y^{(i)}(w^T x^{(i)}+b) >= 1$

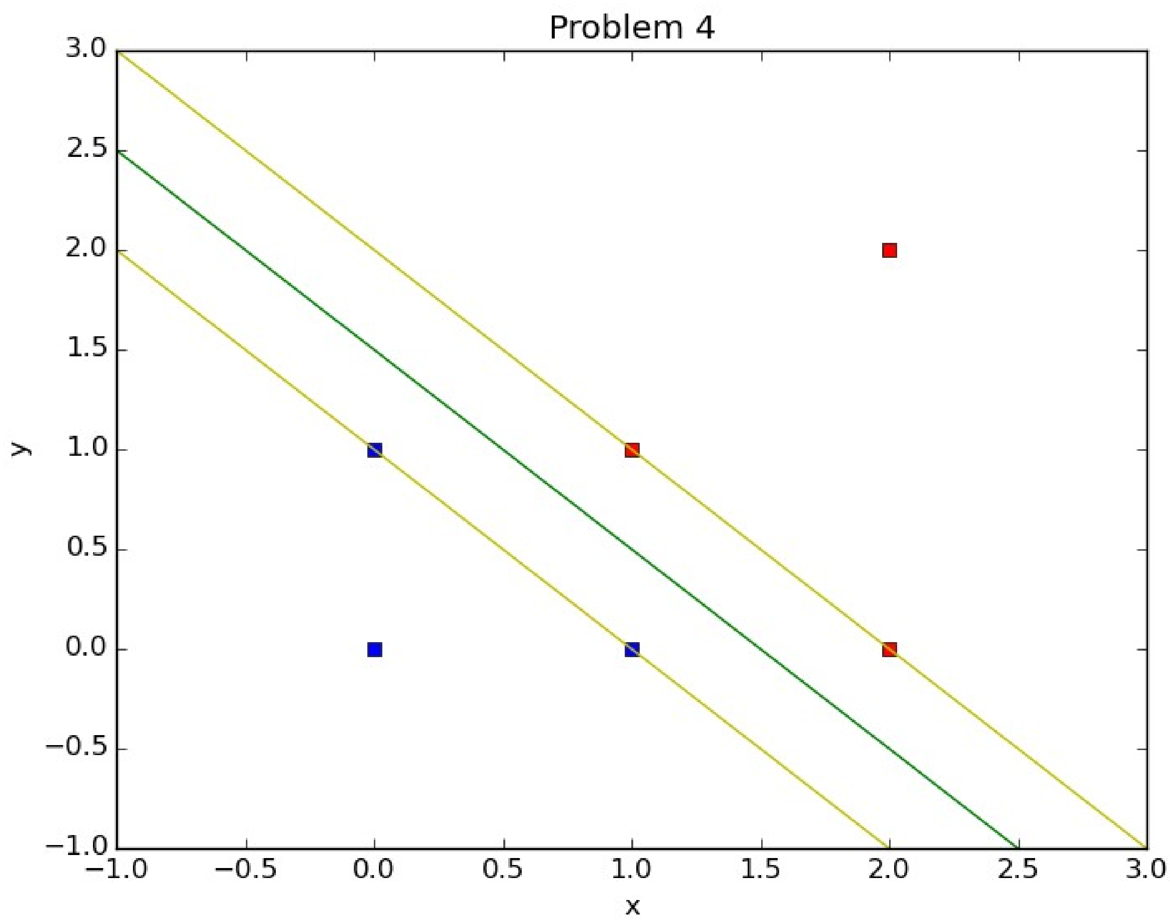
b) $0 < \alpha_i < C \rightarrow y^{(i)}(w^T x^{(i)}+b) = 1$

c) $\alpha_i = C \rightarrow y^{(i)}(w^T x^{(i)}+b) <= 1$

Problem 5.

a)

class 0: (1,1) (2,2) (2,0)
class 1: (0,0) (1,0) (0,1)



Optimal hyperplane:
W = -1.0
b = 1.5

Margin (of the classifier) = maximum width of the band that can be drawn deparating the support vectors of the two classes.
Margin = sqrt(2)/2 = 0.707

b) Support vector points: [0, 1], [1, 0], [2, 0], [1, 1]
*The yellow margin lines in the plot pass through these points.