

6.034 Quiz 4

December 3, 2008

Name	
EMail	

Circle your TA and recitation time, if any, so that we can more easily enter your score in our records and return your quiz to you promptly.

TAs	Thu	Fri
Sam Glidden	Time	Instructor
Mike Klein	1-2	Bob Berwick
Mark Seifter	2-3	Bob Berwick
Norman C. N. Quest	3-4	Bob Berwick
	1-2	Howard Shrobe
	2-3	Howard Shrobe
	3-4	Howard Shrobe

Problem number	Maximum	Score	Grader
1	50		
2	50		
Total	100		

There are 7 pages in this quiz, including this one. In addition, 1 tear-off sheet is provided at the end.

As always, open book, open notes, open just about everything.

Problem 1: Support Vectors (50 points)

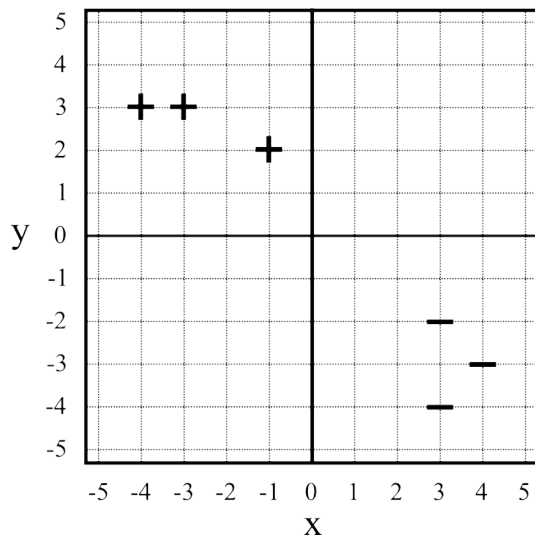
Part A (30 points)

The following three graphs represent the training data for two-dimensional, linear SVMs. + symbols indicate the coordinates of positive examples, which the classifier should give a value of 1 or more, and - symbols represent the coordinates of negative examples, which the classifier should give a value of -1 or less.

For each graph, you should do the following **three** things:

- Circle the support vectors
- Draw the “street” produced by the SVM, with solid lines where the classifier outputs 1 or -1 and a dashed line where the classifier outputs 0
- Describe the classifier function $h(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b$, by giving the values of \mathbf{w} and b

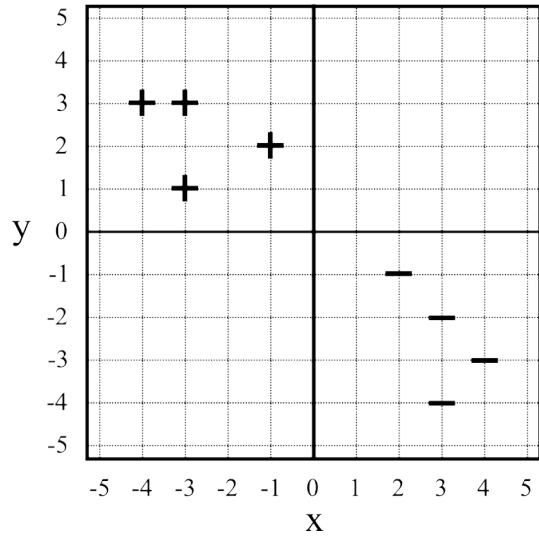
Graph #1



$$\mathbf{w} =$$

$$b =$$

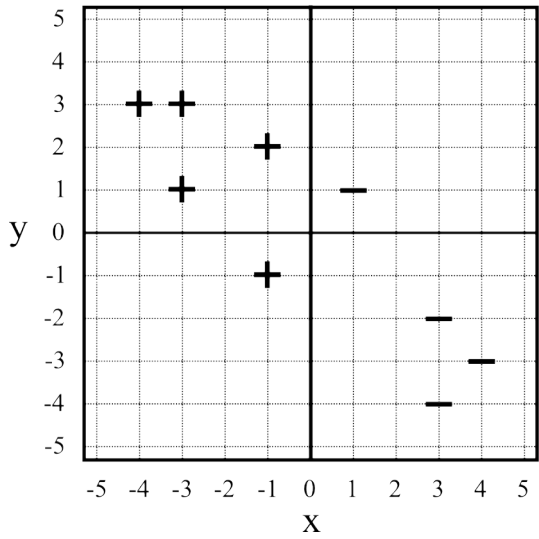
Graph #2



$w =$

$b =$

Graph #3

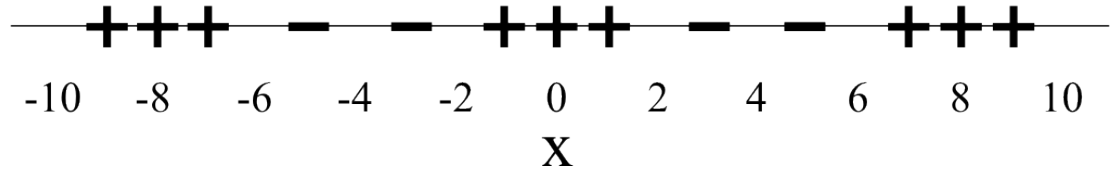


$w =$

$b =$

Part B (20 points)

Susan Q. Randomstudent is frustrated by the data she is supposed to classify with a support vector machine. It's all on the same line, defined only by an x -coordinate, and it switches back and forth from clusters of $+$ points to clusters of $-$ points several times, so it's not at all linearly separable.



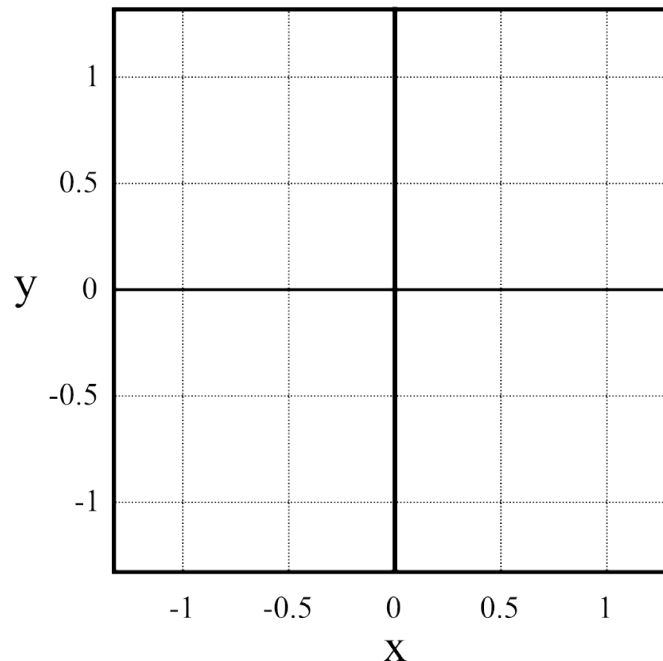
Eventually, though, she figures out a kernel function that she can use to separate the data:

$$K(x_1, x_2) = \cos\left(\frac{\pi}{4}x_1\right)\cos\left(\frac{\pi}{4}x_2\right) + \sin\left(\frac{\pi}{4}x_1\right)\sin\left(\frac{\pi}{4}x_2\right)$$

1. What is the Φ function that projects x into the new space defined by this kernel? (The function should take in x and output a vector with two components.)

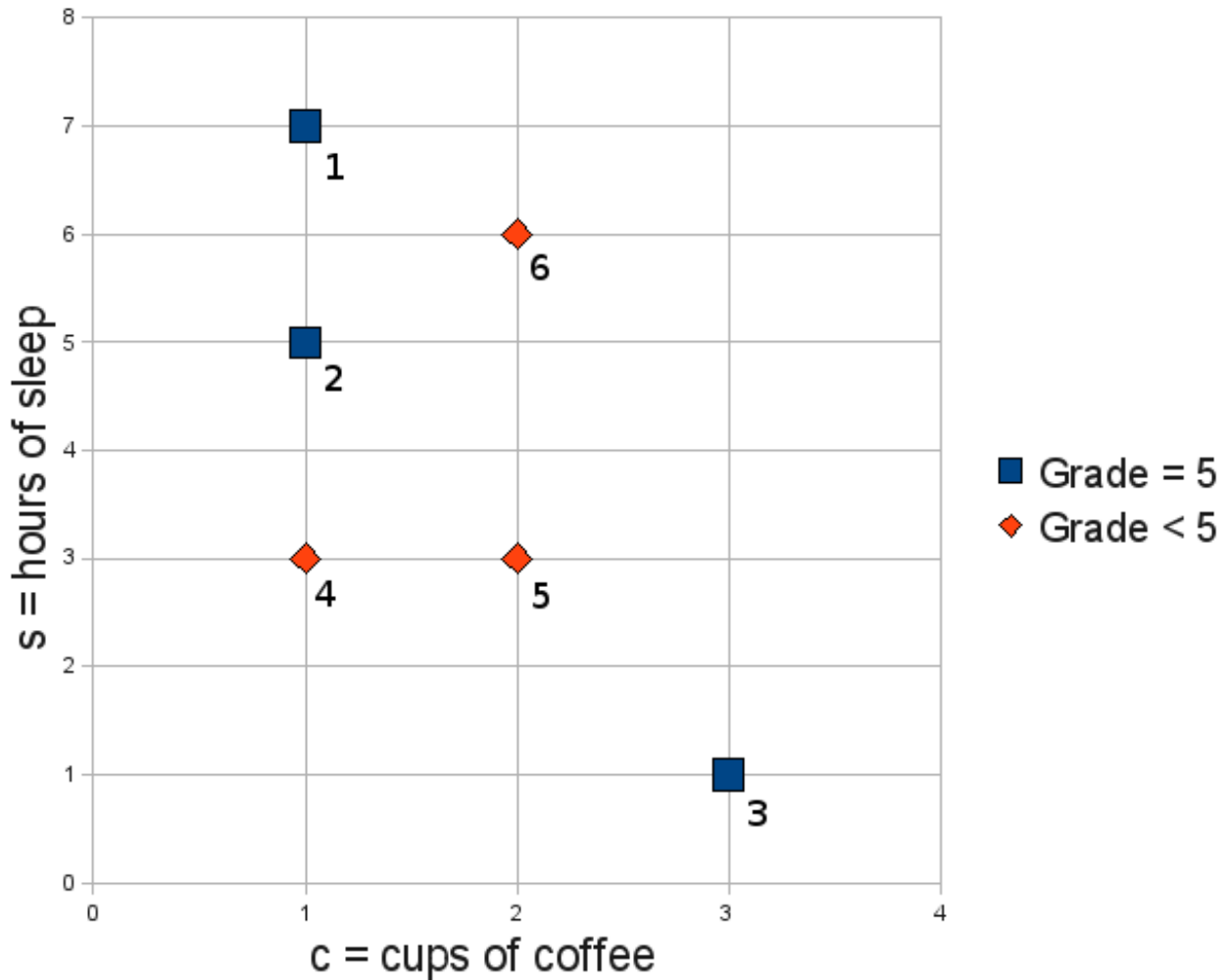
$$\Phi(x) =$$

2. On the grid below, draw the $+$ and $-$ points in the new space, and draw the “street” that linearly separates them.



Problem 2: Boosting (50 points)

Prof. Winston wants you to study the link between sleep, coffee, and 6.034 quiz grades. He gathered 6 data points from the previous quiz for you:



You decide to encode “Grade = 5” as +1, and “Grade < 5” as -1.

For this entire problem, consider only horizontal and vertical classifier stumps halfway between two data points, e.g. $h(c,s) = \text{sign}(c - 2.5)$ or $h(c,s) = \text{sign}(2 - s)$. You may also use $h(c,s) = +1$ and $h(c,s) = -1$.

In case of ties: prefer stumps involving s (horizontal) over stumps involving c (vertical), and prefer lower thresholds ($s < 1.5$) over higher ($s < 2.5$). Blanket classifiers $h(c,s) = +1$ and $h(c,s) = -1$ come last.

Part A (30 points)

Use Adaboost for 3 rounds. First update the weights using the results of the previous round, find the best stump $h(c,s)$ and its error ϵ . The starting weights have been filled in for you.

$$w_1 = 1/6$$

$$w_2 = 1/6$$

$$w_3 = 1/6$$

$$w_4 = 1/6$$

$$w_5 = 1/6$$

$$w_6 = 1/6$$

$$h_1(c,s) =$$

$$\epsilon_1 =$$

$$w_1 =$$

$$w_2 =$$

$$w_3 =$$

$$w_4 =$$

$$w_5 =$$

$$w_6 =$$

$$h_2(c,s) =$$

$$\epsilon_2 =$$

$$w_1 =$$

$$w_2 =$$

$$w_3 =$$

$$w_4 =$$

$$w_5 =$$

$$w_6 =$$

$$h_3(c,s) =$$

$$\epsilon_3 =$$

Part B (15 points)

What is the current Adaboost classifier made from combining these first three stumps? How many errors does it make on the training data?

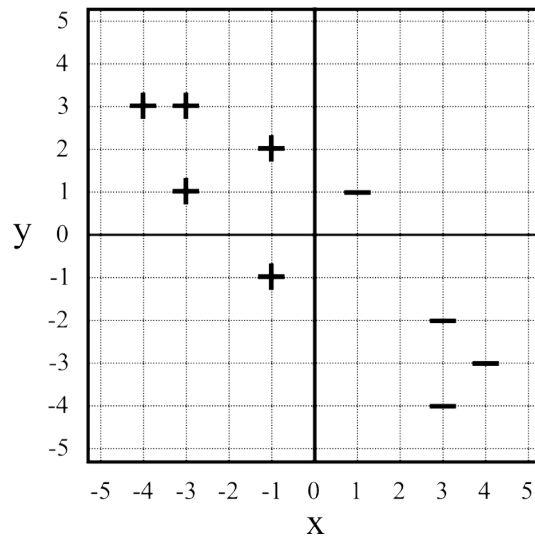
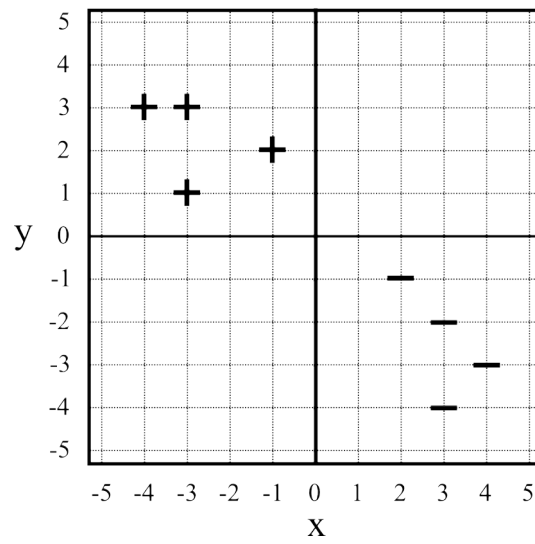
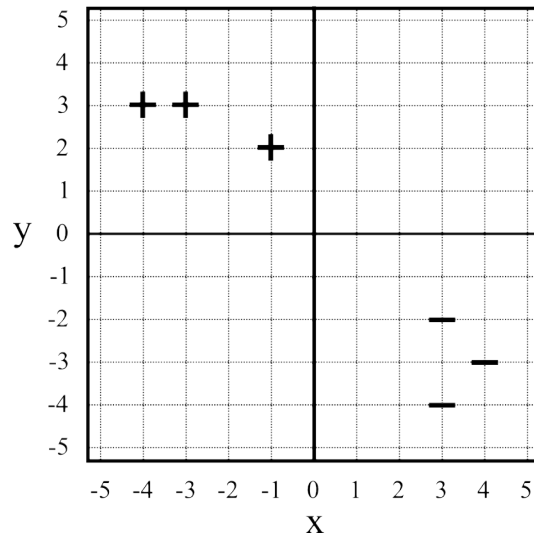
$H(c,s) =$

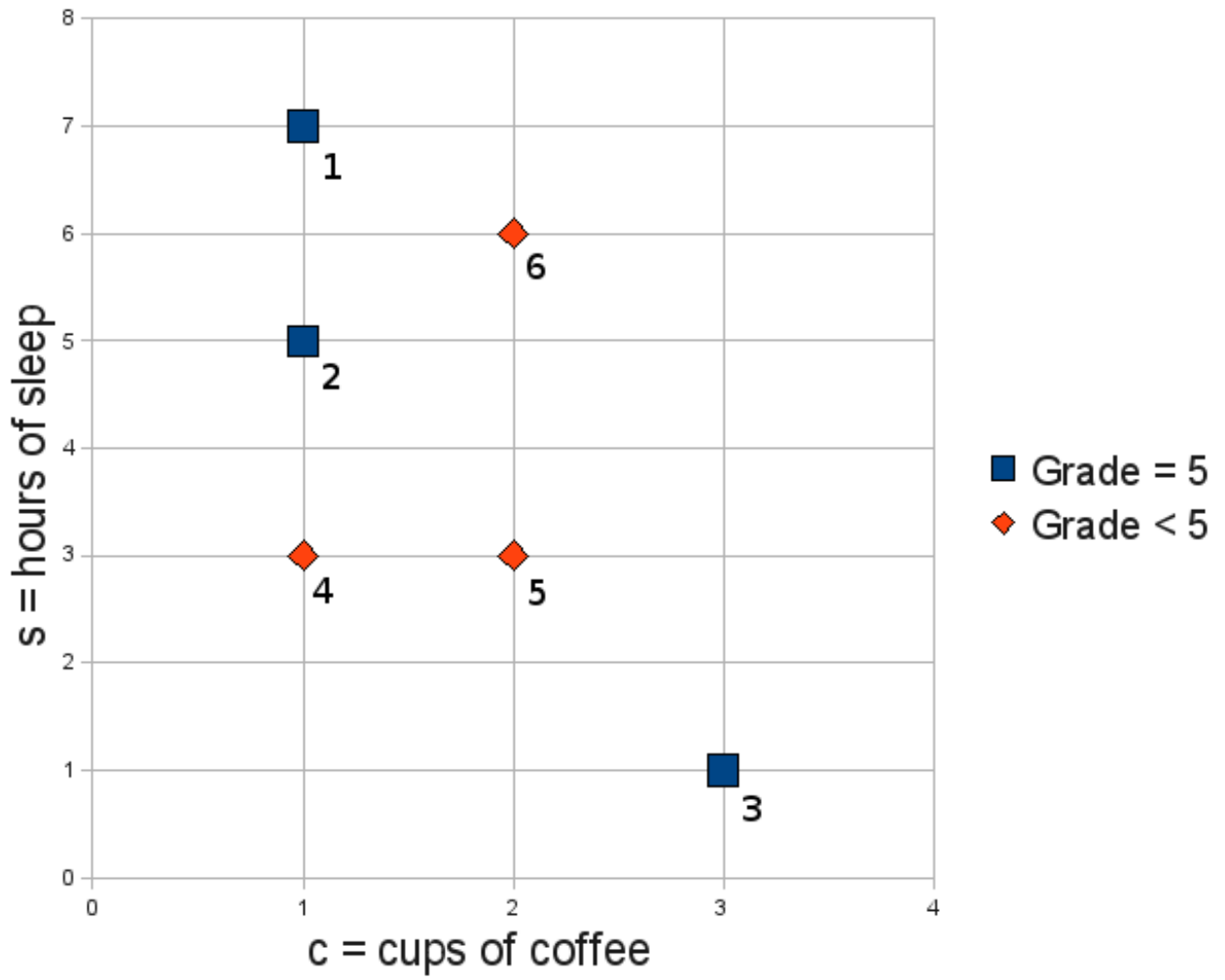
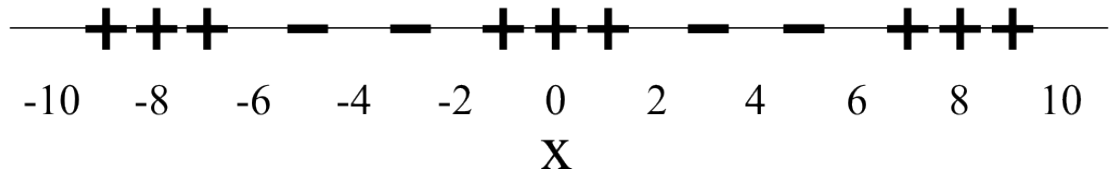
of errors =

Part C (5 points)

If you were to run Adaboost for 100 more rounds, which data point or data points do you expect to have the **least** weight? Explain briefly. (You may write down another 100 rounds as in Part A in lieu of explanation.)

(This page is blank for a very good reason.)





MIT OpenCourseWare
<http://ocw.mit.edu>

6.034 Artificial Intelligence
Fall 2010

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.