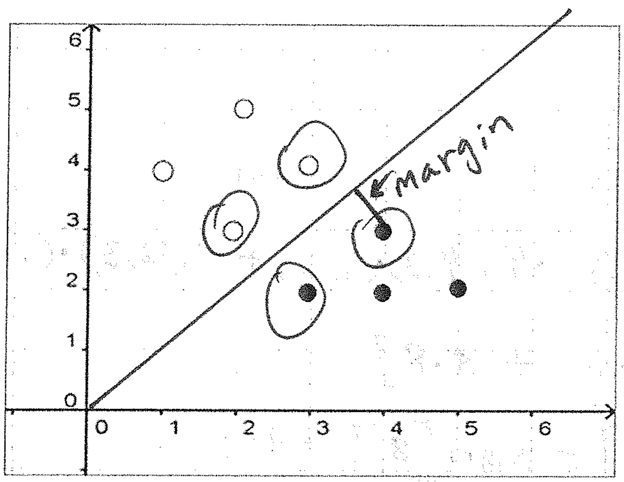
Name Solutions

Quiz 2

Please write all answers on these pages.

1. Given the training set plotted below: Sketch the hyperplane (i.e., line) that maximally separates the two classes (open circle and solid circle). Also sketch a line that indicates the margin, and label it as "margin". Also circle the support vectors.



2. Answer the following in one or two sentences.

(1) What are the inputs to the SVM algorithm?

Training examples

(2) What does the SVM algorithm output?

α_i's, support vectors, bias

3. Consider the following three points, $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$, which have been identified as support vectors for a training set. Here, y_i is the class of the point, and α_i is the support vector coefficient.

$\mathbf{x}_1 = (2, 1)$ $\quad y_1 = -1$ $\quad \alpha_1 = -4$

$\mathbf{x}_2 = (4, 3)$ $\quad y_2 = -1$ $\quad \alpha_2 = -4$

$\mathbf{x}_3 = (2, 3)$ $\quad y_3 = +1$ $\quad \alpha_3 = 8$

The bias is $b = 0$.

(a) Using the formula

$$h(\mathbf{x}) = \operatorname{sgn}\left(\left(\sum_{i=1}^{m} \alpha_i (\mathbf{x}_i \cdot \mathbf{x})\right) + b\right),$$

give the classification of the new instance $\mathbf{x} = (1, 2)$. Show your work.

$$h(1,2) = \operatorname{sgn}\left[-4(2,1)\cdot(1,2) - 4(4,3)\cdot(1,2) + 8(2,3)\cdot(1,2)\right]$$
$$= \operatorname{sgn}\left[-4\cdot 4 - 4\cdot 10 + 8\cdot 8\right]$$
$$= \operatorname{sgn}\left[-56 + 64\right] = \operatorname{sgn}[8] = +1$$

(b) Use the $\mathbf{x}_i$'s and α_i's to find the weight vector $\mathbf{w}$ associated with the separating hyperplane, where $\mathbf{w} = \sum_i \alpha_i \mathbf{x}_i$.

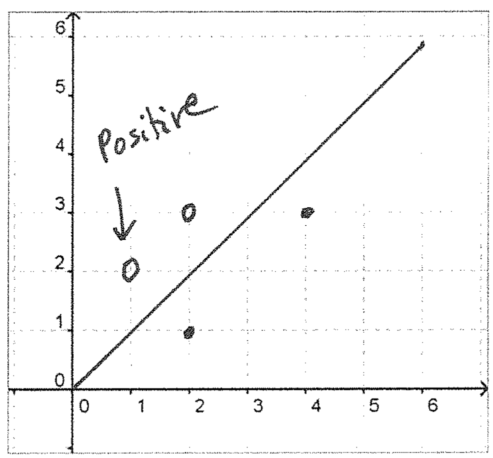
$$w = -4(2,1) - 4(4,3) + 8(2,3)$$
$$= (-8,-4) + (-16,-12) + (16,24)$$
$$= (-8, 8)$$

(c) Using the weight vector you obtained in part (b) and the bias b = 0, find the equation of the separating hyperplane. Give the equation in the slope-intercept form: x_2 = (*slope* * x_1) + *y-intercept.*

$$-8x_1 + 8x_2 = 0$$

$$x_2 = x_1$$

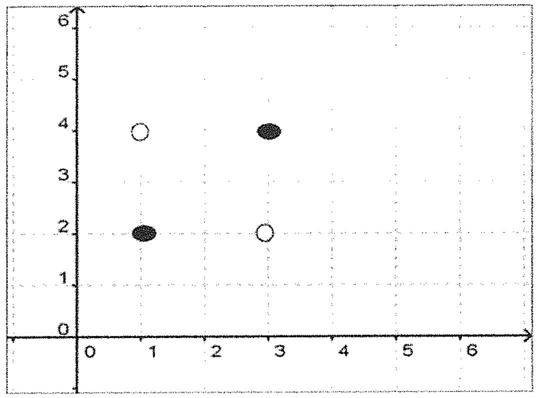
(d) Using the graph below, plot the support vectors and the separating hyperplane. Also draw a line to show the margin. Finally, plot the new point (1, 2) from part (a) to confirm that it is in the class you found in part (a).



4. Consider the points shown in the graph below:

(1, 2), (1, 4), (3, 2), and (3, 4)

where open circles are class +1 and solid ellipses are class −1. Give a mapping Φ ($\mathbf{x}$) from two-dimensional points into three-dimensional points that makes these points linearly separable in three-dimensions.



There are many possible answers. One example:

$\mathbf{x} = (x_1, x_2)$

$$\Phi(\mathbf{x}) = \begin{cases} (x_1, x_2, 1) & \text{if } x_1 + x_2 = 5 \\ (x_1, x_2 -1) & \text{otherwise} \end{cases}$$

5. Suppose you have a training set in which each instance is represented by four integer features: $\mathbf{x} = (x_1, x_2, x_3, x_4)$.

Define a "kernel" function as follows:

$$k(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{4} \min(x_i, y_i)$$

For the following training set, give the Gram matrix for this kernel function.

$\mathbf{x}_1 = (1, 3, 6, 1)$

$\mathbf{x}_2 = (2, 4, 3, 0)$

$\mathbf{x}_3 = (8, 1, 2, 4)$

Recall, the Gram matrix $\mathbf{K}$ is defined as

$$\mathbf{K}_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j), \text{ for } i, j = 1, \ldots, n$$

$$K = \begin{pmatrix} K(X_1,X_1) & K(X_1,X_2) & K(X_1,X_3) \\ K(X_2,X_1) & K(X_2,X_2) & K(X_2,X_3) \\ K(X_3,X_1) & K(X_3,X_2) & K(X_3,X_3 \end{pmatrix}$$

$$= \begin{pmatrix} 11 & 7 & 5 \\ 7 & 9 & 5 \\ 5 & 5 & 15 \end{pmatrix}$$