1. (9 pt) Consider the dataset in Fig 1, with points belonging to two classes, blue squares and red circles. Partial solutions. You should address the parts in red text.

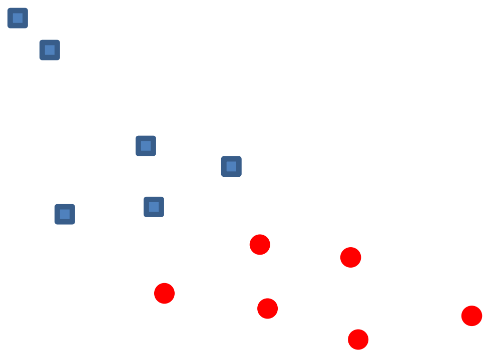
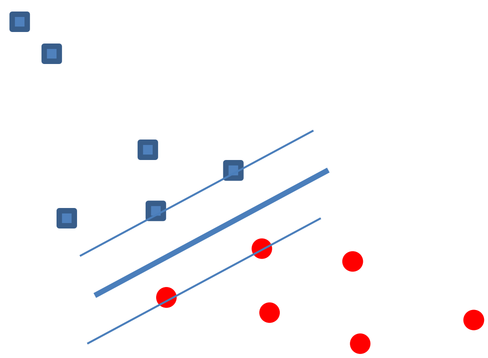


Fig. 1

(a) [1 pt] Draw (approximately) the SVM line separator.



(b) [1 pt] Suppose we find (1/2)***w**^{2} to be 2 in the SVM optimization. What is the margin, i.e. the distance of closest points to the line?

||**w**||=2
Margin = 1/||**w**|| = …

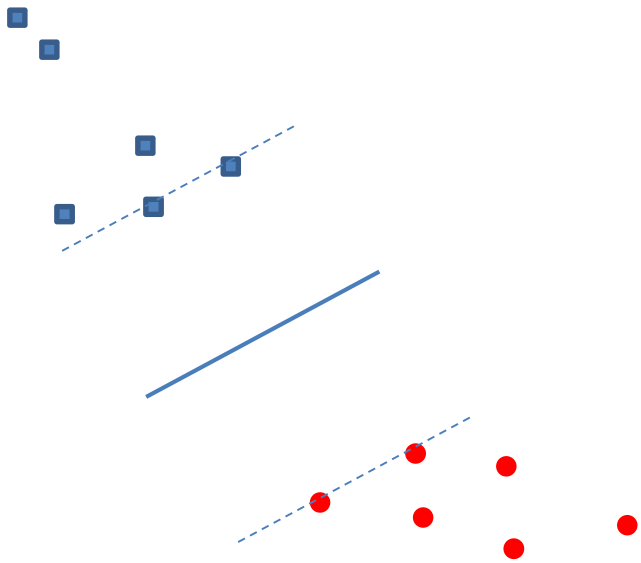


Fig. 2

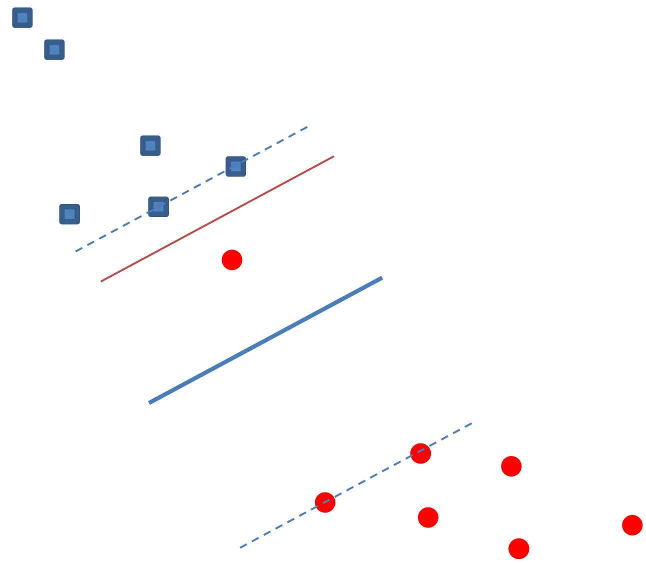


Fig. 3

(c) [1 pt] Now consider the dataset in Fig 2 (the red points are shifted below). Will (1/2)*$\mathbf{w}^2$ be smaller or greater than previously? Explain.

Margin will be greater, so (1/2)*$\mathbf{w}^2$ be ….

(d) [2 pt] Using a ruler, and the fact that (1/2)*$\mathbf{w}^2$ was 2 previously, find (approximately) the magnitude of the new line coefficient vector, **w**'.

See lines in Fig 2.

$1/\|\mathbf{w}'\| = 4*(1/\|\mathbf{w}\|)$, so $\|\mathbf{w}'\|$ = …

(e) [3 pt] Consider the dataset in Fig 3 (with one additional red circle quite close to the blue squares). Assuming optimization using slack variables and C=1, draw a line that does not perfectly separate the points, but which is nonetheless better than the line that perfectly separates the points. (Draw it in the figure, and explain why).

See (bold) blue line in Fig. 3.
The red line that perfectly separates the points will have a higher cost than the blue line.
To see this, observe that ξ for the new red circle will be equal to 1.5 (or slightly less).
[Why? Explain using the discussion in *svm.pdf*]

The cost of the blue line is: (1/2)*$\mathbf{w}'^2+\xi$ (The other ξ's will be zero. Why?)
The cost of the red line is: (1/2)*$\mathbf{w}^2$ (All the ξ's will be zero. Why?)

Now let's plug in the values for ||**w**'||, ||**w**||, and ξ:
TO DO

What conclusion can be derived?

(f) [1 pt] Why would we rather prefer the line in (e) to the line that perfectly separates the points?

TO DO