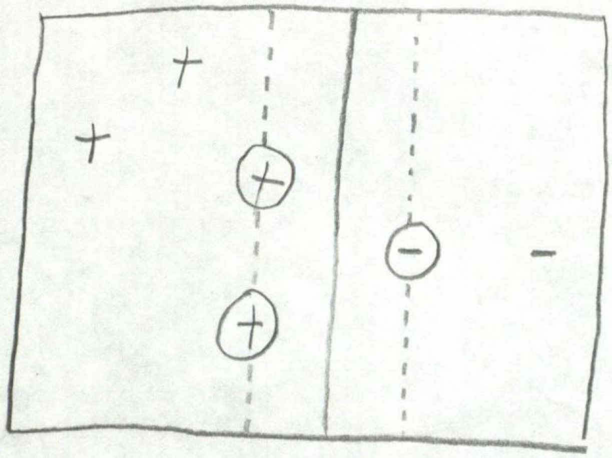
Support Vector Machine terms (SVM)



{+, -} - classes

------ - gutter

——— - SVM boundary, decision boundary, classification function, decision line

+, -, ⊖, ⊕ - Training points

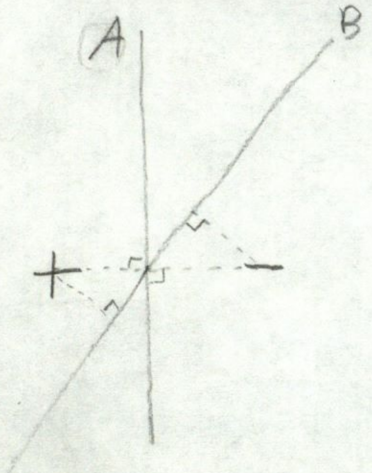
⊖ ⊕ - Support Vectors

Robert ⊙ McIntyre

6.034 ~~N~~ Quiz 3 Notes

SVMs

The point of SVMs is to find a single decision line that separates the training points while maximizing the distance from the decision line to the nearest training point.

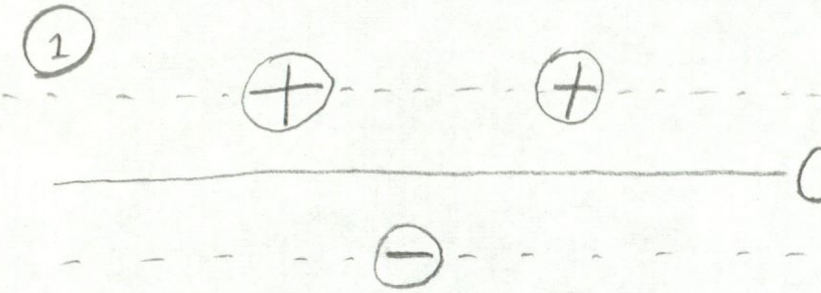


⇒ Both A and B divide the samples, but B has a smaller distance to the training points than A.

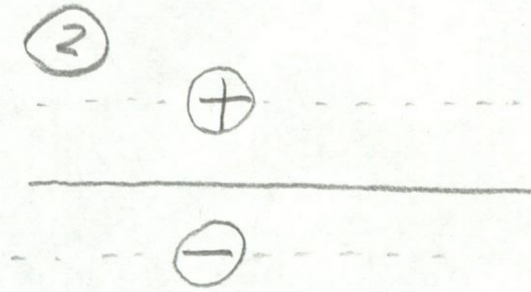
A Support Vector is a training point that, if it was removed, would change the decision line.

In 2D, there are only ever 2 or 3 Support Vectors.

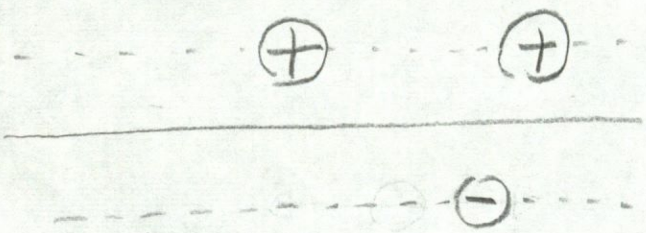
In 2D, there are only 2 different sorts of decision lines you can draw:



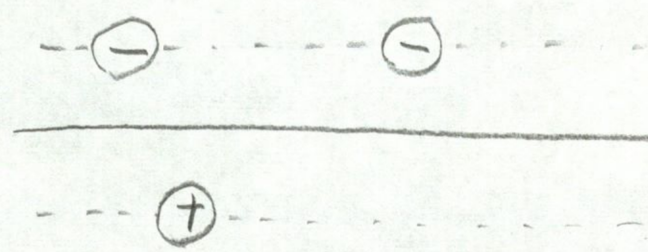
OR



These can vary by switching the classes ("+"→"−" and "−"→"+"), or by moving the "−" point in example 1, as long as it stays "between" the two "+" points.



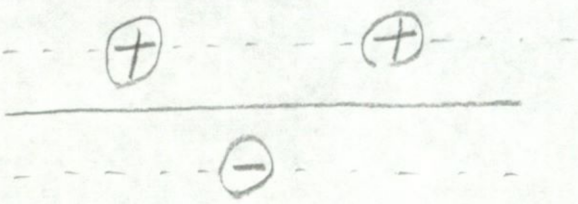
,



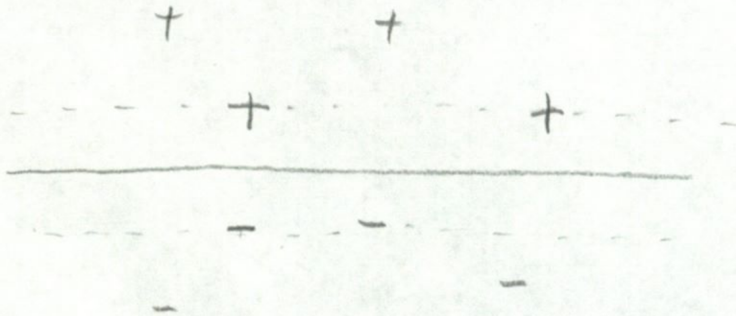
are both variants of 1.

You can also add <u>irrelevant points</u>, that is, points that are no closer to the decision line than the support vectors. Sometimes irrelevant points can make the support vectors <u>ambiguous</u>.

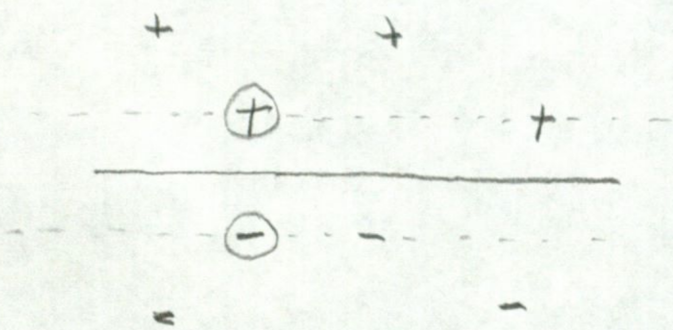
For Example, start with:



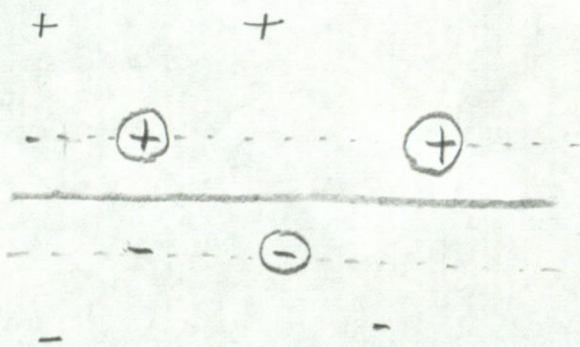
then add irrelevant points.



Now the support vectors are either:



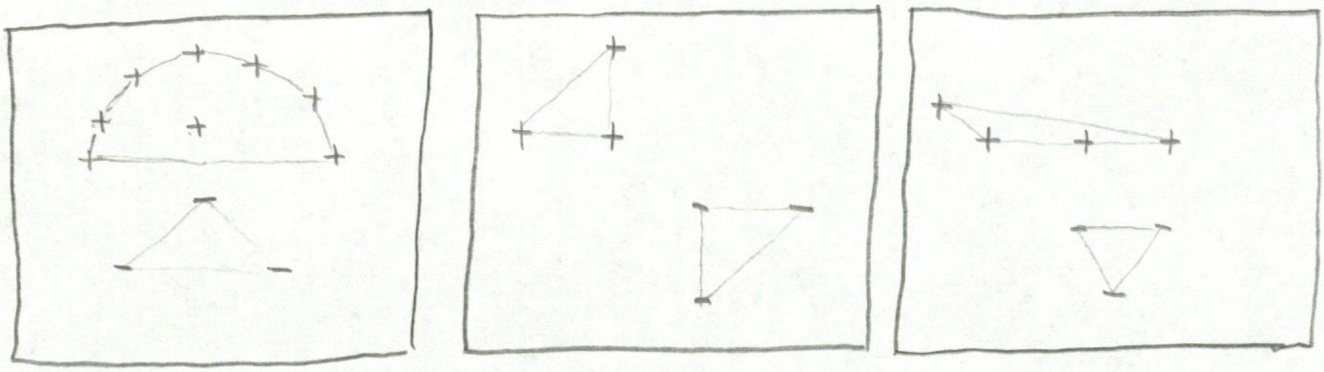
OR



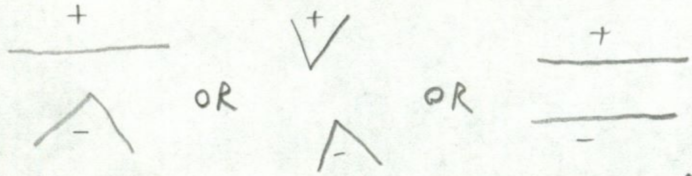
How to draw SVM boundaries* in 2D

* works every time

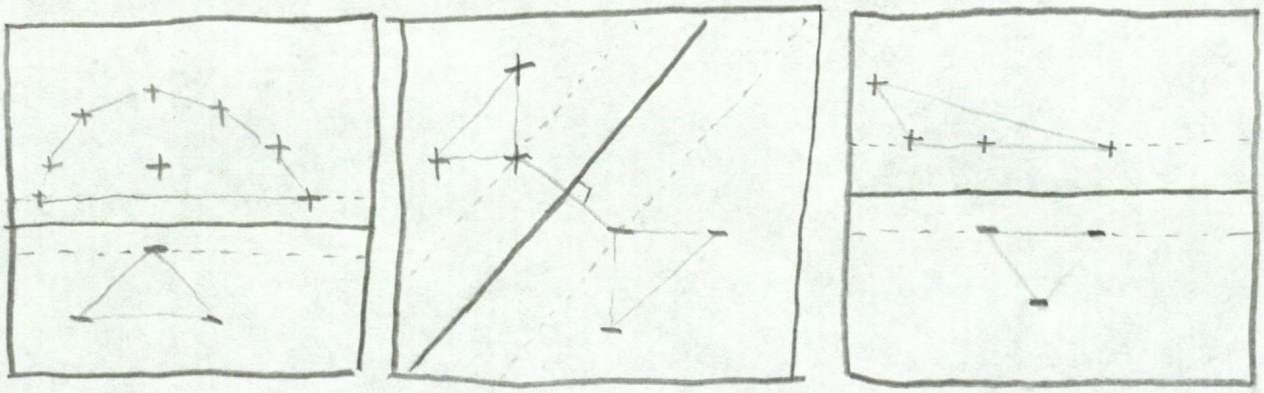
1. Draw the Convex Hulls for the "+" and "−" training points. (A convex hull is the shape you get when you wrap a rubber band around the points and let it contract.)



2. look at the region where the convex hulls are closest. There are three cases:



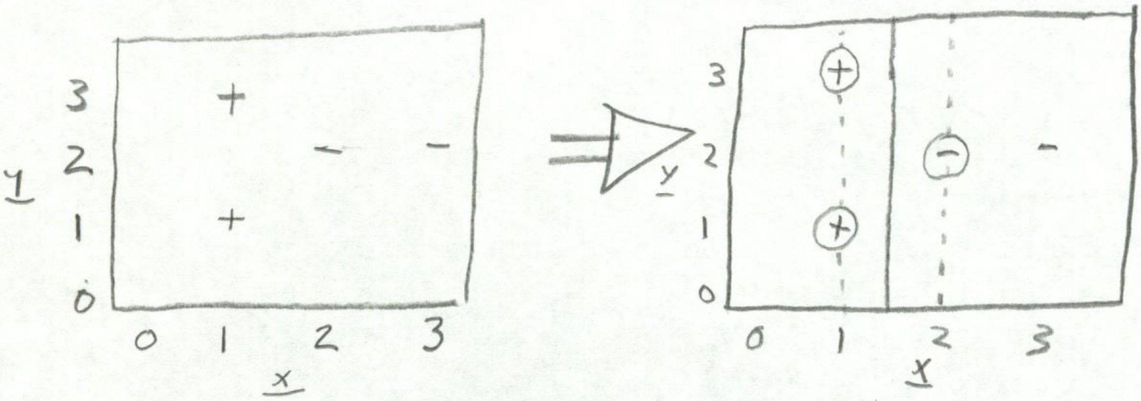
The SVM boundaries in each case look like this:



Finding $\vec{w}$, b, and the α values.

This is the easiest part of a SVM problem!

1. Find the boundary line.



2. Write the equation for the boundary line.

$x = 1.5$

3. Write that equation in the form $\vec{w}\begin{bmatrix} x \\ y \end{bmatrix} + b = 0$

$$x = 1.5 \Rightarrow \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\vec{w}} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\vec{x}} + \underbrace{-1.5}_{b} = 0$$

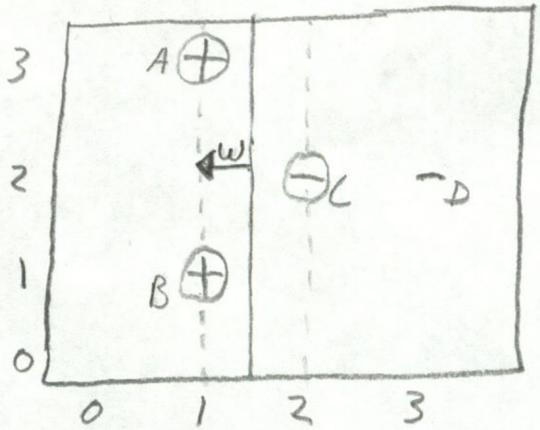
4. The decision line equation must return $\boxed{1}$ for "+" support vectors and $\boxed{-1}$ for "−" support vectors. Multiply the decision line eqn. so this is so.

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} - 1.5 = \frac{1}{2}$, want $\frac{1}{2} \rightarrow -1$, so multiply by -2.

new eqn. is : $\begin{bmatrix} -2 \\ 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + 3 = 0$; $\vec{w} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$; $b = 3$.

Note that if you draw $\vec{w}$ on the decision line, then it should be $\perp$ to the decision line, and point toward the $\oplus$ training points. In our example:

$w = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$



Now that you have $\vec{w}$, find the α values for each training point using the following rules:

① For non-support vectors, $\alpha = 0$

② $\sum \alpha$ values for "+" points $= \sum \alpha$ values for "−" points

③ $$\vec{w} = \sum_{\substack{\vec{P} \in \\ \text{"+" points}}} \alpha_P \vec{P} - \sum_{\substack{\vec{P} \in \\ \text{"−" points}}} \alpha_P \vec{P}$$

In this case,

$\alpha_D = 0$, $\alpha_A + \alpha_B = \alpha_C$, and $\begin{bmatrix} -2 \\ 0 \end{bmatrix} = \alpha_A \begin{bmatrix} 1 \\ 3 \end{bmatrix} + \alpha_B \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \alpha_C \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

so $\alpha_C = 2$, $\alpha_A = \alpha_B = 1$.