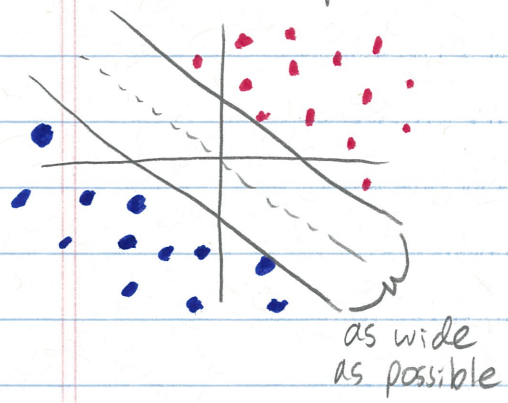
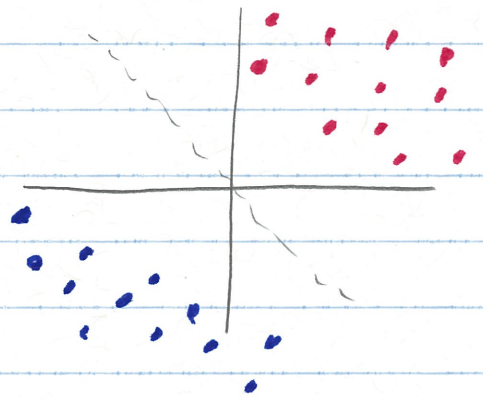
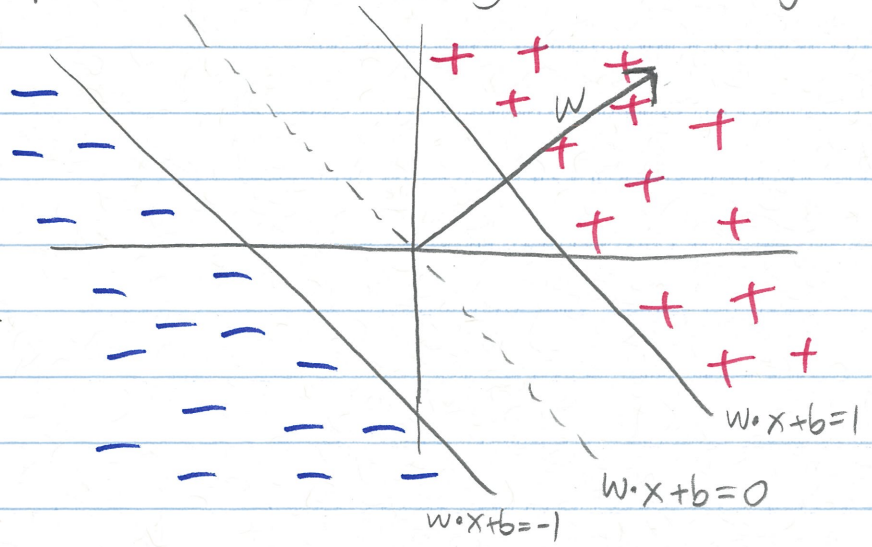
Support Vector Machines (SVMs)

Given two classes, how can we find the best linear separator?





Mathematically Suppose we have data $x_1, x_2, x_3, \ldots \in \mathbb{R}^n$ equipped with labels $y_i = 1$ or $y_i = -1$.



The width of this "road" is inversely proportional to $\|w\|$, the length of w.

We label a new data point as 1 if $w \cdot x + b > 0$
-1 if $w \cdot x + b < 0$.

Solve $\min_{w,b} \frac{1}{2}\|w\|^2$

subject to $y_i(w \cdot x_i + b) \geq 1$ for all i.

Allowing training errors

Called "Boxcount" in Matlab. I suggest $C=1$ over $C=\infty$, which allows no training misclassifications.

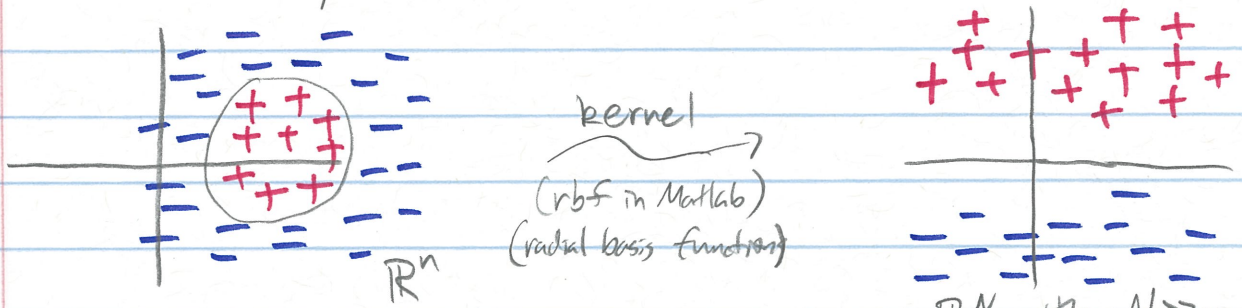
Solve $\min_{w,b,\xi_i \geq 0} \frac{1}{2}\|w\|^2 + C\sum_i \xi_i$

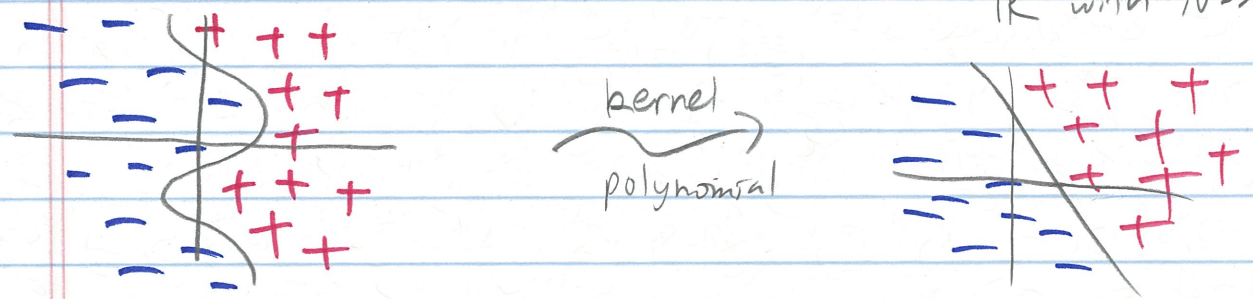
subject to $y_i(w \cdot x_i + b) \geq 1 - \xi_i$ for all i

(If $\xi_i > 1$, then x_i is on the wrong side of hyperplane!)

Nonlinear or kernel SVMs

Pick some $\phi: \mathbb{R}^n \to \mathbb{R}^N$ with $N \gg n$.





Typical kernels are Gaussians (rbf's) or polynomials

More than 2 classes? Possible but more complicated!