Recitation 2

Geometric Derivation of SVMs

Brett Bernstein

CDS at NYU

February 9, 2017

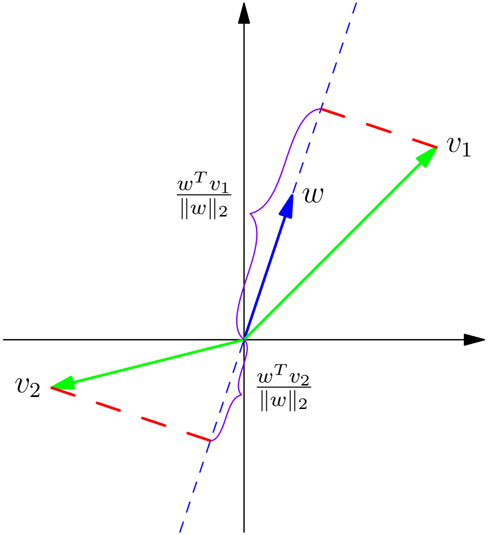
Intro Question

Question

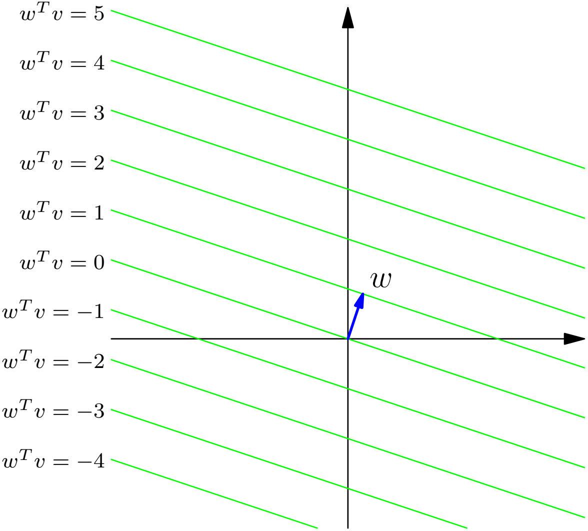
You have been given a data set (x_i, y_i) for $i = 1, \ldots, n$ where $x_i \in \mathbb{R}^d$ and $y_i \in \{-1, 1\}$. Assume $w \in \mathbb{R}^d$ and $a \in \mathbb{R}$.

1. Suppose $y_i(w^T x_i + a) > 0$ for all i. Use a picture to explain what this means when $d = 2$.
2. Fix $M > 0$. Suppose $y_i(w^T x_i + a) \geq M$ for all i. Use a picture to explain what this means when $d = 2$.

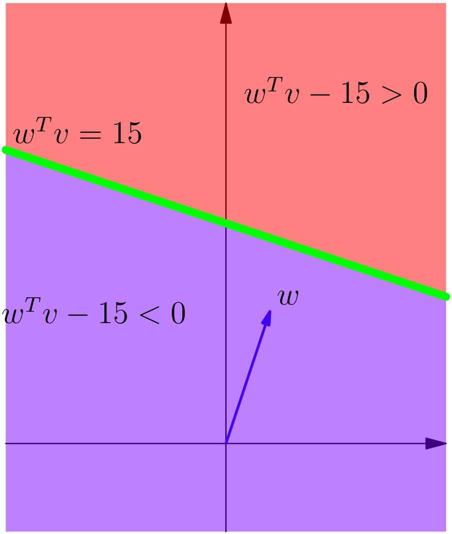
Component of v_1, v_2 in the direction w



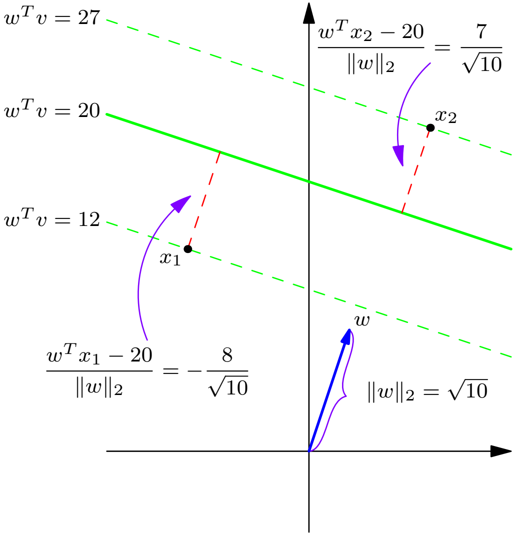
Level Surfaces of $f(v) = w^T v$ with $\|w\|_2 = 1$



Sides of the Hyperplane $w^T v = 15$



Signed Distance from x_1, x_2 to Hyperplane $w^T v = 20$

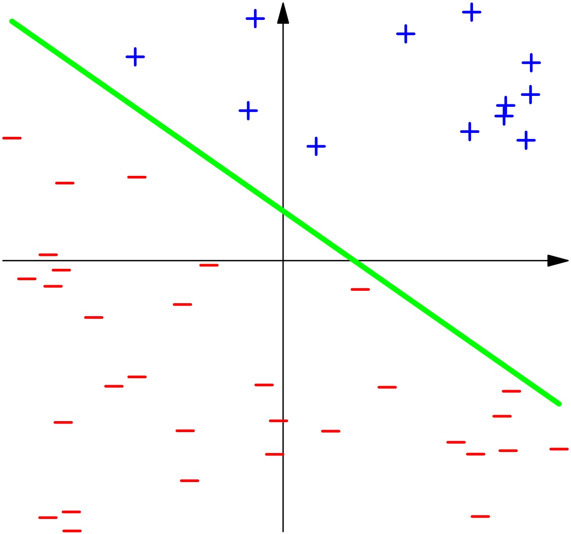


Linearly Separable

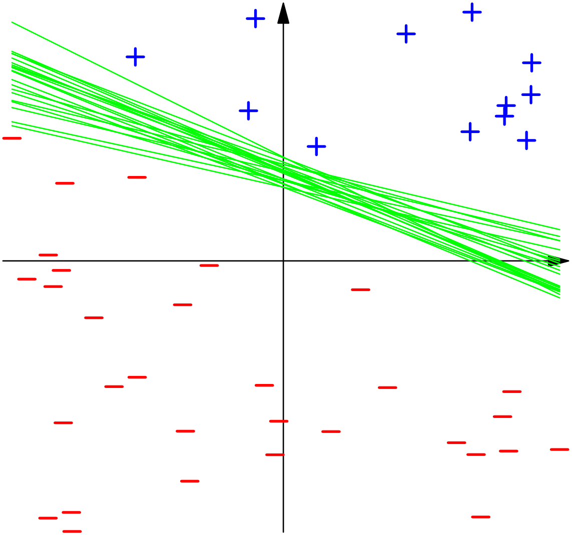
Definition

We say (x_i, y_i) for $i = 1, \ldots, n$ are *linearly separable* if there is a $w \in \mathbb{R}^d$ and $a \in \mathbb{R}$ such that $y_i(w^T x_i + a) > 0$ for all i. The set $\{v \in \mathbb{R}^d \mid w^T v + a = 0\}$ is called a *separating hyperplane*.

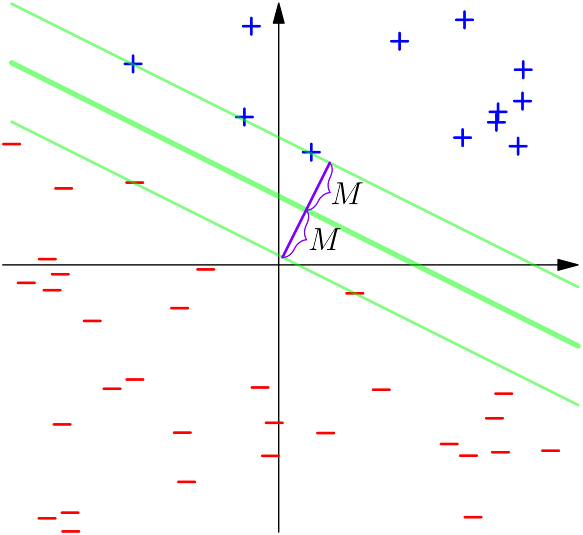
Linearly Separable Data



Many Separating Hyperplanes Exist



Maximum Margin Separating Hyperplane

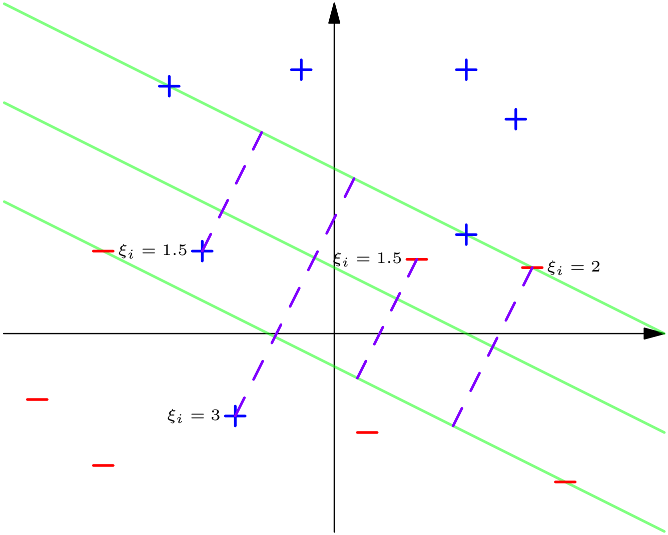


$$\frac{w^Tv+a}{\|w\|_2} = M$$

$$\frac{w^Tv+a}{\|w\|_2} = 0$$

$$\frac{w^Tv+a}{\|w\|_2} = -M$$

Soft Margin SVM (unlabeled points have $\xi_i = 0$)



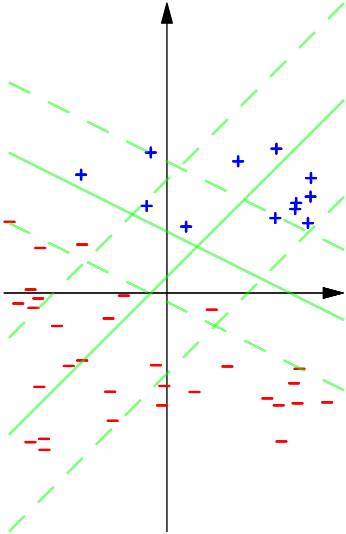
Questions

Questions

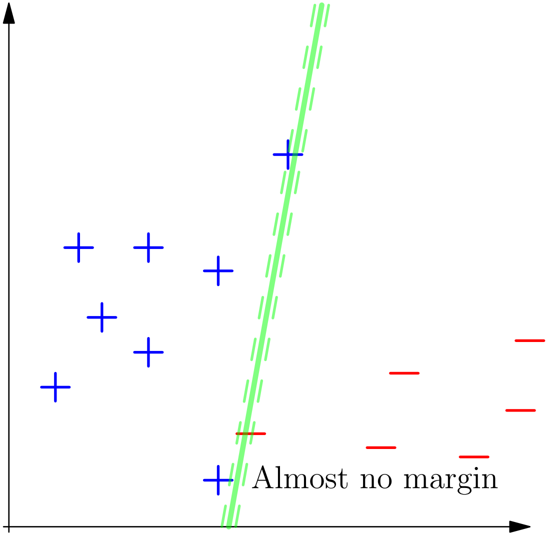
1. If your data is linearly separable, which SVM (hard margin or soft margin) would you use?
2. Explain geometrically what the following optimization problem computes:

$$\begin{array}{ll} \text{minimize}_{w,a,\xi} & \frac{1}{n}\sum_{i=1}^{n}\xi_i \\ \text{subject to} & y_i(w^T x_i + a) \geq 1 - \xi_i \quad \text{for all } i \\ & \|w\|_2^2 \leq r^2 \\ & \xi_i \geq 0 \quad \text{for all } i. \end{array}$$

Optimize Over Cases Where Margin Is At Least $1/r$



Overfitting: Tight Margin With No Misclassifications



Training Error But Large Margin