Recitation 2

Geometric Derivation of SVMs

Brett Bernstein

CDS at NYU

February 5, 2017

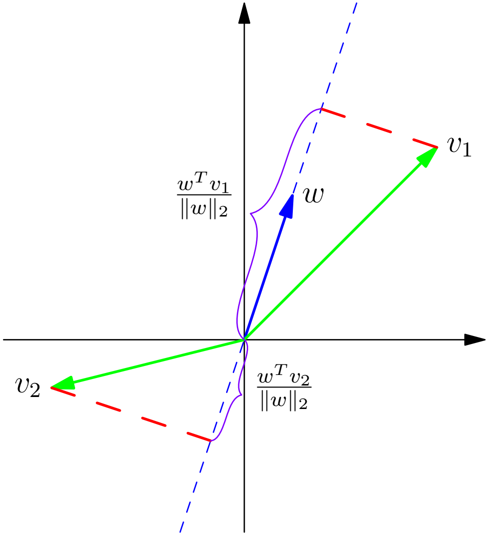
Intro Question

Question

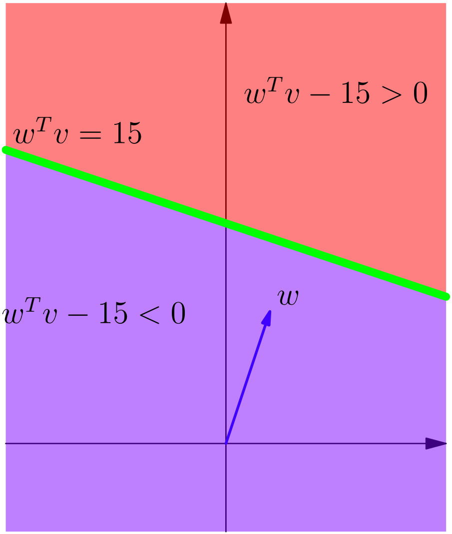
You have been given a data set (x_i, y_i) for $i = 1, \dots, n$ where $x_i \in \mathbb{R}^d$ and $y_i \in \{-1, 1\}$. Assume $w \in \mathbb{R}^n$ and $a \in \mathbb{R}$.

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 $d = 2$.

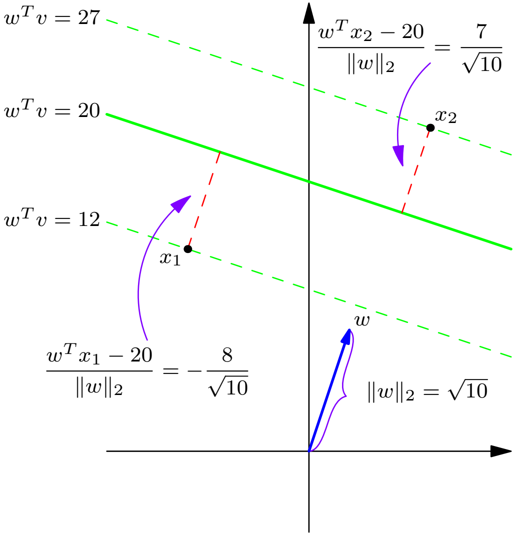
Component of v_1, v_2 in the direction w



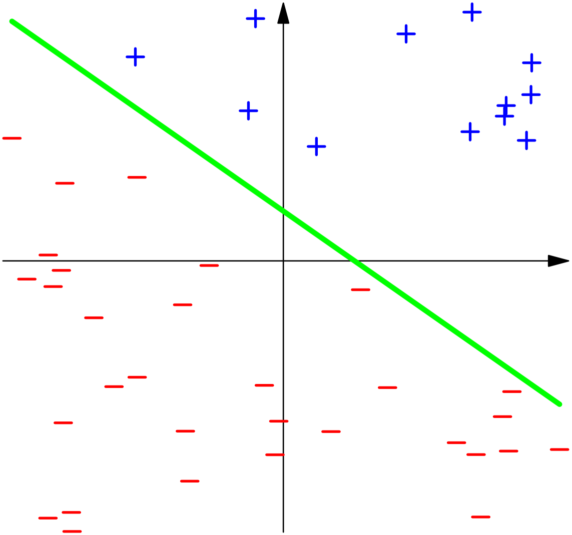
Sides of the Hyperplane $w^T v = 15$



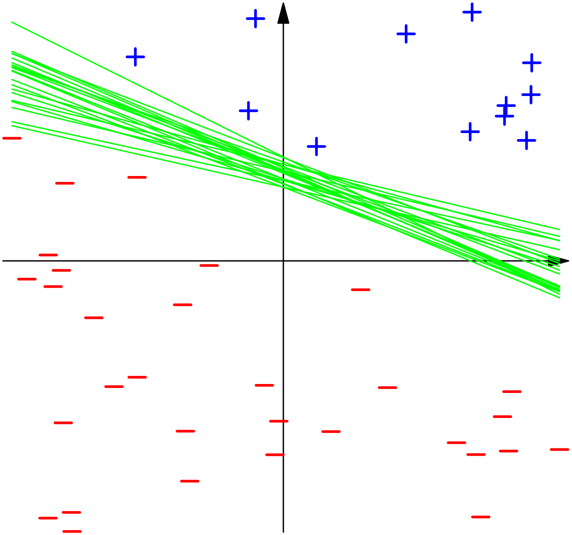
Signed Distance from x_1, x_2 to Hyperplane $w^T v = 20$



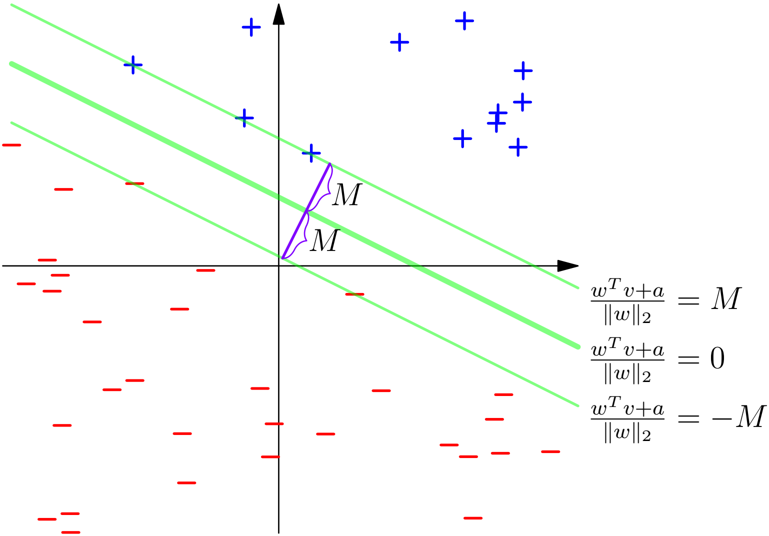
Linearly Separable Data



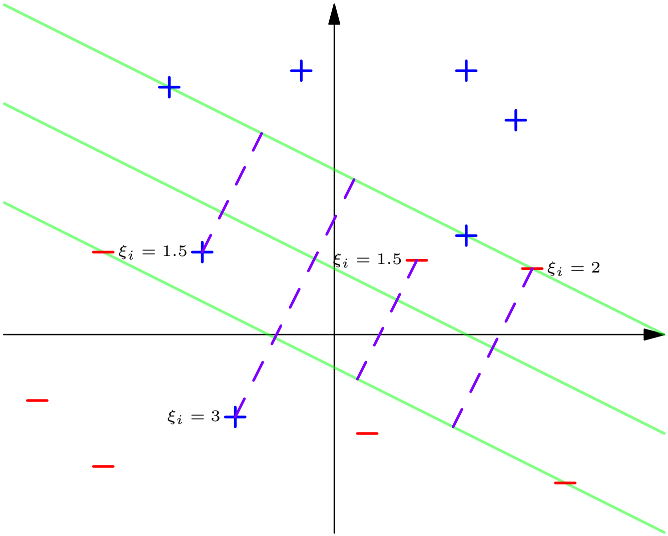
Many Separating Hyperplanes Exist



Maximum Margin Separating Hyperplane



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



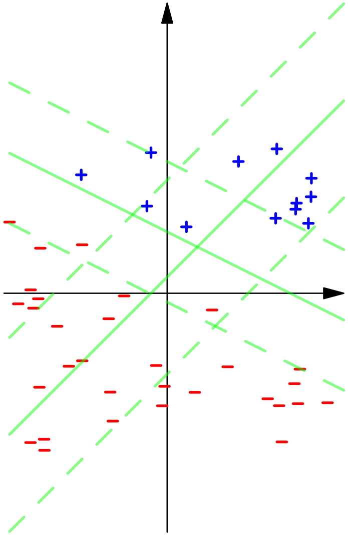
Questions

Questions

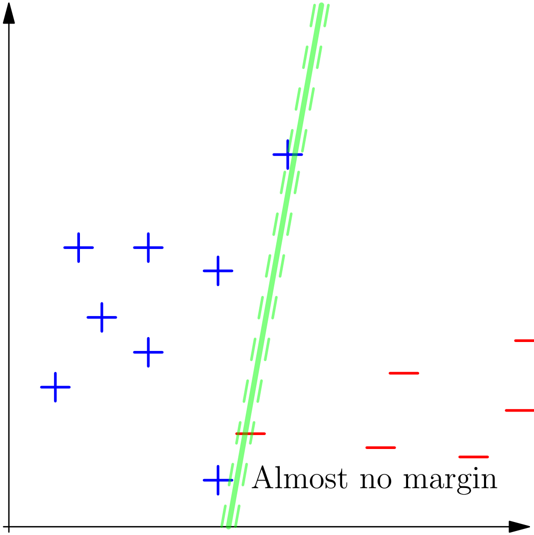
1. If your data is linearly separable, which SVM (hard margin or soft margin) would you use?
2. Explain geometrically what the following optimization problem computes:

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

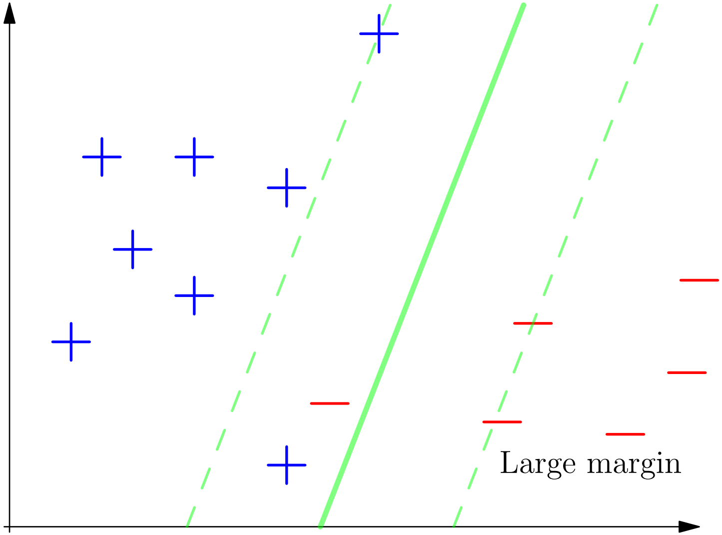
Optimize Over Cases Where Margin Is At Least $1/r$



Overfitting: Tight Margin With No Misclassifications



Some Training Errors But Large Margin



Shapes of Level Curves

Question

For each of the following functions, determine the shape of the given set.

1. $\{x \in \mathbb{R}^2 \mid x^T A x = 1\}$ where $A = \begin{pmatrix} 2 & 2 \\ 2 & 1 \end{pmatrix}$.
2. $\{x \in \mathbb{R}^2 \mid x^T A x = 1\}$ where $A = \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix}$.