Hard Margin SVM: Geometric Approach

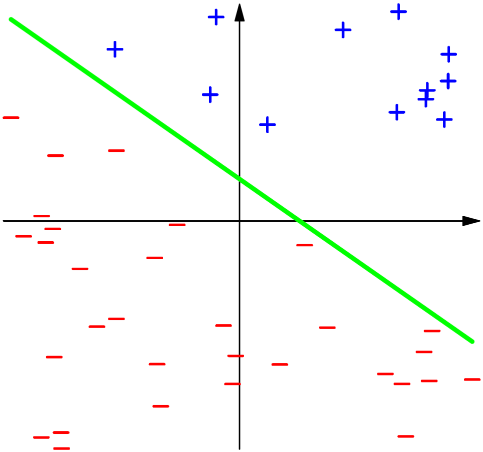
David S. Rosenberg

Bloomberg ML EDU

October 11, 2017

Introduction

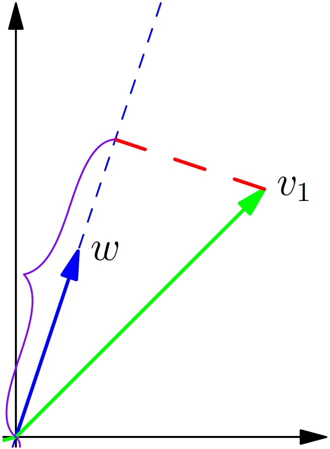
Linearly Separable Data



- The figure represents an input space $\mathcal{X} \in \mathbf{R}^2$.
- The output or **class label** is either $+$ or $-$.
- These data are linearly **separable** if there is a hyperplane (just a line in $\mathbf{R}^2$) that perfectly separates the two classes.
- How can we find such a line?
- What if there are multiple lines?
- If there is no such line, what should we do?

Scalar Projections onto Vectors

Projection of v_1 onto w

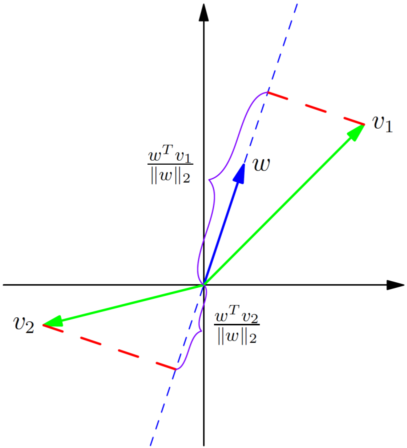


- Want to find **scalar projection** of v_1 onto w.
- Also known as the **component of v_1 in the direction w**.
- It's the length of the segment in the purple curly brace.
- The scalar projection of v_1 onto w is given by

$$\frac{w^T v_1}{\|w\|}.$$

- The scalar projection is a **number**.
- The corresponding **vector projection** is the vector

$$\left(\frac{w^T v_1}{\|w\|}\right)\frac{w}{\|w\|}.$$



- The scalar projection is a **signed length**.
- The component of v_2 in the direction w is **negative.**
- The **vector projection** of v_1 onto w is is the vector

$$\left(\frac{w^T v_1}{\|w\|}\right) \frac{w}{\|w\|}.$$

Interpreting Hyperplanes by Scalar Projections

- You may recall from linear algebra that the set

$$S = \{x \in \mathbf{R}^d \mid w^T x = b\}$$

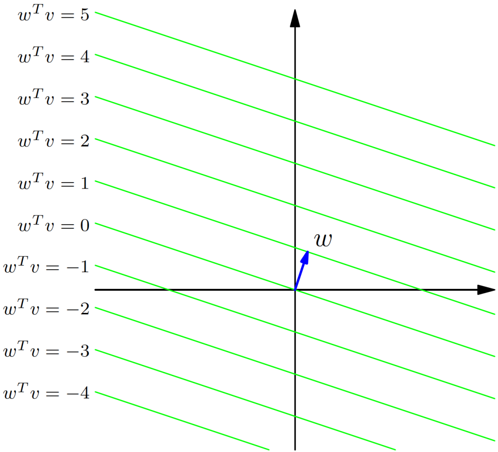
  is a **hyperplane** in $\mathbf{R}^d$, for $w \neq 0$.

- Note that $w^T x = b$ is equivalent to

$$\frac{w^T x}{\|w\|} = \frac{b}{\|w\|}.$$

- So S is set of all x that have the component $b/\|w\|$ in direction w.

Interpreting Hyperplanes by Scalar Projection

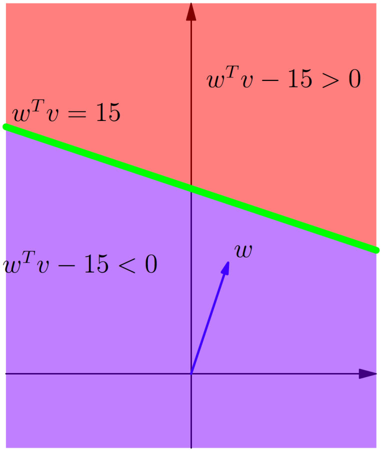


- Take w to be a unit vector.
- Here we have 10 parallel lines in $\mathbf{R}^2$,
  - each with a different component in direction w.
- Each line is a level set of the function $f(v) = w^Tv$.
- What do we get if we consider the points

$$S^- = \{v \mid f(v) < -2\}?$$

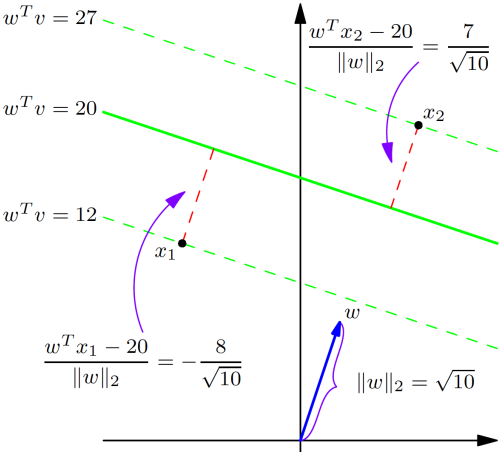
Separating Data with Hyperplanes

Sides of a Hyperplane



- The hyperplane $\{v \mid w^T v = 15\}$ separates the space into 3 parts depending on value of $w^T v - 15$:
  - $w^T v - 15 = 0$ (on the hyperplane)
  - $w^T v - 15 > 0$ (on side w points in)
  - $w^T v - 15 < 0$ (on side $-w$ points in)

Distance from Point to Hyperplanes



- Distance from x_2 to $\{v \mid w^T v = 20\}$.
- Let v be any pointy in $\{v \mid w^T v = 20\}$.
- Distance is difference in components in direction w:

$$\frac{w^T x_2}{\|w\|} - \frac{w^T v}{\|w\|}.$$

- Well almost – this is **signed distance**.
- Positive if x_2 is on the side pointed to by w.
- Negative if x_2 is on the side pointed to by $-w$.