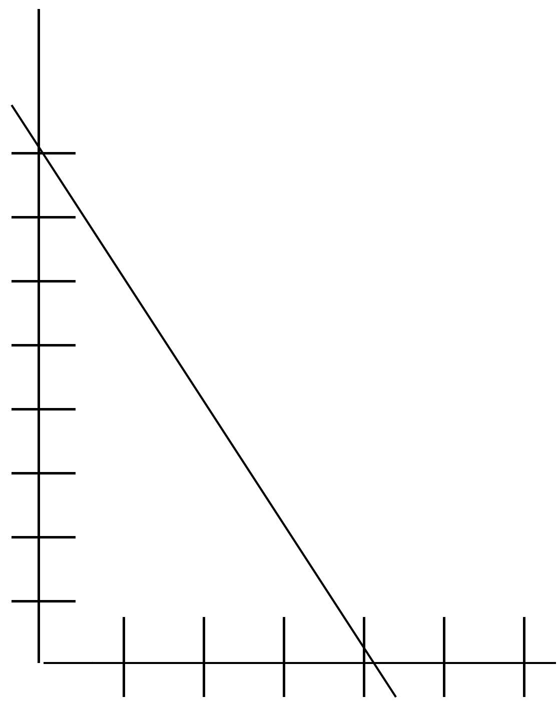
Linear hyperplanes as classifiers

Usman Roshan

Hyperplane separators

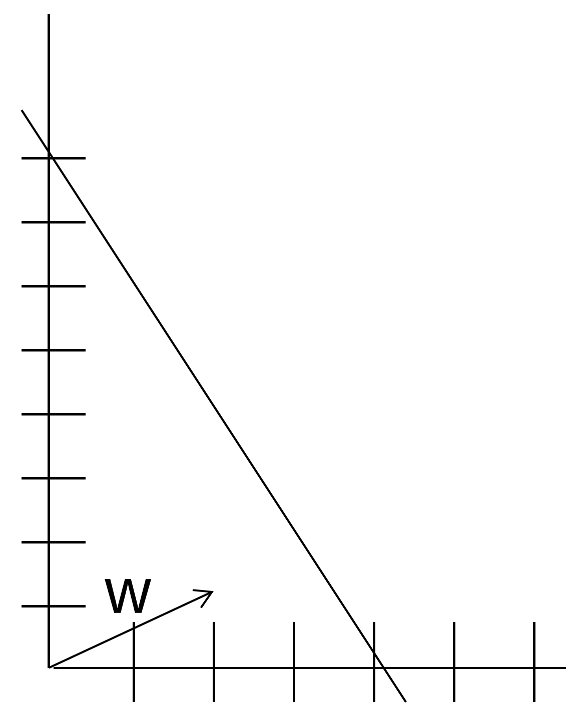
- Consider equation of line shown on right
$$x_1 = -2x_0 + 8$$
- If we let $w = (2,1)$ and $w_0 = -8$ then we can rewrite it as
$$w^T x + w_0 = 0$$



Hyperplane separators

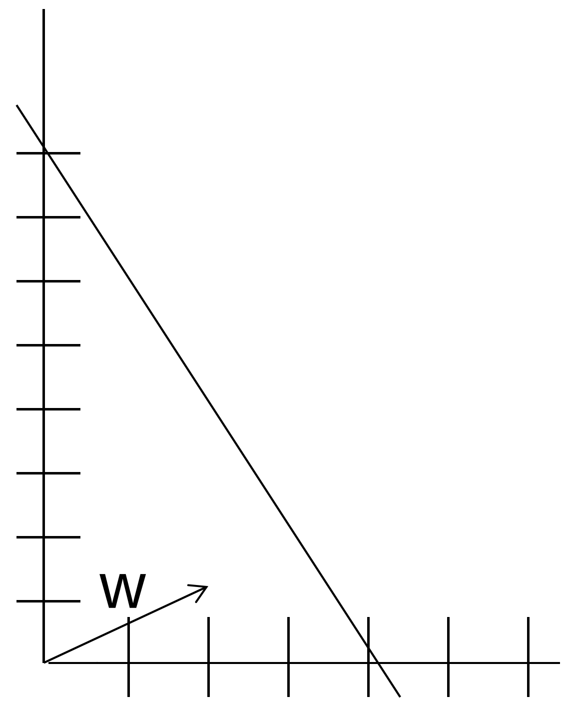
- Shown in the figure is the vector w. Note that w is perpendicular to the plane.
- By definition all points on the line itself satisfy the equation of the line

$$w^T x + w_0 = 0$$



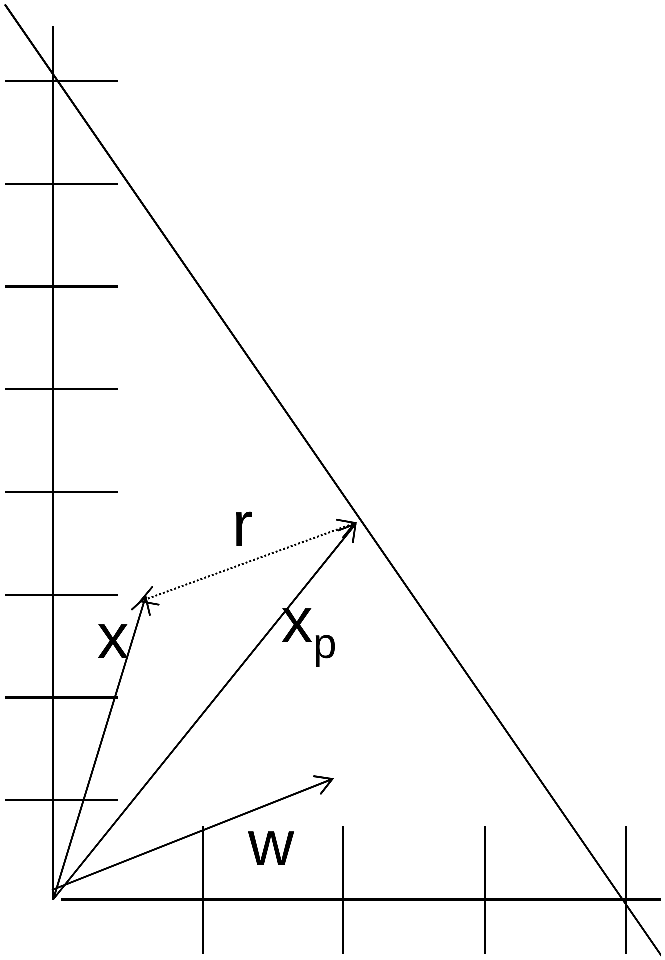
Hyperplane separators

- Let $f(x) = w^T x + w_0$
- For points
  - on the plane $f(x) = 0$
  - to the left of the plane $f(x) < 0$
  - to the right of the plane $f(x) > 0$.
- For example consider $f(x)$ for the points $x = (2,2)$ $x = (4,5)$



Hyperplane separators

- To see this consider a point x and its perpendicular drop on the plane called x_p.
- We can write $x = x_p + r\frac{w}{\|w\|}$ where r < 0
- Applying w^T on both sides
$$w^T x = w^T x_p + r\|w\|$$



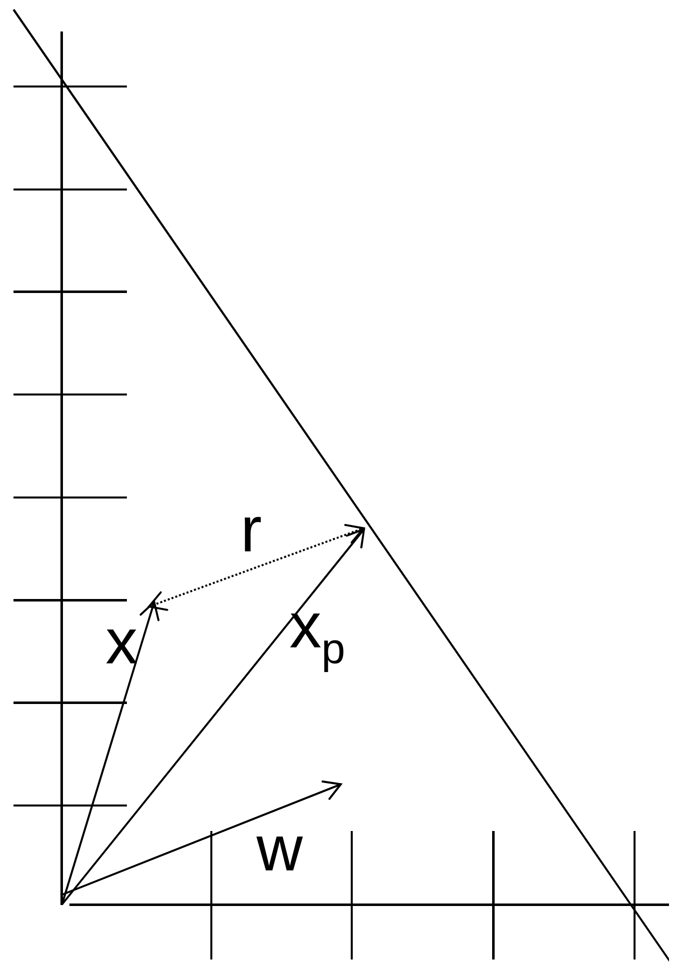
Hyperplane separators

- Since x_p is on the plane and $w^T x_p + w_0 = 0$ we get

$$w^T x = -w_0 + r\|w\|$$

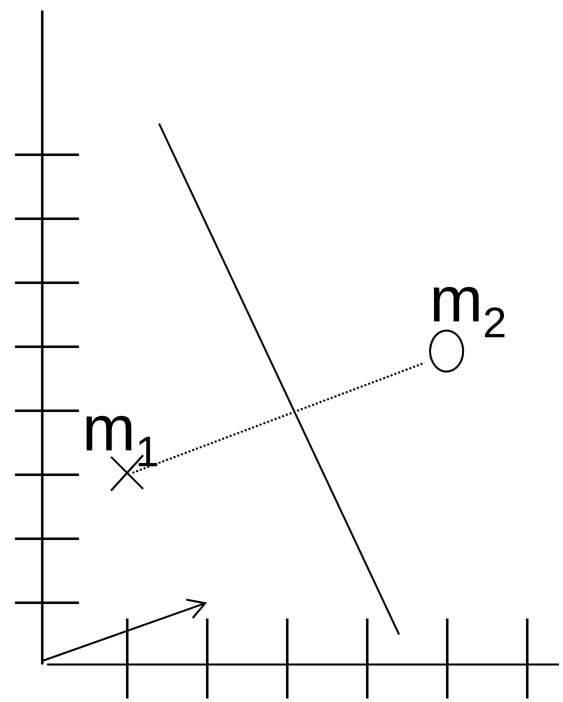
$$\frac{w^T x + w_0}{\|w\|} = r$$

- And so since r < 0 we have $w^T x + w_0 < 0$
- Similarly $w^T x + w_0 > 0$ for points on the other side of the



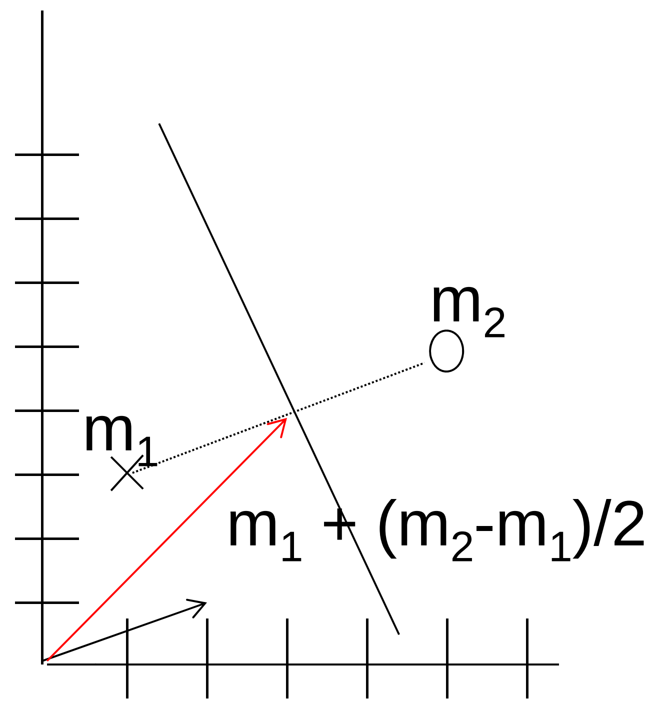
Nearest mean as hyperplane separator

- Now let us cast nearest means as separating hyperplane problem
- The vector $w = m_2 - m_1$ determines the hyperplane direction



Nearest mean as hyperplane separator

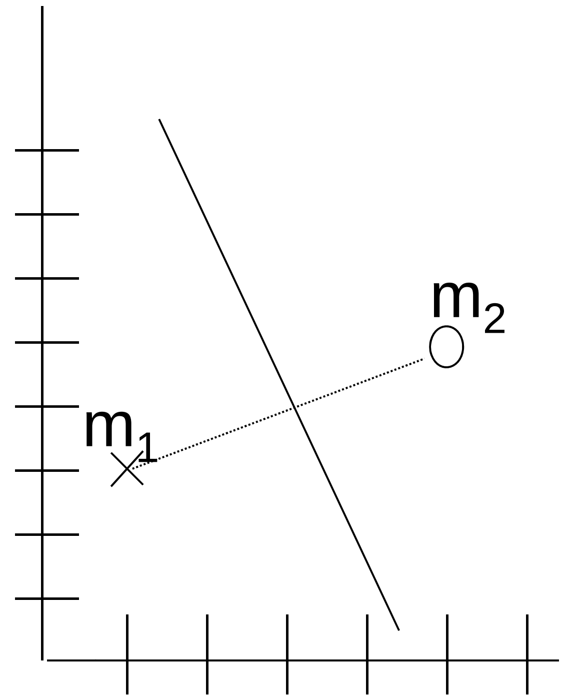
- To determine offset consider the vector $m_1 + \frac{m_2 - m_1}{2}$ that lies on the plane.
- Since it lies on the plane and $w = m_2 - m_1$ we have $(m_2 - m_1)^T \left(m_1 + \frac{m_2 - m_1}{2}\right) + \mathrm{w}_0 = 0$
- Solving for w_0 gives us $w_0 = \frac{1}{2}(\|m_1\|^2 - \|m_2\|^2)$ thus completing the hyperplane parameters



Nearest mean as hyperplane separator

- To classify a point x with nearest means classifier we check the sign of

$$(m_2 - m_1)^T x + \frac{1}{2}(\|m_1\|^2 - \|m_2\|^2)$$



Separating hyperplanes

- Can we write Naïve-Bayes separating hyperplane?
- Yes: $w = \Sigma^{-1}(m_1 - m_2)$

$$w_0 = -\frac{1}{2}(m_1 + m_2)\Sigma^{-1}(m_1 - m_2)$$

- How about finding the separating hyperplane that minimizes?

$$Error = (y_i - w^T x_i)^2$$

for a fixed *i*