Notation and Math

New Notation

Previously we used

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix} \qquad x^{(i)} = \begin{bmatrix} 1 \\ x_1^{(i)} \\ \vdots \\ x_d^{(i)} \end{bmatrix}$$

$$y \in \{0,1\}$$

For SVM, we separate the intercept term from the other weights. The mathematics of this lecture makes easier. We change the notation to make this clearer.

$$w_0 \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{bmatrix} \qquad x^{(i)} = \begin{bmatrix} x_1^{(i)} \\ x_2^{(i)} \\ \vdots \\ x_d^{(i)} \end{bmatrix}$$

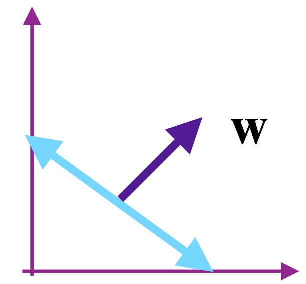
$$y \in \{-1,1\}$$

$$g(\mathbf{x}) = \text{sign}(\mathbf{w}^T\mathbf{x} + w_0) = \begin{cases} +1 & \text{if } \mathbf{w}^T\mathbf{x} + w_0 > 0 \\ -1 & \text{if } \mathbf{w}^T\mathbf{x} + w_0 < 0 \end{cases}$$

Hyperplane

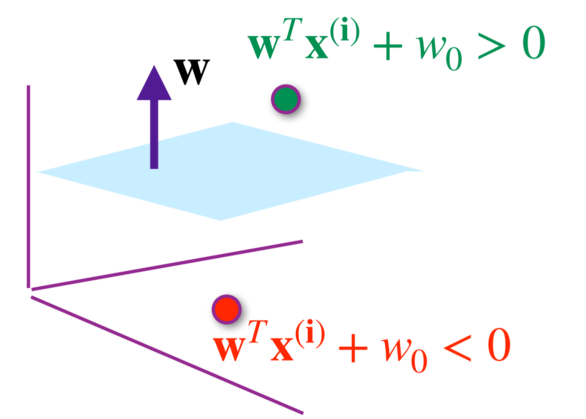
In p-dimensions, a hyperplane is a flat subspace of dimension p-1

The mathematical definition of a hyperplane: $\forall \boldsymbol{x}^{(i)},\ w_0 + \mathbf{w}^T \boldsymbol{x}^{(i)} = 0$



In 2-dimensions, the hyperplane is a line

In 2-dimensions, the hyperplane is defined by $\forall \boldsymbol{x}^{(i)},\ w_0 + w_1 x_1^{(i)} + w_2 x_2^{(i)} = 0$



In 3-dimensions, the hyperplane is a a plane

In 3-dimensions, the hyperplane is defined by $\forall \boldsymbol{x}^{(i)},\ w_0 + w_1 x_1^{(i)} + w_2 x_2^{(i)} + w_3 x_3^{(i)} = 0$

[Proof that w is normal to the plane: https://www.khanacademy.org/math/linear-algebra/vectors-and-spaces/dot-cross-products/v/normal-vector-from-plane-equation](https://www.khanacademy.org/math/linear-algebra/vectors-and-spaces/dot-cross-products/v/normal-vector-from-plane-equation)

Another way to describe a point

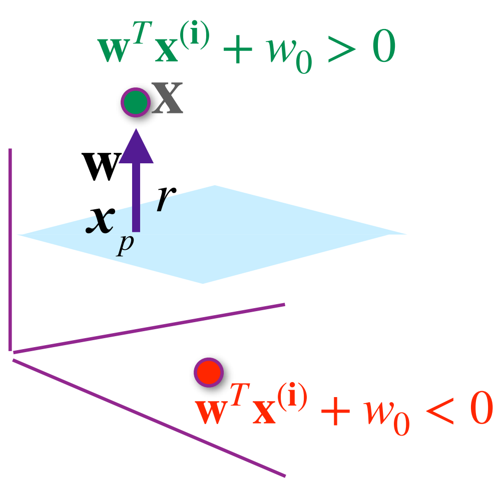
$\mathbf{w} = [w_1, w_2, \ldots, w_d]^T$

$\boldsymbol{x} = \boldsymbol{x}_p + r\dfrac{\mathbf{w}}{\|\mathbf{w}\|_2}$

$\|\mathbf{w}\|_2 = \sqrt{w_1^2 + w_2^2 + \cdots + w_d^2}$ is the length of the vector $\mathbf{w}$

$\mathbf{w}/\|\mathbf{w}\|_2$ converts $\mathbf{w}$ into a unit vector. E.g. $(3,4)^T/5 = (3/5, 4/5)$

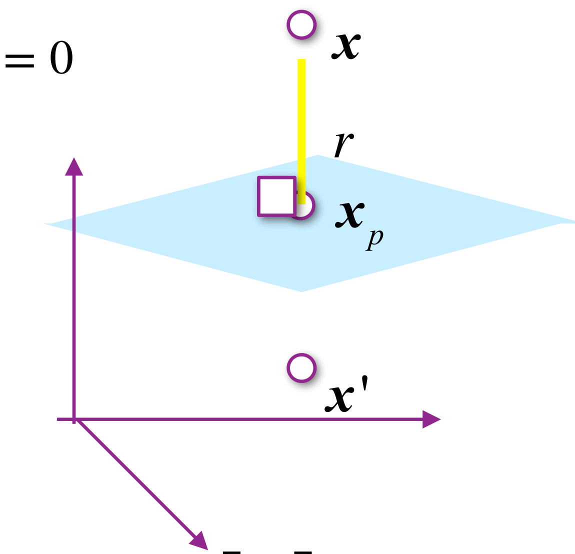
$$\forall \boldsymbol{x}^{(i)},\ w_0 + w_1 x_1^{(i)} + w_2 x_2^{(i)} + w_3 x_3^{(i)} = 0$$



Proof that w is normal to the plane: https://www.khanacademy.org/math/linear-algebra/vectors-and-spaces/dot-cross-products/v/normal-vector-from-plane-equation

Computing the signed distance from a point to the hyperplane

For any $\mathbf{x}$, $\mathbf{w}^T\mathbf{x} + w_0 = 0$



$\mathbf{x}$, a point

x_p, the normal projection of $\mathbf{x}$ onto $\mathbf{w}$,

Note that $\quad \boldsymbol{x} = \boldsymbol{x}_p + r\dfrac{\mathbf{w}}{\|\mathbf{w}\|_2}$

$$z(\boldsymbol{x}) = w_0 + \mathbf{w}^T\boldsymbol{x} = w_0 + \mathbf{w}^T\left(\boldsymbol{x}_p + r\frac{\mathbf{w}}{\|\mathbf{w}\|_2}\right)$$

$$= \underline{w_0 + \mathbf{w}^T\boldsymbol{x}_p} + \mathbf{w}^T r\frac{\mathbf{w}}{\|\mathbf{w}\|_2}$$

observe that $\mathbf{w}^T\mathbf{w} = \|\mathbf{w}\|_2^2 = [w_1 \quad w_2 \quad \cdots \quad w_d]\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{bmatrix} = w_1^2 + w_2^2 + \cdots + w_d^2$

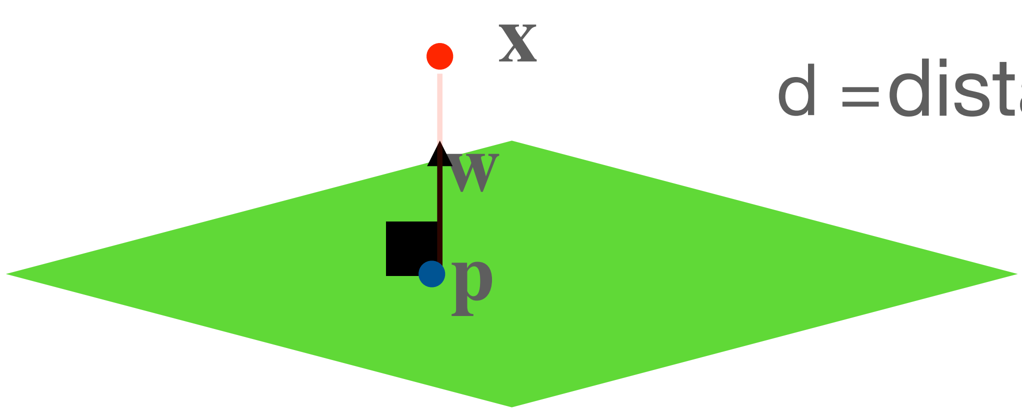
$$= r\|\mathbf{w}\|_2$$

Consequently: $r = \dfrac{z(\boldsymbol{x})}{\|\mathbf{w}\|_2}$

The unsigned distance $\dfrac{|\mathbf{w}^T\mathbf{x} + w_0|}{\|\mathbf{w}\|_2}$ of $\mathbf{x}$ to the hyperplane $\mathbf{w}^T\mathbf{x} + w_0 = 0$

Signed distance

Signed distance of point to hyperplane

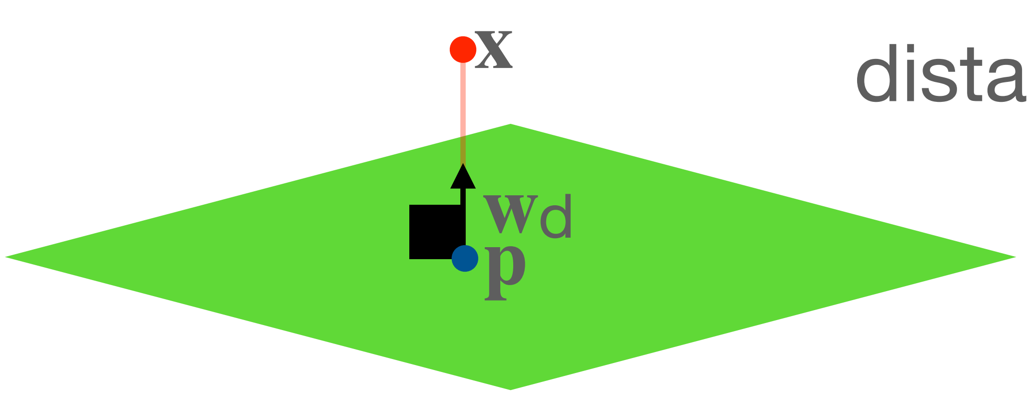


d = distance $= (\mathbf{w}^T\mathbf{x} + w_0)/\|\mathbf{w}\|$

$\mathbf{w}$ as a unit vector

$$\mathbf{w}' = \mathbf{w}/\|\mathbf{w}\| \qquad w_0' = w_0/\|\mathbf{w}\|$$

Signed distance of point to hyperplane



distance $= \mathbf{w}'^T\mathbf{x} + w_0'$