Linear Models [Discriminant Functions]

Recall from Bayes Rule that we choose the class w/ the largest posterior probability

$$\omega^* = \arg\max_{\omega \in \Omega} \boxed{p(x|\omega)P(\omega)} \rightarrow \propto \underbrace{P(\omega|x)}_{\text{posterior}}$$

Two ways of modelling

- **Generative** $\rightarrow p(x|\omega)p(\omega)$ [Difficult to do for complex data]
- **Discriminative** $\rightarrow p(\omega|x)$

$y \in \{\pm 1\}$

$$\boxed{g(x) = w^T x + b}$$

$b \in \mathbb{R}$

$\boxed{x \in \mathbb{R}^d}$

$w \in \mathbb{R}^d$

$y = \text{sign}(g)$

$\boxed{\text{Find } w \,\&\, b}$

The magnitude of w_j is going to stress the importance of component j

Quadratic

$$g(x) = x^T W x + w^T x + b$$

W is a matrix

$X \rightarrow RV$

Linear Regression

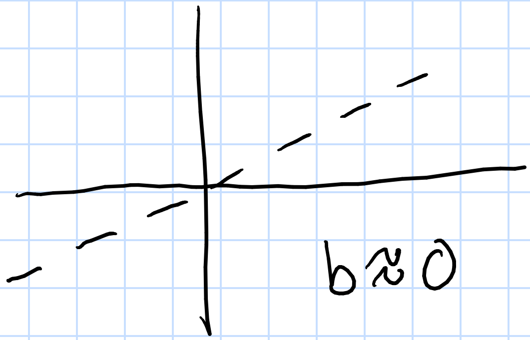
$x \in \mathbb{R}^d$

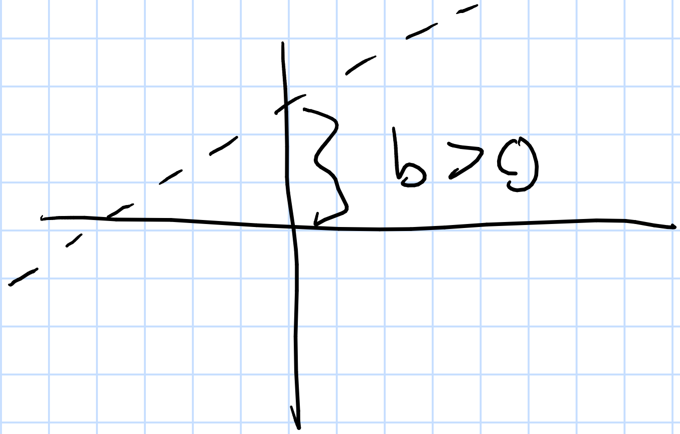
$$\boxed{\hat{y} = w_0 + \sum_{j=1}^{d} w_j x_j = w^T x} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix}^T \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_d \end{bmatrix} \in \mathbb{R}^{d+1}$$

* note of caution

$\|w\|_2^2$

Simple linear model. How do I find the parameters





Least Squares

n → # of samples in my training data

$$L(w) = \sum_{i=1}^{n} (\underbrace{y_i}_{\text{ground truth}} - \underbrace{w^T x_i}_{\text{prediction on } x_i})^2$$

$$= (\vec{y} - \vec{X}\vec{w})^T (\vec{y} - \vec{X}\vec{w})$$

matrix notation

$n \times 1 \quad (n \times d)(d \times 1)$

$\vec{y} \in \mathbb{R}^n$

$\vec{X} \in \mathbb{R}^{n \times d}$

$w \in \mathbb{R}$