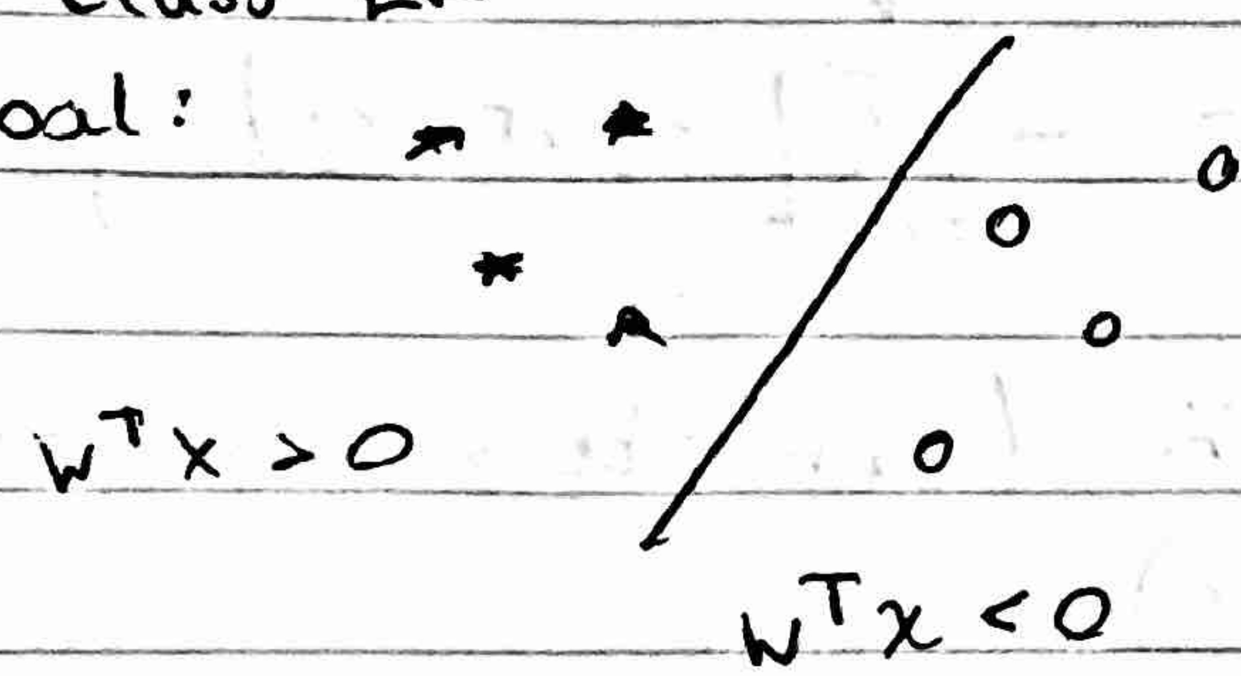
Discussion 6 3/8/19

Logistic Regression

- Logistic regression can be used for 2-class or multiclass problems
- Unlike linear regression, logistic regression tries to find category of belonging w/ continuous outcome
- Can't assume normally distributed as in linear regression

2 class LR

Goal:



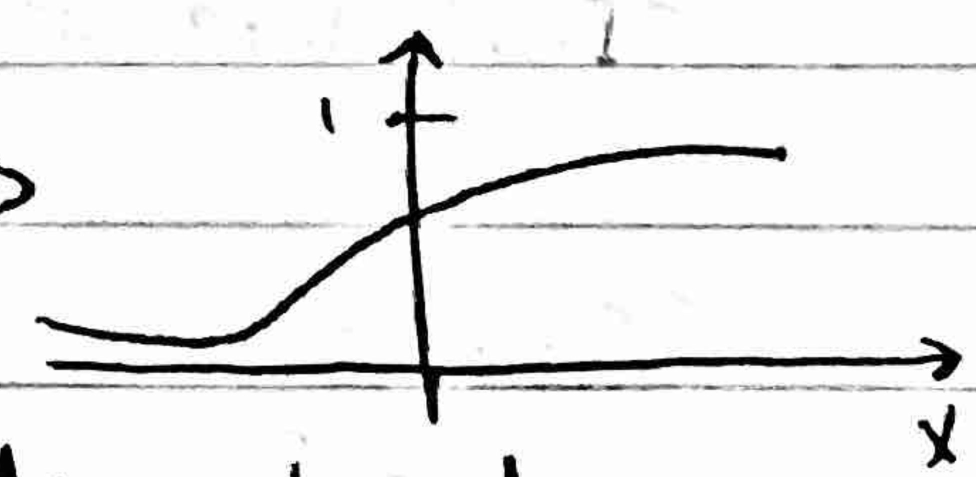
Want to find linear decision boundary

$w^T x + b$

where $x = \begin{bmatrix} x \\ 1 \end{bmatrix}$, $w = \begin{bmatrix} w \\ b \end{bmatrix}$

Need to find $\vec{w}$

$$\sigma(x) = \frac{e^x}{1+e^x} = \frac{1}{1+e^{-x}} \Longrightarrow$$



Converts a real output to probability. Acts as confidence level

For 2 classes,

$$p(y_i = 1 \mid \overbrace{x_i, w}^{\theta}) = \sigma(w^T x_i)$$

$$P(y_i = 0 \mid x_i, w) = 1 - P(y_i = 1 \mid x_i, w) = 1 - \sigma(w^T x_i)$$

Likelihood fn to maximize $= P(y \mid x_i, w) = \prod_{i=1}^{n} P(y_i \mid x_i, w)$

$$\bigstar \quad = \prod_{i=1}^{n} \sigma(w^T x_i)^{y_i} \left(1 - \sigma(w^T x_i)\right)^{1-y_i} \Big\} \text{like Bernoulli}$$

If not $P(y_i = 1)$, then $P(y_i = 0)$

$$w^* \to \underset{w}{\arg\max}\ P(y|x,w) = \underset{w}{\arg\max} \prod_{i=1}^{n} \sigma(w^T x_i)^{y_i} (1-\sigma(w^T x_i))^{1-y_i}$$

Take log likelihood

$$\log(p(y|x,w)) = \sum_{i=1}^{n} \left[y_i \log(\sigma(w^T x)) + (1-y_i) \log(1-\sigma^T(w^T x)) \right]$$

Gradient & set to zero:

$\nabla\sigma(x) = \sigma(x)(1-\sigma(x))$

$\nabla L = x(\sigma(x)-y)$

$$\nabla_w LLH(w, x) = -\sum_{i=1}^{n} \left[\sigma(w^T x_i) - y_i\right] x_i^{T} = 0$$

Not closed form. ☹ (Cannot be evaluated to a finite number of operations)

So we use an iterative approach to find w^*

Initialize w as something random, then choose ε, then:

SGD —

for $i = 0 : n_{epochs}$ OR while not_converged:

 for $j = 1 : n$_training pts:

  $w_t = w_{t-1} - \varepsilon \nabla_{w_{t-1}} LLH(w_{t-1}, x_j)$

Stop criterion:

- $\|w_t - w_{t-1}\|_2^2 < \lambda$ ← something small enough (indicates reaching minimum)
- $\nabla_{w_{t-1}} LLH(w_{t-1}, x) = 0$ (Slope of tangent is 0)
- n_epochs = something "sufficient"