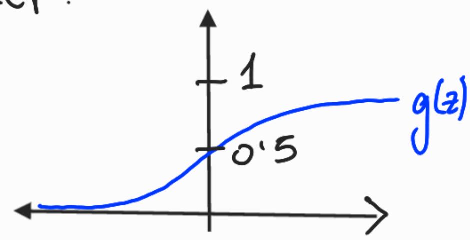
Using $h_\theta(x)$ from linear regression, the hypothesis function can be out of $[0,1]$. So we need another:

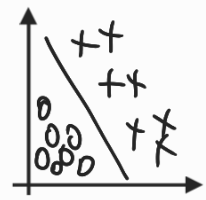


Logistic Regression:

$h_\theta(x) = g(\theta^T x)$; where $g(z) = \dfrac{1}{1+e^{-z}}$

- Interpretation: $h_\theta(x) \Rightarrow$ probability that $y=1$ on input x
  $\Rightarrow P(y=1 \mid x;\theta)$

- Decision boundary
  - $h_\theta(x) \geq 0'5 \equiv \theta^T x \geq 0 \Rightarrow 1$
  - $h_\theta(x) < 0'5 \equiv \theta^T x < 0 \Rightarrow 0$



- Cost function: Choosing θ.

$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y=1 \\ -\log(1-h_\theta(x)) & \text{if } y=0 \end{cases}$$

- Gradient descent: iteratively $\{ -\frac{1}{m}\sum_{i=1}^{m} \text{cost function} \}$

- Advanced optimization: compute $J(\theta)$, $\frac{\partial}{\partial\theta_j} J(\theta)$ $\Big\}$ Conjugate gradient, BFGS, L-BFGS