The Perceptron

- Rosenblatt, 1962 — huge amount of study afterward

Perceptron (τ, D_n)

```
1  θ = [0 0 ... 0]^T
2  θ_0 = 0
3  for t = 1 to τ:
4      for i = 1 to n:
5          if y^(i) (θ^T x^(i) + θ_0) ≤ 0
6              θ = θ + y^(i) x^(i)
7              θ_0 = θ_0 + y
8  return θ, θ_0
```

θ is $d \times 1$

$x^{(i)}$ is $d \times 1$

$y^{(i)}$ is scalar

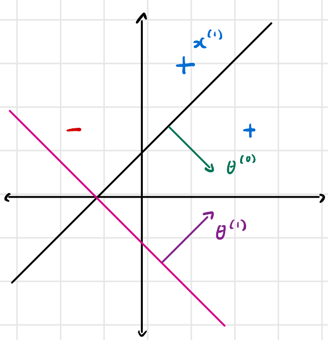
- If alg. goes through 1 iteration of loop on line 4 without an update, it will never make any further updates

- Intuition : On each step — If the current hypothesis θ, θ_0 classifies $x^{(i)}$ correctly $\longrightarrow$ no change

  If it classifies $x^{(i)}$ incorrectly, θ and θ_0 are moved so it's "closer"

Ex. h : $\theta^{(0)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\theta_0^{(0)} = 1$

$x^{(1)} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $y^1 = 1$ (misclassified)



$$y^{(1)}(\theta^T x^{(1)} + \theta_0) = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 1 = -1 < 0$$

After perceptron:

$$\theta^{(1)} = \theta^{(0)} + y^{(1)} x^{(1)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} 1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$\theta_0^{(1)} = \theta_0^{(0)} + y^{(1)} = 1 + 1 = 2$$