p. 44:

$a = \langle w, x \rangle + b$ (prediction)

$y = -1$ (label)

$a > 0$ (prediction is wrong)

$w' = w - x$ (new weights)

$b' = b - 1$ (new bias)

$a' = \langle w', x \rangle + b'$ (new prediction)

$= \langle w - x, x \rangle + b - 1$

$= \langle w, x \rangle + b - \langle x, x \rangle - 1$

$a' = a - (\langle x, x \rangle + 1) < a$

$a = \langle w, x \rangle + b$

$y = 1$

$w' = w + x$

$b' = b + 1$

$a' = \langle w', x \rangle + b'$

$= \langle w + x, x \rangle + b + 1$

$= \langle w, x \rangle + \langle x, x \rangle + b + 1$

$= \langle w, x \rangle + b + \langle x, x \rangle + 1$

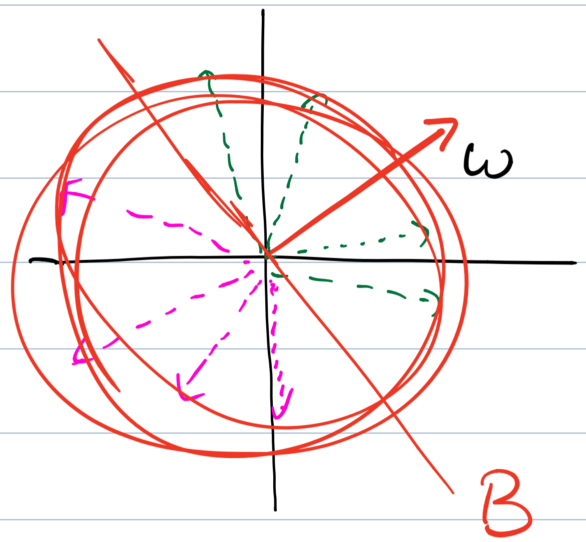
$a' = a + \cdots > a$

4.3 decision boundary of perceptron

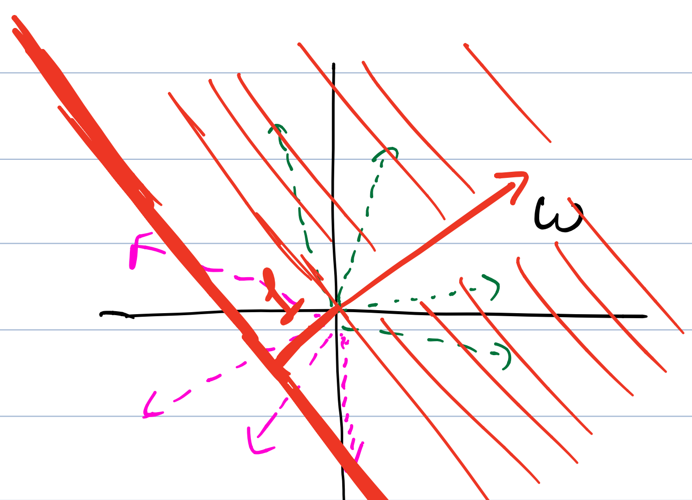
$a = \langle w, x \rangle + b$

First, $b = 0$:

$B = \{ x : \langle w, x \rangle = 0 \}$

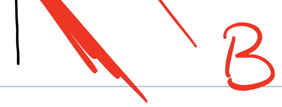


what about $b \neq 0$?



$(b = 1)$

Dot products



Margins and convergence

$$\text{margin}(D,w,b) = \min_{(x,y)\in D} y\left(\langle w,x\rangle + b\right)$$

$$\text{margin}(D) = \sup_{w,b} \text{margin}(D,w,b)$$

γ : the assumed margin of D
w^* : the weight vector which attains the margin γ
$w^{(k)}$: weight after the k-th update

$$\langle w^*, w^{(k)}\rangle = \langle w^*, w^{(k-1)} + yx^{(k)}\rangle$$

$$= \langle w^*, w^{(k-1)}\rangle + \langle w^*, yx^{(k)}\rangle$$

$$= \qquad y\langle w^*, x^{(k)}\rangle$$

$$= \qquad y\left(\langle w^*, x^{(k)}\rangle + b^* - b^*\right)$$

$$\geq \langle w^*, w^{(k-1)}\rangle + \gamma$$

$$\langle w^*, w^{(K)} \rangle \geq K\gamma$$

$$\| w^{(K)} \|^2$$

$$= \| w^{(K-1)} + yx \|^2$$

$$= \| w^{(K-1)} \|^2 + \| yx \|^2 + 2 \langle w^{(K-1)}, yx \rangle$$

$$\leq \quad \| \quad + \quad 1 \quad + \quad 0$$

$$\| w^{(K)} \|^2 \leq K$$

$$\sqrt{K} \geq \| w^{(K)} \|$$

$$\langle w^*, w^{(K)} \rangle \geq K\gamma \qquad (\| w^* \| = 1 \ !)$$

$$\sqrt{K} \geq \| w^{(K)} \| \geq \langle w^*, w^{(K)} \rangle \geq K\gamma$$

$$\sqrt{K} \geq K\gamma$$

$$\frac{1}{\sqrt{K}} \geq \gamma$$

$$K \leq \frac{1}{\gamma^2}$$