$(x_1, y_1) \dots (x_t, y_t)$

$x_i \in [-1, 1]^d$

$y_i \in \{1, -1\}$

$$\boxed{y_i = \text{sign}(\langle w_*, x_i \rangle)}$$

$\Rightarrow \|w_*\|_1 = 1$

$\Rightarrow \forall i \quad y_i \langle w_*, x_i \rangle > \gamma$

To prove $O(\sqrt{T \log d} / \gamma)$

Initialize $w_0 = (1/d, 1/d \dots 1/d)$

for $t = 1, 2, \dots T$:

$\hat{y}_t = \text{sign}(\langle w_t, x_t \rangle)$

On mistake:

$$w_{t+1,i} = \frac{w_{t,i} \, e^{(\eta y_t x_{t,i})}}{Z_t}$$

$$Z_t = \sum_i w_{t,i} e^{\eta y_t x_{t,i}}$$

else

$w_{t+1,i} = w_{t,i}$

$$A(T) \leq (1+\varepsilon) L_*(T) + \frac{\ln d}{\varepsilon}.$$

COROLLARY:

Setting $\varepsilon = \sqrt{\frac{\ln d}{T}}$, we get

$$A(T) \leq L_*(T) + 2\sqrt{T \ln d}$$

$$1-\varepsilon \quad L(t,i) = \frac{e^{\eta y_t x_{t,i}}}{Z_t}$$

$$L(t,i) = \left(1 - \frac{e^{\eta y_t x_{t,i}}}{Z_t}\right) \frac{1}{\varepsilon}$$

$$\boxed{L(t,i) = -y_t x_{t,i}}$$

$$\omega_{t+1} = \underbrace{\left[1 - \varepsilon L(t,i)\right]} \cdot \omega(t-1,i)$$

$$A(T) \leq L_*(T) + 2\sqrt{T \ln d}$$

$$\omega(T) \geq (1-\varepsilon)$$

$$\underbrace{y_i \langle w_*, x_i \rangle \;\; > \gamma}$$

↳ Link this to $L_*(T)$

→ Write $A(T)$ as a function of number of mistakes.

$$\omega_{t,i} = \underbrace{(1 - \varepsilon L(t,i))}\cdot \omega(t-1,i)$$

$$L(t,i) = -y_i x_{t,i}$$

$$\omega_{t,i} = (1 + \varepsilon y_i x_{t,i})\, \omega(t-1,i)$$

$$= \omega(t-1,i) + \varepsilon\, \omega(t-1,i)\, y_i x_{t,i}$$

$$\langle \omega_{t,i}, \omega_* \rangle = \langle \omega(t-1,i), \omega_* \rangle + \varepsilon\, \omega(t-1,i)\, \underbrace{y_i \langle x_{t,i}, \omega_* \rangle}_{\text{+ve}}$$

$$|\langle \omega_*, x_{t,i} \rangle$$

$$> \langle \omega(t-1,i), \omega_* \rangle + \varepsilon\, \omega(t-1,i)\, \gamma$$

$$\omega_{t,i} = \omega(t-1,i) + \varepsilon\, \omega(t-1,i)\, y_i x_{t,i}$$

$$\|\omega_{t,i}\|^2 = \|\omega_{t-1,i}\|^2 + \varepsilon^2 \|\omega t$$

$A(t) = L(1) + L(2) + \cdots + L(T)$

$$A^*(T) = \sum_{t=1}^{T} \sum_{i=1}^{d} L(t,i)$$

$\downarrow$

$-y_t\, x_{t,i}$

$$= \sum_{t=1}^{T} \frac{1}{W(t-1)} \sum_{i=1}^{d} \tilde{w}(t-1,i)\, L(t,i)$$

$\hookrightarrow$ $-y_t\, x_{t,i}$

$y_i \langle w_*, x_i \rangle > \gamma$

$$\underbrace{-[y_t \langle w^*, x_{t,i} \rangle]}_{\text{wrong so -ve}}$$

$[\langle w^*, x_{t,i} \rangle]$

$$> \sum_{t=1}^{T} \frac{1}{\tilde{w}(t-1)} \sum_{i=1}^{d} - \gamma$$

$$> \sum_{t=1}^{T} \frac{1}{W^*(t-1)} (M_a \gamma)$$

$$> M d\gamma$$

$$A(T) \leq L_*(T) + 2\sqrt{T \ln d}$$

$$A(T) \leq M d\gamma + 2\sqrt{T \ln d}$$

$$M \geq \frac{A(T) - 2\sqrt{T \ln d}}{d\gamma}$$

$$A^*(T) = M \times \max_{[0,1]} (-y_t, x_{t,i})$$

$$\langle w^*, A^*(T) \rangle = M \times \max_{[0,1]} \left(-y_t \underbrace{\langle w^*, x_{t,i} \rangle}_{|\langle w^*, x_{t,i} \rangle|}\right)$$

$$< M \times \gamma$$

Expected mistakes $= \sum_t F_t \leq \hat{M}(1+\varepsilon) + \frac{\ln d}{\varepsilon}$

$$M \leq M^* + 2\sqrt{T \ln d}$$

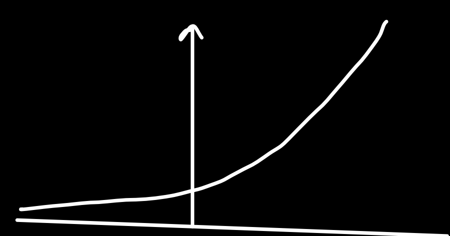
$$M^* \leq T$$

$$L(T) = -y_i x_i$$

$$w_{t+1} = (1 - \varepsilon L(T))\, w_t$$

$$\boxed{L(T) = e^{-y_i x_i}}$$

$y_i = -1$



$$A^*(T) = M \times \max_{[0,1]}$$

$$= M \times \max\left(e^{-y_i x_{t,i}}\right)$$