Lecture 6. Perceptron.

- online learning:

(1) make predictions while learning.

$(x_1, y_1) \cdots (x_m, y_m)$

total number of error.

(2) perceptron:

$h_\theta(x) = g(\theta^T x)$ $\quad g(x) = \begin{cases} 1 & z \geq 0 \\ -1 & z < 0 \end{cases}$

$\theta := \theta + yx$ $\quad (y \in \{-1, 1\})$ (online)

(3) Theorem:

sequence: $(x_1, y_1) \cdots (x_m, y_m)$

$\|x_i\| \leq D$ $\quad \exists u: \|u\| = 1.$ (margin)

$y_i \cdot (u^T x_i) \geq \gamma.$

$\sum_{i=1}^{m} \varepsilon_i \leq \left(\frac{D}{\gamma}\right)^2$

Proof: θ_k: k-th mistake.

$x_i^T \theta_k y_i \leq 0$

Considering the updating rule:

$(\theta_{k+1})^T u = \theta_k^T u + y_i x_i^T u \geq \theta_k^T u + \gamma$

$\geq \theta_{k-1}^T u + 2\gamma \cdots \geq k\gamma.$

$\|\theta_{k+1}\|^2 \leq \|\theta_k\|^2 + D^2 \Rightarrow \|\theta_{k+1}\|^2 \leq kD^2$

$\Leftrightarrow k^2\gamma^2 \leq kD^2 \Rightarrow k \leq \left(\frac{D}{\gamma}\right)^2$ $\square$