MACHINE LEARNING ASSIGNMENT 3

Part 1

1) $\epsilon_i(x) = f(x) - h_i(x)$

$$E(\epsilon_i(x)^2) = E[(f(x) - h_i(x))^2]$$

$$E_{avg} = \frac{1}{M}\sum_{i=1}^{M} E(\epsilon_i(x)^2)$$

$$h_{agg}(x) = \frac{1}{M}\sum_{i=1}^{M} h_i(x)$$

$$E_{agg}(x) = E\left[\left\{\frac{1}{M}\sum_{i=1}^{M} h_i(x) - f(x)\right\}^2\right]$$

$$E_{agg}(x) = E\left[\left\{\frac{1}{M}\sum_{i=1}^{M} \epsilon_i(x)\right\}^2\right].$$

Assumptions:

i) $E(\epsilon_i(x)) = 0$

ii) $E(\epsilon_i(x)\epsilon_j(x)) = 0$ for all $i \neq j$.

$$\left(\sum_{i=0}^{n-1} z_i\right)^2 = \sum_{i=0}^{n-1} z_i \times \sum_{j=0}^{n-1} z_j$$

$$= \sum_{i=0}^{n-1} z_i\left(z_i + \sum_{i \neq j} z_j\right)$$

$$= \sum_{i=0}^{n-1} z_i^2 + \sum_{i \neq j} z_i z_j. \quad -(1)$$

$$E_{agg} = \cancel{E\left[\frac{1}{M}\sum_{i=1}^{M}\right.} \; E\left[\left\{\frac{1}{M}\sum_{i=1}^{M}\epsilon_i(x)\right\}^2\right]$$

Expanding

$$\left\{\frac{1}{M}\sum_{i=1}^{M}\epsilon_i(x)\right\}^2 = \frac{1}{M^2}\left[\sum_{i=1}^{M}\epsilon_i(x)^2 + \sum_{i\neq j}\epsilon_i(x)\epsilon_j(x)\right]$$

From ①.

$$= \frac{1}{M^2}\sum_{i=1}^{M}\epsilon_i(x)^2 \quad \text{as } \epsilon_i(x)\epsilon_j(x) = 0 \text{ for } i\neq j.$$

$$E_{agg} = E\left[\frac{1}{M^2}\sum_{i=1}^{M}\epsilon_i(x)^2\right]$$

$$= \frac{1}{M^2}\sum_{i=1}^{M}E[\epsilon_i(x)^2] \quad \text{as } \frac{1}{M^2} \text{ is constant}$$

$$E_{avg} = \frac{1}{M}\sum_{i=1}^{M}E[\epsilon_i(x)^2]$$

$$\therefore \boxed{E_{agg} = \frac{1}{M}E_{avg}}$$

Hence proved.

2)

$$f\left(\sum_{i=1}^{M}\lambda_i x_i\right) \leq \sum_{i=1}^{M}\lambda_i f(x_i) \quad -①.$$

$$E_{avg} = \frac{1}{M}\sum_{i=1}^{M}E[\epsilon_i(x)^2]$$

$$\cancel{E_{agg}} = \cancel{\frac{1}{M^2}\sum} \; E_{agg}(x) = E\left[\left\{\frac{1}{M}\sum_{i=1}^{M}\epsilon_i(x)\right\}^2\right]$$

From equation 1

$$E_{agg} = E\left[\left\{\frac{1}{M}\sum_{i=1}^{M}\epsilon_i(x)\right\}^2\right] \leq \sum_{i=1}^{M}\frac{1}{M^2}E[\{\epsilon_i(x)\}^2]$$

$$= \frac{1}{M}\times\frac{1}{M}\sum_{i=1}^{M}E[\{\epsilon_i(x)\}^2] \quad -2$$

We know

$$E_{avg} = \frac{1}{M}\sum_{i=1}^{M}E[\{\epsilon_i(x)\}^2] \quad -3$$

From 2 & 3

$$E_{agg} \leq \frac{1}{M}E_{avg}$$

$\boxed{E_{agg} \leq E_{avg}}$ as M always ≥ 1

Hence proved

3) Given:

$$H(x) = \text{sign}\left(\sum_{t=1}^{T}\alpha_t h_t(x)\right) \quad -1$$

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times e^{-\alpha_t h_t(i) y(i)}$$

y_i & $h_t(x_i)$ are in $\{-1, 1\}$

$$D_{t+1}(i) = D_t(i) \times \frac{e^{-\alpha_1 h_1(i) y(i)}}{Z_1} \times \frac{e^{-\alpha_2 h_2(i) y(i)}}{Z_2} \times \cdots\cdots \times \cancel{e^{-\alpha_t h_t(i)}} \cdots \times \frac{e^{-\alpha_t h_t(i) y(i)}}{Z_t}$$

$$\approx \frac{1}{N} \cancel{D_t} \frac{e^{-y(i) \sum_t \alpha_t h_t(i)}}{\pi_t Z_t}$$

$$= \frac{1}{N} \frac{e^{-y(i) f(x_i)}}{\pi_t Z_t}$$

$$\text{Training error} = \frac{1}{N} \sum_i \begin{cases} 1 & \text{if } y(i) \neq H(x_i) \\ 0 & \text{or else} \end{cases} \quad -②$$

The above iegn is from training error definitivo.

From ① we know that

$H(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$ where we can write it as

$$H(x) = \text{Sign}(f(x)) \quad -③$$

$\therefore$ We can rewrite our condition in ② as

$$\text{Training error} = \frac{1}{N} \sum_i \begin{cases} 1 & \text{if } y(i) f(x_i) \leq 0 \\ 0 & \text{or else} \end{cases}$$

$$\leq \frac{1}{N} \sum_i e^{-y(i) f(x_i)} \quad [e^{-t} \geq 1 \text{ if } t \leq 0$$

$$= \sum_i D_{t+1}(i) \pi_t Z_t .$$

$$= \pi_t Z_t .$$

Scanned with CamScanner

$$Z_t = \sum_i D_t(i) \cdot \begin{cases} e^{-\alpha_t} & \text{if } h_t(i) = y(i) \\ e^{\alpha_t} & \text{if } h_t(i) \neq y(i) \end{cases}$$

$$= \sum_{i\, h_t(i) = y(i)} D_t(i) \times e^{-\alpha_t} + \sum_{i\, h_t(i) \neq y(i)} D_t(i) \times e^{\alpha_t}$$

$$= e^{-\alpha_t}(1 - \epsilon_t) + e^{\alpha_t} \epsilon_t .$$

$$= 2\sqrt{\epsilon_t (1 - \epsilon_t)}$$

We know

$$\epsilon_t = \frac{1}{2} - \gamma_t .$$

$$Z_t = 2\sqrt{\left(\frac{1}{2} - \gamma_t\right)\left(\gamma_t + \frac{1}{2}\right)}$$

$$= 2\sqrt{\frac{1}{4} - \gamma_t^2}$$

$$= \sqrt{1 - 4\gamma_t^2}$$

which is

$$\leq e^{-2\gamma_t^2} = \cancel{\exp(-2\gamma_t^2)} \equiv \exp\left(-2 \sum_{t=1}^{T} \gamma_t^2 .\right)$$

Hence proved.