Aadi Kothari

CS 4375 Assignment 3

Part 1

1) $X \to y$

Error:

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

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

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

Aggregate model:

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

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

To prove: $E_{agg} = \frac{1}{M} E_{avg}$

Assumption: 0 mean errors $\Rightarrow E(\epsilon_i(x)) = 0$ for all i

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

Ans: $E_{avg} = \frac{1}{M} \sum_{i=1}^{M} E(\epsilon_i(x)^2) = \frac{1}{M} \sum_{i=1}^{M} E\left(\left(f(x) - h_i(x)\right)^2\right)$

$$= \frac{1}{M}\sum_{i=1}^{M} E\left(f(x)^2 + h_i(x)^2 - 2\, f(x)\, h_i(x)\right)$$

$$= \frac{1}{M}\sum_{i=1}^{M}\left(E\left(f(x)^2\right) + E\left(h_i(x)^2\right) - 2\, E\left(f(x)\, h_i(x)\right.\right.$$

$$= E\left(f(x)^2\right) + \frac{1}{M}\sum_{i=1}^{M} E\left(h_i(x)^2\right) - \frac{2}{M}\sum_{i=1}^{M} E(\ldots h_i$$

Since, $E\left(\epsilon_i(x)\right) = 0$, $E\left(f(x)\right) = E\left(h_i(x)\right)$

$$\therefore E_{avg} = E\left(f(x)^2\right) + \frac{1}{M}\sum_{i=1}^{M}\left(E\left(h_i(x)^2\right) - 2\, E(h_i(\ldots\right.$$

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

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

$$= E\left(\frac{1}{M^2}\sum \epsilon_i(x)^2\right) \qquad \left(\because \text{ since } E(\epsilon_i(x)\, \epsilon_j(\ldots = (\ldots\right.$$

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

$$= \frac{1}{M} E_{avg}$$

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

$$E_{agg}(x) = E\left(\frac{1}{M^2}\left(\epsilon_1(x) + \epsilon_2(x) + \dots + \epsilon_m(x)\right)^2 \right)$$

$$= \frac{1}{m} E\left(\frac{1}{m}\left(\epsilon_1(x) + \epsilon_2(x) + \dots \epsilon_M(x)\right)^2 \right)$$

$$\le \frac{1}{m} E\left(\epsilon_1^2(x) + \epsilon_2^2(x) + \dots \epsilon_m^2(x) \right)$$

$$= \frac{1}{m}\left(E\left(\sum_{i=1}^{M} \epsilon_i(x)^2 \right)\right) \le \frac{1}{m} \sum_{i=1}^{m} E\left(\epsilon_i(x)^2\right)$$

$$\le \quad E_{agg} \quad \le \quad E_{avg}$$

proof $\quad 2ab \le a^2 + b^2$

$$a^2 - 2ab + b^2 \Rightarrow (a-b)^2 \ge 0$$ squares can't be negative

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

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

$$M\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{pmatrix} + M\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{pmatrix} = 2M(\epsilon_1 + \epsilon_2 + \dots \epsilon_m)$$

$$\frac{2}{M} \sum_{i=1}^{M} \sum_{j=1}^{M} \epsilon_i(x)\epsilon_j(x) \le \frac{2\cancel{M}}{\cancel{M}} \sum_{i=1}^{M} \epsilon_i(x)$$

$$\frac{1}{M}\left(\sum_{i=1}^{m} \epsilon_i(x)\right)^2 \le \sum_{i=1}^{M} \epsilon_i(x)$$

OR

use cauchy-schwarz $\quad |\langle u,v\rangle|^2 \le \langle u,u\rangle \cdot \langle v,v\rangle$

$$\left|\left\langle \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{bmatrix}, \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \right\rangle\right|^2 \le \left(\overbrace{1+1+\dots 1}^{m}\right)\left(\epsilon_1^2 + \epsilon_2^2 + \dots \epsilon_m^2\right)$$

$u = \langle \epsilon_1, \epsilon_2, \dots \epsilon_m \rangle$
$v = \langle 1, 1, \dots, 1 \rangle$

$$\left(\epsilon_1 + \epsilon_2 + \dots + \epsilon_m\right)^2 \le m\left(\epsilon_1^2 + \epsilon_2^2 + \dots \epsilon_m^2\right)$$

$$\frac{1}{m}\left(\sum_{i=1}^{m} \epsilon_i(x)\right)^2 \le \sum_{i=1}^{M} \epsilon_i(x)$$

(3)

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

Referece: Princeton COS 402 Fall '08 readings boosting.PDF

$$D_1 = \frac{1}{N}$$

$$D_2 = \frac{1}{N Z_1} e^{-\alpha_1 h_1(i) y(i)}$$

$$D_3 = \frac{1}{N Z_1 Z_2} e^{-\alpha_1 h_1(i) y(i)}$$

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

$$D_{t+1}(i) = \frac{1}{N(Z_1 Z_2 .. Z_t)} e^{-y(i)\sum_{j=1}^{t} \alpha_j h_j(i)}$$

$$\text{Error} = \frac{1}{N}\left(\sum^{i} h_t(x_i) \neq y_i\right)$$

$$\frac{1}{N}\left(\sum^{i} \mathbb{I}\left(y(i)\left(\sum^{t} \alpha_j h_j(i)\right)\right)\right) \leq \frac{1}{N}\sum^{i} \exp\left(-y(i)\left(\sum^{t} \alpha_j h_j(i)\right)\right)$$

$$\{e^x > 1 \iff x > 0\}$$

$$= (Z_1 Z_2 \ldots Z_t)\left(\sum^{i} D_{T+1}(i)\right)$$

$D_{T+1}(i)$'s are normalizeel and add to be 1

$$= Z_1 Z_2 \ldots Z_t$$

$$H(x) \text{ Error} \leq \prod_{j=1}^{t} Z_j$$

$$= \prod_{j=1}^{t} Z_j = \prod_{j=1}^{t}\left(\sum^{i} D_j(i) \exp(-\alpha_j I(h_j(x_i) = y_i))\right)$$

$$= \prod_{j=1}^{t}\left(\sum_{\text{correct}} D_j(i) e^{-\alpha_j} + \sum_{\text{incorrect}} D_j(i) e^{\alpha_j}\right)$$

$$= \prod_{j=1}^{t}\left(e^{\alpha_j}(1-\epsilon_j) + e^{y_t}\epsilon_j\right) = \prod_{j=1}^{t}\left(e^{-\left(\frac{1}{2}\ln\left(\frac{1-\epsilon_j}{\epsilon_j}\right)\right)}(1-\epsilon_j) + e^{\left(\frac{1}{2}\ln\left(\frac{1-\epsilon_j}{\epsilon_j}\right)\right)}\right)$$

$$= \prod_{j=1}^{t}\left(\sqrt{\frac{\epsilon_j}{1-\epsilon_j}}(1-\epsilon_j) + \sqrt{\frac{1-\epsilon_j}{\epsilon_j}}\right) = \prod_{j=1}^{t}\left(\frac{1-\epsilon_j}{\sqrt{\frac{1-\epsilon_j}{\epsilon_j}}} + \sqrt{\frac{1-\epsilon_j}{\epsilon_j}}\epsilon_j\right) = \prod_{j=1}^{t} 2\sqrt{\epsilon_j(1-\epsilon_j)}$$

$$= \prod_{j=1}^{t} 2\sqrt{\left(\tfrac{1}{2}-\gamma_j\right)\left(\tfrac{1}{2}+\gamma_j\right)} = \prod_{j=1}^{t}\sqrt{1-4\gamma_j^2} \leq \prod_{j=1}^{t}\exp(-2\gamma_j^2) = \exp\left(-2\sum_{t=1}^{t}\gamma_t^2\right)$$

$$1 + x \leq e^{x} \qquad e^a e^b = e^{a+b}$$

$$\text{Error} \leq \exp\left(-2\sum_{t=1}^{t}\gamma_t^2\right)$$