William Minchew

Daniel Le

Prove that

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

provided you make the following assumptions:

1. Each of the errors have a 0 mean
$$E(\epsilon_i(x)) = 0 \text{ for all i}$$
2. Errors are uncorrelated
$$E\left(\epsilon_i(x)\epsilon_j(x)\right) = 0 \text{ for all i} \neq \text{j}$$

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

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

$$= E \frac{1}{M^2}\left[\sum_{i=1}^{M} \varepsilon_i(x)^2 + \underbrace{\sum_{j\neq i}^{M} \varepsilon_i(x)\, \varepsilon_j(x)}_{0 \text{ (assumption)}}\right]$$

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

$$= \frac{1}{M} E_{avg} \quad \text{(proved)}$$

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

Part 2

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

Prove $E_{agg} \leq E_{avg}$

Where $f(x) = E(x^2)$

$\lambda_i = \frac{1}{M}$

$x_i = \epsilon_i(x)$

Replace variables

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

Which are the same as our E_{agg} and E_{avg} proving

$E_{agg} \leq E_{avg}$

Part 3

Step 1: $D_{T+1}(i) = D_1(i) \cdot \frac{\exp(-\alpha_1 y_i h_1(x_i))}{Z_1} \cdots \frac{\exp(-\alpha_T y_i h_T(x_i))}{Z_T}$

$$= \frac{1}{N} \cdot \frac{\exp(-y_i \sum_t \alpha_t h_t(x_i))}{\prod_t Z_t}$$

Step 2: Show training error of the final classifier H is at most $\prod_{t=1}^{T} Z_t$.

Proof:

$$\text{training error}(H) = \frac{1}{N}\sum_i \begin{cases} 1 & \text{if } y_i \neq H(x_i) \\ 0 & \text{else} \end{cases}$$

$$= \frac{1}{N}\sum_i \begin{cases} 1 & \text{if } y_i f(x_i) \leq 0 \\ 0 & \text{else} \end{cases} \quad \left(H(x) = sign(\sum_{t=1}^{T} \alpha_t h_t(x)) \right)$$

$$\leq \frac{1}{N}\sum_i \exp(-y_i f(x_i)) \quad \text{since } e^{-z} \geq 1 \text{ if } z \leq 0$$

$$= \sum_i D_{T+1}(i) \prod_t Z_t \quad (\text{step } 1)$$

$$= \prod_t Z_t \quad (1)$$

Step 3:

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

$$= \sum_{i: h_t(x_i) = y_i} D_t(i) e^{-\alpha_t} + \sum_{i: h_t(x_i) \neq y_i} D_t(i) e^{\alpha_t}$$

$$= e^{-\alpha_t}(1 - \varepsilon_t) + e^{\alpha_t}\varepsilon_t \quad \left(\varepsilon_t = \frac{1}{2} - \gamma_t \right)$$

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

$$\leq e^{-2\gamma_t^2} \quad (2)$$

$$(1), (2) \Rightarrow H \leq e^{-2\gamma_t^2}$$