Homework 3

Zhong Yun 2016K8009915009

Oct 8，2018

1 Answer Sheet

Theorem 1 *AdaBoost Algorithm*

$$\frac{1}{N}\sum_{i=1}^{N} I(f(x_i) \neq y_i) \leq exp(-2\sum_{k=1}^{K}\gamma_k^2)$$

$where\ \gamma_k = \frac{1}{2} - \epsilon^{(k)}$

$$I(f(x_i) \neq y_i) = \begin{cases} 0, while f(x_i) = y_i \\ 1, while f(x_i) \neq y_i \end{cases}$$

Lemma 1

$$\prod_{k=1}^{K}\sqrt{1-4\gamma_k^2} \leq e^{-2\sum_{k=1}^{K}\gamma_k^2}$$

$where\ \gamma_k \geq 0$

Proof 1 *1.* We need to clarify some definitions:

$$\alpha^{(k)} = \frac{1}{2}log(\frac{1-\epsilon^{(k)}}{\epsilon^{(k)}})$$

$$d_i^{(k+1)} = \frac{d_i^k}{Z^{(k)}}exp(-\alpha^{(k)}y_i f(x_i))$$

$$\epsilon^{(k)} = \sum_{i=1}^{N} d_i^{k-1}[y_i \neq f^{(k)}(x_i)]$$

$$Z^{(k)} = \sum_{i=1}^{N} d_i^k exp(-\alpha^{(k)}y_i f(x_i))$$

$$F(x) = \sum_{k=1}^{K}\alpha^{(k)}f^{(k)}(x)$$

$$f(x) = sgn[F(x)] = sgn[\sum_k \alpha^{(k)}f^{(k)}(x)]$$

2. According to Lemma, we only need to prove the following inequality:

$$\frac{1}{N}\sum_{i=1}^{N} I(f(x_i) \neq y_i) \leq \prod_{k=1}^{K}\sqrt{1-4\gamma_k^2}$$

3. First, we prove this inequality

$$\sum_{i=1}^{N} I(f(x_i) \neq y_i) \leq \sum_{i=1}^{N} exp(-y_i F(x_i))$$

Proof:

If $f(x_i) = y_i$, then $I(f(x_i) \neq y_i) = 0$ and $y_i F(x_i) \geq 0$, so $exp(-y_i F(x_i)) \geq 0$

We have $I(f(x_i) \neq y_i) \leq exp(-y_i F(x_i))$

If $f(x_i) \neq y_i$, then $I(f(x_i) \neq y_i) = 1$ and $y_i F(x_i) \leq 0$, so $exp(-y_i F(x_i)) \geq 1$

Still, we have $I(f(x_i) \neq y_i) \leq exp(-y_i F(x_i))$

4. We prove that:

$$\frac{1}{N} \sum_{i=1}^{N} exp(-y_i F(x_i)) = \prod_{k=1}^{K} Z^{(k)}$$

Proof:
According to the definition in part1, we have that

$$\alpha^{(k)} = \frac{1}{2} log(\frac{1-\epsilon^{(k)}}{\epsilon^{(k)}})$$

$$d_i^{(k+1)} = \frac{d_i^k}{Z^{(k)}} exp(-\alpha^{(k)} y_i f(x_i))$$

$$\epsilon^{(k)} = \sum_{i=1}^{N} d_i^{k-1} [y_i \neq f^{(k)}(x_i)]$$

$$Z^{(k)} = \sum_{i=1}^{N} d_i^k exp(-\alpha^{(k)} y_i f(x_i))$$

$$\begin{aligned}
\frac{1}{N} \sum_{i=1}^{N} exp(-y_i F(x_i)) &= \frac{1}{N} \sum_{i=1}^{N} exp(-\sum_{k=1}^{K} \alpha^{(k)} y_i f^{(k)}(x_i) \\
&= \sum_{i=1}^{N} \frac{1}{N} \prod_{k=1}^{K} exp(\alpha^{(k)} y_i f^{(k)}(x_i)) \\
&= \sum_{i=1}^{N} d_i^{(0)} \prod_{k=1}^{K} exp(\alpha^{(k)} y_i f^{(k)}(x_i)) \\
&= Z^{(1)} \sum_{i=1}^{N} d_i^{(1)} \prod_{k=2}^{K} exp(\alpha^{(k)} y_i f^{(k)}(x_i)) \\
&= Z^{(1)} Z^{(2)} \sum_{i=1}^{N} d_i^{(2)} \prod_{k=3}^{K} exp(\alpha^{(k)} y_i f^{(k)}(x_i)) \\
&= \cdots \\
&= Z^{(1)} Z^{(2)} \cdots Z^{(K-1)} \sum_{i=1}^{N} d_i^{(K)} exp(\alpha^{(K)} y_i f^{(K)}(x_K)) \\
&= Z^{(1)} Z^{(2)} \cdots Z^{(K)} \\
&= \prod_{k=1}^{K} Z^{(k)}
\end{aligned}$$

5. We prove the following equality:

$$Z^{(k)} = \sqrt{1 - 4\gamma_k^2}$$

where $\gamma_k = \frac{1}{2} - \epsilon^{(k)}$

Proof:

$$\begin{aligned} Z^{(k)} &= \sum_{i=1}^{N} d_i^k exp(-\alpha^{(k)} y_i f^{(k)}(x_i)) \\ &= \sum_{f^{(k)}(x_i)=y_i} d_i^k exp(-\alpha^{(k)}) + \sum_{f^{(k)}(x_i)\neq y_i} d_i^k exp(\alpha^{(k)}) \\ &= (1 - \epsilon^{(k)}) exp(-\alpha^{(k)}) + (\epsilon^{(k)}) exp(\alpha^{(k)}) \\ &= \sqrt{1 - 4\gamma_k^2} \end{aligned}$$

where $\gamma_k = \frac{1}{2} - \epsilon^{(k)}$

6. Taking 1, 2, 3, 4, 5 into acount, We can prove that:

$$\frac{1}{N} \sum_{i=1}^{N} I(f(x_i) \neq y_i) \leq exp(-2 \sum_{k=1}^{K} \gamma_k^2)$$