清華大學
TSINGHUA UNIVERSITY

1. $D_{t+1}(i) = \dfrac{D_t(i)\exp(-\alpha_t y_i h_t(x_i))}{Z_t}$

$D_{T+1}(i) = \dfrac{1}{m} \dfrac{\exp(-y_i f(x_i))}{\prod_{t=1}^{T} Z_t}$

$\hat{\epsilon}(h) \leq \frac{1}{m}\sum_{i=1}^{m} 1[y_i f(x_i) \leq 0] = \frac{1}{m}\sum_{i=1}^{m} e^{-y_i f(x_i)}$

$= \sum_{i=1}^{m} \prod_{t=1}^{T} Z_t D_{T+1}(i) = \prod_{t=1}^{T} Z_t \sum_{i=1}^{m} D_{T+1}(i) = \prod_{t=1}^{T} Z_t$

待证 $\prod_{t=1}^{T} Z_t \leq e^{-2\sum_{t=1}^{T}(\frac{1}{2}-\epsilon_t)^2} =$

即 $\sum_{t=1}^{T} \log Z_t \leq \sum_{t=1}^{T} -2(\frac{1}{2}-\epsilon_t)^2$

$d(\epsilon_t) = \log Z_t + 2(\frac{1}{2}-\epsilon_t)^2$

$= 4\epsilon_t^2 + 4\epsilon_t + 1 + \ln 2 + \frac{1}{2}\ln \epsilon_t + \frac{1}{2}\ln(1-\epsilon_t)$

$\dfrac{\partial d(\epsilon_t)}{\partial \epsilon_t} = \left(8 - \dfrac{2}{\epsilon_t(1-\epsilon_t)}\right)(\epsilon_t - \tfrac{1}{2})$

$d(\epsilon_t)$ 在 $(0, \frac{1}{2})$ 单调增，$(\frac{1}{2}, 1)$ 单调减

$\frac{1}{2}$ 处取最大值 0

$\therefore (0,1)$ 上 $d(\epsilon_t) \leq 0$.

$\therefore$ 有 $\hat{\epsilon}(h) \leq \prod_{t=1}^{T} Z_t \leq e^{-2\sum_{t=1}^{T}(\frac{1}{2}-\epsilon_t)^2}$，得证

$\mathcal{H} \subseteq \{-1,+1\}^X$

for $i \leftarrow 1$ to m do

$D_1(i) \leftarrow \frac{w_i}{\sqrt{\sum_{i=1}^{n} w_i^2}}$

for $t \leftarrow 1$ to T do

$h_t \leftarrow$ base classifier in H with $\epsilon_t = \Pr_{i \sim D_t}[h_t(x_i) \neq y_i]$

$\alpha_t \leftarrow \frac{1}{2}\log\frac{1-\epsilon_t}{\epsilon_t}$

$Z_t \leftarrow 2[\epsilon_t(1-\epsilon_t)]^{\frac{1}{2}}$

for $i \leftarrow 1$ to m do

$D_{t+1}(i) \leftarrow \frac{D_t(i)\exp(-\alpha_t y_i h_t(x_i))}{Z_t}$

$f_t \leftarrow \sum_{s=1}^{t}\alpha_s h_s$

return $h = \mathrm{sgn}(f_T)$

证明，令 $\overline{D_t}(i) = \frac{\overline{w_i} e^{-y_i \sum_{j=1}^{t} \alpha_{t-1} h_j(x_i)}}{\overline{Z_t}}$

$\cancel{D_{t+1}(i) = \frac{D_{t-1}(i)\, e^{-\alpha_{t-1} y_i h_{t-1}(x_i)}\, e^{-\alpha_t y_i h_t(x_i)}}{Z_{t-1} Z_t}}$

$\cancel{= \frac{w_i}{n}\frac{e^{-y_i f_t(x_i)}}{\prod_{s=1}^{t} Z_s} = \frac{w_i}{n}\frac{e^{-y_i \sum_{s=1}^{t}\alpha_s h_s(x_i)}}{\prod_{s=1}^{t} Z_t}}$

设 $\overline{w_i} = \frac{w_i}{\sqrt{\sum_{i=1}^{n} w_i^2}}$

$F(\overline{\alpha}) = \sum_{i=1}^{n} \overline{w_i} e^{-y_i f(x_i)}$

$= \sum_{i=1}^{n} \overline{w_i} e^{-y_i \sum_{j=1}^{N} \alpha_j h_j(x_i)}$

$F(\overline{\alpha}_{t-1}, e_k) = \sum_{i=1}^{n} \overline{w_i} e^{-y_i \sum_{j=1}^{N} \alpha_j h_j(x_i) - \eta y_i h_k(x_i)}$

$F'(\overline{\alpha}_{t-1}, e_k) = -\sum_{i=1}^{n} y_i h_k(x_i)\, \overline{w_i} e^{-y_i \sum_{j=1}^{N} \alpha_{t-1} h_j(x_i)}$

$= -\sum_{i=1}^{n} y_i h_k(x_i)\, \overline{D_t}(i)\, \overline{Z_t}$

$= -\left[\sum_{i=1}^{n} \overline{D_t}(i)\, 1[y_i h_k(x_i) = 1] - \sum_{i=1}^{n} \overline{D_t}(i)\, 1[y_i h_k(x_i) = -1]\right] \overline{Z_t}$

$= -[(1-\overline{\epsilon}_{t,k}) - \overline{\epsilon}_{t,k}]\, \overline{Z_t}$

$= [2\overline{\epsilon}_{t,k} - 1]\, \overline{Z_t}$

$\frac{\partial F(\overline{\alpha}_{t-1}, e_k)}{\partial \eta} = \sum_{i=1}^{n} y_i h_k(x_i)\, \overline{D_t}(i)\, \overline{Z_t}\, e^{-\eta y_i h_k(x_i)} = 0$

$= \sum_{i=1}^{n} \overline{D_t}(i)\, 1[y_i h_k(x_i) = 1] e^{-\eta} - \sum_{i=1}^{n} \overline{D_t}(i)\, 1[y_i h_k(x_i) = -1] e^{\eta} = 0$

$\Leftrightarrow [(1-\overline{\epsilon}_{t,k}) e^{-\eta} - \overline{\epsilon}_{t,k} e^{\eta}] = 0$

$\eta = \frac{1}{2}\log\frac{1-\overline{\epsilon}_{t,k}}{\overline{\epsilon}_{t,k}}$