AdaBoost continued

$$L_m(\phi) = \left(e^{\beta} - e^{-\beta}\right) \sum_{i=1}^{N} w_{i,m}\, \mathbb{I}\left(\tilde{y}_i \neq \phi(x_i)\right) + e^{-\beta} \sum_{i=1}^{N} w_{i,m}$$

<u>Substitute ϕ_m into L_m:</u>

$$L_m(\phi_m) = \left(e^{\beta} - e^{-\beta}\right) \sum_{i=1}^{N} w_{i,m}\, \mathbb{I}\left(\tilde{y}_i \neq \phi_m(x_i)\right) + e^{-\beta} \sum_{i=1}^{N} w_{i,m}$$

<u>Solve for β_m that minimizes $L_m(\phi_m)$</u>

$$\frac{\partial}{\partial \beta_m} L_m(\phi_m) = 0$$

$$\left(e^{\beta_m} + e^{-\beta_m}\right) \sum_{i=1}^{N} w_{i,m}\, \mathbb{I}\left(\tilde{y}_i \neq \phi_m(x_i)\right) - e^{-\beta_m} \sum_{i=1}^{N} w_{i,m} = 0$$

$$\left(e^{\beta_m} + \frac{1}{e^{\beta_m}}\right) \sum_{i=1}^{N} w_{i,m}\, \mathbb{I}\left(\tilde{y}_i \neq \phi_m(x_i)\right) = \frac{1}{e^{\beta_m}} \sum_{i=1}^{N} w_{i,m}$$

$$\left(\frac{e^{2\beta_m}+1}{\cancel{e^{\beta_m}}}\right) \cdot \sum_{i=1}^{N} w_{i,m}\, \mathbb{I}\left(\tilde{y}_i \neq \phi_m(x_i)\right) = \frac{1}{\cancel{e^{\beta_m}}} \sum_{i=1}^{N} w_{i,m}$$

$$\underbrace{\boxed{\frac{\sum_{i=1}^{N} w_{i,m}\, \mathbb{I}\left(\tilde{y}_i \neq \phi_m(x_i)\right)}{\sum_{i=1}^{N} w_{i,m}}}}_{err_m} = \frac{1}{e^{2\beta_m}+1}$$

$$err_m = \frac{1}{e^{2\beta_m}+1}$$

$$e^{2\beta_m} \cdot err_m + err_m = 1$$

$$e^{2\beta_m} \cdot err_m = 1 - err_m$$

$$e^{2\beta_m} = \frac{1-err_m}{err_m}$$

$$\log\left(e^{2\beta_m}\right) = \log \frac{1-err_m}{err_m}$$

$$2\beta_m = \log \frac{1-err_m}{err_m}$$

$$\beta_m = \frac{1}{2} \log \frac{1-err_m}{err_m}$$

<u>The overall update is:</u>

$$f_m(x) = f_{m-1}(x) + \beta_m \phi_m(x_i)$$

<u>The weights at the next iteration:</u>

$$w_{i,m+1} = \exp\left(-\tilde{y}_i f_m(x_i)\right)$$

$$= \exp\left[-\tilde{y}_i \left(f_{m-1}(x_i) + \beta_m \phi_m(x_i)\right)\right]$$

$$= \exp\left[-\tilde{y}_i \cdot f_{m-1}(x_i) - \beta_m \tilde{y}_i \phi_m(x_i)\right]$$

$$= \exp(-\tilde{y}_i \cdot f_{m-1}(x_i)) \cdot \exp(-\beta_m \cdot \tilde{y}_i \, \phi_m(x_i))$$

$$= w_{i,m} \cdot \exp(-\beta_m \cdot \tilde{y}_i \, \phi_m(x_i))$$

$$= w_{i,m} \cdot \exp(\beta_m \cdot (2\mathbb{I}(\tilde{y}_i \neq \phi_m(x_i)) - 1))$$

$$= w_{i,m} \cdot \exp(2\beta_m \mathbb{I}(\tilde{y}_i \neq \phi_m(x_i))) \cdot e^{-\beta_m}$$

We can drop $e^{-\beta_m}$ as it will cancel out in the normalijation step.

If $\tilde{y}_i \neq \phi_m(x_i)$ $\tilde{y}_i \, \phi_m(x_i) = -1$, else $\tilde{y}_i \, \phi_m(x_i) = 1$

We can write

$$\tilde{y}_i \cdot \phi_m(x_i) = 1 - 2\mathbb{I}(\tilde{y}_i \neq \phi_m(x_i))$$

LogitBoost

$$f^{x} = \frac{1}{2} \log \frac{\pi_i}{1-\pi_i}$$

$$f(x_i) = \frac{1}{2} \log \frac{\pi_i}{1-\pi_j}$$

$$2f(x_i) = \log \frac{\pi_i}{1-\pi_i}$$

$$e^{2f(x_i)} = \frac{\pi_i}{1-\pi_i}$$

$$e^{2f(x_i)} - \pi_i e^{2f(x_i)} = \pi_i$$

$$e^{2f(x_i)} = \pi_i + \pi_i \cdot e^{2f(x_i)}$$

$$e^{2f(x_i)} = \pi_i \left(1 + e^{2f(x_i)}\right)$$

$$\pi_i = \frac{e^{2f(x_i)}}{1 + e^{2f(x_i)}} = \frac{1}{e^{-2f(x_i)} + 1}$$