For given loss function, get (β_m, G_m) **and** $f_m(x)$

$$L(y, f(x)) = \exp(-y\, f(x)).$$

$$(\beta_m, G_m) = \arg\min_{\beta, G} \sum_{i=1}^{N} \exp[-y_i(f_{m-1}(x_i) + \beta\, G(x_i))]$$

basis function : individual classifier (stump) $G_m(x) \in \{-1,1\}$

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

1. Let $w_i^{(m)} = e^{-y_i f_{m-1}(x_i)}$. We know that $w_i^{(m)}$ does not depend on β or G.

$\therefore (\beta_m, G_m) = \arg\min_{\beta, G} \sum_{i=1}^{N} w_i^{(m)} e^{-\beta y_i G(x_i)}$. For a given $\beta > 0$, $G_m = \arg\min_{G} \sum_{i=1}^{N} w_i^{(m)} e^{-y_i G_i(x)}$. Since $w_i^{(m)} \geqq 0$ for all i, and does not depend on G, $G_m = \arg\min_{G} \sum_{i=1}^{N} w_i^{(m)} I(y_i \neq G(x_i))$.

Then, $\beta_m = \arg\min_{\beta} \sum_{i=1}^{N} w_i^{(m)} e^{-\beta y_i G_m(x_i)}$. $\sum_{i=1}^{N} w_i^{(m)} e^{-\beta y_i G_m(x_i)} = \sum_{y_i = G(x_i)} w_i^{(m)} e^{-\beta} + \sum_{y \neq G(x_i)} w_i^{(m)} e^{\beta} = \left(\sum_{i=1}^{N} w_i^{(m)} e^{-\beta} - \sum_{y_i \neq G(x_i)} w_i^{(m)} e^{-\beta}\right) + \sum_{y \neq G(x_i)} w_i^{(m)} e^{\beta} = (e^{\beta} - e^{-\beta}) \sum_{y \neq G(x_i)} w_i^{(m)} + e^{-\beta} \sum_{i=1}^{N} w_i^{(m)} = \left(\sum_{i=1}^{N} w_i^{(m)}\right)\left((e^{\beta} - e^{-\beta}) \frac{\sum_{y \neq G(x_i)} w_i^{(m)}}{\sum_{i=1}^{N} w_i^{(m)}} + e^{-\beta}\right) := \sum_{i=1}^{N} w_i^{(m)} \left((e^{\beta} - e^{-\beta}) err_m + e^{-\beta}\right)$

$\therefore \arg\min_{\beta} \sum_{i=1}^{N} w_i^{(m)} e^{-\beta y_i G_m(x_i)} = \arg\min_{\beta} (e^{\beta} - e^{-\beta}) err_m + e^{-\beta} := \arg\min_{\beta} f(\beta; m)$

$\frac{\partial}{\partial \beta} f(\beta; m) = (e^{\beta} + e^{-\beta}) err_m - e^{-\beta}$ $\quad \therefore (e^{\beta_m} + e^{-\beta_m}) err_m - e^{-\beta_m} = 0$. $\quad e^{\beta_m} = e^{-\beta_m} \frac{(1 - err_m)}{err_m}$. $\quad e^{2\beta_m} = \frac{1 - err_m}{err_m}$.

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

2. $f_m(x) = f_{m-1}(x) + \beta_m G_m(x)$.

$$w_i^{(m+1)} = e^{-y_i f_m(x)} = e^{-y_i f_{m-1}(x_i)} e^{-\beta_m y_i G_m(x_i)}.$$