Idea: choose α_t to minimize a bound on training error!

$$\frac{1}{m}\sum_{i=1}^{m}\delta(H(x^i)\neq y^i)\leq\frac{1}{m}\sum_{i=1}^{m}D_t(i)\exp(-y^i f(x^i)) \;=\prod_t Z_t$$

Where

$$f(x)=\sum_t \alpha_t h_t(x);\, H(x)=sign(f(x))$$

And

$$Z_t=\sum_{i=1}^{m}D_t(i)\exp(-\alpha_t y^i h_t(x^i))$$

This equality isn't obvious! Can be shown with algebra (telescoping sums)!

If we minimize $\prod_t Z_t$, we minimize our training error!!!

- We can tighten this bound greedily, by choosing α_t and h_t on each iteration to minimize Z_t.
- h_t is estimated as a black box, but can we solve for α_t?