Discussion 9.

1. Singularity Issue in EM

Example with 2 Gaussian Components

$$\max_{\substack{\mu_1\mu_2\\ \sigma_1\sigma_2}} P = \prod_{i=1}^{N}\left[\pi_1 \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{\sigma_1^2}(x_n-\mu_1)^2} + \pi_2 \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{\sigma_2^2}(x_n-\mu_2)^2}\right]$$

First find μ_1, μ_2 and the $\sigma_1\ \sigma_2$

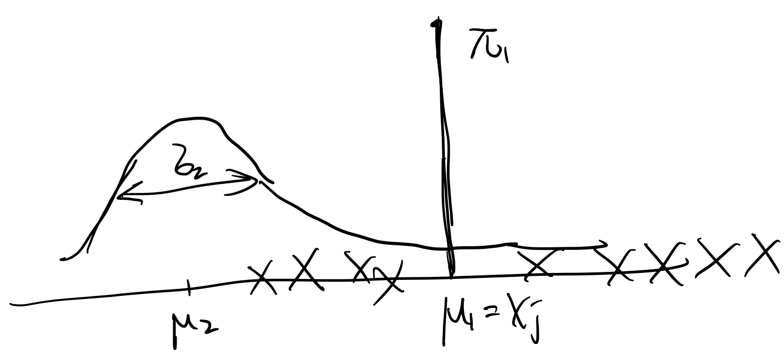
What happens when $x_j = \mu_1$?

$$P = \prod_{\substack{i=1\\ i\neq j}}^{N}\Bigg[\underbrace{\boxed{\pi_1 \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{\sigma_1^2}(x_n-\mu_1)^2}}}_{\to 0} + \underbrace{\pi_2 \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{\sigma_2^2}(x_n-\mu_2)^2}}_{>0}\Bigg]$$

$$\times \Bigg[\underbrace{\boxed{\pi_1 \frac{1}{\sqrt{2\pi}\sigma_1}}}_{\to \infty} \cancel{e} + \underbrace{\pi_2 \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{\sigma_2^2}(x_j-\mu_2)^2}}_{>0}\Bigg]$$

$\sigma_1 \to 0$

$$\lim_{\sigma_1\to 0} P = \infty$$



Singularity Issue

We can avoid this by checking $x_j = \mu_1$ or $x_j = \mu_2$.

If true, change μ_1 or μ_2

2. Adaboost as maximizing an exponential loss function

In class, we see how $w_n^{(k)}$, α_k and $y_k(x)$ are obtained

why ?

$$E = \sum_{n=1}^{N} e^{-t_n f_k(x_n)} \qquad t_n \in \{-1, 1\}$$

$$f_k(x) = \frac{1}{2} \sum_{l=1}^{k} \alpha_l \boxed{y_l(x)} \longrightarrow \text{weak classifiers}$$

Now we have $y_1(x) \ldots\ldots y_{k-1}(x)$ and $\alpha_1 \ldots\ldots \alpha_{k-1}$

Wish to find α_k, $y_k(x)$

$$E = \sum_{n=1}^{N} e^{-t_n \frac{1}{2} \sum_{l=1}^{k} \alpha_l y_l(x)}$$

$$= \sum_{n=1}^{N} e^{-t_n \underbrace{\frac{1}{2} \sum_{l=1}^{k-1} \alpha_l y_l(x_n)}_{f_{k-1}(x_n)}} \times e^{-\frac{1}{2} t_n \alpha_k y_k(x_n)}$$

$$= \sum_{n=1}^{N} w_n^{(k)} e^{-\frac{1}{2} t_n \alpha_k y_k(x_n)}$$

I_k : set of index that are correctly classified by $y_k(x)$

M_k : - - ... ... misclassified - - - - - -

$$E = \sum_{n=1}^{N} w_n^{(k)} e^{-\frac{1}{2} t_n \alpha_k y_k(x_n)}$$

$$= e^{-\alpha_k/2} \sum_{n \in I_k} w_n^{(k)} + e^{\alpha_m/2} \sum_{n \in M_k} w_n^{(k)}$$

$$= \underbrace{\left(e^{\alpha_k/2} - e^{-\alpha_k/2}\right)}_{\text{constant}} \sum_{n=1}^{N} w_n^{(k)} \mathbb{1}\left(y_k(x_n) \neq t_n\right) + e^{-\alpha_k/2} \underbrace{\sum_{n=1}^{N} w_n^{(k)}}_{\text{constant}}$$

① Suppose α_k is known

$$\min E \iff \min \sum_{n=1}^{N} w_n^{(k)} \mathbb{1}\left(y_k(x_n) \neq t_n\right)$$

② Find α_k

$$\frac{\partial E}{\partial \alpha_k} = 0 = \left(\frac{1}{2} \underbrace{\boxed{e^{\alpha_k/2}}}_{x} + \frac{1}{2} e^{-\alpha_k/2}\right) \underbrace{\boxed{\sum_{n=1}^{N} w_n^{(k)} \mathbb{1}\left(y_k(x_n) \neq t_n\right)}}_{a} - \frac{1}{2} e^{-\alpha_k/2} \underbrace{\boxed{\sum_{n=1}^{N} w_n^{(k)}}}_{b}$$

$$\left(x + \frac{1}{x}\right) a - \frac{b}{x} = 0 \Rightarrow x^2 = \frac{b-a}{a}$$

$$e^{\alpha_k} = \frac{b-a}{a}$$

$$\alpha_k = \ln\left(\frac{b-a}{a}\right)$$

In class, $\varepsilon_k = \dfrac{\sum_{n=1}^{N} w_n^{(k)} \mathbb{1}\left(y_k(x_n) \neq t_n\right)}{\sum_{n=1}^{N} w_n^{(k)}} = \dfrac{a}{b}$

$$\alpha_k = \ln\left(\frac{1-\varepsilon_k}{\varepsilon_k}\right) = \ln\left(\frac{1-\frac{a}{b}}{\frac{a}{b}}\right) = \ln\left(\frac{b-a}{a}\right)$$

③ How to update the weights

$$w_n^{(k+1)} = w^{(k)} e^{-\frac{1}{2} t_n \alpha_k y_k(x_n)}$$

$$t_n y_k(x_n) = 1 - 2\mathbb{1}(y_k(x) \neq t_n)$$

$$w_n^{(k+1)} = w_n^{(k)} e^{-\frac{1}{2}\alpha_k} e^{\alpha_k \mathbb{1}(y_k(x_n) \neq t_n)}$$