EM Derivation.

Seylom Ayivi
HW4
CS 543.

$$E(x|m,\theta) = \sum_m \log\left(p(x,m|\theta)\right) \cdot P(m|x,\theta^+)$$

$$= \sum_m \log\left(p(x|m,\theta)\cdot p(m|\theta)\right)\cdot P(m|x,\theta^+)$$

$j \in S_i$
Set of Annotators for image i

$$= \sum_m \left[\sum_i \sum_{\substack{j \\ j\in S_i}} \log\left(p(x_{ij}|m_j=k,\theta)\cdot p(m_j|\theta)\right)\right]\cdot P(m|x,\theta^+)$$

$$= \sum_i \sum_j \sum_{m_1} \cdots \sum_{m_n} \sum_{m_j} \left[\log\left(p(x_{ij}|m_j=k,\theta)\cdot p(m_j|\theta)\right)\right]\cdot P(m|x,\theta^+)$$

$$= \sum_i \sum_j \sum_{m_j} \sum_{\substack{m_q \\ q\neq j}} \left[\log\left(p(x_{ij}|m_j=k,\theta)\cdot p(m_j=k|\theta)\right)\right]\cdot P(m|x,\theta^+)$$

$$E(x|m,\theta) = \sum_i \sum_j \sum_{m_j} \log\left(p(x_{ij}|m_j=k,\theta)\cdot p(m_j=k|\theta)\right) \sum_{\substack{m_q \\ q\neq j}} P(m|x,\theta^+) \quad (1)$$

Let's expand the term $\sum_{\substack{m_q \\ q\neq j}} P(m|x,\theta^+)$

$$\sum_{\substack{m_q \\ q\neq j}} P(m|x,\theta^+) = \sum_{\substack{m_q \\ q\neq j}} \prod_i \prod_j P(m_j|x_{ij},\theta^+)$$

$$= \prod_i \prod_j \sum_{\substack{m_q \\ q\neq j}} P(m_j|x_{ij},\theta^+)$$

$$= \prod_i \prod_q \sum_{\substack{m_q \\ q\neq j}} P(m_q|x_{iq},\theta^+)$$ ← just for clarity, I changed the indice to q to show what will happen when the j^{th} term is extracted

Seylom

$$\sum_{\substack{m_q \\ q\neq j}} P(m|x,\theta^t) = \sum_{\substack{m_q \\ q\neq j}} \cdots\cdots$$

$$= \prod_i \prod_q \sum_{\substack{m_q \\ q\neq j}} P(m_q|x_{iq},\theta^t)$$

$$= \prod_i \prod_{q\neq j} \left(\sum_{\substack{m_q \\ q\neq j}} P(m_q|x_{iq},\theta^t)\right) \cdot P(m_j|x_{ij},\theta^t)$$

We just extracted the jth term from the product.

$$\sum_{\substack{m_q \\ q\neq j}} P(m_q|x_{iq},\theta^t) = 1 \quad \text{So replacing in that equation}$$

$$\sum_{\substack{m_q \\ q\neq j}} P(m|x,\theta^t) = \prod_i \prod_q (\times 1) \cdot P(m_j|x_{ij},\theta^t) = P(m_j|x_{ij},\theta^t)$$

replacing in (1) $\Rightarrow$

$$E(x|\theta,m) = \sum_i \sum_{\substack{j \\ j\in S_i}} \sum_{m_j} \log\left(P(x_{ij}|m_j=k,\theta) \cdot P(m_j=k|\theta)\right) P(m_j|x_{ij},\theta^t)$$

$m_j \in \{0,1\} \Rightarrow$

$$\boxed{E(x|m,\theta) = \sum_i \sum_{\substack{j \\ j\in S_i}} \sum_{k=0}^{1} \log\left(P(x_{ij}|m_j=k,\theta) \cdot P(m_j=k,\theta)\right) \cdot P(m_j=k|x_{ij},\theta^t)}$$

Here $j \in S_i$, Set of Annotators for image j

Seylom
CS 54
HW4

$$E = \sum_i \sum_j \sum_{m_j} \log\left(P(x_{ij} \mid m_j = k, \theta) \cdot P(m_j = k \mid \theta)\right) \cdot \underbrace{P(m_j = k \mid x_{ij}, \theta^t)}_{\alpha_{jk}}$$

$$E = \sum_i \sum_j \left[\alpha_{j1} \log\left(\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_{ij}-\mu_i)^2}{2\sigma^2}\right) \beta\right) + \log\left(\frac{1-\beta}{10}\right)\alpha_{j0}\right]$$

$$\alpha_{j1} = P(m_j = 1 \mid x_{ij}, \theta^t) \quad \text{and} \quad \alpha_{j0} = P(m_j = 0 \mid x_{ij}, \theta^t)$$

$$\frac{\partial E}{\partial \mu_i} = \sum_j \alpha_{j1}\left[\frac{x_{ij}-\mu_i}{\sigma^2}\right] = 0$$

($\sum_i$ disappears since we only derive with respect to μ_i only the ith term is left.)

$$\Rightarrow \frac{\partial E}{\partial \mu_i} = \sum_i \sum_j x_{ij}\alpha_{j1} - \sum_i \sum_j \alpha_{j1}\mu_i = 0$$

$$\boxed{\mu_i = \left(\sum_{j \in S_i} x_{ij}\alpha_{j1}\right) \cdot \frac{1}{\sum_{j \in S_i} \alpha_{j1}}}$$

$j \in S_i$
S_i: annotators for image i

$$\frac{\partial E}{\partial \sigma} = \sum_i \sum_j \left[\left(-\frac{1}{\sigma} + \frac{(x_{ij}-\mu_i)^2}{\sigma^3}\right) \cdot \alpha_{j1}\right]$$

$$\sum_i \sigma^2 \sum_j \alpha_{j1} = \sum_i \left(\sum_j \alpha_{j1}(x_{ij}-\mu_i)^2\right)$$

$$\sum_i \sigma^2 \sum_j \alpha_{j1} = \sum_i \sum_j \alpha_{j1}(x_{ij}-\mu_i)^2$$

$$\sigma^2 = \frac{\sum_i \sum_j (x_{ij}-\mu_i)^2 \alpha_{j1}}{\sum_i \sum_j \alpha_{j1}}$$

$$\rightarrow \boxed{\sigma = \left[\frac{\sum_i \sum_j (x_{ij}-\mu_i)^2 \alpha_{j1}}{\sum_i \left(\sum_j \alpha_{j1}\right)}\right]^{1/2}}$$

$j \in S_i$:
S_i annotators for image i

Seylom
HW4

$$\frac{\partial E}{\partial \beta} = \sum_i \sum_j \left[\frac{\alpha_{j1}}{\beta} - \frac{1}{1-\beta}\alpha_{j0} \right] = 0$$

$$\Rightarrow \sum_i \sum_j \frac{\alpha_{j1}}{\beta} - \sum_i \sum_j \frac{\alpha_{j0}}{1-\beta} = 0$$

$$\sum_i \sum_j (1-\beta)\alpha_{j1} - \sum_i \sum_j \beta \cdot \alpha_{j0} = 0$$

$$\sum_i \sum_j (\alpha_{j0} + \alpha_{j1})\beta = \sum_i \sum_j \alpha_{j1}$$

$$\boxed{\beta = \frac{\sum_i \sum_j (\alpha_{j1})}{\sum_i \sum_j (\alpha_{j0} + \alpha_{j1})}}$$

$j \in S_i$: S_i the set of annotators for image i

$$\alpha_{j0} = P(m_j = 0 \mid x_{ij}, \theta^t) = \frac{P(x_{ij} \mid m_j = 0, \theta^t) \cdot P(m_j = 0 \mid \theta^t)}{P(m_j = 0 \mid \theta^t) \cdot P(x_{ij} \mid m_j = 0, \theta^t) + P(m_j = 1 \mid \theta^t) \cdot P(x_{ij} \mid m_j = 1, \theta^t)}$$

Seyhan

Similarly

$$\alpha_{j1} = P(m_j = 1 \mid x_{ij}, \theta^t) = \frac{P(x_{ij} \mid m_j = 0, \theta^t) \cdot P(m_j = 0 \mid \theta^t)}{P(m_j = 0 \mid \theta^t) \cdot P(x_{ij} \mid m_j = 0, \theta^t) + P(m_j = 1 \mid \theta^t) \cdot P(x_{ij} \mid m_j = 1, \theta^t)}$$

$$\boxed{P(m_j = 0 \mid x_{ij}, \theta^t) = 1 - P(m_j = 1 \mid x_{ij}, \theta^t)}$$

$$\boxed{P(m_j = 1 \mid x_{ij}, \theta^t) = \frac{\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_{ij} - \mu_i)^2}{\sigma^2}\right) \cdot \beta}{\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_{ij} - \mu_i)^2}{\sigma^2}\right)\beta + (1-\beta) \cdot 1/10}}$$