Jensens Inequality

$$f(\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3) \geq \alpha_1 f(x_1) + \alpha_2 f(x_2) + \alpha_3 f(x_3)$$

Kullback - Leibler

$$KL(q \| p) = \int q(x) \log \frac{q(x)}{p(x)} dx$$

1. $KL(q \| p) \neq KL(p \| q)$
2. $KL(q \| q) = 0$
3. $KL(q \| p) \geq 0$

$$p(X \mid \theta) = \prod_{i=1}^{N} p(x_i \mid \theta)$$

$$\log p(X \mid \theta) = \sum_{i=1}^{N} \log p(x_i \mid \theta)$$

$$= \sum_{i=1}^{N} \log \sum_{c=1}^{3} p(x_i, t_i = c \mid \theta) \geq \underbrace{\mathcal{L}(\theta)}_{\text{lower bound}}$$

rewrite...

$$= \sum_{i=1}^{N} \log \sum_{c=1}^{3} \frac{q(t_i = c)}{q(t_i = c)} p(x_i, t_i = c \mid \theta)$$

Jensens

$$\geq \sum_{i=1}^{N} \sum_{c=1}^{3} q(t_i = c) \log \frac{p(x_i, t_i = c \mid \theta)}{q(t_i = c)} = \mathcal{L}(\theta, q)$$

E-step

close gap btw $p(X|\theta)$ and lower bound $\mathcal{L}(\theta, q)$

$$\sum_{i=1}^{N} \log p(x_i|\theta) \quad - \sum_{i=1}^{N}\sum_{c=1}^{3} q(t_i=c) \log \frac{p(x_i, t_i=c|\theta)}{q(t_i=c)}$$

$$\sum_{i=1}^{N} \log p(x_i|\theta) \cdot \sum_{t=1}^{3} q(t_i=c) - \sum_{i=1}^{N}\sum_{c=1}^{3} q(t_i=c) \log \frac{p(x_i, t_i=c|\theta)}{q(t_i=c)}$$

$$\sum_{n=1}^{N}\sum_{t=1}^{3} q(t_i=c) \left(\log p(x_i|\theta) - \log \frac{p(x_i, t_i=c|\theta)}{q(t_i=c)} \right)$$

$$\sum_{n=1}^{N}\sum_{t=1}^{3} q(t_i=c) \left(\log \frac{p(x_i|\theta) q(t_i=c)}{p(x_i, t_i=c|\theta)} \right)$$

$$\sum_{n=1}^{N}\sum_{t=1}^{3} q(t_i=c) \left(\log \frac{p(x_i|\theta) q(t_i=c)}{p(t_i=c|x_i, \theta) p(x_i|\theta)} \right)$$

$$\sum_{n=1}^{N}\sum_{t=1}^{3} q(t_i=c) \log \left(\frac{q(t_i=c)}{p(t_i=c|x_i, \theta)} \right)$$

$$\sum_{n=1}^{N} KL(q(t_i) \,||\, p(t_i|x_i, \theta))$$

To lower... minimum at $q(t_i) = p(t_i|x_i, \theta)$

M-step

$$\mathcal{L}(\theta, q) = \sum_i \sum_c q(t_i=c) \log \frac{p(x_i, t_i=c|\theta)}{q(t_i=c)}$$

$$= \sum_i \sum_c q(t_i=c) \log p(x_i, t_i=c|\theta) - \sum_i \sum_c q(t_i=c) \log q(t_i=c)$$

$$= \sum_i \sum_c q(t_i=c) \log p(x_i, t_i=c|\theta) - \text{constant}$$

?

$$= E_q \log p(X, T|\theta) + \text{constant}$$

$$p(x_i|t_i=c)\, p(t_i=c)$$

Convergence guarantee

$$\log p(X|\theta^{k+1}) \geq \mathcal{L}(\theta^{k+1}, q^{k+1}) \geq \mathcal{L}(\theta^{k}, q^{k+1}) = \log p(X|\theta^{k})$$

GMM

$$p(x_i \mid t_i = c, \theta) = \mathcal{N}(\Pi_c, \Sigma_c)$$
$$P(t_i = c) = \pi_c$$

$$\sum \log \underline{p(x_i \mid t_i = c, \theta)\, p(t_i = c)}$$

$$\sum_{c=1}^{3} \sum_{i=1}^{N} g(t_i = c)\left(\log p(x_i \mid t_i = c) + \log p(t_i = c)\right)$$

$$\frac{d}{d\Pi} \quad \sum_{i=1}^{N} g(t_i = 1)\frac{1}{2}(x_i - \Pi_1)^T \Sigma^{-1} (x_i - \Pi_1) = 0$$

$$\sum_{i=1}^{N} -\frac{1}{2} g(t_i = 1)\, 2\, \Sigma^{-1} (x_i - \Pi_1) = 0$$

$$\sum_{i=1}^{N} \cancel{+1}\, g(t_i = 1) \Sigma^{-1} x_i - \sum_{i=1}^{N} g(t_i = 1) \Sigma^{-1} \Pi_1$$

$$\sum_{i=1}^{N} g(t_i = 1) x_i - \sum_{i=1}^{N} g(t_i = 1) \Pi_1 = 0$$

$$\Pi_1 = \frac{\sum_{i=1}^{N} g(t_i = 1) x_i}{\sum_{i=1}^{N} g(t_i = 1)}$$

$$g(t_i = 1) = p(t_i = 1 \mid x_i, \theta)$$

$$\boxed{\Pi_1 = \frac{\sum_i p(t_i = 1 \mid x_i, \theta) x_i}{\sum_i p(t_i = 1 \mid x_i, \theta)}}$$

$$\frac{d}{d\Sigma^{-1}} \quad \sum_{i=1}^{N} g(t_i = 1)\left(\frac{1}{2} \log |\Sigma^{-1}|\right) - g(t_i = 1) \frac{1}{2} (x_i - \Pi_1)^T \Sigma^{-1} (x_i - \Pi_1)$$

$$\sum_{i=1}^{N} g(t_i = 1) \frac{1}{|\Sigma^{-1}|} |\Sigma^{-1}| \Sigma^{-1^{-1^T}} - g(t_i = 1)(x_i - \Pi_1)^T (x_i - \Pi_1) = 0$$

$$\sum_{i=1}^{N} g(t_i = 1) \Sigma - \sum_{i=1}^{N} g(t_i = 1)(x_i - \Pi_1)^T (x_i - \Pi_1)$$

$$\boxed{\Sigma = \frac{\sum_{i=1}^{N} g(t_i = 1)(x_i - \Pi_1)^T (x_i - \Pi_1)}{\sum_{i=1}^{N} g(t_i = 1)}}$$

$$\pi_1 + \pi_2 + \pi_3 = 1$$

$$\sum_{c=1}^{3} \sum_{i=1}^{N} g(t_i = c)\left(\log(p(x_i \mid t_i = c) + \log p(t_i = c)\right) + \lambda(\pi_1 + \pi_2 + \pi_3 - 1)$$

$$\frac{d}{d\pi_1} \quad \sum_{i=1}^{N} g(t_i = 1) \frac{1}{\pi_1} + \lambda = 0$$

$$\frac{d}{d\pi_2} \quad \sum_{i=1}^{N} g(t_i = 2) \frac{1}{\pi_2} + \lambda = 0$$

$$\frac{d}{d\pi_3} \quad \sum_{i=1}^{N} g(t_i = 3) \frac{1}{\pi_3} + \lambda = 0$$

$$\pi_1 + \pi_2 + \pi_3 = 1$$

$$\pi_1 = -\frac{\sum_{i=1}^{N} g(t_i = 1)}{\lambda}$$

$$-\frac{\sum_{i=1}^{N} g(t_i = 1)}{\lambda} - \frac{\sum_{i=1}^{N} g(t_i = 2)}{\lambda} - \frac{\sum_{i=1}^{N} g(t_i = 3)}{\lambda} = 1$$

$$\lambda = -\left(\sum_{i=1}^{N} g(t_i = 1) + g(t_i = 2) + g(t_i = 3)\right)$$

$$\pi_1 = \frac{\sum_{i=1}^{N} g(t_i = 1)}{\sum_{i=1}^{N} g(t_i = 1) + g(t_i = 2) + g(t_i = 3)} = \boxed{\frac{\sum_{i=1}^{N} g(t_i = 1)}{N} = \pi_1}$$