PS 4 1) 1/3 ~~4~~

1.) EM for MAP

$$\ell(\theta) = \sum_{i=1}^{m} \log \sum_{z^{(i)}} p(x^{(i)}, z^{(i)} | \theta) \cdot p(\theta)$$

$$= m \cdot \log(p(\theta)) + \cancel{\log} \sum_{i=1}^{m} \log \sum_{z^{(i)}} p(x^{(i)} | z^{(i)}, \theta) \cdot p(z^{(i)} | \theta)$$

New start:

$$\sum_i \log p(x^{(i)}; \theta) \cdot p(\theta) = \sum_i \log \sum_{z^{(i)}} p(x^{(i)} | z^{(i)}; \theta) \cdot p(z^{(i)} | \theta) \cdot p(\theta)$$

Now: $\sum_z Q_i(z) = 1$ (some distr over z)

$$= \sum_i \log \sum_{z^{(i)}} Q_i(z^{(i)}) \frac{p(x^{(i)}, z^{(i)}; \theta) \cdot p(\theta)}{Q_i(z^{(i)})}$$

W/ Jensen: ~~$f(E(X)) \geq E(f(x))$~~ $E(f(X)) \geq f(E(x))$ if f convex

since log concave:

$$\geq \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)}; \theta) \cdot p(\theta)}{Q_i(z^{(i)})}$$

$\Rightarrow$ to make lower bound tight: $Q_i(z^{(i)}) = p(x^{(i)}, z^{(i)}; \theta) \cdot \cancel{p(\theta)}$

(so log term is const)

$$\frac{p(x^{(i)}, z^{(i)}; \theta) \cdot p\theta}{Q_i(z^{(i)})}$$

PS4 1 2/3

needs to be constant for all $z^{(i)}$

$$\Rightarrow Q_i(z^{(i)}) = \frac{p(x^{(i)}, z^{(i)}; \theta)}{\sum_z p(x^{(i)}, z; \theta)} = p(z^{(i)} | x^{(i)}; \theta)$$

$$\Rightarrow \quad = \sum_z p(x^{(i)}, z | \theta) \cdot p(\theta) = \text{const for any } z^{(i)}$$

$\Rightarrow$ E-Step: for each i:

$$Q_i(z^{(i)}) = p(z^{(i)} | x^{(i)}; \theta)$$

M-Step

$$\theta := \arg\max_\theta \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \cdot \log \frac{p(x^{(i)}, z^{(i)}; \theta) \cdot p(\theta)}{Q_i(z^{(i)})}$$

$$= \arg\max_\theta \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \cdot \left(\log \frac{p(x^{(i)}, z^{(i)}; \theta)}{Q_i(z^{(i)})} + \log(p(\theta)) \right)$$

$\Rightarrow$ M-Step = tractable, because

lin combo of $\log p(x, z | \theta)$ and $\log p(\theta)$

Monoton increase because:

~~$\ell(\theta^{(t+1)}) \geq \sum_i \sum_{z^{(i)}} Q_i$~~ PSh 1 3/3

Monoton incr. prove:

$$\ell(\theta^{(t)}) = \sum_i \sum_{z^{(i)}} Q_i^{(t)}(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)} \mid \theta^{(t)}) \cdot p(\theta)}{Q_i^{(t)}(z^{(i)})}$$

in $t+1$ we maximize right side w/ θ

$\Rightarrow$

$$\ell(\theta^{(t+1)}) \geq \sum_i \sum_{z^{(i)}} Q_i^{(t)}(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)} \mid \theta^{(t+1)})}{Q_i^{(t)}(z^{(i)})} \cdot p(\theta)$$

holds because $\ell(\theta) \geq \sum_i \sum_{z_i} \dots$ for any θ and Q_i

$$\geq \sum_i \sum_{z^{(i)}} Q_i^{(t)}(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)}; \theta^{(t)})}{Q_i^{(t)}(z^{(i)})} \cdot p(\theta)$$

$= \ell(\theta^{(t)})$, holds because we set $Q_i^{(t)}$ that way