Convergence Guarantee

Goal: maximizing "Incomplete" likelihood: $L(\theta) = \log p(X|\theta)$

I.e., choosing $\theta^{(n)}$, so that $L(\theta^{(n)}) - L(\theta^{(n-1)}) \geq 0$

Note that, since $p(X,H|\theta) = p(H|X,\theta) P(X|\theta)$, $L(\theta) = Lc(\theta) - \log p(H|X,\theta)$

$$L(\theta^{(n)}) - L(\theta^{(n-1)}) = Lc(\theta^{(n)}) - Lc(\theta^{(n-1)}) + \log [p(H|X,\theta^{(n-1)}) / p(H|X,\theta^{(n)})]$$

Taking expectation w.r.t. $p(H|X,\theta^{(n-1)})$,

$$L(\theta^{(n)}) - L(\theta^{(n-1)}) = Q(\theta^{(n)}; \theta^{(n-1)}) - Q(\theta^{(n-1)}; \theta^{(n-1)}) + D(p(H|X,\theta^{(n-1)}) || p(H|X,\theta^{(n)}))$$

Doesn't contain H

EM chooses $\theta^{(n)}$ to maximize Q

KL-divergence, always non-negative

Therefore, $L(\theta^{(n)}) \geq L(\theta^{(n-1)})$!