MLE: $P(X|\theta)$

$\theta_{MLE} = \arg\max_{\theta} \log P(X|\theta) \to$ log likelihood

$$P(A|B) = \frac{P(B|A)\cdot P(A)}{P(B)}$$

$$\theta^{(t+1)} = \arg\max_{\theta} \underbrace{\int_z \log P(x,z|\theta)\cdot P(z|x,\theta^{(t)})\,dz}_{E_{z|x,\theta^{(t)}}[\log P(x,z|\theta)]}$$

$\theta^{(t)} \to \theta^{(t+1)}$

要证 $\log P(x|\theta^{(t)}) \le \log P(x|\theta^{(t+1)})$

$$= P(x) = \frac{P(x,z)}{P(z|x)}$$

$$\log P(x|\theta) = \log P(x,z|\theta) - \log P(z|x,\theta)$$

left: $= \int_z P(z|x,\theta^{(t)})\cdot \log P(x|\theta)\,dz$

$= \log P(x|\theta) \underbrace{\int_z P(z|x,\theta^{(t)})\,dz}_{1}$

$= \log P(x|\theta)$

right $= \underbrace{\int_z P(z|x,\theta^{(t)})\cdot \log P(x,z|\theta)}_{Q(\theta,\theta^{(t)})} - \underbrace{\int_z P(z|x,\theta^{(t)})}_{H(\theta,\theta^{(t)})}$

$$D_{KL}(P||Q) = \int_{-\infty}^{\infty} P(x)\log\left(\frac{P(x)}{q(x)}\right)dx$$

$$D_{KL}(P||Q) = \int_{-\infty}^{\infty} P(x)\log\left(\frac{q(x)}{P(x)}\right)dx$$

$Q(\theta^{(t+1)},\theta^{(t)}) \ge Q(\theta^{(t)},\theta^{(t)})$

$H(\theta^{(t+1)},\theta^{(t)}) \le H(\theta^{(t)},\theta^{(t)})$

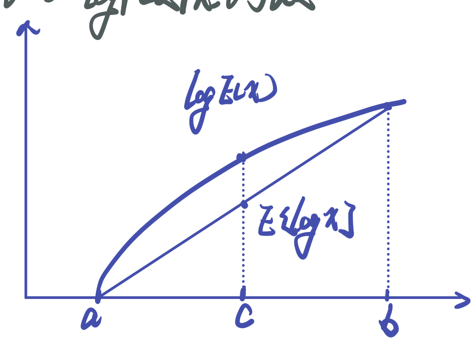
$H(\theta^{(t+1)},\theta^{(t)}) - H(\theta^{(t)},\theta^{(t)})$

$= \int_z P(z|x,\theta^{(t)})\cdot \log P(z|x,\theta^{(t+1)})\,dz - \int_z P(z|x,\theta^{(t)})\cdot \log P(z|x,\theta^{(t)})\,dz$

$= \int_z P(z|x,\theta^{(t)})\cdot \log\left(\frac{P(z|x,\theta^{(t+1)})}{P(z|x,\theta^{(t)})}\right)dz$

$= -KL(P(z|x,\theta^{(t)})||P(z|x,\theta^{(t+1)})) \le 0$

$E[\log x] \le \log E(x) \le \underbrace{\log}_{0} \underbrace{\int_z P(z|x,\theta^{(t+1)})\,dz}_{1} = \log 1 = 0$



$$= \int_z P(z|x,\theta^{(t)}) \cdot \log\left(\frac{P(z|x,\theta^{(t+1)})}{P(z|x,\theta^{(t)})}\right) dz$$

$$\le \log \int_z \cancel{P(z|x,\theta^{(t)})} \cdot \log\left(\frac{P(z|x,\theta^{(t+1)})}{\cancel{P(z|x,\theta^{(t)})}}\right) dz = \log \underbrace{\int_z P(z|x,\theta^{(t+1)}) dz}_{=\log 1 = 0}$$

EM公式:

$$\theta^{(t+1)} = \arg\max_{\theta} \int_z \log P(x,z|\theta) \cdot P(z|x,\theta^{(t)}) dz$$

E-step: $P(z|x,\theta^{(t)}) \longrightarrow E_{z|x,\theta^{(t)}}[\log P(x,z|\theta)]$

M-step: $\theta^{(t+1)} = \arg\max_{\theta} E_{z|x,\theta^{(t)}}[\log P(x,z|\theta)]$

$$\log P(x|\theta) = \log P(x,z|\theta) - \log P(z|x,\theta)$$
$$= \log \frac{P(x,z|\theta)}{q(z)} - \log \frac{P(z|x,\theta)}{q(z)}$$

左边 $= \int_z q(z) \cdot \log P(x|\theta) dz = \log P(x|\theta) \cdot \overbrace{\int_z q(z) dz}^{1} = \log P(x|\theta)$

右边 $= \underbrace{\int_z q(z) \log \frac{P(x,z|\theta)}{q(z)} dz}_{\substack{ELBO \\ = \text{evidence lower bound}}} - \underbrace{\int_z q(z) \log \frac{P(z|x,\theta)}{q(z)} dz}_{KL(q(z)\|P(z|x,\theta))}$

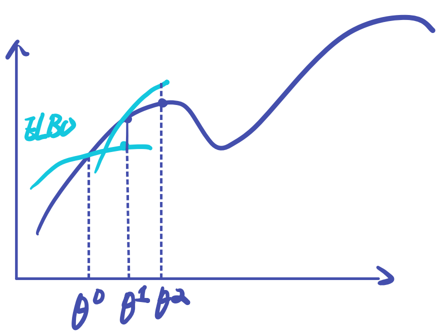
$$\log P(x|\theta) = ELBO + \underline{KL(q\|\overset{\text{posterior}}{P})} \ge 0 \iff \text{“=”}\ q = P$$

$\log P(x|\theta) \ge ELBO$

$\because$

$\therefore$ 取ELBO的最大值可以使$\log P(x|\theta)$最大

$$\hat{\theta} = \arg\max_{\theta} ELBO$$
$$= \arg\max_{\theta} \int q(z) \log \frac{P(x,z|\theta)}{q(z)} dz$$



$$= \arg\max_{\theta} \int P(z|x,\theta^{(t)}) \log \frac{P(x,z|\theta)}{P(z|x,\theta^{(t)})} dz$$

$$= \arg\max_{\theta} \int P(z|x,\theta^{(t)}) [\log P(x,z|\theta) - \underbrace{\log P(z|x,\theta^{(t)})}_{\text{与}\theta\text{无关}}] dz$$

$$= \arg\max_{\theta} \int P(z|x,\theta^{(t)}) \cdot \log P(x,z|\theta)$$