Math 789 - VAE Homework Assignment

April 13th, 2020 hv8322@rit.edu

Q.1) Given:

$$\log p_\theta(x^{(i)}) = D_{KL}(q_\phi(z|x^{(i)}) \,\|\, p_\theta(z|x^{(i)})) + \mathcal{L}(\theta, \phi; x^{(i)})$$

We know about KL Divergence that,

$$D_{KL}(P\|Q) = H(P,Q) - H(P)$$

where, $H(P,Q) = E_{x\sim p}[-\log Q(x)]$

$$H(P) = E_{x\sim p}[-\log P(x)]$$

Therefore,

$$D_{KL}(P\|Q) = E_{x\sim p}[-\log Q(x)] - E_{x\sim p}[-\log P(x)]$$

$$= E_{x\sim p}[-\log Q(x) + \log P(x)]$$

$$D_{KL}(P\|Q) = E_{x\sim p}\left[\log \frac{P(x)}{Q(x)}\right] \quad \text{——— (1)}$$

Consider $\log p_\theta(x^{(i)}) = 0$ in given equation

Therefore,

$$\mathcal{L}(\theta, \phi; x^{(i)}) = -D_{KL}(q_\phi(z|x^{(i)}) \,\|\, p_\theta(z|x^{(i)}))$$

~~So~~ Using eqn (1) in above equation,

$$\mathcal{L}(\theta, \phi; x^{(i)}) = -E\left[\log \frac{q_\phi(z|x^{(i)})}{p_\theta(z|x^{(i)})}\right]$$

$$= E\left[-\log q_\phi(z|x^{(i)}) + \log p_\theta(z|x^{(i)})\right]$$

$$= E\left[-\log q_\phi(z|x^{(i)}) + \log \frac{p_\theta(z, x^{(i)})}{p(x^{(i)})}\right]$$

$$\boxed{\mathcal{L}(\theta, \phi; x^{(i)}) = E\left[-\log q_\phi(z|x^{(i)}) + \log p_\theta(x, z)\right]} \quad \text{--- (2)}$$

$\because$ $\mathcal{L}(\theta, \phi, x^{(i)})$ is lower bound,

$\therefore$ $\log p_\theta(x^{(i)}) \geq \mathcal{L}(\theta, \phi; x^{(i)})$

adding & subtracting $\log p_\theta(z)$ in (2),

$$\mathcal{L}(\theta, \phi; x^{(i)}) = E\left[-\log \frac{q_\phi(z|x^{(i)})}{p_\theta(z)} - \log p_\theta(z) + \log p_\theta(x, z)\right]$$

$$= E\left[-\log \frac{q_\phi(z|x^{(i)})}{p_\theta(z)}\right] + E\left[\log \frac{p_\theta(x, z)}{p_\theta(z)}\right]$$

from eq^n (1)

$$\boxed{\mathcal{L}(\theta, \phi; x^{(i)}) = -D_{KL}(q_\phi(z|x^{(i)}) \| p_\theta(z)) + E_{q_\phi(z|x^{(i)})}\left[\log p_\theta(x^{(i)}|z)\right]}$$