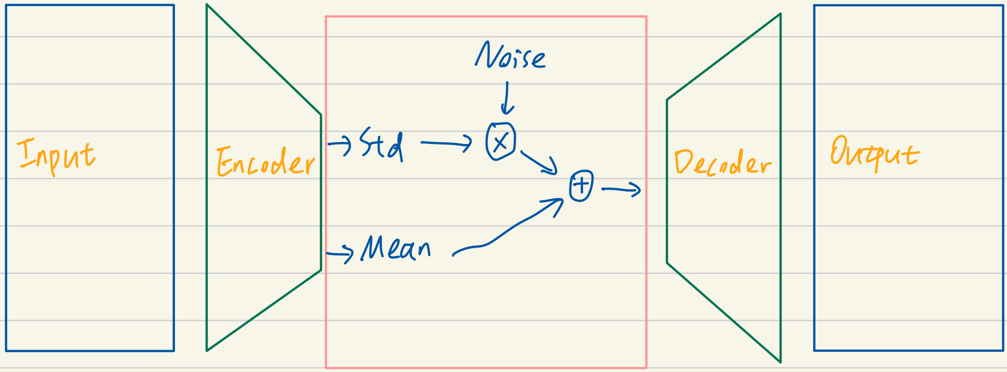


posterior distribution

observed data

$$p(z \mid x) = \frac{p(x \mid z)\, p(z)}{p(x)}$$

$p(x)$ → no closed form

hidden random variable

$$p(x) = \int p(x, z)\, dz = \int p(x \mid z)\, p(z)\, dz.$$

Find $q(z \mid \varphi) \approx p(z \mid x)$

KL divergence: $KL\left(q(z) \,\|\, p(z \mid x)\right) = E_{z \sim q}\left[\log \frac{q(z)}{p(z \mid x)}\right]$

$$q^* = \arg\min_q KL\left(q(z) \,\|\, p(z \mid x)\right)$$

ELBO $\quad \log P(x;\theta) \geq \boxed{E_{z\sim q}\left[\log \frac{p(x,z;\theta)}{q(z)}\right]}$

$$\log P(x;\theta) = \log \int p(x,z;\theta)\,dz$$

$$= \log \int p(x,z;\theta)\frac{q(z)}{q(z)}\,dz$$

Jensen's inequality

concave

$$= \log E_{z\sim q}\left[\frac{p(x,z)}{q(z)}\right]$$

$$\varphi(E[x]) \leq E[\varphi(x)]$$

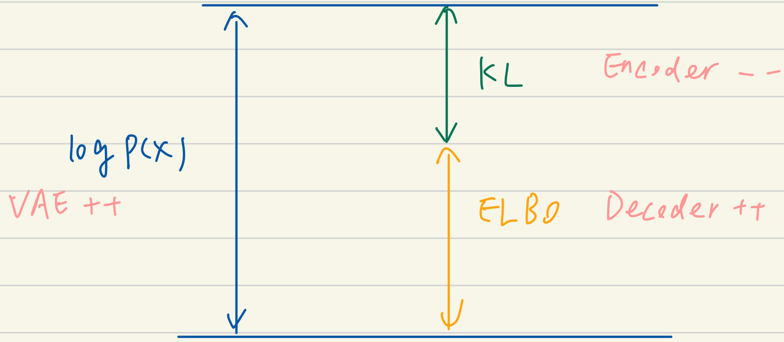
$$\geq E_{z\sim q}\left[\log \frac{p(x,z)}{q(z)}\right]$$

$$KL(q(z)\,\|\,p(z|x;\theta)) = E_{z\sim q}\left[\log \frac{q(z)}{p(z|x;\theta)}\right]$$

$$= E_{z\sim q}[\log q(z)] - E_{z\sim q}\left[\log \frac{p(x,z;\theta)}{p(x;\theta)}\right]$$

$$= E_{z\sim q}[\log q(z)] - E_{z\sim q}[\log p(x,z;\theta)] + E_{z\sim q}[\log p(x;\theta)]$$

$$= \underbrace{\log p(x;\theta)}_{\text{evidence}} - \underbrace{E_{z\sim q}\left[\log \frac{p(x,z;\theta)}{p(z)}\right]}_{\text{ELBO}}$$



$$\mathrm{ELBO} = E_{z\sim q}[\log P(x, z)] - E_{z\sim q}[\log q(z)]$$

$$q^* = \arg\max_q \mathrm{ELBO}(q) \Rightarrow q^* = \arg\min_q KL(q(z) \| q(z|x))$$

$$E[V(\varphi)] = \nabla_\varphi \mathrm{ELBO}(\varphi)$$

$$v \sim V(\varphi_t) \quad , \quad \varphi_{t+1} = \varphi_t + \gamma v$$

$$q_\varphi = N(\mu, \sigma^2) \quad , \quad \varphi = \{\mu, \sigma^2\}$$

$$\epsilon \sim N(0,1) , \; z = \mu + \sigma\epsilon \Longrightarrow z \sim N(\mu, \sigma^2) = q_\varphi(z)$$

$$\mathrm{ELBO}(\varphi) = E_{\epsilon\sim\psi}[\log P(x, g_\varphi(\epsilon) - \log q_\varphi(g_\varphi(\epsilon))]$$

$$\epsilon_1' \ldots \epsilon_L' \sim \psi$$

$$\mathrm{ELBO}(\varphi) = \frac{1}{L}\sum_l [\log P(x, g_\varphi(\epsilon_l')) - \log q_\varphi(g_\varphi(\epsilon_l'))]$$

$$\nabla_\varphi \mathrm{ELBO}(\varphi) = \nabla_\varphi \frac{1}{L}\sum_l [\log P(x, g_\varphi(\epsilon_l')) - \log q_\varphi(g_\varphi(\epsilon_l'))]$$

$$E[\nabla_\varphi \mathrm{ELBO}(\varphi)] = \nabla_\varphi \mathrm{ELBO}(\varphi)$$

$$\mathcal{L}(\theta, \phi; x^{(i)}) \approx \frac{1}{2}\sum_j \left(1 + \log((\sigma_j^{(i)})^2) - (\mu_j^{(i)})^2 - (\sigma_j^{(i)})^2\right)$$

$$+ \frac{1}{L}\sum_l \log p_\theta(x^{(i)} \mid z^{(i,l)})$$

$$z^{(i,l)} = \mu^{(i)} + \sigma^{(i)} \odot \epsilon^{(l)}, \quad \epsilon^{(l)} \sim \mathcal{N}(0, I)$$

$$KL(p_1 \| p_2) = \frac{1}{2}\left[\log\frac{|\Sigma_2|}{|\Sigma_1|} - n + tr\{\Sigma_2^{-1}\Sigma_1\} + (\mu_2 - \mu_1)^T \Sigma_2^{-1}(\mu_2 - \mu_1)\right]$$

$p_1 = q(z|x)$, $p_2 = p(z)$.

$\mu_1 = \mu$, $\Sigma_1 = \Sigma$

$\mu_2 = 0$, $\Sigma_2 = I$

$$= \frac{1}{2}\left[\log\frac{|I|}{|\Sigma|} - n + tr\{I^{-1}\Sigma\} + (0-\mu)^T I^{-1}(0-\mu)\right]$$

$$= \frac{1}{2}\left[-\log|\Sigma| - n + tr\{\Sigma\} + \mu^T\mu\right]$$

$$= \frac{1}{2}\left[-\log\prod_i \sigma_i^2 - n + \sum_i \sigma_i^2 + \sum_i \mu_i^2\right]$$

$$= \frac{1}{2}\left[-\sum_i \log\sigma_i^2 - n + \sum_i \sigma_i^2 + \sum_i \mu_i^2\right]$$

$$= \frac{1}{2}\left[-\sum_i(\log\sigma_i^2 + 1) + \sum_i \sigma_i^2 + \sum_i \mu_i^2\right]$$