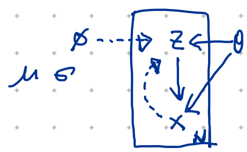
[VAE: Variational Auto-Encoder]



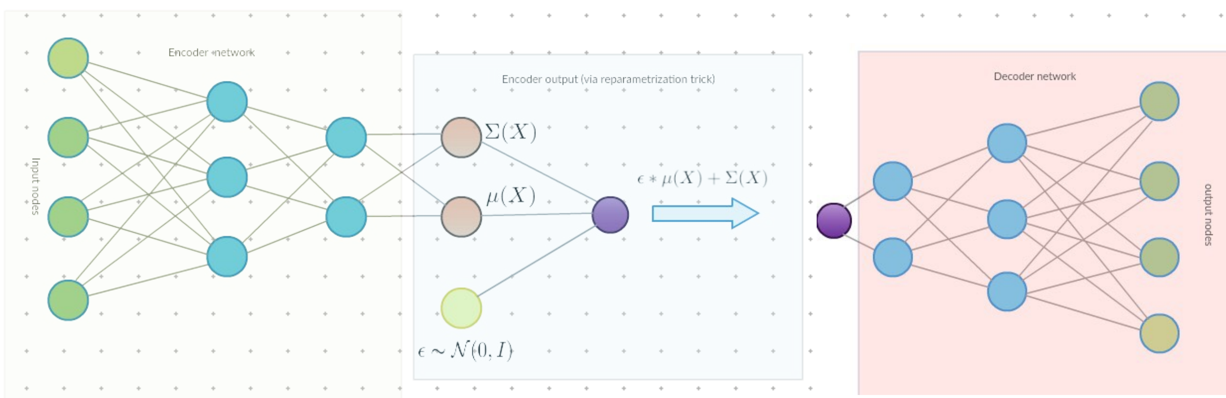
$$P(z|x) = \frac{P(x|z)\cdot P(z)}{P(x)}$$

① variation inference idea → optimization

② posterior $q(z|x)$ ⟺ likelihood $p(x|z)$ → auto encoder

$$D_{KL}\begin{cases} P_\theta(z|x) & \text{posterior} \\ q_\phi(z|x) & \text{variation posterior (variation approximation of posterior)} \end{cases}$$

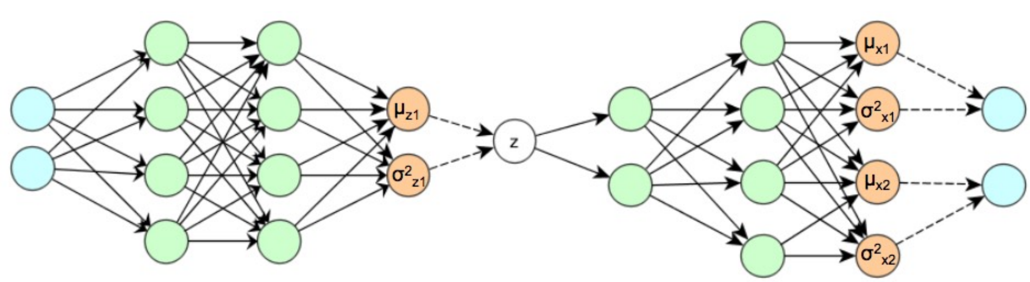
$$= N(z\,;\,\mu_z(x), \sigma_z(x))$$



$$D_{KL}\left(N(\text{Encoder}_\phi(X)) \,\|\, N(0,1)\right) \qquad \left\| X - \text{Decoder}_\theta\left(f(\text{Encoder}_\phi(X))\right)\right\|^2$$

where $\text{Encoder}_\phi(X) = (\mu(X), \Sigma(X))$

and $f(\mu(X), \Sigma(X)) = \mu(X) + \Sigma(X) * \varepsilon$



Cost: Regularisation

$$-D_{KL}\left(q(z|x^{(i)})\|p(z)\right) = \frac{1}{2}\sum_{j=1}^{J}\left(1 + \log(\sigma_{z_j}^{(i)^2}) - \mu_{z_j}^{(i)^2} - \sigma_{z_j}^{(i)^2}\right)$$

Cost: Reproduction

$$-\log\left(p(x^{(i)}|z^{(i)})\right) = \sum_{j=1}^{D}\frac{1}{2}\log(\sigma_{x_j}^2) + \frac{(x_j^{(i)} - \mu_{x_j})^2}{2\sigma_{x_j}^2}$$

$$D_{KL}\left(q_\phi(z|x) \,\|\, p_\theta(z|x)\right)$$

$$= \int q_\phi(z|x) \log \frac{q_\phi(z|x)}{p_\theta(z|x)} \qquad\qquad p_\theta(z|x) = \frac{p_\theta(x|z)\, p_\theta(z)}{p_\theta(x)}$$

$$= \int q_\phi(z|x) \log \frac{p_\theta(x)\; q_\phi(z|x)}{p_\theta(x|z) \cdot p_\theta(z)}$$

$$= \int q_\phi(z|x) \log p_\theta(x)\, dz - \int q_\phi(z|x) \cdot p_\theta(x|z)\, dz + \int q_\phi(z|x) \log \frac{q_\phi(z|x)}{p_\theta(z)}\, dz$$

$$= \log p_\theta(x) - \mathbb{E}_{q_\phi(z|x)}\left[p_\theta(x|z)\right] + D_{KL}\left(q_\phi(z|x) \,\|\, p_\theta(z)\right)$$

$$\therefore \log p_\theta(x) = \boxed{\mathbb{E}_{q_\phi(z|x)}\left[p_\theta(x|z)\right] - D_{KL}\left(q_\phi(z|x) \,\|\, p_\theta(z)\right)} + \boxed{D_{KL}\left(q_\phi(z|x) \,\|\, p_\theta(z|x)\right)}$$

ELBO : Evidence Lower BOund $\Uparrow$ maximization

≥ 0 $\Downarrow$ minimization

$$\mathcal{L}(\theta, \phi : x)$$

$$(\theta^*, \phi^*) = \underset{\theta\ \phi}{\operatorname{argmax}}\,(\theta, \phi : x)$$

$$\therefore \mathcal{L}(\theta, \phi : x) = \underbrace{\boxed{\mathbb{E}_{q_\phi(z|x)}\left[p_\theta(x|z)\right]}}_{\text{Reconstruction}} \underbrace{\boxed{-D_{KL}\left(q_\phi(z|x) \,\|\, p_\theta(z)\right)}}_{\text{Regularization}}$$

$\log(1)$, if $x^{(i)}$ gets always reconstructed perfectly (z produces $x^{(i)}$)

$$\approx \frac{1}{L} \sum_{l=1}^{L} \log p\left(x^{(i)} \mid z^{(i,l)}\right)$$

$$\approx \log p\left(x^{(i)} \mid z^{(i,l)}\right)$$

$$\simeq \sum_{j=1}^{D} \frac{1}{2} \log\left(\sigma_{x_j}^2\right) + \frac{\left(x_j^{(i)} - \mu_{x_j}\right)^2}{2\sigma_{x_j}^2}$$

$P(z)$ is always a simple prior $N(0,1)$

$$\therefore -D_{KL}\left[N(\mu(x), \Sigma(x)) \,\|\, N(0,1)\right]$$

$$= \frac{1}{2} \sum_{j=1}^{J} \left(1 + \log\left(\sigma_{z_j}^{(i)^2}\right) - \mu_{z_j}^{(i)^2} - \sigma_{z_j}^{(i)^2}\right)$$

(cf Numerical Approximation

$$\int p(x) \cdot f(x)\, dx = \mathbb{E}_{x \sim p(x)}[f(x)] \approx \frac{1}{k} \sum_{i=0}^{k} \left[f(x_i)\right]_{x_i \sim p(x)}$$
)