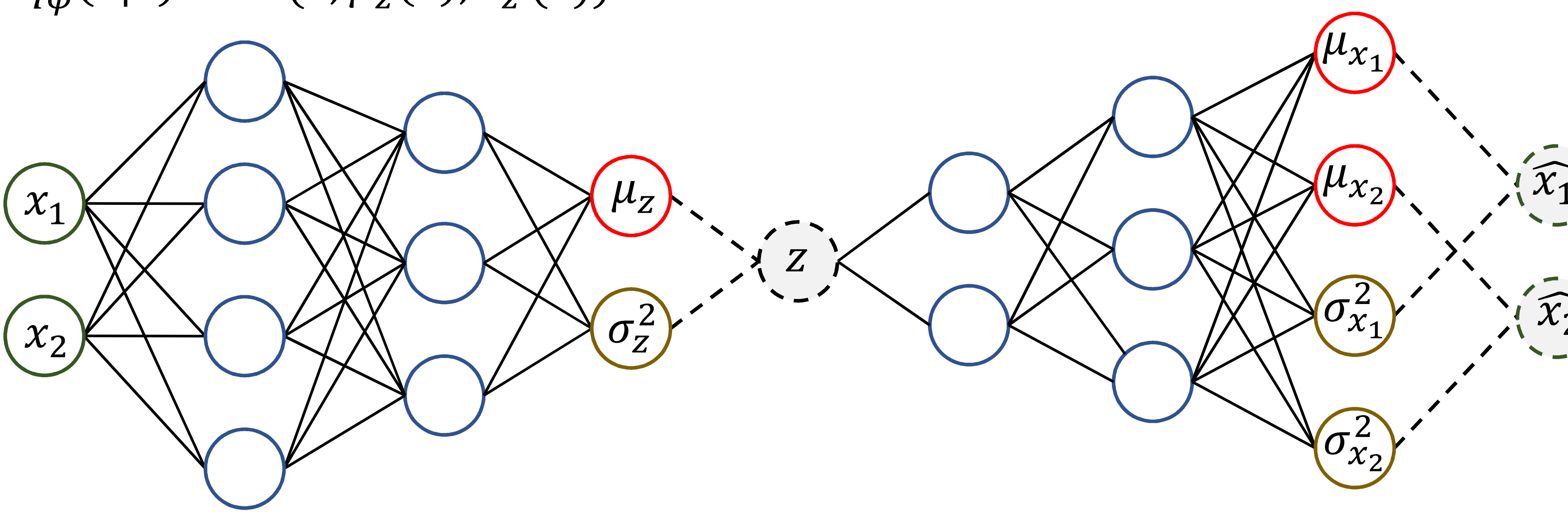
Lecture 23: Variational Autoencoder

Course Teacher: Md. Shariful Islam Bhuyan

Variational Autoencoder

$$q_\phi(z|\mathbf{x}) = \mathcal{N}(z; \mu_z(\mathbf{x}), \sigma_z^2(\mathbf{x}))$$

$$p_\theta(\mathbf{x}|z) = \mathcal{N}(\mathbf{x}; \mu_{\mathbf{x}}(z), \sigma_{\mathbf{x}}^2(z))$$



Loss Function

$$L = \log(p(x)) = \sum_z q(z|x) \log(p(x)) = \sum_z q(z|x) \log\left(\frac{p(z,x)}{p(z|x)}\right)$$
$$= \sum_z q(z|x) \log\left(\frac{p(z,x)}{q(z|x)} \frac{q(z|x)}{p(z|x)}\right)$$
$$= \sum_z q(z|x) \log\left(\frac{p(z,x)}{q(z|x)}\right) + \sum_z q(z|x) \log\left(\frac{q(z|x)}{p(z|x)}\right)$$
$$= \sum_z q(z|x) \log\left(\frac{p(x|z)p(z)}{q(z|x)}\right) + D_{KL}(q(z|x)||p(z|x))$$
$$= \sum_z q(z|x) \log\left(\frac{p(z)}{q(z|x)}\right) + \sum_z q(z|x) \log\big(p(x|z)\big) + D_{KL}(q(z|x)||p(z|x))$$
$$= -D_{KL}(q(z|x)||p(z)) + \mathbb{E}_{q(Z|x)} \log\big(p(x|z)\big) + D_{KL}(q(z|x)||p(z|x))$$

Implementation

$$-\mathbb{E}_{q(\mathbf{z}|\mathbf{x}_i)}\log\big(p(\mathbf{x}_i|\mathbf{z})\big) + D_{KL}(q(\mathbf{z}|\mathbf{x}_i)||p(\mathbf{z}))$$
$$= -\log(p(\mathbf{x}_i)) + D_{KL}(q(\mathbf{z}|\mathbf{x}_i)||p(\mathbf{z}|\mathbf{x}_i))$$

$$-D_{KL}(q(\mathbf{z}|\mathbf{x}_i)\|p(\mathbf{z})\big) = \frac{1}{2}\sum_{d}\big(1 + \log\sigma^2_{z_d,i} - \mu^2_{z_d,i} - \sigma^2_{z_d,i}\big) \quad p(\mathbf{z}) = \mathcal{N}(\mathbf{0},\mathbf{I})$$

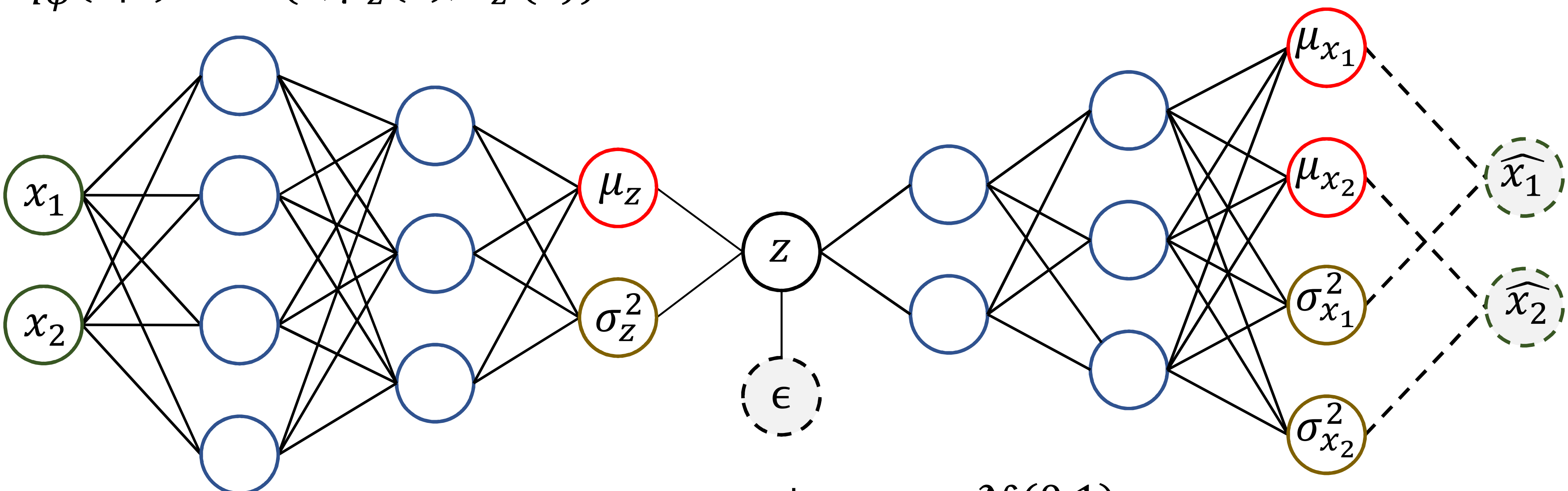
$$\mathbb{E}_{q(\mathbf{z}|\mathbf{x}_i)}\log\big(p(\mathbf{x}_i|\mathbf{z})\big) = \frac{1}{n_i}\sum_{j=1}^{n_i}\log\Big(p\big(\mathbf{x}_i|\mathbf{z}_j\big)\Big) \quad \mathbf{z}_j\sim\mathcal{N}\big(\boldsymbol{\mu}_{z,i},\boldsymbol{\sigma}^2_{z,i}\big), n_i \approx 1$$

$$\log\big(p(\mathbf{x}_i|\mathbf{z})\big) = \frac{1}{2}\sum_{d}\left(\log\sigma^2_{x_{i,d}} + \frac{\big(x_{i,d} - \mu^2_{x_{i,d}}\big)^2}{\sigma^2_{x_{i,d}}}\right)$$

Reparameterization Trick

$q_\phi(z|\mathbf{x}) = \mathcal{N}(z; \mu_z(\mathbf{x}), \sigma_z^2(\mathbf{x}))$

$p_\theta(\mathbf{x}|z) = \mathcal{N}(\mathbf{x}; \mu_{\mathbf{x}}(z), \sigma_{\mathbf{x}}^2(z))$



$z = \mu_{\mathbf{z}} + \epsilon\sigma_z;\ \epsilon \sim \mathcal{N}(0,1)$