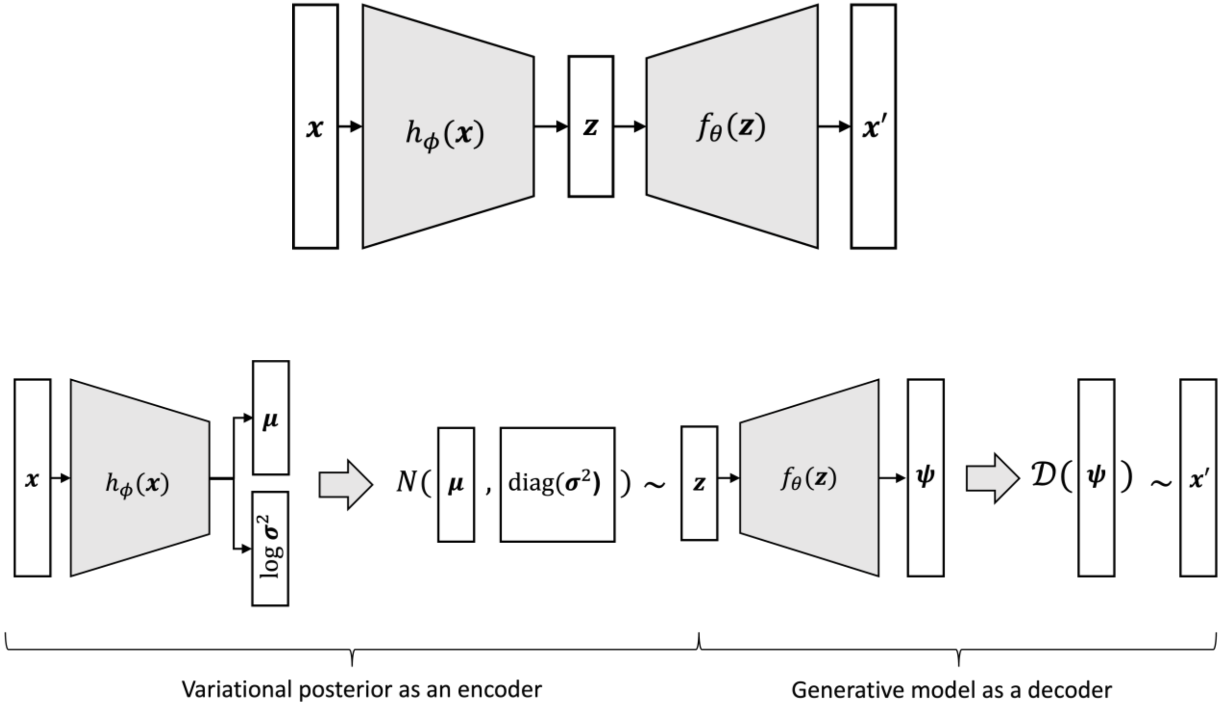
Fundamentals of Machine Learning

VARIATIONAL AUTOENCODERS

Amit K Roy-Chowdhury

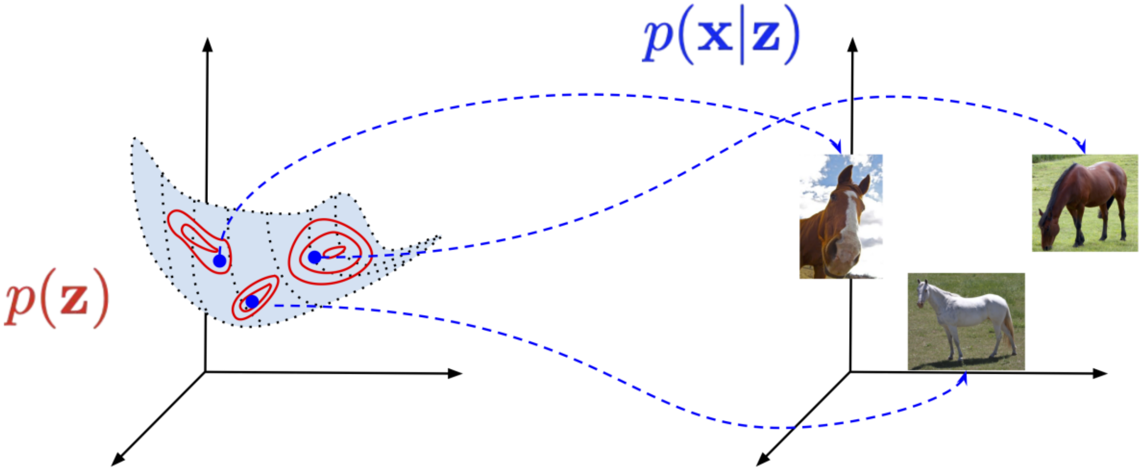
Acknowledgments: https://jmtomczak.github.io/blog/4/4_VAE.html

Autoencoders vs VAEs



Acknowledgments: https://mbernste.github.io/posts/vae/

VAEs as Generative Models



$$
\begin{aligned}
p(\mathbf{x}) &= \int p(\mathbf{x}|\mathbf{z})\, p(\mathbf{z})\, \mathrm{d}\mathbf{z} \\
&= \mathbb{E}_{\mathbf{z}\sim p(\mathbf{z})}\left[p(\mathbf{x}|\mathbf{z})\right] \\
&\approx \frac{1}{K}\sum_k p(\mathbf{x}|\mathbf{z}_k)
\end{aligned}
$$

curse of dimensionality

Evidence Lower Bound (ELBO)

$$\begin{aligned}
\ln p(\mathbf{x}) &= \ln \int p(\mathbf{x}|\mathbf{z})p(\mathbf{z})\, d\mathbf{z} \\
&= \ln \int \frac{q_\phi(\mathbf{z})}{q_\phi(\mathbf{z})} p(\mathbf{x}|\mathbf{z})p(\mathbf{z})\, d\mathbf{z} \\
&= \ln \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z})}\left[\frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{q_\phi(\mathbf{z})}\right] \\
&\geq \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z})} \ln\left[\frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{q_\phi(\mathbf{z})}\right] \\
&= \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z})}\left[\ln p(\mathbf{x}|\mathbf{z}) + \ln p(\mathbf{z}) - \ln q_\phi(\mathbf{z})\right] \\
&= \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z})}\left[\ln p(\mathbf{x}|\mathbf{z})\right] - \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z})}\left[\ln q_\phi(\mathbf{z}) - \ln p(\mathbf{z})\right]
\end{aligned}$$

Jensen's Inequality

convex function f

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i)$$

$\lambda_i \geq 0$ and $\sum_{i=1}^{n} \lambda_i = 1$

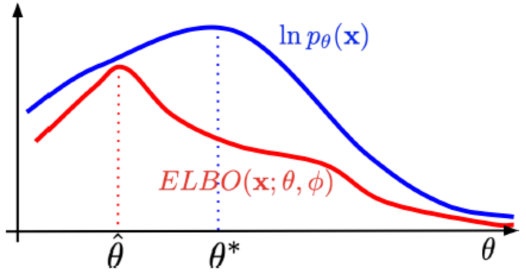
Evidence Lower Bound (ELBO)

Considering $q_\phi(\mathbf{z}|\mathbf{x})$ instead of $q_\phi(\mathbf{z})$

$$\ln p(\mathbf{x}) \geq \underbrace{\mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\left[\ln p(\mathbf{x}|\mathbf{z})\right]}_{\text{Reconstruction Error}} - \underbrace{\mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\left[\ln q_\phi(\mathbf{z}|\mathbf{x}) - \ln p(\mathbf{z})\right]}_{\text{KL Divergence}}$$

ELBO - Interpretation

$$
\begin{aligned}
\ln p(\mathbf{x}) &= \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\left[\ln p(\mathbf{x})\right] \\
&= \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\left[\ln \frac{p(\mathbf{z}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{z}|\mathbf{x})}\right] \\
&= \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\left[\ln \frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{z}|\mathbf{x})}\right] \\
&= \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\left[\ln \frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{z}|\mathbf{x})}\frac{q_\phi(\mathbf{z}|\mathbf{x})}{q_\phi(\mathbf{z}|\mathbf{x})}\right] \\
&= \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\left[\ln p(\mathbf{x}|\mathbf{z})\frac{p(\mathbf{z})}{q_\phi(\mathbf{z}|\mathbf{x})}\frac{q_\phi(\mathbf{z}|\mathbf{x})}{p(\mathbf{z}|\mathbf{x})}\right] \\
&= \mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\left[\ln p(\mathbf{x}|\mathbf{z}) - \ln \frac{q_\phi(\mathbf{z}|\mathbf{x})}{p(\mathbf{z})} + \ln \frac{q_\phi(\mathbf{z}|\mathbf{x})}{p(\mathbf{z}|\mathbf{x})}\right] \\
&= \underbrace{\mathbb{E}_{\mathbf{z}\sim q_\phi(\mathbf{z}|\mathbf{x})}\left[\ln p(\mathbf{x}|\mathbf{z})\right] - KL\left[q_\phi(\mathbf{z}|\mathbf{x})\|p(\mathbf{z})\right]}_{\text{ELBO}} + \underbrace{KL\left[q_\phi(\mathbf{z}|\mathbf{x})\|p(\mathbf{z}|\mathbf{x})\right]}_{\geq 0}
\end{aligned}
$$



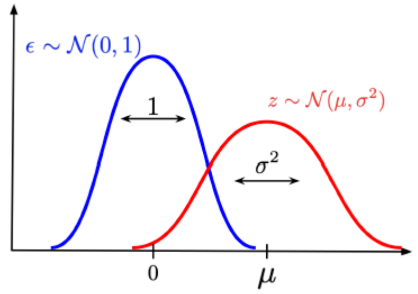
Reparametrization

How to choose the distribution of latent variables?

$$q_\phi(\mathbf{z}|\mathbf{x}) = \mathcal{N}\left(\mathbf{z}|\mu_\phi(\mathbf{x}), \mathrm{diag}\left[\sigma^2_\phi(\mathbf{x})\right]\right)$$
$$p(\mathbf{z}) = \mathcal{N}\left(\mathbf{z}|0, \mathbf{I}\right)$$

How to sample from $\mathcal{N}(z|\mu, \sigma)$?

$$\epsilon \sim \mathcal{N}(\epsilon|0, 1) \Rightarrow z = \mu + \sigma \cdot \epsilon$$



Computation of VAE