Lecture 21 VAE
$z \sim P(z; \mathcal{N}(0, 1))$

$x \sim P(x; f(z; \theta))$

$x \sim \mathcal{N}(\mu, \sigma^2)$

$\mu = f(z; \theta)$

$\sigma^2 = f(z; \theta)$
\[ P(x) = \int_{z} P(z) P(x|z) \]

\[ \mathbb{Q}(z) = \begin{cases} P(z|x) & \text{Gaussian} \\ \mathbb{P}(z) P(x|z) & \text{Gaussian} \end{cases} \]

\[ E_{\mathbb{Q}(z)} \log P(x, z; \theta) \]

\[ x \sim \text{NCHM, UC} \]

\[ \mu = f(z) \]
\[ \log P(x) \geq \text{ELBO} \]

unrelated to \( \alpha \)

\[ \Omega(z) = P(z|x) \]

\[ \log P(x) = \text{ELBO} \]

\[ \text{ELBO} = \log P(x) - KL E_\phi(z|x) \| P(z|x) \]

\[ \arg\max \text{ELBO} = \arg\min \text{KL} E_\phi(z|x) \| P(z|x) \]
we cannot compute $P(z|x)$

$\arg \max \text{ELBO} \quad q(c|x) \rightarrow P(z|x)$

Variational EM
\[ \text{arg max } \log \frac{p(x, z; \theta)}{q(z|x; \phi)} \]

\[ z \sim q(z|x; \phi) \]

\[ z = \arg \max \quad E_z \sim q(z|x; \phi) \log \frac{p(x, z; \theta)}{q(z|x; \phi)} \]

\[ z \sim q(z|x; \phi) \log \frac{p(x, z; \theta)}{q(z|x; \phi)} \]
\( \varepsilon \sim N(0, 1) \)

\[ z = \mu + \varepsilon \]

\[ z = \mu + \sigma \varepsilon \]

Element-wise multiplication:

\[ z = \mu + 6 \varepsilon \]

What is the distribution of \( z \)?

\[ z \sim N(\mu, 6) \]

\[ p(z|x) \]

\[ z = \mu + 6 \varepsilon \]
\[ z \sim \mathcal{N}(\mu, \phi) \]

\[ z = g(x, \phi, \psi) \]

Not discrete \( g(\mu, \phi) \) Gaussian

Gumbel / Softmax
\[ \text{ELBO} = \left( \log P(x) \right) - \text{KL} \quad \mathcal{Q}(z|x) \parallel P(z|x) \]

\[ E_{z \sim \mathcal{Q}(z|x)} \left[ \log P(x|z) \right] \]

\[ \text{KL} \quad \mathcal{Q}(z|x) \parallel P(z|\mathcal{X}) \]

\[ \mathcal{N}(\mu, \Sigma) \]
E-step

\[ q(\theta|X) \approx \hat{p}(\theta|X) \]

Optimize \( q(\theta|x; \phi) \) till convergence

\[ p(\theta, x) \]

\[ x \sim p(x|\theta) \]
VAE model $P_{x|z}$, $P_{z|z}$

VAE inference

$L = \text{ELBO}$
\[ \log P(x) \]

E-step:

\[ (z|\xi, x) \sim P(z|x) \]
Loykox does not depend practically on $t$. 

$P_c(x; \omega)$

$\log P_c(x)$

$\text{ELBO}$
\[ E_2 \sim \phi(z_2|x) \left[ \log P_{\theta}(x|z_2) \right] \]

Autoencoder

\[ p(z_2) \sim \mathcal{N}(0, 1) \]

\[ \hat{z} = \text{how to sample } z \]

\[ \mathcal{N}(0, 1) \]
$z \sim \mathcal{N}(0, 1)$

$P(Z)$

$P_c(x)$
VAE:
\[ x \sim \mathcal{N}(\mu, \sigma^2) \]
\[ \mu, \sigma^2 = f(c \pm; \phi) \]

GAN:
\[ P(x) = \int_{\mathbb{R}^d} P(x | z) \, P(z) \]
\[ x = G(z) \]
\[ x \sim \mathcal{N}(\mu, \sigma^2) \]
\[ E_x \sim \eta(x), \ \log P(x | z) \]

\[ \mathbb{E}_{x \sim \eta(x)} \log P(x | z) \]

\[ \mathbb{E}_{x \sim \eta(x)} \log P(x | z) \xrightarrow{1} 1 - \infty \]