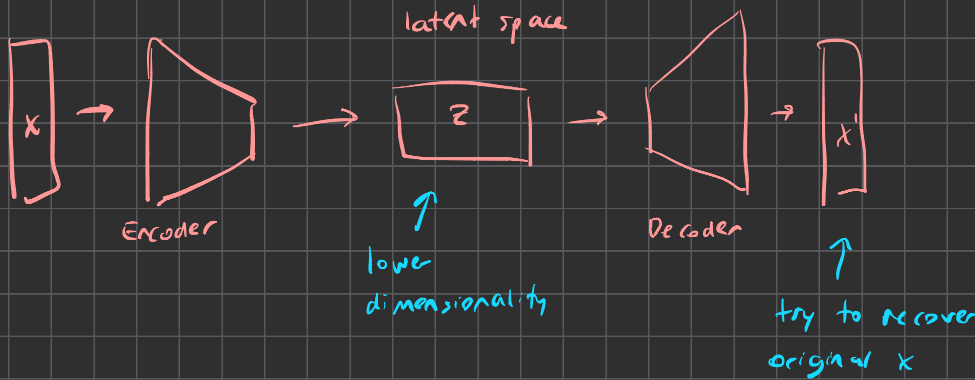
VAE



- good autoencoder → reconstruct input as close as possible to original input

- good "variational" autoencoder - capture semantic relationship b/t data
    - ↳ maps x to z, but it also learns relational information between different regions in z
    - ↳ similar x should take up a similar area of space in z

Math concepts

Expectation of a random variable:

$$E_x[f(x)] = \int x \, f(x) \, dx$$

Chain rule of probability $\quad p(x, y) = p(x \mid y)\, p(y)$

Bayes theorem: $\quad p(x \mid y) = \dfrac{p(y \mid x)\, p(x)}{p(y)}$

Kullback - Leibler divergence (KL - divergence)

$$D_{KL}(P \,\|\, Q) = \int p(x) \log \left(\frac{p(x)}{q(x)} \right) dx$$

→ a measure of how far apart 2 probability distributions are

→ not symmetric: divergence of P from Q ≠ divergence of Q from P

→ always ≥ 0, 0 when 2 distributions are the same

In VAEs, we want to maximize likelihood of our observed data for the best fit model

encoder: $q(z \mid x)$

decoder: $p(x \mid z)$

latent space variable

"likelihood" as $p(x) = \int p(x, z)\, dz$

since we integrate over z, this is intractable

$$p(x, z) = p(z \mid x)\, p(x)$$

probability of observing both at same time

conditional probability of z given x

we want this

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

we don't have this

since we have 2 unknowns, we need an approximation

$$p_\theta(z \mid x) \approx q_\phi(z, x)$$

true posterior   approximate

log likelihood

$$\log p_\theta(x) = \log p_\theta(x) \underbrace{\int q_\phi(z \mid x)\, dz}_{=1}$$

=1 (integrating over all probabilities of z)

$$= \int \log p_\theta(x)\, q_\phi(z \mid x)\, dz$$

$$= E_{q_\phi(z \mid x)}\left[\log p_\theta(x)\right]$$

definition of expectation value

$$E[f(z)] = \int f(z) \cdot q(z)\, dz$$

definition of expectation value

$$= E_{q_\phi(z \mid x)}\left[\log \frac{p_\theta(x, z)}{p_\theta(z \mid x)}\right]$$

apply chain rule of probability

$$= E_{q_\phi(z \mid x)}\left[\log \frac{p_\theta(x, z)}{p_\theta(z \mid x)} \cdot \underbrace{\frac{q_\phi(z \mid x)}{q_\phi(z \mid x)}}_{}\right]$$

multiply by 1

$$\log p(x) = \underbrace{E_{q_\phi(z \mid x)}\left[\log \frac{p_\theta(x, z)}{q_\phi(z \mid x)}\right]}_{\text{ELBO}} + \underbrace{E_{q_\phi(z \mid x)}\left[\log \frac{q_\phi(z \mid x)}{p_\theta(z \mid x)}\right]}_{\text{KL-divergence}}$$

ELBO

Evidence lower bound

KL-divergence

≥ 0

$$\log p_\theta(x) = \underbrace{E_{q_\phi(z|x)}\left[\log p_\theta(x|z)\right]}_{\text{maximizing reconstruction likelihood of decoder}} - \underbrace{D_{KL}\left[q_\phi(z|x) \,||\, p_\theta(z)\right]}_{\text{KL divergence}}$$