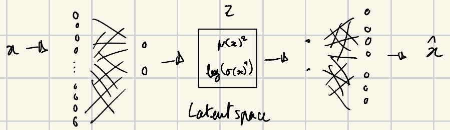
VAE



intractability in VAEs

The true posterior distribution

$$p(z|x) = \frac{p(x|z)\,p(z)}{p(x)} \leftarrow \qquad p(x) = \int p(x|z)\,p(z)\,dz \quad \text{over all } z$$

Problem scenario

* We can define the likelihood of our data as: $p(x) = \int p(x,z)\,dz$

→ it's intractable, we would need to evaluate this integral over all latents z.

* We could use $p(x) = \dfrac{p(x,z)}{p(z|x)} \leftarrow$ we don't have this ground truth, this is sth we want to learn.

Solution

We want to approximate our

true posterior $p_\theta(z|x) \approx q_\varphi(z|x)$ approximate posterior

$$\log p_\theta(x) = \log p_\theta(x) \quad , \quad \text{why log? log-likelihood, turns products into sum}$$

$$= \log p_\theta(x) \underbrace{\int q_\varphi(z|x)\,dz}_{=1,\ \text{integral over the domain}}$$

probability distribution over z

$$= \int \log(p_\theta(x))\, q_\varphi(z|x)\,dz \qquad E_q[c] = \int c \cdot q(z)\,dz = c\int q(z)\,dz = c$$

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

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

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

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

$$\log p_\theta(x) = \underbrace{E_{q_\varphi(z|x)}\left[\log \frac{p_\theta(x,z)}{q_\varphi(z|x)}\right]}_{\text{ELBO}} + \underbrace{D_{KL}\left(q_\varphi(z|x)\,\|\,p_\theta(z|x)\right)}_{\geq 0}$$

We know: $\log p_\theta(x) \geq E_{q_\varphi(z|x)}\left[\log \dfrac{p_\theta(x,z)}{q_\varphi(z|x)}\right] = E_{q_\varphi(z|x)}\left[\log \dfrac{p_\theta(x|z)\,p(z)}{q_\varphi(z|x)}\right]$

maximizing this = maximizing this

$$= E_{q_\varphi(z|x)}\left[\log p_\theta(x|z)\right] + E_{q_\varphi(z|x)}\left[\log \frac{p(z)}{q_\varphi(z|x)}\right]$$

$$= \underbrace{E_{q_\varphi(z|x)}\left[\log p_\theta(x|z)\right]}_{\text{maximize}} - \underbrace{D_{KL}\left(q(z|x)\,\|\,p(z)\right)}_{\text{minimize}}$$

$$D_{KL}(q(z|x)\,\|\,p(z)) = \int q(z|x) \log \frac{q(z|x)}{p(z)}$$

$$= E\left[\log \frac{q(z|x)}{p(z)}\right]$$

$$= -E\left[\log \frac{p(z)}{q(z|x)}\right]$$

Reparameterization Trick

Introducing new estimator, bc we can't run backpropagation over a stochastic process $(q_\varphi(z|x))$.

$$\rightarrow \mathcal{L}(\theta, \varphi, x) = \underbrace{\mathbb{E}_{q_\varphi(z|x)}\left[\log \frac{p_\theta(x,z)}{q_\varphi(z|x)}\right]}_{\text{ELBO}} = \mathbb{E}_{p(\epsilon)}\left[\log \frac{p_\theta(x,z)}{q_\varphi(z|x)}\right]$$

$\epsilon \approx p(\epsilon)$, $z = g(\varphi, x, \epsilon)$

$\epsilon \sim \mathcal{N}(0, I)$

Loss function

$$\mathcal{L}(\theta, \phi; x) \simeq \frac{1}{2}\sum_{j=1}^{J}\left(1 + \log\left((\sigma_j^{(i)})^2\right) - (\mu_j^{(i)})^2 - (\sigma_j^{(i)})^2\right) + \frac{1}{2}\sum_{l=1}^{L}\log p_\theta(x^{(i)} | z^{(i,l)})$$

where $z^{(i,l)} = \mu^{(i)} + \sigma^{(i)} \odot \epsilon^{(l)}$, $\epsilon^{(l)} \sim \mathcal{N}(0, I)$