09/11/23

Deep Learning

- Recap: Generative Modeling, VAE Setup – Encoder and Decoder
- Problem formulation
- Solution
- Reparameterization trick
  - $\mathcal{X}$: data space, Θ: parameter space; $\mathcal{Z}$: latent space

    Decoder: $f(z;\theta): \mathcal{Z} \times \Theta \to \mathcal{X}$.

  - $$P(X) = \int P(X|z).\, P(z)\, dz \quad \text{(Marginal)} \quad -①$$

  - Connecting the expression in ① to a ML model.

    $- \; P(X|z) \sim \mathcal{N}(\mu(z;\theta) ; \sigma^2(z).I)$ or $\mathcal{N}(\underbrace{f(z;\theta)}_{\text{ML model or decoder}}, \sigma^2 I)$

  - Goal: Find $f(z;\theta)$ such that $P(f(z;\theta))$ is maximized i.e. the likelihood that a generated sample comes from $P(X)$ is maximized.

  - The set of samples needed to approximate $P(X) \approx \frac{1}{n} \sum P(X|z)$ is extremely large.

  - Can we find an alternative? Yes. Let $Q(z|X)$ be a proxy to $P(z)$.

  - Let's look at the KL divergence between $Q(z|X)$ and $P(z|X)$

    Recall: $D(p(x) \| q(x)) = E_{x \sim p(x)} [\log p(x) - \log q(x)]$

$$D(Q(z|X) \| P(z|X)) = E_{z \sim Q(z|X)} [\log Q(z|X) - \log P(z|X)]$$

$$= E_{z \sim Q(z|X)} [\log \underbrace{Q(z|X)}_{\text{Encoder}} - (\log P(X,z) - \log P(X))]$$

$$= E_{z\sim Q(z|x)}\left[\log Q(z|x) - (\log P(X|z) + \log P(z)) - \log P(X)\right]$$

$$= E_{z\sim Q(z|x)}\left[\log Q(z|x) - \log P(z) - \log P(X|z) + \log P(X)\right]$$

$$= D(Q(z|x)\,\|\,P(z)) + \log P(X) - E_{z\sim Q(z|x)} \log P(x|z)$$

Rearranging:

$$\underbrace{\log P(X)}_{\text{unknown}} - D\big(Q(z|x)\,\|\,\underbrace{P(z|x)}_{\text{unknown}}\big) = E_{z\sim Q(z|x)} \log \underbrace{P_\theta(x|z)}_{\mathcal{N}(\mu(z;\theta);\ \sigma^2 I)} - D\big(Q_\phi(z|x)\,\|\,\underbrace{P(z)}_{\mathcal{N}(0,I)}\big)$$

Assume: $Q(z|x) \sim \mathcal{N}(\mu_\phi(X), \Sigma_\phi(X))$

Observe that the terms on the RHS are all known.