Let $x \in \mathbb{R}^D$, Cifar 10, $D = 3 * 28 * 28 = 2532$

We want to model $p(x)$ $x \in 2532$ 

We're trying to capture the joint probability distribution over all pixels in an image

$$p(x) = p(x_1, x_2, \text{——}, x_D)$$ Dependency assumption

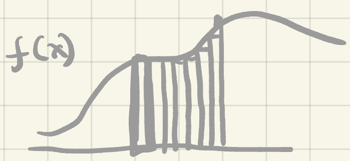
Modeling $p(x)$ directly is insanely complex. So we introduce a latent variable $z \in \mathbb{R}^L$, where $L \ll D$, say 32

Now we try to model $p(x) = \int p(x|z)\,p(z)\,dz$

Let's say $p(x|z)$ a decoder and a prior $p(z)$ over latent codes z

we try

i) Even with this approach and $D=32$, reasonable for a lower dimensional space, $p(x)$ is intractable



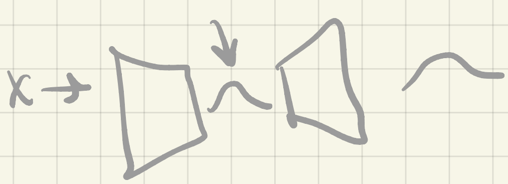
imagine in 32 dimensions

$O(y^D)$ exponential

ii) Sampling $z \sim p(z) = N(0, I) \rightarrow p(x|z)p(z) = \hat{x}$ will never lead to any useful $\hat{x}$.

If z not useful $\Rightarrow$ whatif we have $p(z|x)$? Ti wiwh

$x \rightarrow z \rightarrow \hat{x}$ $\quad p(z|x) \rightarrow p(x|z)$



$$p(z|x) = \frac{p(x|z)\,p(z)}{p(x)} = \frac{p(x|z)\,p(z)}{\int p(x|z)\,p(z)\,dz}$$

intractable

So then we say what if we have $q_\phi(z|x)$?

$$\mathbb{E}_{q(z)}[f(z)] = \int f(z) q(z) dz$$

$$p(x) = \int p(x|z) p(z) dz$$

$$= \int \frac{p(x|z) p(z)}{q_\phi(z|x)} q_\phi(z|x) dz$$

Jensen's Inequality

$$f(\mathbb{E}[x]) \geq \mathbb{E}[f(x)]$$

$$\lg p(x) = \lg \int \frac{p(x|z) p(z)}{q_\phi(z|x)} q_\phi(z|x) dz$$

$$= \lg \mathbb{E}_{q_\phi(z|x)}\left[\frac{p(x|z)p(z)}{q_\phi(z|x)}\right] \geq \mathbb{E}_{q_\phi(z|x)}\left[\lg \frac{p(x|z)p(z)}{q_\phi(z|x)}\right]$$

$$= \mathbb{E}_{q_\phi} \lg p(x|z) + \mathbb{E}_{q_\phi} \lg p(z) - \mathbb{E}_{q_\phi} \lg q(z|x)$$

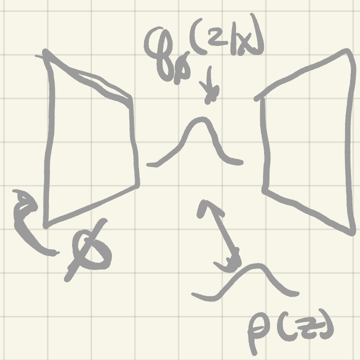
$$= \mathbb{E}_{q_\phi} \lg p(x|z) - \left(\mathbb{E}_{q_\phi} \lg q(z|x) - \mathbb{E}_{q_\phi} \lg p(z)\right)$$

KL Divergence

$$D_{KL}(P(x) \| Q(x)) = \mathbb{E}_{P(x)} \log \frac{P(x)}{Q(x)}$$

$$= \mathbb{E}_{q_\phi} \lg p(x|z) - \mathbb{E}_{q_\phi(z|x)}\left[\lg \frac{q_\phi(z|x)}{p(z)}\right]$$

$$\Rightarrow \boxed{\underbrace{\mathbb{E}_{q_\phi} \lg p(x|z)}_{\text{reconstruction}} - \underbrace{D_{KL}(q_\phi(z|x) \| p(z))}_{\text{regularization}}}$$



Multivariate Gaussian PDF

$$\frac{1}{2\pi^{D/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

Take an image $x \in [0,1]^{D=2532}$ iid $\rightarrow$ The predictive distribution of the model $p(x|z)$ is a multivariate Gaussian with likelihood

$$p(x|z) = \prod_{i=0}^{D} N(x_i; \hat{x}_i, \sigma^2 I) \qquad \hat{x}_i = f_\phi(z) \qquad \overset{\text{iid}}{\Sigma} = \sigma^2 I$$

$$\text{Gauss}(x_i; \mu, \Sigma) \qquad \Sigma^{-1} = \frac{1}{\sigma^2} I$$

$$|\Sigma| = \sigma^{2D}$$

Then $p(x|z) = \dfrac{1}{2\pi^{D/2}\sigma^{2D/2}} \exp\left(-\frac{1}{2}(x-\hat{x})^T \frac{1}{\sigma^2} I (x-\hat{x})\right)$

$$= \frac{1}{2\pi^{D/2}\sigma^D} \exp\left(-\frac{1}{2\sigma^2}(x-\hat{x})^T I (x-\hat{x})\right)$$

$$\log p(x|z) = \underbrace{\cancel{\log \frac{1}{2\pi^{D/2}\sigma^D}}}_{\text{constant}} - \frac{1}{2\sigma^2}\|x-\hat{x}\|_2^2$$

objective

$$-\frac{1}{2\sigma^2}\|x-\hat{x}\|_2^2 - D_{KL}(q_{\theta\phi}(z|x) \| p(z))$$

in practice we minimize

$$\underset{\theta}{\text{argmin}} \frac{1}{2\sigma^2}\|x-\hat{x}\|_2^2 + D_{KL}(q_{\theta\phi}(z|x) \| p(z))$$