# Autoencoders and Variational Autoencoders

## Autoencoder

Input: $x \in \mathbb{R}^d$

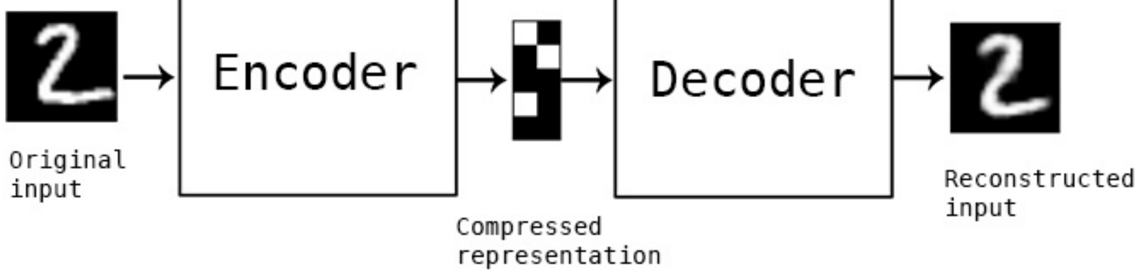
Encoder: $f : \mathbb{R}^d \to \mathbb{R}^m$

Decoder: $g : \mathbb{R}^m \to \mathbb{R}^d$

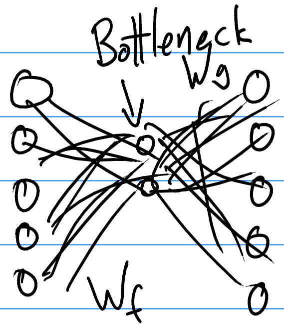
If $m = d$,
$f(x) = x$
$g(x) = x$

Goal output $g(f(x)) = x$

Interesting if $m \ll d$. $(d = 1000, m = 10)$




## Linear Autoencoder

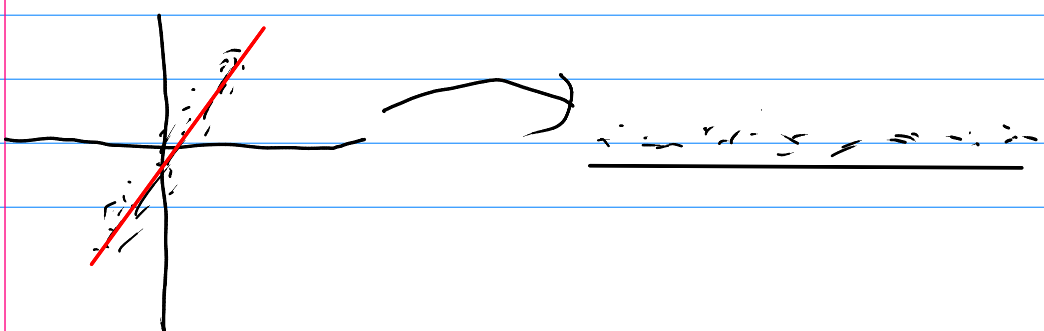



Output $= W_g W_f x$

Objective: $\min\limits_{W_f, W_g} \sum\limits_i \frac{1}{2} \| W_g W_f x_i - x_i \|_2^2$

$W_f x$ is a compression of $x$

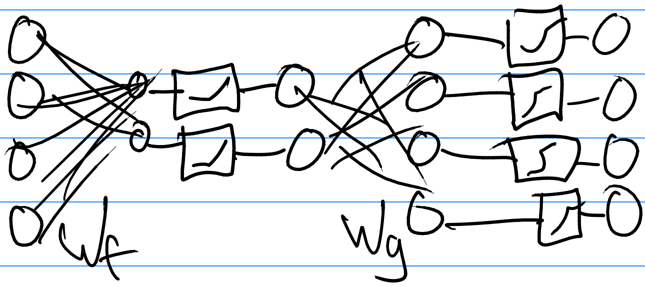
## Principal Component Analysis (PCA)



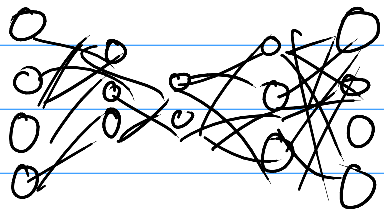
# Nonlinear autoencoder

- f, g are nonlinear

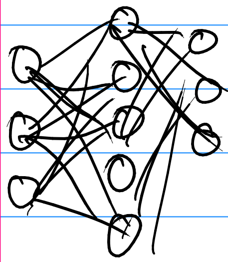
$$\min_{W_f, W_g} \sum \frac{1}{2} \| g(f(x_i)) - x_i \|_2^2$$




# Deep autoencoders



# Sparse Autoencoder

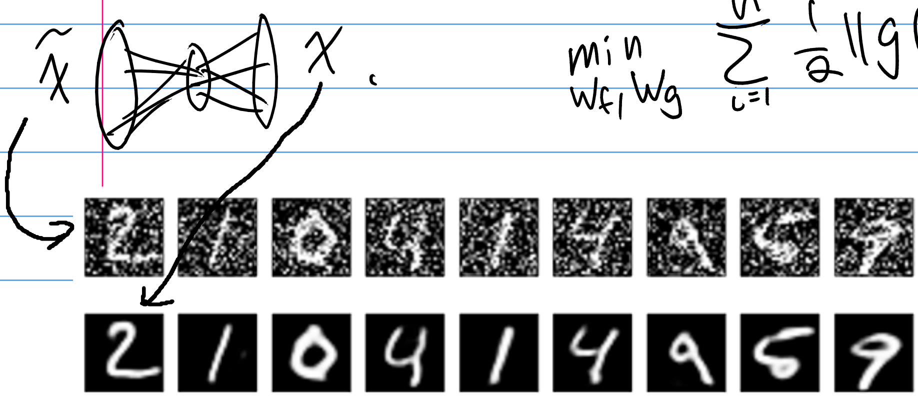


$$\min_{W_f, W_g} \sum \frac{1}{2} \| g(f(x_i)) - x_i \|_2^2 + \lambda \| f(x_i) \|_1$$

$f(x_i)$ will be sparse

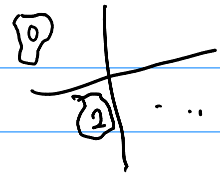
# Denoising autoencoder

Given noised input $\tilde{x}$, produce $x$ as output




$$\min_{W_f, W_g} \sum_{i=1}^{n} \frac{1}{2} \| g(f(\tilde{x}_i)) - x_i \|_2^2$$

# Variational Autoencoder




## Generative Modelling

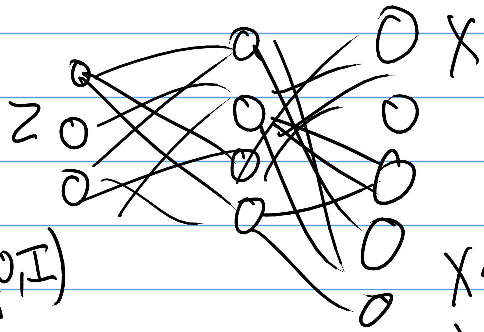
$X_1, \dots, X_n \sim D \Rightarrow$ Generate $X_{n+1}, X_{n+2}, \dots$

D may be complex (moreso than a GMM)

Say D is the distribution of Handwritten numbers or even Imagenet

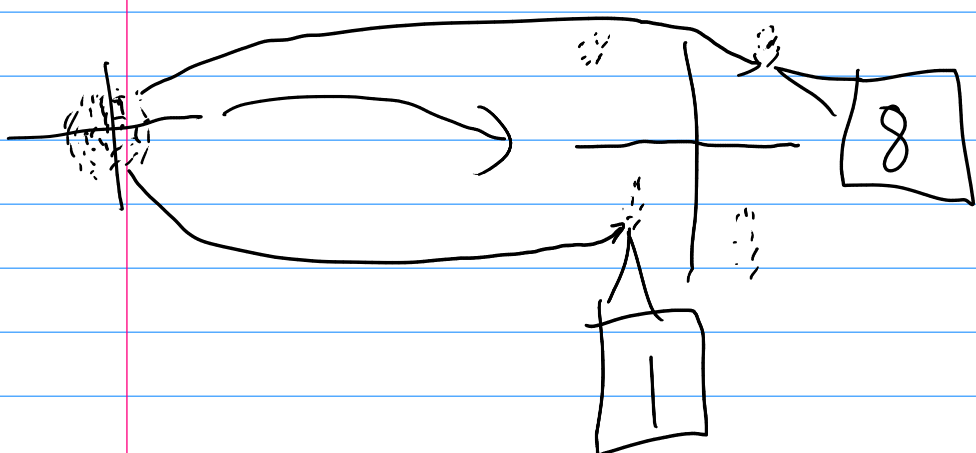
Soln: Use NN to do the work!

Draw sample from $N(0,I)$, map it to a sample from D




$Z \sim N(0,I)$
$Z \in \mathbb{R}^m$

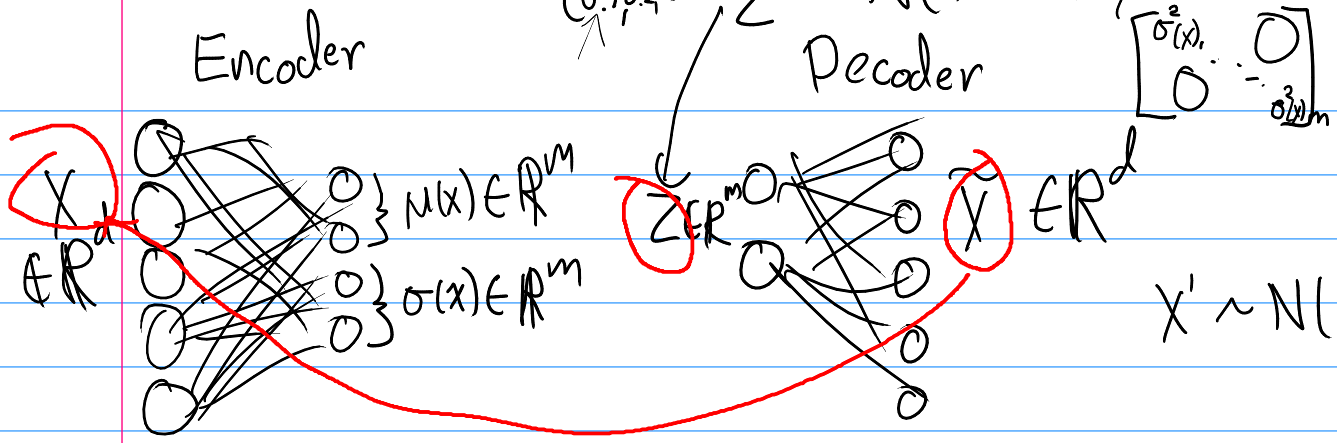
$X \sim D$
$X \in \mathbb{R}^d$




(VAE)

Actually: Use Variational autoencoder

$Z \sim N(\mu(X), \mathrm{diag}(\sigma^2(X))) \in \mathbb{R}^m$



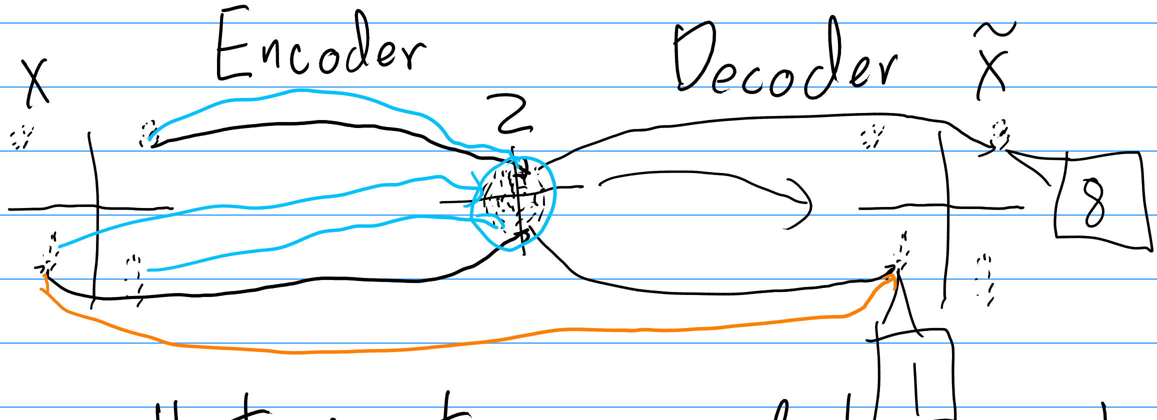

$X' \sim N(\mu_\theta(X|z), I)$

Some notation: $p_\theta(x)$ is density of decoder network.

$p_\theta(x|z)$ is a dist over decoder's outputs, given a fixed $z$.

$p_\theta(z)$ is dist over encoder's input (us: $N(0,I)$)

$q_\phi(z|x)$




- Ensure that input image dist maps to $N(0,I)$

Minimize $KL(q_\phi(z|x) \| p_\theta(z)) = KL(q_\phi(z|x) \| N(0,I))$

- Inputs get encoded and then decoded back

Maximize $\underset{z \sim q_\phi(z|x)}{E} [\log p_\theta(x|z)]$

Claim: $\log p_\theta(x) \geq \underset{z \sim q_\phi(z|x)}{E} [\log p_\theta(x|z)] - KL(q_\phi(z|x) \| N(0,I))$

Related to EM

Variational Inference

$$\log p_\theta(x) \geq \mathop{E}_{z\sim q_\phi(z|x)}\left[\log p_\theta(x|z)\right] - KL(q_\phi(z|x)\,||\,N(0,I))$$

Optimize: look at one term at a time

1. $KL(q_\phi(z|x)\,||\,N(0,I)) = KL(N(\mu_\phi(x), \operatorname{diag}(\sigma_\phi^2(x)))\,||\,N(0,I))$

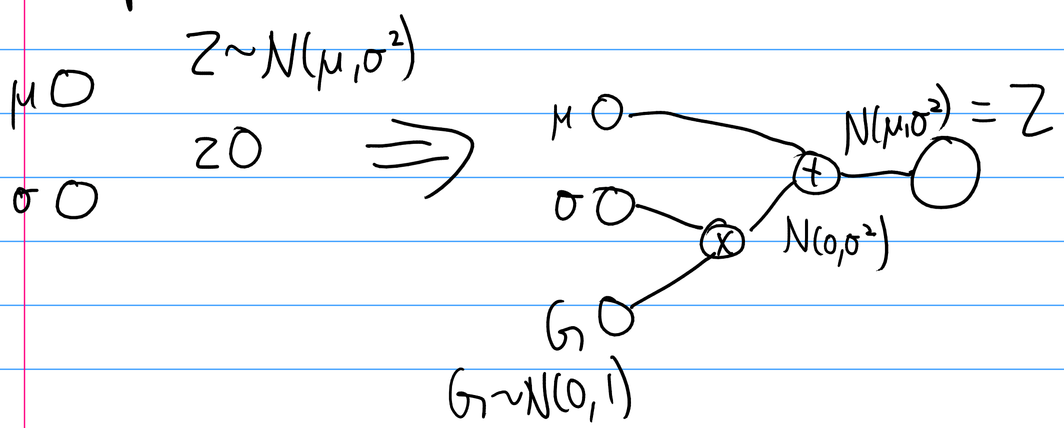
$$= \frac{1}{2}\Big(\underbrace{\|\mu_\phi(x)\|_2^2}_{0} \underbrace{- m}_{-m} + \sum_{j=1}^{m}\big(\sigma_\phi^2(x)_j - \log(\sigma_\phi^2(x)_j)\big)\Big)$$

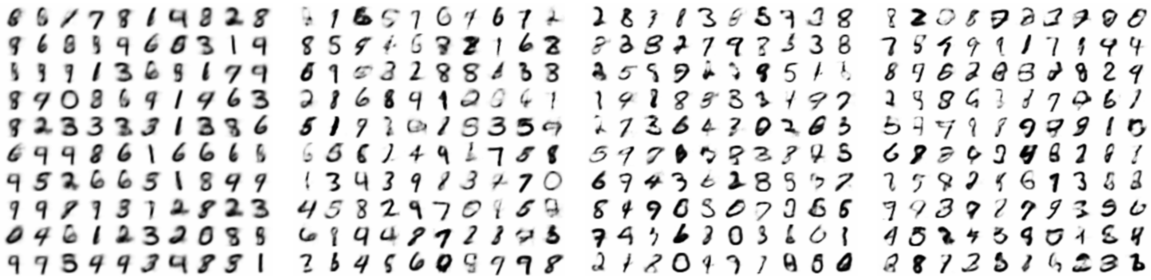
$\mu_\phi(x)=0,\ \sigma^2_\phi(x)_j=1\ \forall j \qquad \sum(1-0) \qquad = 0$

2. $\mathop{E}_{z\sim q_\phi(z|x)}\left[\log p_\theta(x|z)\right] \qquad p_\theta(x|z) \sim N(\mu_\theta(x|z), I)$

$$\mathop{E}_{z\sim q_\phi(z|x)}\left[\|x-\mu_\theta(x|z)\|_2^2\right] \cancel{- d\log(\sqrt{2\pi})}$$

## Reparametrization Trick





(a) 2-D latent space (b) 5-D latent space (c) 10-D latent space (d) 20-D latent space