Lecture 11 cluster, EM

$|x_i - y_i|^2$ $\cos(\vec{x}, \vec{y})$

K

hyperparameters

1. assign $x^{(i)}$ to its closest center

based on L^2 distance

$x^{(i)} - \mu_1$

$\|x^{(i)} - \mu_2\|^2$

2. reestimate $\mu = \dfrac{x_1 + \cdots x_m}{m}$

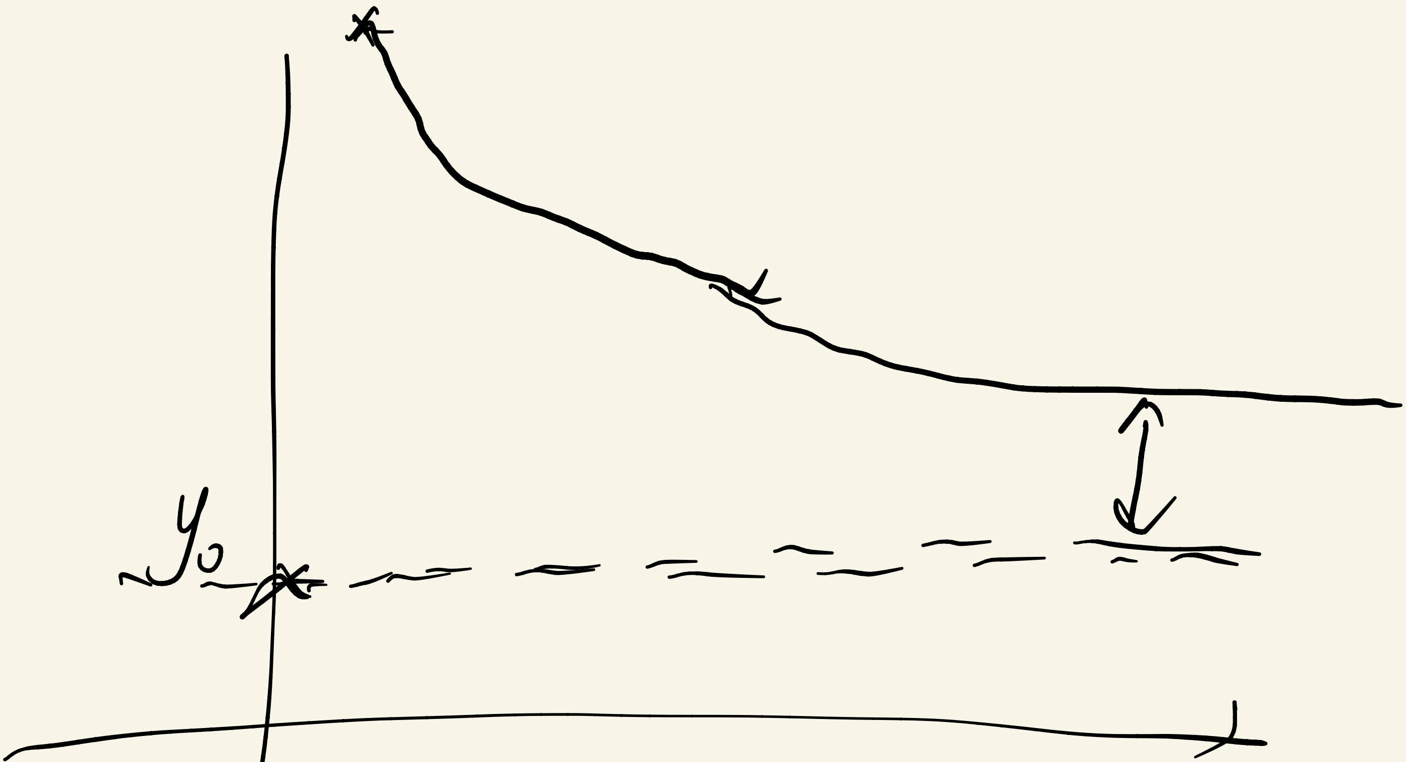
$J(\text{cluster } 1, \mu_1)$

$$= \sum_{i=1}^{m} \| x^{(i)} - \mu_1 \|^2$$

assignment is fixed

$$\mu_1^* = \underbrace{\arg\min_{\mu_1}} \sum_{i=1}^{m} \| x^{(i)} - \mu_1 \|^2$$

$$\mu_1^* = \frac{x_1 + x_2 \cdots x_m}{m}$$



$K = 7$

unsupervised metric

$P_{data}(x)$ $P_{model}(x)$

$P(z)$

$(x, \cancel{z})$

$\log P(x, \cancel{z})$ latent variable

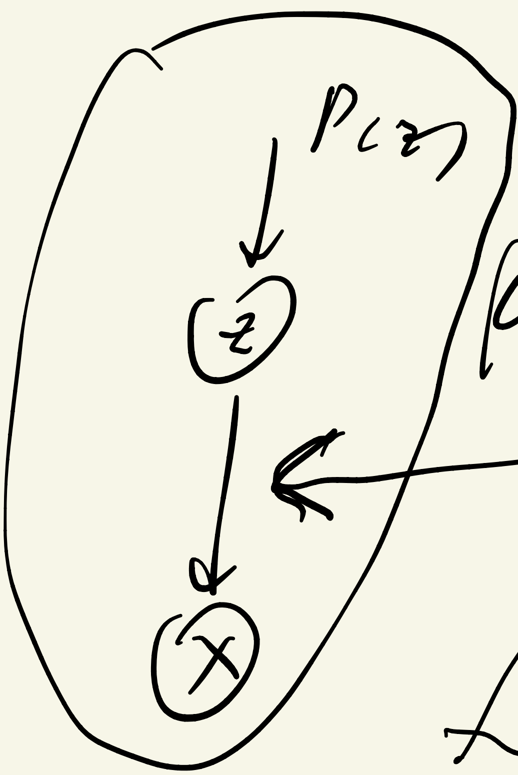
latent variable model

VAE GAN Diffusion

$$P(x) = \sum_{z} P(x,z)$$

$$P(x|z)\,P(z)$$

$k=3$



z

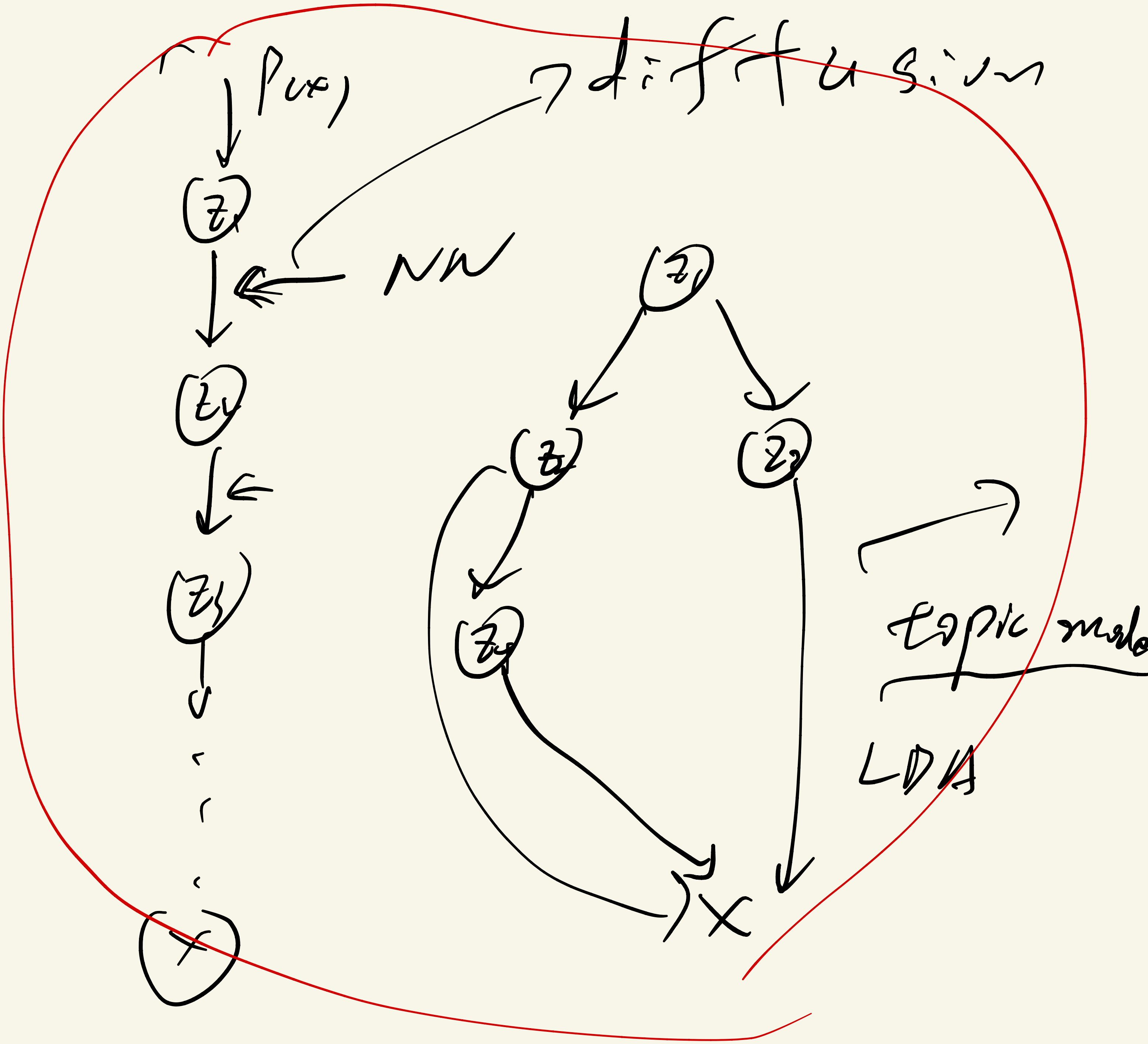
assumption

f

neural network

$x = f(z) \longrightarrow$ GAN

$x \sim D[f(z)] \longrightarrow$ VAE



p(x)
diffusion
NN
topic model
LDA

$$\underbrace{P(z\mid x)}_{\downarrow}$$

$$\frac{\mathbb{1}(z=j)}{P(z=j\mid x)}$$

$$\boxed{P(x;\theta) = \sum_z P(x, z;\theta)}$$

$P(x, z)$ joint

$P(x)$

$$\boxed{\sum_z Q(z) = 1}$$

$$\log P(x;\theta) = \log \sum_z P(x,z;\theta)$$

$$= \log \sum_z Q(z) \frac{P(x,z;\theta)}{Q(z)}$$

$$\geq \sum_z Q(z) \log \frac{P(x,z;\theta)}{Q(z)}$$

$E(x)$

$= \sum_x P(x)\, x$

Jensen inequality

$\underset{x \sim p_\alpha}{E(x)}$

$\alpha_1\, x_1 \quad \alpha_2 x_2$

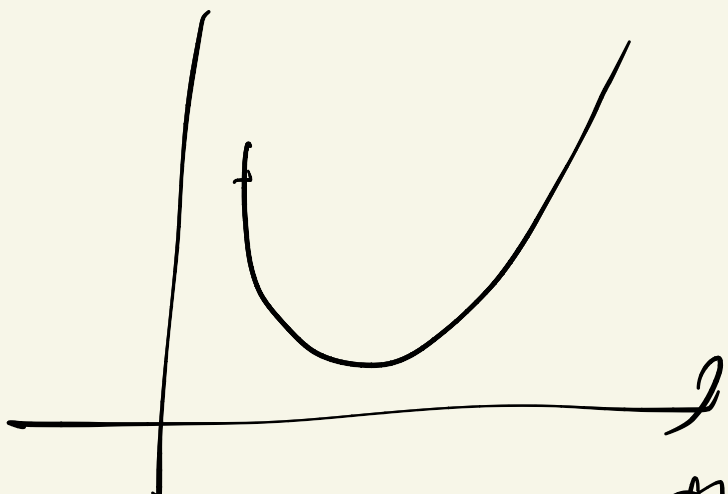
$\alpha_1 + \alpha_2 = 1$

$\boxed{\alpha_1 + \alpha_2 \cdots \alpha_n = 1}$

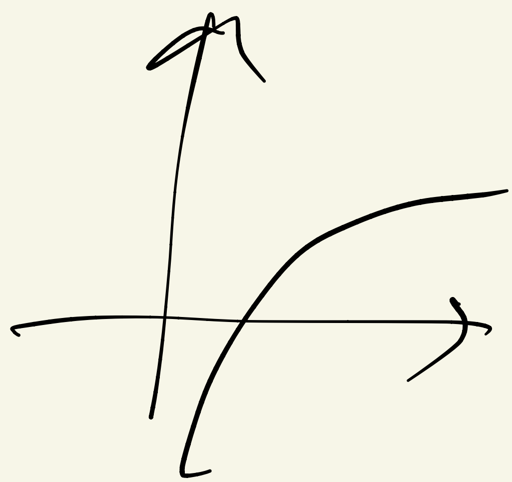
$f(\alpha_1 x_1 + \alpha_2 x_2 \cdots \alpha_n x_n) \leq$

$\alpha_1 f(x_1) + \alpha_2 f(x_2) \cdots$ $E[f(x)]$

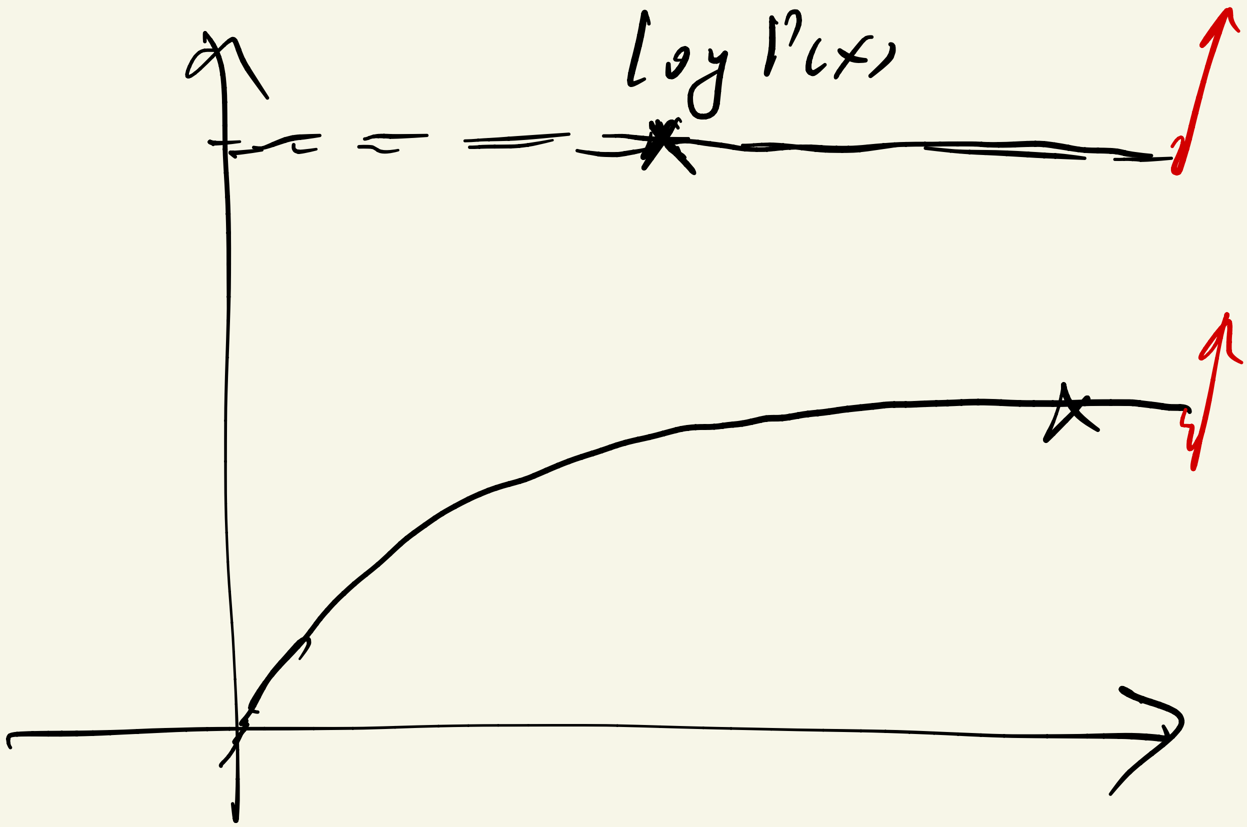
$$f(E(x)) \leq E[f(x)]$$



$$\log \sum_z Q(z) \frac{p(x,z;\theta)}{Q(z)}$$



$$\geq \sum_z Q(z) \log \frac{p(x,z;\theta)}{Q(z)}$$



equality

$$P(x,z) = P(z)\,P(x|z)$$

$$P(z|x) = \frac{P(z)\,P(x|z)}{P(x)} \propto P(z)\,P(x|z)$$

$$\frac{P(x,z)}{Q(z)} \propto \text{constant}$$

$P(x)$

$P(z|x)$

$$f(t_1 x_1 + t_2 x_2 + t_3 x_3)$$

$$\le t_1 f(x_1) + t_2 f(x_2) \cdots$$

$$- \quad -$$