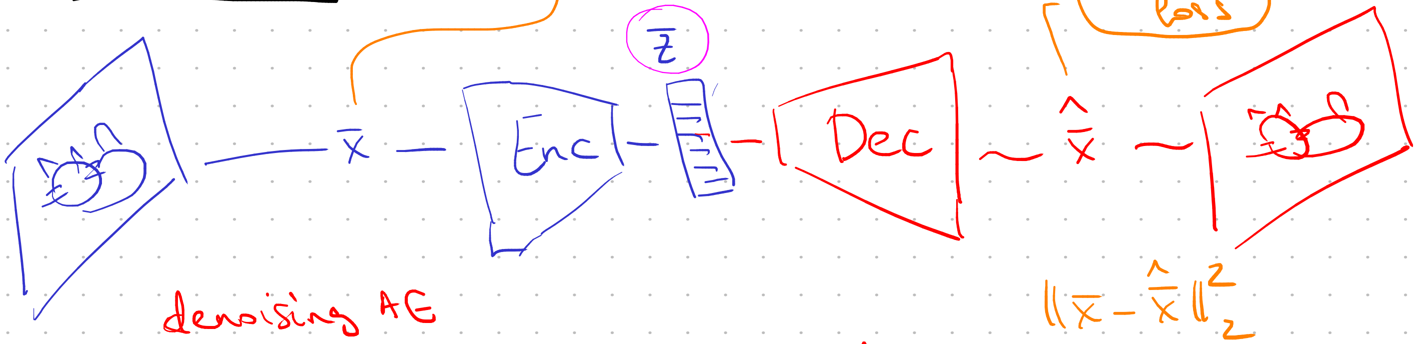
Autoencoders

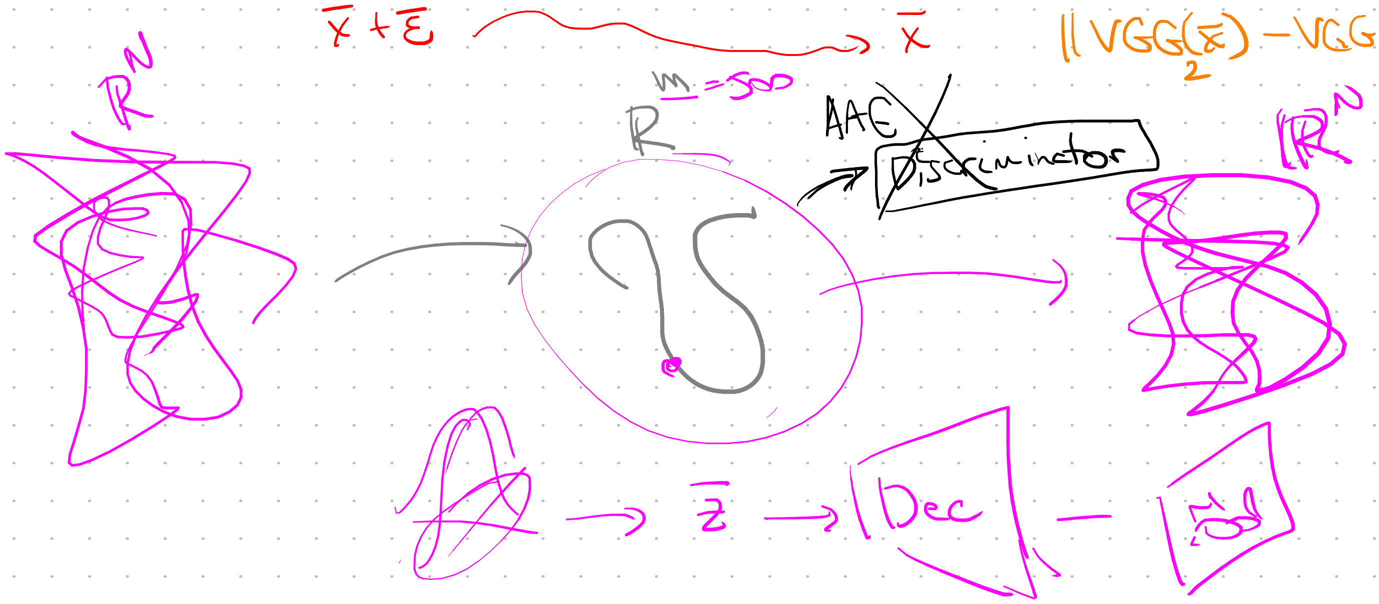


$\|\bar{x} - \hat{\bar{x}}\|_2^2$

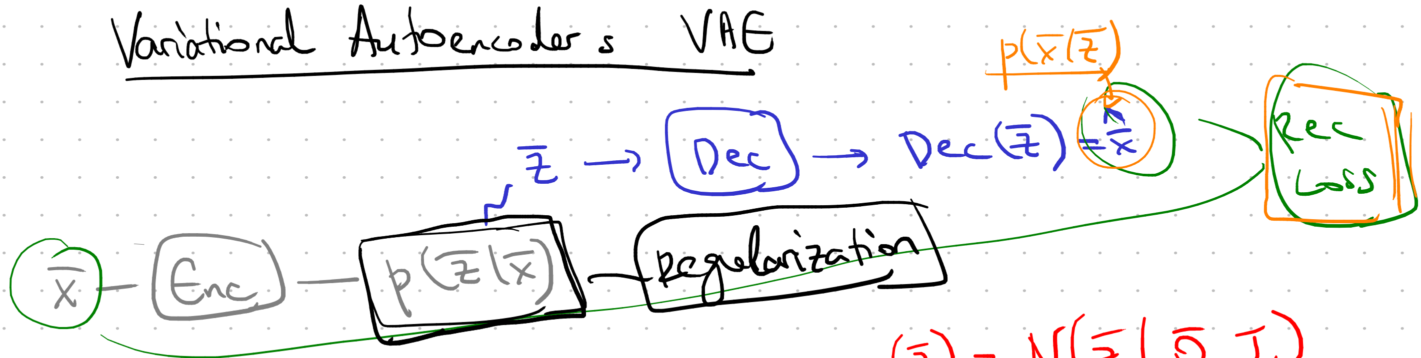
denoising AE

$\bar{x} + \bar{\varepsilon} \rightsquigarrow \hat{\bar{x}}$

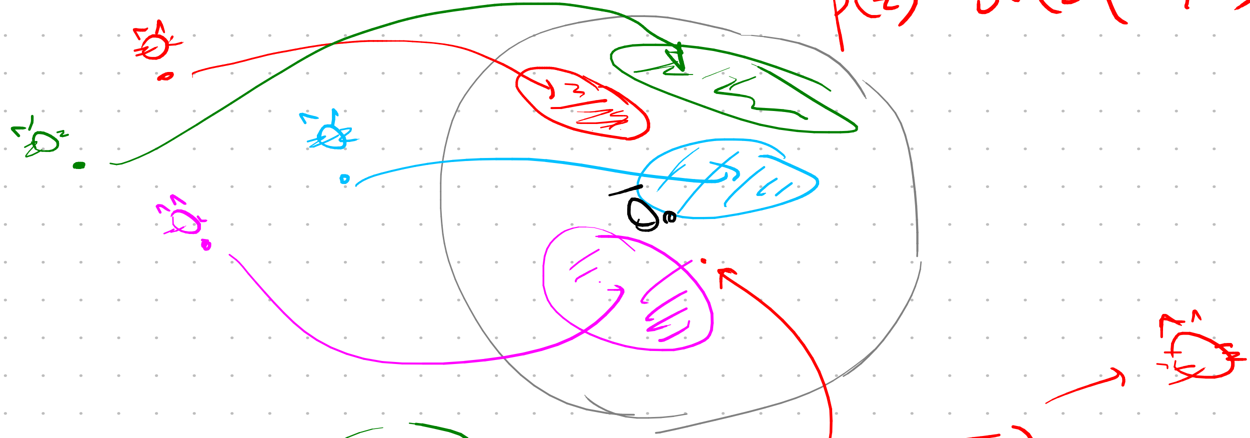
$\|VGG_2(\bar{x}) - VGG_2(\hat{\bar{x}})\|_2^2$



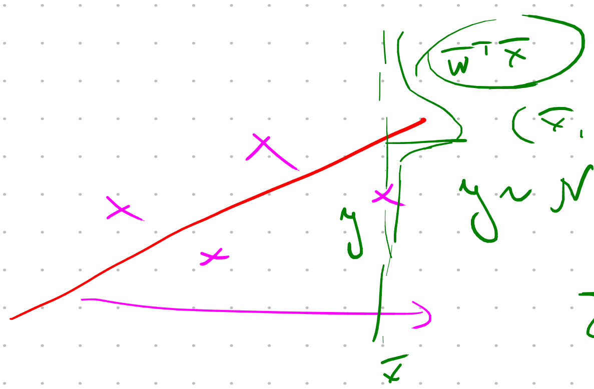
Variational Autoencoders VAE



$p(\bar{z}) = \mathcal{N}(\bar{z}|\bar{0}, I)$



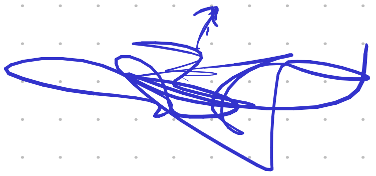
$\bar{z} \sim p(\bar{z})$



$y \sim \mathcal{N}(y \mid \bar{w}^T\bar{x}, \sigma^2)$

$\sum (y_n - \bar{w}^T\bar{x}_n)^2 \to \min$

$$\underbrace{p(\bar z)}_{\mathcal{N}(\bar z|\bar 0, I)}\ \underbrace{\overbrace{p(\bar x|\bar z)}^{\text{Decoder}}}_{\mathcal{N}} = \underbrace{p(\bar x, \bar z)}_{\mathcal{N}} = \underbrace{p(\bar x)}\ \overbrace{p(\bar z|\bar x)}^{\text{Encoder}} \approx q(\bar z)$$



$\mathbb{R}^N$

Variational approximation

$$p(\bar x|\bar z) = \mathcal{N}(\bar x \mid \underset{\text{Decoder}(\bar z)}{f(\bar z)},\ cI)$$

$$\log p(\bar x, \bar z) = \log p(\bar x)\, p(\bar z|\bar x)$$

$$\underline{\log p(\bar x)} = \log p(\bar x, \bar z) - \log p(\bar z|\bar x) \quad \Big\}\ \mathbb{E}_{q(\bar z)}$$

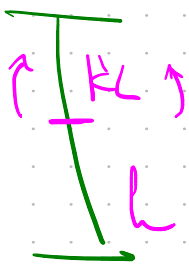
$$\log p(\bar x) = \mathbb{E}_q[\log p(\bar x, \bar z)] - \mathbb{E}_q[\log p(\bar z|\bar x)] \pm \mathbb{E}_q[\log q(\bar z)]$$

$$\boxed{\log p(\bar x)} = \underbrace{\mathbb{E}_{q(\bar z)}\left[\log \frac{p(\bar x, \bar z)}{q(\bar z)}\right]}_{\mathcal{L}(q)} - \underbrace{\mathbb{E}_{q(\bar z)}\left[\log \frac{p(\bar z|\bar x)}{q(\bar z)}\right]}_{=\int q(\bar z) \log \frac{q(\bar z)}{p(\bar z|\bar x)}\, d\bar z = KL(q\|p(\bar z|\bar x))}$$

Variational lower bound

ELBO

$$\text{Const} = \underbrace{\overbrace{\mathcal{L}(q)}}_{\to \max} + \underbrace{KL(q(\bar z)\,\|\,p(\bar z|\bar x))}_{\to \min}$$



$$\mathcal{L}(q) = \int q(\bar z) \log \frac{\overset{= p(\bar z)p(\bar x|\bar z)}{p(\bar x, \bar z)}}{q(\bar z)}\, d\bar z =$$

$$= \underbrace{\int q(\bar z) \log p(\bar x|\bar z)\, d\bar z}_{=} - \underbrace{\int q(\bar z) \log \frac{q(\bar z)}{p(\bar z)}\, d\bar z}_{=}$$

$$\mathbb{E}_{q(\bar z)}\left[\log \mathcal{N}(\bar x | f(\bar z), cI)\right] \qquad KL(q(\bar z) \| p(\bar z))$$

$$= \mathbb{E}_q\left[\text{Const} - \frac{1}{2c} \cdot \|\bar x - f(\bar z)\|^2\right]$$

$$\approx \text{const} - \frac{1}{2c} \cdot \sum_n \|\bar x_n - \underbrace{f}_{\text{Dec}}(\text{Enc}(\bar x_n))\|^2$$

$$\mathcal{N}(\bar z | \bar 0, I)$$

$$h(q) = \text{Const} - \left(\underbrace{\frac{1}{2c} \cdot \mathbb{E}_q\left[\|\bar x - f(\bar z)\|^2\right]}_{\text{Reconstruction loss}} + \underbrace{KL(q(\bar z) \| p(\bar z))}_{\text{Regularizer}}\right)$$

$$q(\bar z) = \mathcal{N}(\bar z | \bar\mu_x, \bar\sigma_x) \qquad p(\bar z) = \mathcal{N}(\bar z | \bar 0, I)$$

$$\bar\sigma_x = \begin{pmatrix} \sigma_{x1} & & 0 \\ & \ddots & \\ 0 & & \sigma_{xd} \end{pmatrix}$$

$$KL(q \| p) = KL\left(\prod_i q_i(z_i) \,\Big\|\, \prod_i p_i(z_i)\right) =$$

$$= \sum_{i=1}^{d} KL(q_i(z_i) \| p_i(z_i))$$

$$KL\left(\underline{\mathcal{N}(z_i | \mu_{x_i}, \sigma_{x_i}^2)} \,\|\, \underline{\mathcal{N}(z_i | 0, 1)}\right) =$$

$$= \int \underline{\frac{1}{\sqrt{2\pi\sigma_{x_i}^2}} e^{-\frac{1}{2\sigma_{x_i}^2}(z_i - \mu_{x_i})^2}} \cdot \log \frac{\mathcal{N}(\_)}{\mathcal{N}(\_)}\, dz_i =$$

$$= \mathbb{E}_{\mathcal{N}(z_i | \mu_{x_i}, \sigma_{x_i}^2)}\left[-\frac{1}{2}\log \sigma_{x_i}^2 - \frac{1}{2\sigma_{x_i}^2}(z_i - \mu_{x_i})^2 + \frac{1}{2} z_i^2\right] =$$

$$= \mathbb{E}_{\mathcal{N}(z_i|\mu_{x_i},\sigma_{x_i}^2)}\left[\left(-\frac{1}{2\sigma_{x_i}^2}+\frac{1}{2}\right)\overset{\mu_{x_i}^2+\sigma_{x_i}^2}{z_i^2}+\frac{\mu_{x_i}}{\sigma_{x_i}^2}\overset{\mu_{x_i}}{z_i}\right]-\frac{1}{2}\log\sigma_{x_i}^2-\frac{\mu_{x_i}^2}{2\sigma_{x_i}^2}$$

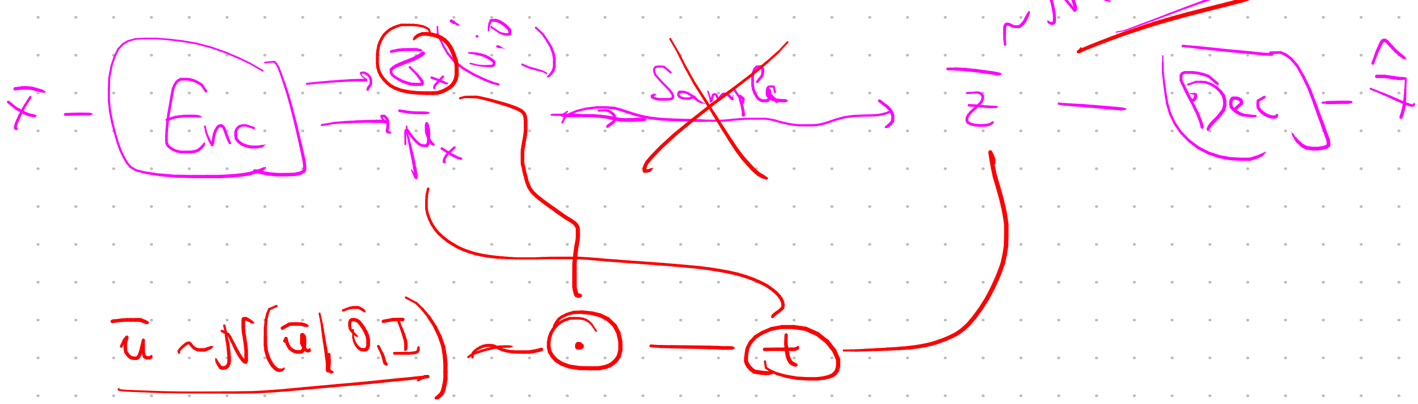
$$\mathbb{E}[z_i^2] = \mathbb{E}[z_i]^2 + \mathrm{Var}[z_i]$$

$$= \left(\frac{1}{2}-\frac{1}{2\sigma_{x_i}^2}\right)\left(\mu_{x_i}^2+\sigma_{x_i}^2\right)+\cancel{\frac{\mu_{x_i}^2}{2\sigma_{x_i}^2}}-\frac{1}{2}\log\sigma_{x_i}^2-\cancel{\frac{\mu_{x_i}^2}{2\sigma_{x_i}^2}} =$$

$$= \frac{1}{2}\mu_{x_i}^2+\frac{1}{2}\sigma_{x_i}^2-\frac{1}{2}\log\sigma_{x_i}^2-\frac{1}{2}$$

$$KL(q\,||\,p) = \sum_{i=1}^{d}\frac{1}{2}\left(\mu_{x_i}^2+\sigma_{x_i}^2-\log\sigma_{x_i}^2-1\right)$$

Reparametrization trick



$$\Rightarrow \quad \bar{u}\odot\bar{\sigma}_x+\bar{\mu}_x \sim \mathcal{N}(\cdot\,|\bar{\mu}_x,\bar{\sigma}_x)$$