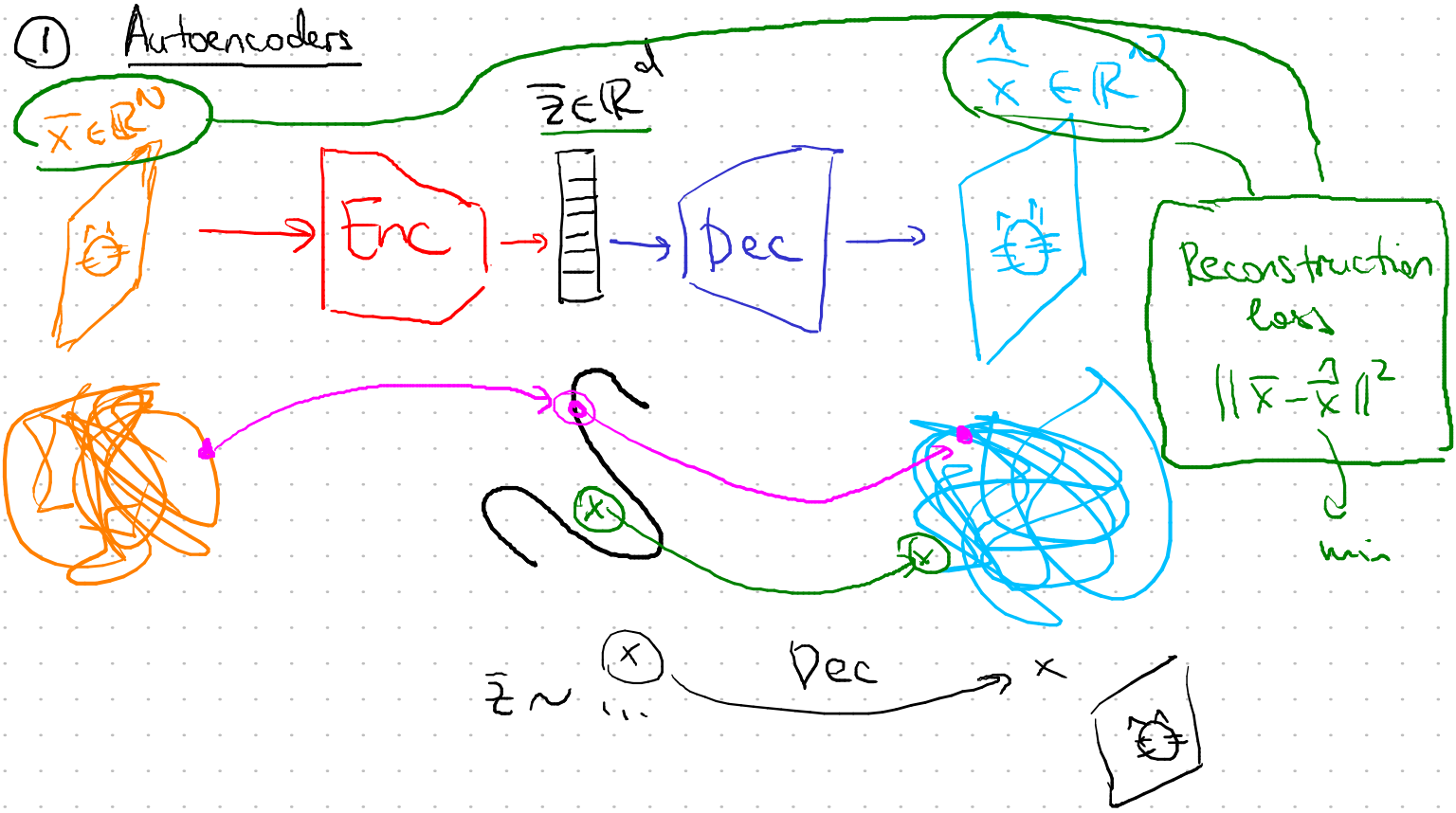
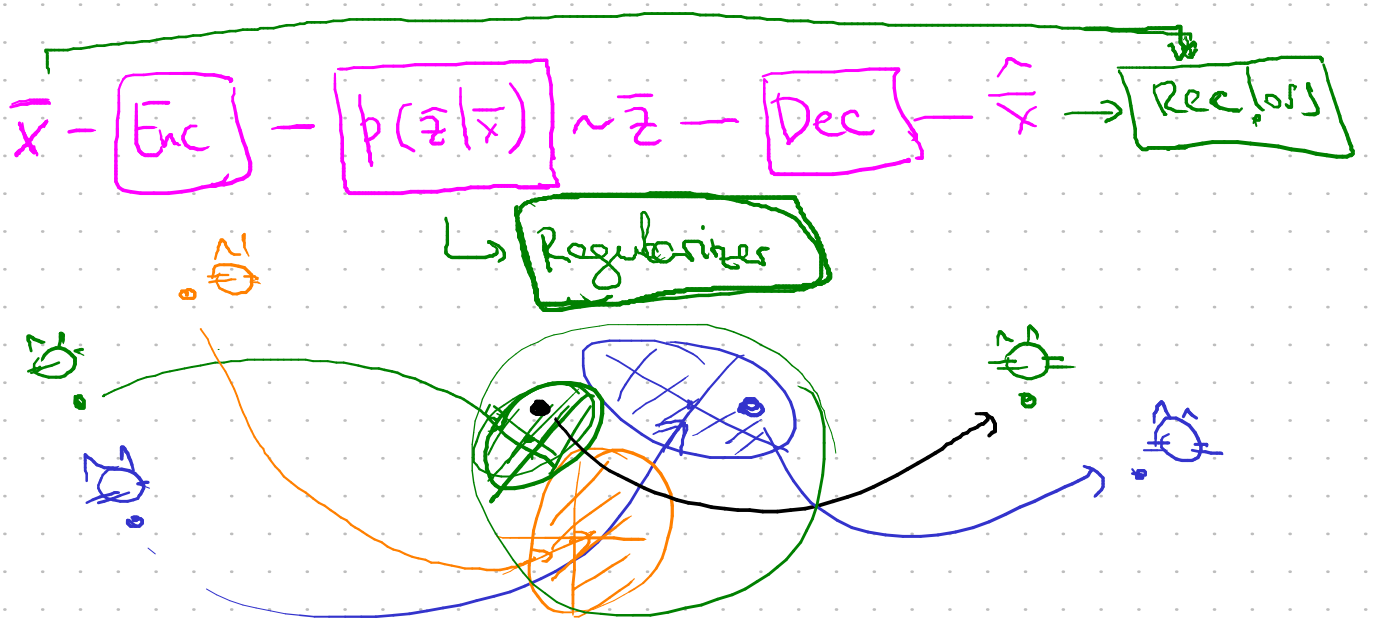
① Autoencoders

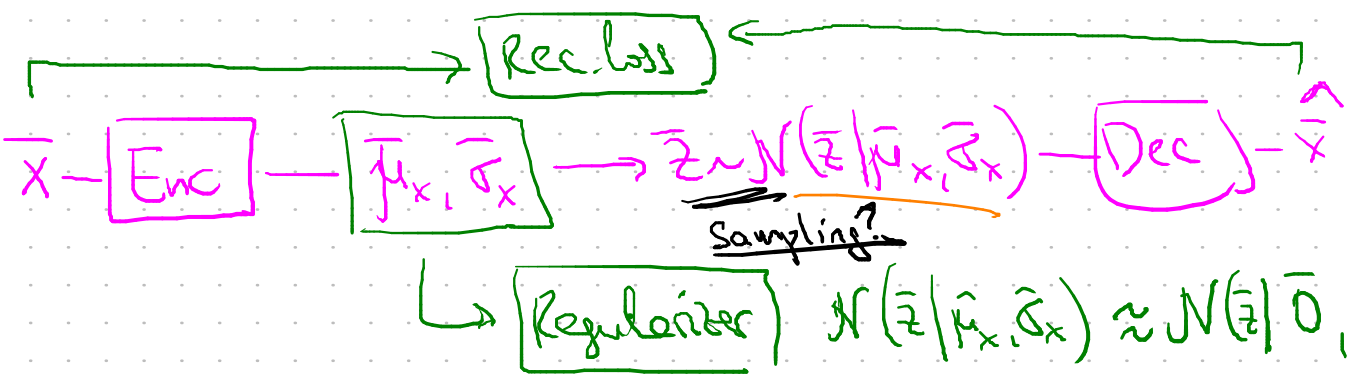


② Variational auto-encoders

$\bar{x}$ — Enc — $\bar{z}$ — Dec — $\hat{\bar{x}}$



Reg: $\bar{\mu}_{\bar{x}} \approx \bar{0}$, $\bar{\sigma}_{\bar{x}} \approx \bar{1}$



③ VAE, take 2

generation

$$p(\bar{z})\,p(\bar{x}\mid\bar{z}) = p(\bar{x},\bar{z}) = p(\bar{x})\,p(\bar{z}\mid\bar{x})$$

$p(\bar{z})$: $\mathcal{N}(\bar{z}\mid\bar{0},\bar{I})$ — decoder: $p(\bar{x}\mid\bar{z})$ — $\mathcal{N}(\bar{x}\mid f(\bar{z}), c\cdot\bar{I})$

encoder: $p(\bar{z}\mid\bar{x})$

$$p(\bar{x})\cdot p(\bar{z}\mid\bar{x}) \approx q(\bar{z}) = \mathcal{N}(\bar{z}\mid\bar{\mu}_x,\bar{\sigma}_x)$$

$$\log p(\bar{x},\bar{z}) = \log p(\bar{x}) + \log p(\bar{z}\mid\bar{x})$$

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

$$\log p(\bar{x}) = \mathbb{E}_q\left[\log p(\bar{x},\bar{z}) - \log p(\bar{z}\mid\bar{x})\right] + \mathbb{E}_q\left[\log q(\bar{z})\right]$$

$p(\bar{z})\,p(\bar{x}\mid\bar{z}) = p(\bar{x},\bar{z})$

$$\log p(\bar{x}) = \mathbb{E}_{q(\bar{z})}\left[\log \frac{p(\bar{x},\bar{z})}{q(\bar{z})}\right] + \mathbb{E}_{q(\bar{z})}\left[\log \frac{q(\bar{z})}{p(\bar{z}\mid\bar{x})}\right]$$

$KL(q\|p)$, $KL(p\|q)$

$$\text{Const} = \mathcal{L}(q) + KL\left(q(\bar{z})\,\|\,p(\bar{z}\mid\bar{x})\right)$$

Variational lower bound

ELBO

q → min

$\Leftrightarrow$

q → max

$$L(q) = \mathbb{E}_q\left[\log \frac{p(\bar{x},\bar{z})}{q(\bar{z})}\right] = \int q(\bar{z}) \log \frac{p(\bar{x},\bar{z})}{q(\bar{z})} d\bar{z} =$$

max

$$\left(p(\bar{x},\bar{z}) = p(\bar{z})\, p(\bar{x}\mid\bar{z}) \right)$$

$$= \int q(\bar{z}) \left(\log p(\bar{z}) + \log p(\bar{x}\mid\bar{z}) - \log q(\bar{z}) \right) d\bar{z}$$

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

$$= \underbrace{\mathbb{E}_{q(\bar{z})}\left[\log p(\bar{x}\mid\bar{z})\right]}_{\text{decoder}} - \underbrace{KL\left(q(\bar{z}) \,\|\, p(\bar{z})\right)}_{L_{reg}}$$

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

$\to \min$

$q(\bar{z}) \approx \mathcal{N}(\bar{z}\mid\bar{0},\bar{I})$

$$\mathbb{E}_{q(\bar{z})}\left[\log \mathcal{N}\left(\bar{x} \mid f(\bar{z}), c\cdot\bar{I}\right)\right] =$$

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

$$\approx \text{Const} - \text{Const} \cdot \underbrace{\sum_{n=1}^{N} \|\bar{x}_n - f(Enc(\bar{x}_n))\|^2}_{L_{rec}}$$

$$q(\bar{z}) = \mathcal{N}(\bar{z}\mid\bar{\mu}_x,\bar{\sigma}_x) = \prod_i \underbrace{\mathcal{N}(z_i\mid\mu_i,\sigma_i)}_{q_i}$$

$$p(\bar{z}) = \prod_i \underbrace{\mathcal{N}(z_i\mid 0,1)}_{p_i}$$

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

$$KL\left(N(z_i|\mu_{x_i},\sigma_{x_i}) \,\|\, N(z_i|0,1)\right) = \dots \log\frac{N_x}{N}$$

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

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

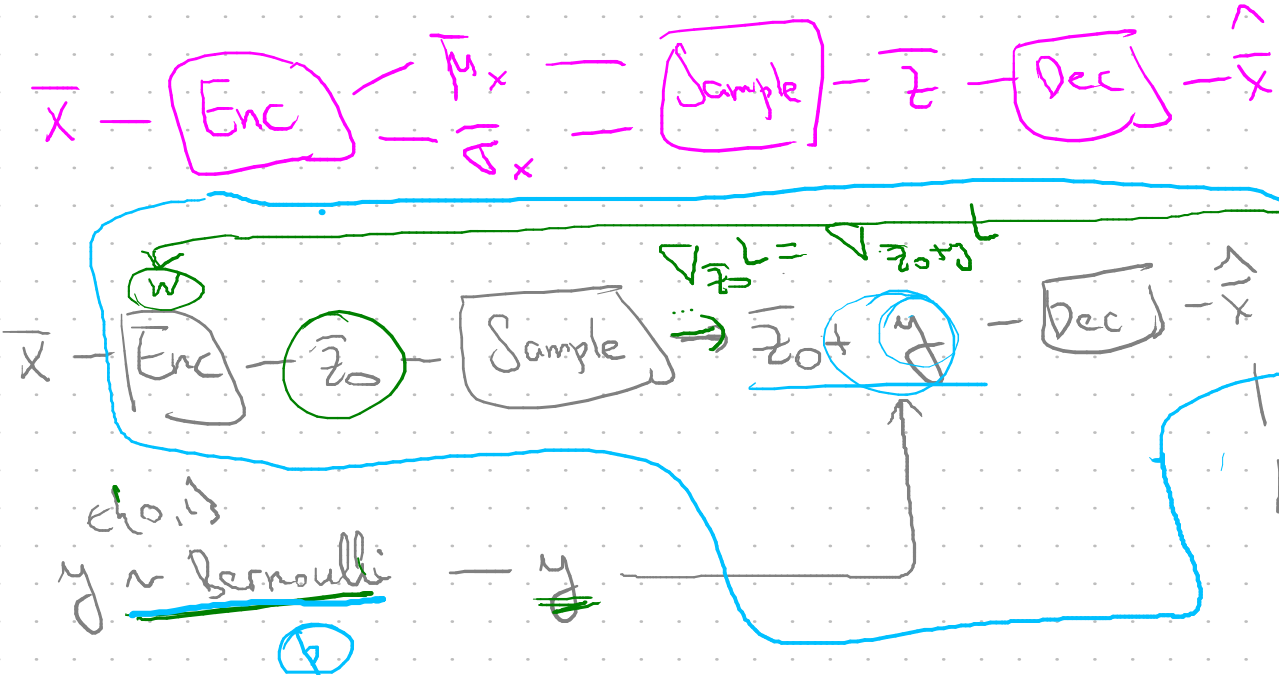
$$\mathbb{E}_{N(x|\mu,\sigma)}[x] = \mu, \quad \mathbb{E}[x^2] = \mu^2+\sigma^2$$

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

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

$$L_{reg} = \sum_{i=1}^{d}\left(\mu_{x_i}^2 + \sigma_{x_i}^2 - \log\sigma_{x_i}^2\right)$$

(4) Reparametrization trick



$$\bar{z} \sim \mathcal{N}(\bar{z} \mid \bar{\mu}_x, \bar{\sigma}_x) \Leftrightarrow \begin{array}{l} \bar{u} \sim \mathcal{N}(\bar{u} \mid \bar{0}, \bar{I}), \\ \bar{z} = \bar{u} \odot \bar{\sigma}_x + \bar{\mu}_x \end{array}$$

$$z_i \sim \mathcal{N}(z_i \mid \mu_{x_i}, \sigma_{x_i}) \Leftrightarrow \begin{array}{l} u \sim \mathcal{N}(u \mid 0, 1) \\ z_i = u \cdot \sigma_{x_i} + \mu_{x_i} \end{array}$$

