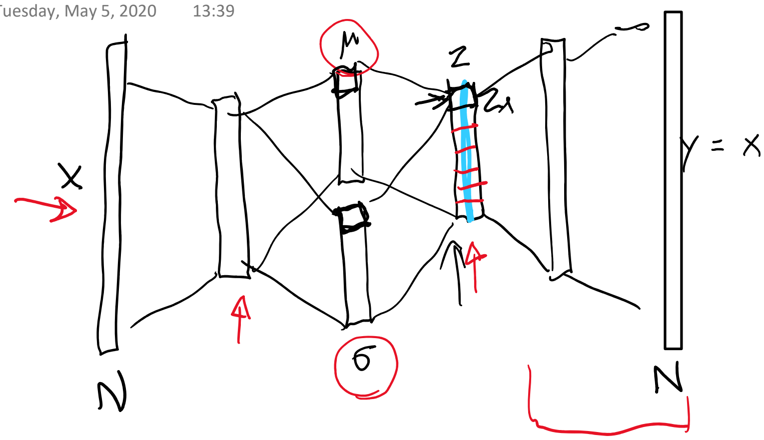
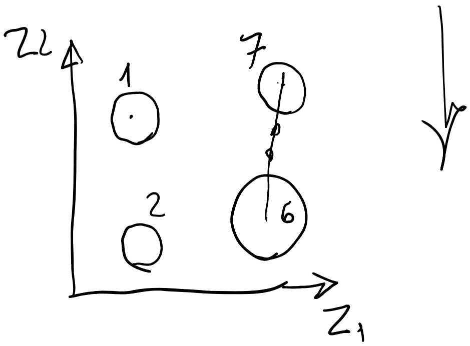
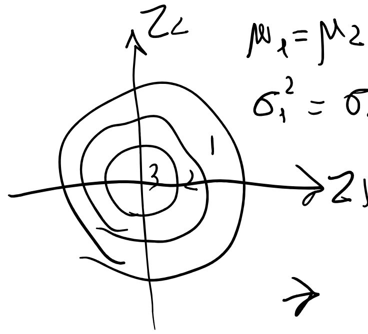
Recurrent (lecture)

Tuesday, May 5, 2020 13:39



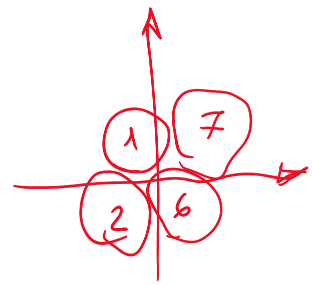
$$L = \sum_{i=1}^{N} \sum_{\mu=1}^{P} \left(y_i^\mu - x_i^\mu \right)^2$$

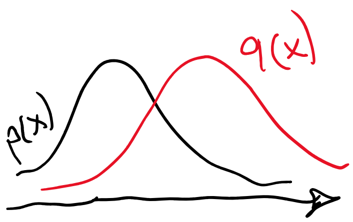




$\mu_1 = \mu_2 = \mu_N$

$\sigma_1^2 = \sigma_2^2 = \sigma_k^2 = \dots = 1$





$$D_{KL}(p \| q) = \int p(x) \left(\log p(x) - \log q(x) \right)$$

$$E_p\left(\log p(x) - \log q(x) \right) =$$

$$= E_p\left(\log \frac{p(x)}{q(x)} \right)$$

$$H = - \int p(x) \log p(x)$$

$$D_{KL}(p \| q) \neq D_{KL}(q \| p)$$

$$p(z) = \frac{e^{-\frac{(z-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \qquad q(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}$$

$$E_p\left(\log \frac{p}{q} \right) = E_p\left(\log \frac{e^{-\frac{(z-\mu)^2}{2\sigma^2} + \frac{z^2}{2}} \cancel{\sqrt{2\pi}}}{\dots} \right) =$$

$$E_p\left(\log\frac{p}{q}\right) = E_p\left(\log \frac{e^{-\frac{(z-\mu)^2}{2\sigma^2}+\frac{z^2}{2}}}{\cancel{\sqrt{2\pi\sigma^2}}}\cancel{\sqrt{2\pi}}\right)$$

$$= E_p\left(\frac{-z^2+2\mu z-\mu^2+\sigma^2 z^2}{2\sigma^2}\right) - \log\sigma^2 =$$

$$= \frac{1}{2\sigma^2}\left(-\langle z^2\rangle + \underbrace{2\mu\overset{\mu}{\langle z\rangle} - \mu^2}_{\mu^2} + \sigma^2\langle z^2\rangle\right) = \frac{1}{2}\left(\mu^2+\sigma^2-1\right) - \log\sigma^2$$

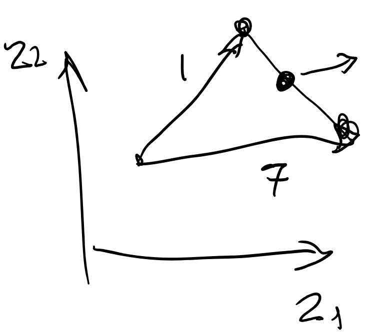
$$\langle z^2\rangle(\sigma^2-1)+\mu^2$$

$$\left(\underbrace{\langle z^2\rangle-\mu^2}_{\sigma^2}+\mu^2\right)(\sigma^2-1)+\mu^2$$

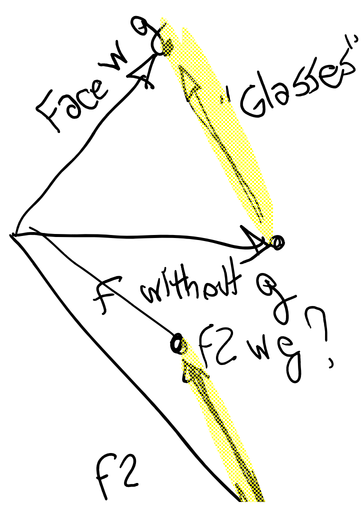
$$D_{KL} = \frac{1}{2}\left(\mu^2+\sigma^2-1\right) - \log\sigma^2 \qquad \mu=0,\ \sigma^2=1$$

$$\boxed{D_{KL} = \sum_{1}^{N_z}{}_{k}\left(\mu_k^2+\sigma_k^2-1\right) - \log\sigma_k^2}$$

$$L = \text{Rec.} + D_{KL}$$



$$\frac{\mu_1+\mu_7}{2}$$



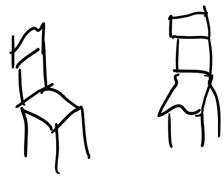
f2

βVAE

$$\boxed{L = \text{Rec.} + \beta D_{KL}}$$

$\beta = 1 \Rightarrow$ Standard VAE

$\beta > 1 \Rightarrow$ βVAE



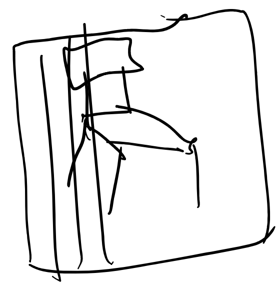
$\tilde{z}$

Azimuth

size

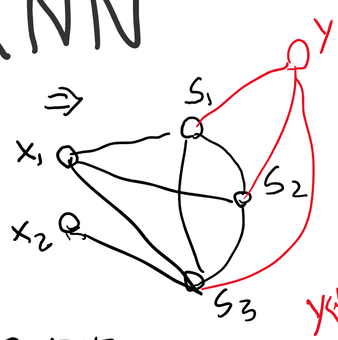
style

$\Rightarrow$



$\Rightarrow$ βVAE

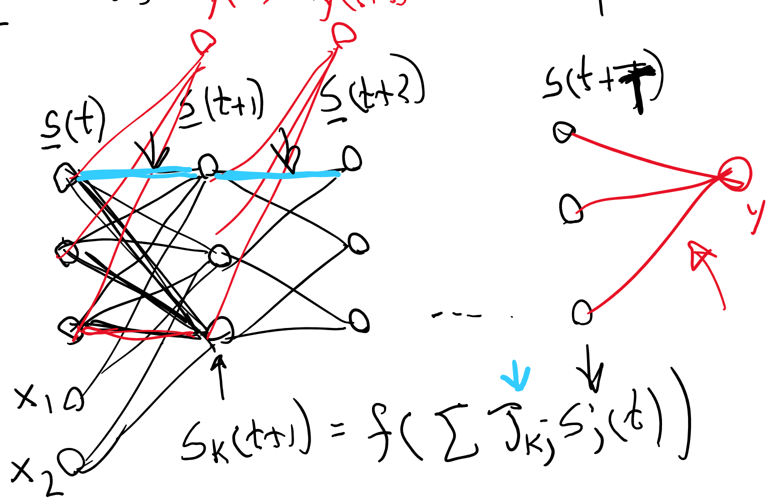
RNN



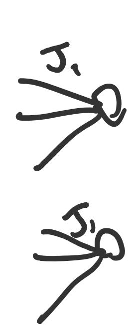
$$s_k(t+1) = f\left(\sum_j^N J_{kj} s_j(t) + \sum_j^N \tilde{J}_{kj} x_j(t)\right)$$

$$\tau \frac{ds_k}{dt} = -s_k + f(\ \ldots\ )$$

BPTT



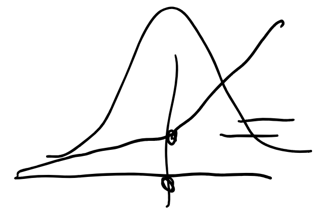
$$s_k(t+1) = f\left(\sum J_{kj} s_j(t)\right)$$

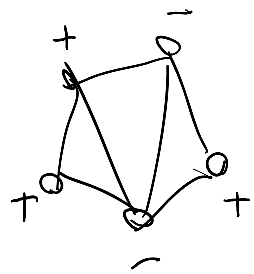


$\Delta J_{ij}(t)$

$$\boxed{\Delta J_{11} = \sum_{t}^{T} \Delta J_{11}(t)}$$

$\lambda_k < 1$ $\forall k$ $\quad (J^T)^{\mathbf{T}} \underline{\delta} \to 0$





$$J_{ij} = \sum_{\mu}^{p} \xi_i^{\mu} \xi_j^{\mu}$$