Wednesday, April 14, 2021 10:31 AM

Lecture 22

Variational Autoencoders (2013)

Last time: Vanilla AE not suitable for generative modeling

- latent space doesn't have a fixed dist'n, so how to sample?
- latent space has gaps

VAE fixes these problems and can be used for generation.

Idea: add term to loss fn that enforces latent space regularity & known dist'n

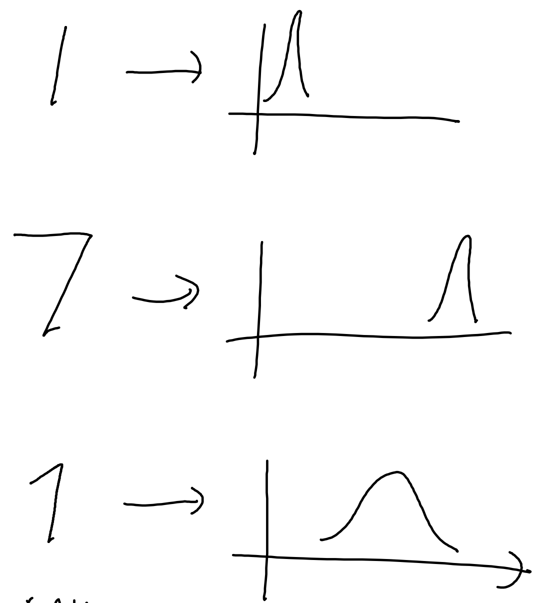
Idea: Vanilla AE | VAE

$x \longrightarrow z$ | $x \longrightarrow$  μ,σ

every pt in latent is smeared out and nearby pts in z are similar in x.

Wednesday, April 14, 2021 10:39 AM

close in z (as measured by σ) $\longrightarrow$ close in x



width of gaussian can encode how "distinct" an input is

Statistical Treatment

latent variables

Given x from data, we want to infer z statistically

Want: $\underline{p(z|x)} = \dfrac{p(x|z)\,p(z)}{p(x)}$ (Bayes Thm)

need to know full prob. dist'n of data $\rightarrow$ difficult, often intractable.

Instead: try to learn tractable $q(z|x)$ to be as close as possible to $p(z|x)$. "variational inference"

Wednesday, April 14, 2021 10:54 AM

$$\min KL\left(q(z|x) \,||\, p(z|x)\right)$$

$$\hookrightarrow \int dz\, q(z|x) \log \frac{q(z|x)}{p(z|x)} \overset{\text{(Bayes Thm)}}{=} \int dz\, q(z|x) \log \frac{q(z|x)\, p(x)}{p(x|z)\, p(z)}$$

$$\int dz\, q(z|x) \log p(x) = \log p(x) = \text{const.}$$

$$= -\left[E_{q(z|x)}\left[\log p(x|z)\right] - KL\left(q(z|x) \,||\, p(z)\right) + \text{const} \right]$$

"reconstruction likelihood"

promotes $q(z|x)$ to be similar to prior dist'n on latent space $p(z)$. for every x.

promote smooth, regular latent space!

So min KL(...)

$$= \max\left[\text{reconstruction likelihood} - KL\right]$$

Wednesday, April 14, 2021 11:06 AM

Implementation:

"prior"

Assume: • $p(z) = \underset{\text{PDF}}{\mathcal{N}(0,1)}(z)$

• $q(\vec{z}|x) = \mathcal{N}(\vec{\mu}(x), \vec{\sigma}(x))$ (assume uncorrelated latent space)

"variational hypothesis"

can be a source of suboptimality.

2nd term (KL) becomes tractable!

$$KL\left(\mathcal{N}(\mu(x), \sigma(x)) \,\|\, \mathcal{N}(0,1)\right) = \frac{1}{2}\left[-1 + \mu(x)^2 + \sigma(x)^2 - \log\sigma(x)^2\right]$$

Assume: $p(x|z) \sim e^{-\frac{\|x - f(z)\|^2}{2\beta}}$

decoder!
separate fn to minimize over.

Then 1st term is tractable

Wednesday, April 14, 2021 11:17 AM

$$\max\left(E_{q(z|x)} \log p(x|z) - KL \cdots \right)$$

encoder: $\mu(x), \sigma(x)$
decoder: $f(z)$.
$\Big\}$ all NNs in practice.

$$= \frac{1}{2\beta} \min\left(E_{q(z|x)}\left[\|x - f(z)\|^2 \right] + \beta\left[-1 + \mu(x)^2 + \sigma(x)^2 - \log \sigma(x)^2 \right] \right)$$

(β-) VAE loss

need to sample from $q(z|x) = N(\mu(x), \sigma(x))$

but also need to be able to differentiate wrt to parameters of $\mu(x), \sigma(x)$.

"Reparametrization trick"

instead of $z \sim N(\mu(x), \sigma(x))$

$g \sim N(0,1)$ $\quad z = \mu(x) + g \cdot \sigma(x)$.

Now $f(z) = f(\mu(x) + g\sigma(x))$ is backpropable wrt. par. of $\mu(x)$ & $\sigma(x)$.

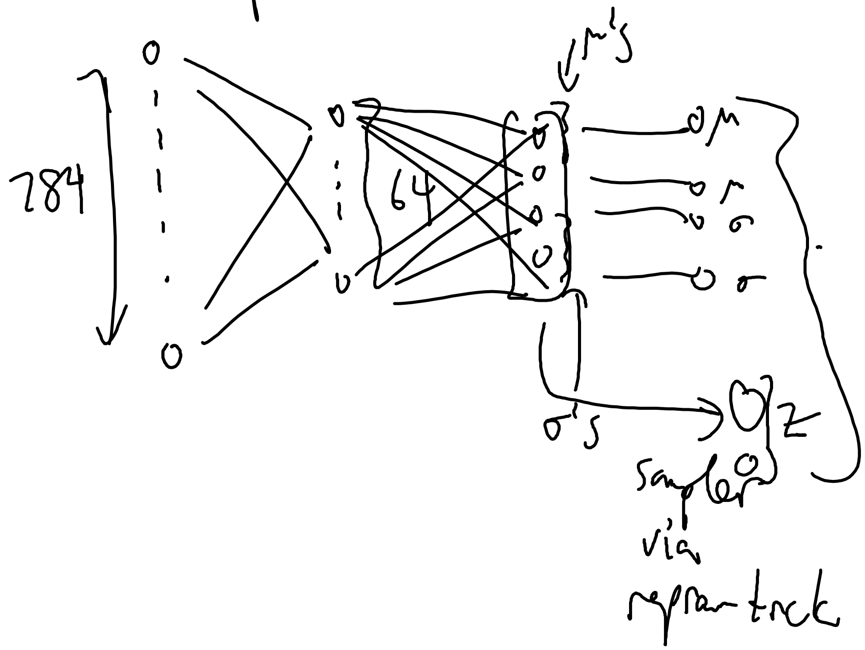
regularization

higher $\beta \rightarrow$ smoother latent space!

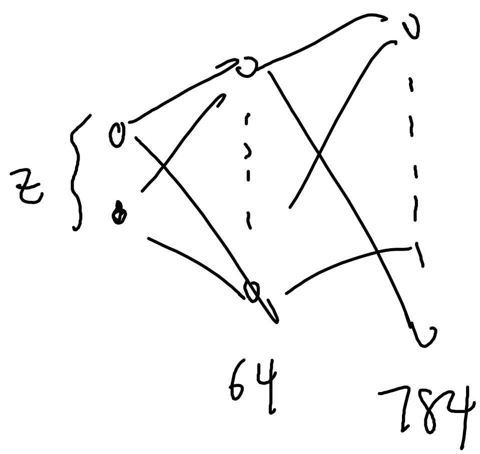
often have to tune β to get good performance.

Wednesday, April 14, 2021 11:29 AM

Example VAE - MNIST from Keras blog



encoder



decoder