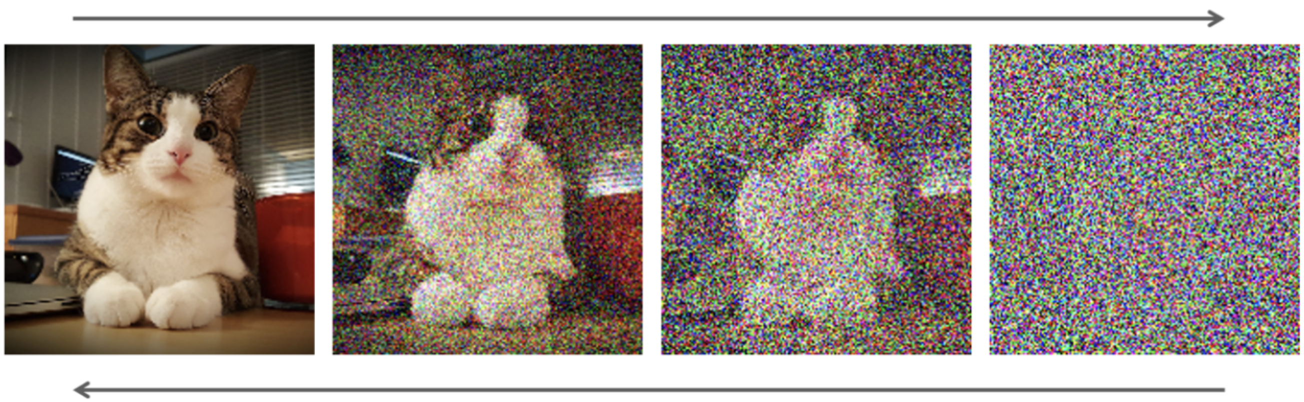
TTIC 31230, Fundamentals of Deep Learning

David McAllester, Autumn 2022

Markovian VAEs

Markovian VAEs

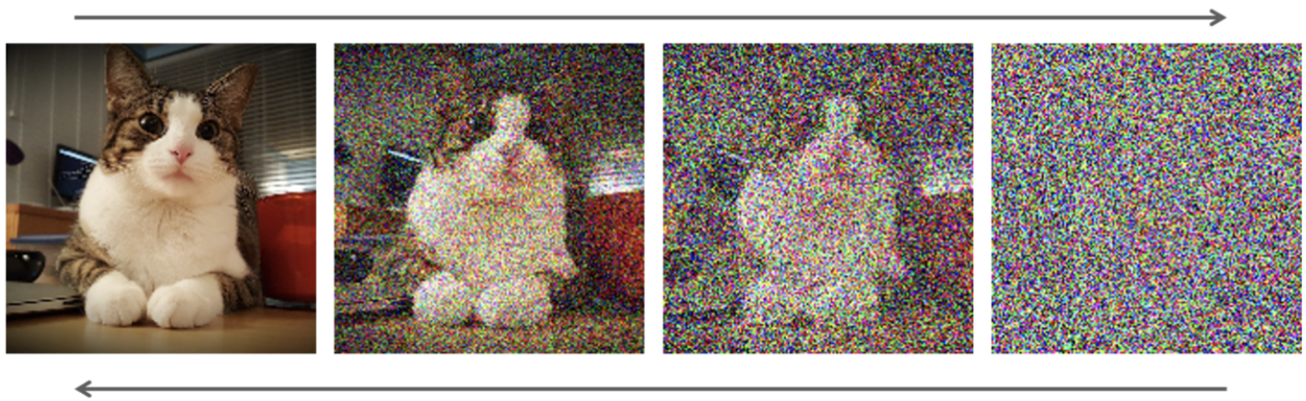
A diffusion models computes and inverts a sequence



So does an autoregressive language model

[Sally talked to John] $\rightleftarrows$ [Sally talked to] $\rightleftarrows$ [Sally talked] $\rightleftarrows$ [Sally]

Markovian VAEs



[Sally talked to John] $\rightleftarrows$ [Sally talked to] $\rightleftarrows$ [Sally talked] $\rightleftarrows$ [Sally]

$$y \rightleftarrows z_1 \rightleftarrows \cdots \rightleftarrows z_N$$

Markovian VAEs

$$y \overset{\rightarrow}{\leftarrow} z_1 \overset{\rightarrow}{\leftarrow} \cdots \overset{\rightarrow}{\leftarrow} z_N$$

Encoder: $\mathrm{Pop}(y)$, $P_{\mathrm{enc}}(z_1|y)$, and $P_{\mathrm{enc}}(z_{\ell+1}|z_\ell)$.

Generator: $P_{\mathrm{pri}}(z_N)$, $P_{\mathrm{gen}}(z_{\ell-1}|z_\ell)$, $P_{\mathrm{gen}}(y|z_1)$.

The encoder and the decoder define distributions $P_{\mathrm{enc}}(y, \ldots, z_N)$ and $P_{\mathrm{gen}}(y, \ldots, z_N)$ respectively.

VAE Review

A variational autoencoder (VAE) has only y and z.

$$P_{\mathrm{enc}}(y, z) \;=\; \mathrm{Pop}(y) P_{\mathrm{enc}}(z|y)$$

$$P_{\mathrm{gen}}(y, z) \;=\; P_{\mathrm{pri}}(z) P_{\mathrm{gen}}(y|z)$$

The Single Layer ELBO

$$
\begin{aligned}
H_{\mathrm{Pop}}(y) \;&=\; E_{\mathrm{enc}}\left[-\ln \frac{P_{\mathrm{enc}}(y)P_{\mathrm{enc}}(z|y)}{P_{\mathrm{enc}}(z|y)}\right] \\
&=\; E_{\mathrm{enc}}\left[-\ln \frac{P_{\color{red}{\mathrm{enc}}}(z)P_{\color{red}{\mathrm{enc}}}(y|z)}{P_{\mathrm{enc}}(z|y)}\right] \\
&\color{red}{\leq}\; E_{\mathrm{enc}}\left[-\ln \frac{P_{\color{red}{\mathrm{gen}}}(z)P_{\color{red}{\mathrm{gen}}}(y|z)}{P_{\mathrm{enc}}(z|y)}\right] \quad \text{cross-entropy bounds entropy} \\
&=\; E_{\mathrm{enc}}\;\; KL(P_{\mathrm{enc}}(z|y),\; P_{\mathrm{gen}}(z)) + E_{\mathrm{enc}}[-\ln P_{\mathrm{gen}}(y|z)]
\end{aligned}
$$

The Markovian ELBO

$$
\begin{aligned}
H(y) &= E_{\mathrm{enc}}\left[-\ln \frac{P_{\mathrm{enc}}(y)P_{\mathrm{enc}}(z_1,\ldots,z_N|y)}{P_{\mathrm{enc}}(z_1,\ldots,z_N|y)}\right] \\
&= E_{\mathrm{enc}}\left[-\ln \frac{P_{\mathrm{enc}}(y|z_1)P_{\mathrm{enc}}(z_1|z_2)\cdots P_{\mathrm{enc}}(z_{N-1}|z_N)P_{\mathrm{enc}}(z_N)}{P_{\mathrm{enc}}(z_1|z_2,y)\cdots P_{\mathrm{enc}}(z_{N-1}|z_N,y)P_{\mathrm{enc}}(z_N|y)}\right] \\
&\leq E_{\mathrm{enc}}\left[-\ln \frac{P_{\mathrm{gen}}(y|z_1)P_{\mathrm{gen}}(z_1|z_2)\cdots P_{\mathrm{gen}}(z_{N-1}|z_N)P_{\mathrm{gen}}(z_N)}{P_{\mathrm{enc}}(z_1|z_2,y)\cdots P_{\mathrm{enc}}(z_{N-1}|z_N,y)P_{\mathrm{enc}}(z_N|y)}\right] \\
&= \left\{\begin{array}{l}
E_{\mathrm{enc}}\left[-\ln P_{\mathrm{gen}}(y|z_1)\right] \\
\\
+\sum_{i=2}^{N} \; E_{\mathrm{enc}}\; KL(P_{\mathrm{enc}}(z_{i-1}|z_i,y),\;\; P_{\mathrm{gen}}(z_{i-1}|z_i)) \\
\\
+E_{\mathrm{enc}}\; KL(P_{\mathrm{enc}}(Z_N|y), p_{\mathrm{gen}}(Z_N))
\end{array}\right.
\end{aligned}
$$

END