TTIC 31230, Fundamentals of Deep Learning

David McAllester, Autumn 2021

Progressive VAEs

Progressive VAEs

I am going to describe a general method of constructing multi-layer VQ-VAEs that is different from the method described in the VQ-VAE paper but seems (to me) better theoretically motivated.

The method described here trains a progressive GAN generator with a VAE objective rather than a discriminator.

Progressive VAEs

We consider a general VAE with a sequence of latent variables $z_1, \ldots, z_n$ where we want to model an observable variable s.

In the image VQ-VAE $z_1, \ldots, z_N$ is a sequence of symbolic images of increasing spacial dimension.

A progressive VAE has the following structure on the prior and decoder.

$$P_\Phi(z_1, \ldots, z_N) = P_\Phi(z_1) \prod_{i=2}^{N} P_\Phi(z_i \mid z_{i-1})$$

$$P_\Theta(s|z_1, \ldots, z_N) = P_\Theta(s|z_N)$$

This copies the structure of a progressive GAN generator.

Progressive VAE Training

The discriminator of the progressive GAN is replaced by a series of encoders that go up from a given image to create the latent layers.

The encoder is analogous to a GAN inverter.

But here, recognizing the automony of the encoder for VAEs, we train the generator to match the inverter rather than other way around.

Progressive VAE Training

First we train $P_\Psi(z_N|s)$ and $P_\Theta(s|z_N)$ under a fixed simple prior on z_N as in phase 1 of a two phase training of a single layer.

Next we train $P_\Psi(z_{N-1}|z_N)$ and $P_\Theta(z_N|z_{N-1})$ under a fixed simple prior on z_{N-1} as in single layer phase 1 training.

We continue in this way training $P_\Psi(z_{i-1}|z_i)$ and $P_\Theta(z_i|z_{i-1})$

Progressive VAE Training

We now have an encoder $P_\Psi(z_N, \ldots, z_1|s)$ and appropriately strong decoders $P_\Theta(s|z_N)$ and $P_\Theta(z_{i-1}|z_i)$.

For each s we how have a highest latent variable $z_1(s)$ and we can train a strong auto-regressive prior $P_\Phi(z_1)$.

We then define the prior by

$$P_\Phi(z_1, \ldots, z_n) = P_\Phi(z_1) \prod_{i=2}^{N} P_\Theta(z_i \mid z_{i-1})$$

Progressive VAE Motivation

As in two-phase VAE training, the progressive VAE training is motivated by the autonomy of encoder.

The encoder (the GAN inverter) is less important than the prior and the decoder (the GAN generator) as long as the prior and decoder are trained on latent structure defined by the encoder.

END