Variational Autoencoders: an introduction

Talk given at the Alma AI research center
University of Bologna
April 28, 2021

Andrea Asperti

DISI - Department of Informatics: Science and Engineering
University of Bologna
Mura Anteo Zamboni 7, 40127, Bologna, ITALY
andrea.asperti@unibo.it

Generative modeling

The goal of generative modeling is to learn the probability distribution P(X) of a given dataset (the likelihood that X belongs to the dataset).

Usually, we try to learn the distribution in an implicit way, training a generator to sample data according to P(X).

There are two main classes of geneative models:

- Generative Adversarial Networks (GANs)
- Variational Autoencoders (VAEs)

In this talk we focus on VAEs.

Latent variables models

In latent variable models we express the probability of a data point X through marginalization over a vector of latent variables:

$$P(X) = \int P(X|z)P(z)dz \approx \mathbb{E}_{z \sim P(z)} P(X|z) \qquad (1)$$

The term $P(X|z)$ can be understood as the generator sampling X given the latent encoding z.

GANs and VAEs differ in the way the generator is trained.

Exploiting an auxiliary distribution

The variational approach renounce to directly sample from the prior distribution $P(z)$, exploiting instead an auxiliary distribution $Q(z)$ (sometimes called inference distribution):

$$\mathbb{E}_{z \sim Q(z)} P(X|z) \tag{2}$$

But why does this help, and how it is related to $P(X)$?

KL at work

In order to answer to the previous questions, it is convenient to start from the Kullback-Leibler divergence betwenn $Q(z)$ and $P(z|X)$:

$$KL(Q(z)||P(z|X)) = \mathbb{E}_{z \sim Q(z)}\, log \frac{Q(z)}{P(z|X)} \tag{3}$$

or also, exploiting Bayes rule,

$$KL(Q(z)||P(z|X)) = \mathbb{E}_{z \sim Q(z)}\, log \frac{Q(z)P(X)}{P(X|z)P(z)} \tag{4}$$

KL at work

In equation (4), the term $P(X)$ does not depend on z and may come out of the expectation:

$$KL(Q(z)||P(z|X)) = log(P(X)) + \mathbb{E}_{z\sim Q(z)}\, log\frac{Q(z)}{P(X|z)P(z)} \quad (5)$$

and via a few additional manipulations:

$$\begin{aligned}
&= log(P(X)) + \mathbb{E}_{z\sim Q(z)}\, log\tfrac{Q(z)}{P(X|z)P(z)} \\
&= log(P(X)) + \mathbb{E}_{z\sim Q(z)}\, log\tfrac{1}{P(X|z)} + \mathbb{E}_{z\sim Q(z)}\, \tfrac{Q(z)}{P(z)} \\
&= log(P(X)) - \mathbb{E}_{z\sim Q(z)}\, log(P(X|z)) + KL(Q(z)||P(z))
\end{aligned} \quad (6)$$

ELBO

Summing up:

$$log(P(X)) - KL(Q(z)||P(z|X)) = \underbrace{\mathbb{E}_{z\sim Q(z)}\, log(P(X|z)) - KL(Q(z)||P(z))}_{ELBO} \tag{7}$$

Since the Kullback-Leibler divergence is always positive, the term on the right is a lower bound to the loglikelihood $P(X)$, known as Evidence Lower Bound (ELBO).

Introducing the encoder

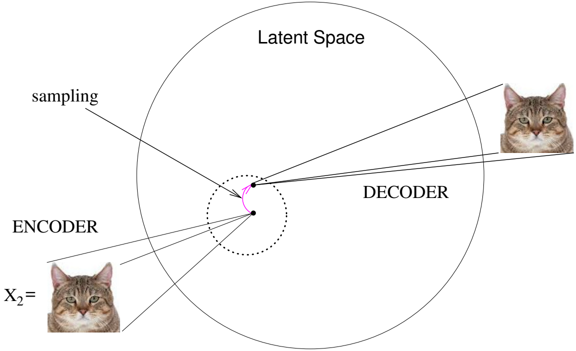
In Equation 7, $Q(z)$ can be any distribution; the trick is to take one depending on X: the "encoder" $Q(z|X)$ encoding X to its latent representation z.

$Q(z|X)$ should hopefully resemble $P(z|X)$, so that the quantity $KL(Q(z|X)||P(z|X))$ is small.

In this case the loglikelihood $P(X)$ is close to the Evidence Lower Bound; our learning objective is its maximization:

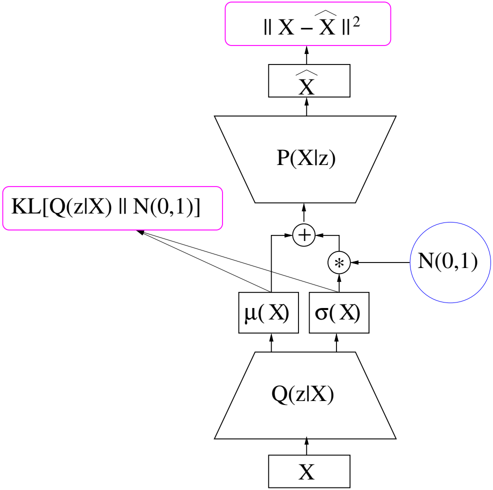
$$log(P(X)) \approx \underbrace{\mathbb{E}_{z \sim Q(z|X)} log(P(X|z))}_{loglikelihood} - \underbrace{KL(Q(z|X)||P(z))}_{regularizer} \quad (8)$$

Sampling in the latent space



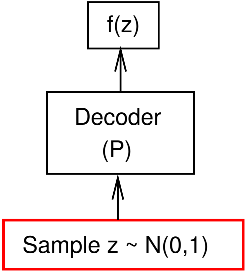
Among other things, sampling add noise to the encoding, improving its robustness.

The full picture



Generation of new samples

$\mu(X)$ and $\Sigma(X)$ are not used to generate new samples from the input domain (we have no X)



Problems with VAE

- balancing loglikelihood and KL regularizer in the loss function
- variable collapse phenomenon
- aggregate inference vs prior mismatch
- blurriness (aka variance loss)

Additional links

- Turorial on Variational Autoencoders by C.Doersch
- Lesson on Variational Autoencoders by R.Yeh, J.Lou, T.Li
- A survey on Variational Autoencoders from a GreenAI perspective by A.Asperti, D.Evangelista, E.Piccolomini.