Variational Auto-Encoders

Background

Let x denote the observable data and z denote the corresponding latent variable that can describe the data. The generative process of the Variational Auto-Encoders (VAEs) with Gaussian prior of z and Bernoulli likelihood of x is defined as follows:

$$
\begin{aligned}
z &\sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \\
\mu_x &= f(z) \\
x &\sim \mathcal{B}\left(\mu_x\right),
\end{aligned}
$$

where $\mathcal{N}(\mathbf{0}, \mathbf{I})$ is the standard Gaussian distribution, f is parameterized by a deep neural network and $\mathcal{B}(\mu_x)$ is the multivariate Bernoulli distribution. As in our case where x is multi-dimensional, each dimension of x is sampled independently given the corresponding dimension of μ_x following one-dimensional Bernoulli distribution.

To infer the latent variable z given observable x, VAEs build a recognition model as follows:

$$
\begin{aligned}
\mu_z &= g_\mu(x) \\
\sigma_z &= g_\sigma(x) \\
z &\sim \mathcal{N}\left(\mu_z, \sigma_z\right),
\end{aligned}
$$

where g_μ and g_σ are parameterized as deep neural networks, which can share most of the parameters. $\mathcal{N}(\mu_z, \sigma_z)$ is a multivariate distribution with diagonal covariance given μ_z and σ_z.

To jointly learn the parameters in all networks, VAEs defines a variational lower-bound of the marginal data likelihood for per data, i.e. $p(x)$ and sums over them.

Install ZhuSuan and Get the Code Files

ZhuSuan is a python library for Generative Models, built upon Tensorflow. Unlike existing deep learning libraries, which are mainly designed for supervised tasks, ZhuSuan is featured for its deep root into Bayesian Inference, thus supporting various kinds of generative models: both the traditional hierarchical Bayesian models and recent deep generative models.

Run the following comment to download ZhuSuan from Github:

```
git clone https://github.com/thu-ml/zhusuan.git
```

Run the following commands to install ZhuSuan:

```
cd zhusuan/
git reset --hard 10ead0
export PYTHONPATH=[YourPathToZhusuan]:$PYTHONPATH
```

Preparation

- In this homework, we use Tensorflow version==**1.0.0**
- Case if you are not familiar with Tensorflow, learn it by these tutorials:
  - https://www.tensorflow.org/get_started/mnist/beginners
  - https://www.tensorflow.org/get_started/mnist/pros
  - https://www.tensorflow.org/get_started/mnist/mechanics
  - https://www.tensorflow.org/programmers_guide/dims_types
  - https://www.tensorflow.org/programmers_guide/variables
  - https://www.tensorflow.org/programmers_guide/variable_scope
- **Attention**: We don't use the latest version of ZhuSuan for this homework. Look for the examples and tutorials on the github two years ago.

Report

- Implement the VAEs on ZhuSuan. Please see detailed instructions in the code file named by `vae_basic.py`. You need fill some code in the space with the `TODO` comments. Note that VAEs use thousands of epochs to converge completely while training tens of epochs is sufficient in this assignment due to the time limit.
- Randomly sample 100 latent variables z from Gaussian prior and plot their corresponding images using the trained generator in a 10×10 grid layout.
- Change the likelihood function to Gaussian with diagonal covariance matrix. That is, the generative process should be:

$$
\begin{aligned}
z &\sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \\
\mu_x &= f_\mu(z) \\
\sigma_x &= f_\sigma(z) \\
x &\sim \mathcal{N}\left(\mu_x, \sigma_x\right)
\end{aligned}
$$

  Please see detailed instructions in the code file named by `vae_gaussian.py`. Also plot 100 sampled images and discuss the difference compared to the above results.

Submission Guideline:

You need to submit both report and codes, which are:

- **report**: well formatted and readable summary including your results, discussions and ideas. Source codes should *not* be included in report writing. Only some essential lines of codes are permitted for explaining complicated thoughts.

- **codes**: organized source code files with README for extra modifications or specific usage. Ensure that others can successfully *reproduce* your results following your instructions. **DO NOT include model weights/raw data/compiled objects/unrelated stuff over 50MB**

Deadline: Jun. 12

TA contact info: Yulong Wang (王宇龙）, wang-yl15@mails.tsinghua.edu.cn