Exercise 3: Learning as Bayesian inference

Please upload to moodle by 27.10.2020 (23:59pm) a pdf with your written solutions (preferably in LaTeX) **and** upload the modified file `utils.py` for the coding part of the assignment.
Always provide your name, student ID and email address at the top of the files.

1 Bayesian updating

Bayesian inference, an application of the famous Bayes' theorem, is a way to (recursively) update a prior belief about a *hypothesis* y with the help of a *model* $p(x|y)$ for the observed data x. This results in updating the prior $p(y)$ into the posterior $p(y|x)$. The posterior distribution incorporates the additional evidence that comes from observing x.

(a) Assuming that $p(x|y) = \mathcal{N}(x; y, \sigma^2)$, where

$$\mathcal{N}(x; y, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-y)^2}{2\sigma^2}} \tag{1}$$

and that $p(y) = \mathcal{N}(y; \mu, \phi^2)$, obtain $p(y|x)$ (or at least show what is proportional to it.)

Hint: Assume that the resulting posterior is again Gaussian, and use the results that we obtained in the exercise for multiplying two univariate Gaussians.

(b) Consider limit situations ($\sigma^2 \to 0, \sigma^2 \to \infty, \phi^2 \to 0, \phi^2 \to \infty$) and interpret the results.

2 Variational inference

In the first exercise above, we analytically computed the posterior distribution given the model likelihood and the prior. As discussed in the exercise, this procedure is intractable for more complex distributions. Variational inference offers a way to approximate this posterior we wish to obtain.

2.1 Re-derive ELBO

In the exercise, we derived the *evidence lower bound.* This is an important quantity we use to approximate an intractable posterior such as $p(W|X)$ with W the weights of a neural network.

Re-derive the ELBO by starting with $KL(q(W)||p(W|X))$.

2.2 Analytical form for the KL divergence between two Gaussians

Given two univariate normal distributions $\mathcal{N}(x; \mu_1, \sigma_1^2)$ and $\mathcal{N}(x; \mu_2, \sigma_2^2)$, compute the Kullback–Leibler divergence between the two distributions, i.e., $\mathrm{KL}(\mathcal{N}(x; \mu_1, \sigma_1^2) \;||\; \mathcal{N}(x; \mu_2, \sigma_2^2))$.

Why is the Kullback–Leibler divergence also called the relative entropy?

2.2.1 Prior matching term for a simple prior

In one of the two terms of the ELBO, we need to compute the Kullback–Leibler divergence of two distribution - this term is called the *prior-matching* term, $KL(q(W)||p(W))$. What is the analytical form of this term when the prior $p(W)$ is a standard univariate normal distribution, and $q(W)$ is an arbitrary univariate Gaussian distribution?

You will need this expression of the KLD for its implementation in Pytorch.

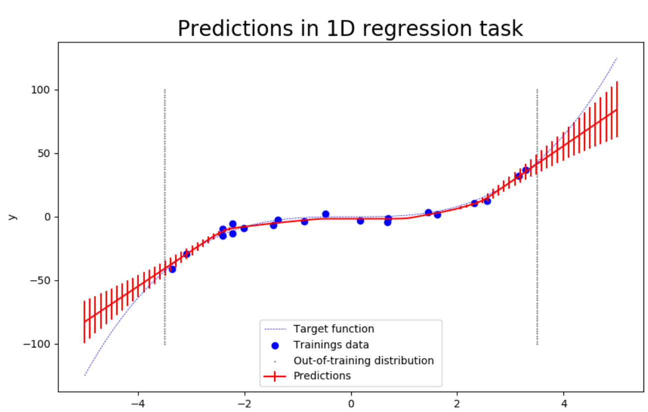
3 Bayes-by-backprop

Please download and unzip the corresponding source files, which contain all implementations as well as the documentation.
You should start by **opening the documentation**. Therefore, you have to open the file `docs/index.html` in your webbrowser.
Furthermore, you have to **setup your Python environment**. You can do so manually (e.g., by installing all missing packages) or use the conda environment that we provide with the file `tutorial3_env.yml`. The provided conda environment (same as for the last exercise) can be installed via[1]

```
$ conda env create -f tutorial3_env.yml
$ conda activate tutorial3_env
```

The goal of this exercise is to implement the missing code in the `computeELBO()`, `sampleGaussian()` and `computeKLD()` functions in module `lib/utils.py`.
The corresponding computational rules were derived in the tutorial session and are partly described in the documentation (that you opened in your webbrowser).

There are no test cases for this exercise.
You should be able to see high variance / uncertainty in the predictions outside the training data range, and low variance within. Similar to the following:



[1] See here for more details: `https://docs.conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html#creating-an-environment-from-an-environment-yml-file`