Deep Learning 2024 Assignment 4

This is an **individual assignment** and its deadline is **Friday, January 17, 2025, 22:00**. You must submit your solution electronically via the Absalon home page.

Diffusion models such as the DDPM and VDM are formulated as hierarchical latent variable models. Arguably, it is more natural to formulate these models through differential equations as is common practice in the closely related score and flow matching methods. In this assignment we will study ordinary and stochastic differential equations (ODEs and SDEs) likelihood approaches to generative modeling.

1 ODE

In this exercise, you will work with ordinary differential equations for density modeling in one dimension. A notebook `ODE.ipynb` implements a class to simulate training and validation data, models, model training and visualization. Your tasks will be about analyzing the approach and try different variations. Specifically, your tasks are:

1.1. **Background on the methodology:** State the generative model equations (see notebook) and log likelihood objective equations (see lecture slides). Just the equations nothing else.

1.2. **Analyze the program:** List the different parts of the notebook with brief descriptions of the function of each part (1-2 sentences) and its relation to the methodology. (Some functions and options in the code are not used. These of course should not be described.

1.3. **Run training with Gaussian training data:** Run the program without changing anything. From what distribution and with what parameters are the training data generated? Include and describe the plot you obtain. Can the model learn the task?

In the program we generate data $y \sim p(y)$ from a distribution $p(y)$ that we can also evaluate exactly. We can therefore calculate the true expected log likelihood: $\frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \log p(y_i)$. This is the number that is returned from the dataset generator program. Compare this number with log likelihood that the final iteration gives. Include in the report the expected log likelihood that the dataset generator returns and the final log likelihood you achieve by running the program. Are the trained models close to the ideal value? Does this agree with the visualizations of the densities?

1.4. **Run training with mixture of Laplace training data:** Now change the program to run with mixture of Laplacians data instead. Do the same as above for this new dataset. Try to change the drift network. Can you improve the performance?

The simplest numerical integrator method is Euler (euler). The program also implements two Runge-Kutta methods (rk2 and rk4). Try to change the integration method. Does this change anything?

1.5. **Complexity:** The calculation of the log likelihood requires numerical integration of the ODE with T integration steps. ($T = 100$ in the program.) How will the complexity scale with T? How will back propagation for parameter estimation scale with T?

2 SDE

In this exercise, you will work with stochastic differential equations for density modeling in one dimension. A notebook `SDE.ipynb` implements a class to simulate training and validation data, models, model training and visualization. Your tasks are:

2.1. **Background on the methodology:** State the generative model (see notebook) and log likelihood lower bound (evidence lower bound, ELBO) objective equations (see lecture slides). Just the equations nothing else.

2.2. **Analyze the program:** List the different parts of the notebook with brief descriptions of the function of each part (1-2 sentences) and its relation to the methodology. No need to describe the code also found in the ODE notebook or code not used.

2.3. **Run training:** Run the program without changing anything. Include and describe the plot you have obtained. Also include in the report the expected log likelihood that the dataset generator returns and the final log likelihood you achieve by running the program.

2.4. **Change parameterization of the model components:** Try to change the network that specifies the variational distribution $q_t(x_t|y) = \mathcal{N}(x_t|\alpha(y,t), \beta^2(y,t))$ and generative model $f(x,t)$ and $\sigma(x,t)$. Some possible parameterizations are given in the notebook. Do you observe any changes to the behavior?

2.5. **Expected log likelihood:** Compare the ELBO values you get to the ODE case. Are they more or less the same? Does the SDE fit to the distributions look good visually? Why do you think that the SDE objective fluctuate more than the ODE objective? Hint: The SDE objective depends upon sampling random t-values.

2.6. **Complexity:** The ELBO objective is so-called simulation-free because it does not require integration of the differential equation. What does that mean for the complexity compared to the ODE approach?