Variational Inference with Normalizing Flows

ATML Group 12

Department of Computer Science, University of Oxford

1032626, 1034125, 1034129, 1036969

UNIVERSITY OF OXFORD

Introduction

- Calculating the true posterior distribution of inference tasks is in most cases an intractable problem.
- Lots of research on approaches for efficient approximation of the posterior, however the resulting classes prove to be of limited expressiveness.
- The authors in [1] introduce the notion of normalizing flows, sequences of invertible transformations applied to a simple initial density, to efficiently create more expressive families of candidate posteriors to be used for variational inference.
- We compare the performance of different types of normalizing flows on the MNIST dataset.

Our Work

- Reproduced experiment on MNIST using Linear Normalizing Flows
- Reproduced experiment on MNIST using NICE
- Extended the ideas of the paper and experimented with Invertible Convolutional Flows
- Created open-source Github repository with code and results: `github.com/ATML-Group-12/normalising_flows`

Theoretical Background

Normalizing flows are sequences of invertible, smooth mappings $f : \mathbb{R}^d \to \mathbb{R}^d$. We define a flow of length K to be K such transformations as follows:

$$\mathbf{z_K} = f_K \circ \ldots \circ f_2 \circ f_1(\mathbf{z_0})$$

where $\mathbf{z_0}$ is a random variable that has a simple initial distribution, and $\mathbf{z_K}$ is the corresponding random variable after applying the transformations. The properties of normalizing flows allow us to calculate the log-density of $\mathbf{z_K}$ efficiently, using the change of variable theorem:

$$\ln q_K(\mathbf{z_K}) = \ln q_0(\mathbf{z_0}) - \sum_{k=1}^{K} \ln \left| \det \frac{\partial f_k}{\partial \mathbf{z_{k-1}}} \right|$$

The normalizing flow is defined as the path of the successive distributions q_K.

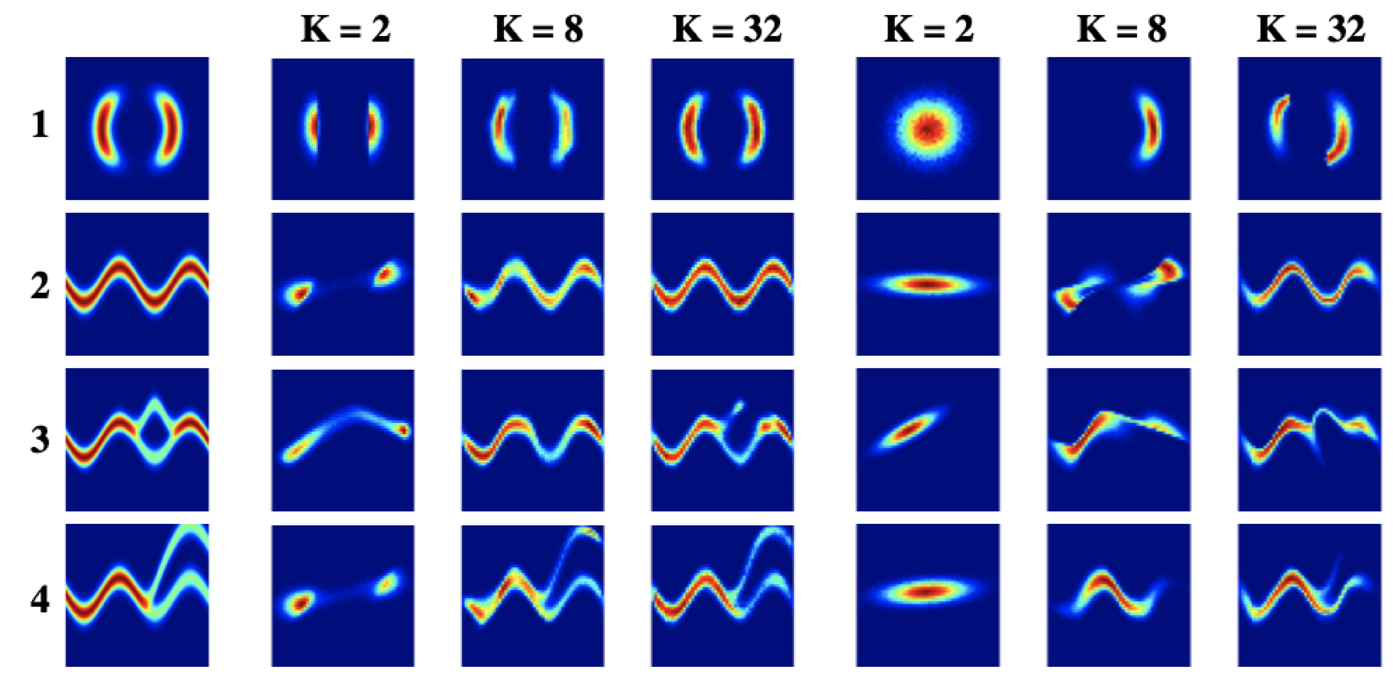
The first class of flows that we use in our experiments is that of linear flows. More specifically, we consider **planar** and **radial** flows, which perform series of contractions and expansions in the direction perpendicular to a fixed hyperplane and around a reference point respectively.

Theoretical Background

TODO We formulate... $\mathcal{S} = \{s_1, s_2, s_3, s_4, s_5, s_6\}$

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Quisque faucibus scelerisque sapien, non commodo ipsum viverra nec. Donec euismod, ipsum at tincidunt viverra, erat metus vestibulum eros, ut venenatis ligula risus vel enim. Vestibulum pulvinar pretium nisi, vel consequat mauris ultrices vitae. Nam a suscipit est. Nunc viverra justo nec justo dignissim posuere. Integer condimentum iaculis est eu posuere. Aliquam egestas viverra sagittis. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Donec auctor sollicitudin tortor, vel sagittis arcu tincidunt sit amet. Mauris sed elementum turpis. Nunc id venenatis nunc.

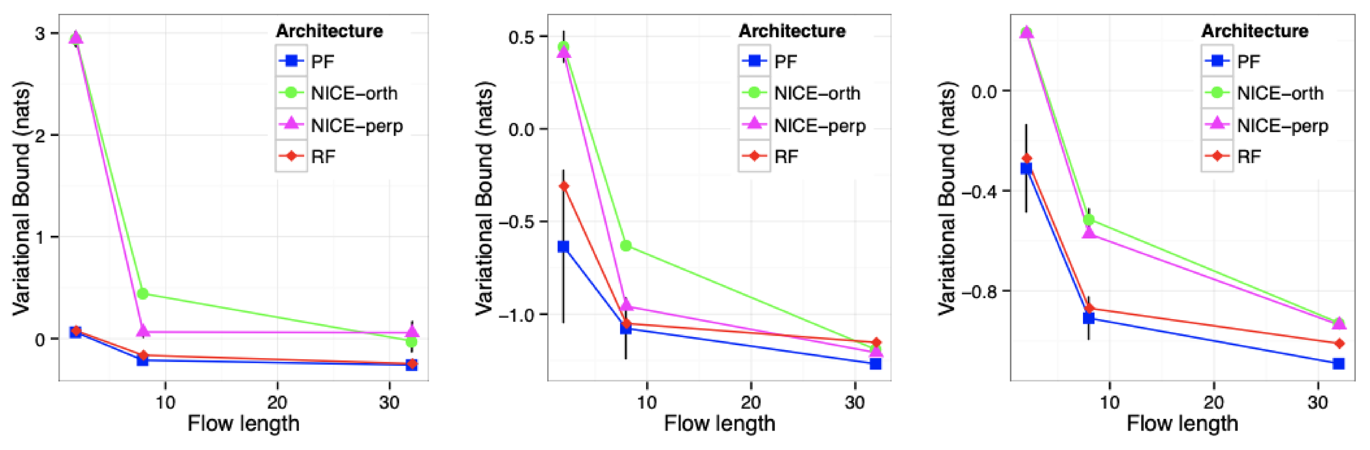
Image example:



(a) (b) Norm. Flow (c) NICE

Experiments

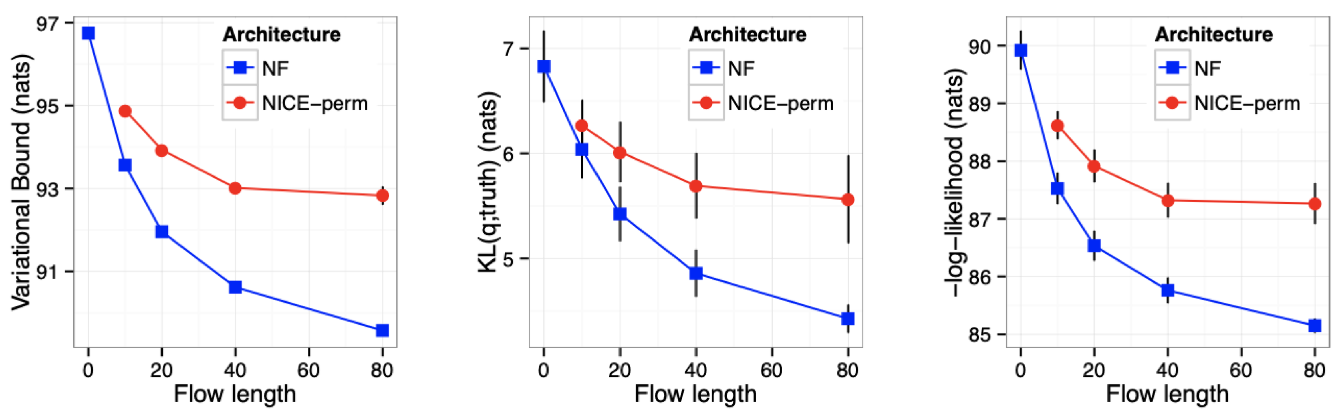
- DLGM + NF
- DLGM + NICE



(d) Comparison of KL-divergences.

Results

TODO!!! Results...



(a) Bound $\mathcal{F}(\mathbf{x})$ (b) $\mathbb{D}_{\mathrm{KL}}(q; p(z|x))$ (c) $-\ln p(\mathbf{x})$

Figure 4. Effect of the flow-length on MNIST.

Our Improvements and Extensions

TODO Pellentesque vitae dui velit. Aenean tincidunt eros facilisis turpis tincidunt, non mollis ipsum venenatis. Praesent consectetur venenatis est, quis rutrum justo faucibus vitae. Nam id orci ex. Aenean id finibus libero. Nam tristique pellentesque eros et mattis. Proin vel nunc accumsan, aliquet leo ut, consectetur sem. Ut ut elit libero. Donec aliquet nulla ac venenatis egestas. Maecenas eu nunc hendrerit turpis dictum laoreet at ut velit. Phasellus tempus tellus id leo bibendum, ac rhoncus turpis molestie. Quisque commodo, quam vitae elementum fermentum, erat dolor hendrerit quam, bibendum malesuada lectus lacus sit amet nisi. In dignissim nisl elit. Aenean vitae enim ut ligula congue vehicula sed non lacus.

References

[1] Danilo Jimenez Rezende and Shakir Mohamed. Variational inference with normalizing flows, 2015.

[2] Laurent Dinh, David Krueger, and Yoshua Bengio. Nice: Non-linear independent components estimation, 2014.