Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2021-0087

Vol. **31**, No. 4, pp. 1049-1082
April 2022

VAE-KRnet and its Applications to Variational Bayes

Xiaoliang Wan[1,*] and Shuangqing Wei[2]

[1] *Department of Mathematics, Center for Computation and Technology, Louisiana State University, Baton Rouge 70803, USA.*
[2] *Division of Electrical & Computer Engineering, Louisiana State University, Baton Rouge 70803, USA.*

Received 21 April 2021; Accepted (in revised version) 12 December 2021

Abstract. In this work, we have proposed a generative model, called VAE-KRnet, for density estimation or approximation, which combines the canonical variational autoencoder (VAE) with our recently developed flow-based generative model, called KRnet. VAE is used as a dimension reduction technique to capture the latent space, and KRnet is used to model the distribution of the latent variable. Using a linear model between the data and the latent variable, we show that VAE-KRnet can be more effective and robust than the canonical VAE. VAE-KRnet can be used as a density model to approximate either data distribution or an arbitrary probability density function (PDF) known up to a constant. VAE-KRnet is flexible in terms of dimensionality. When the number of dimensions is relatively small, KRnet can effectively approximate the distribution in terms of the original random variable. For high-dimensional cases, we may use VAE-KRnet to incorporate dimension reduction. One important application of VAE-KRnet is the variational Bayes for the approximation of the posterior distribution. The variational Bayes approaches are usually based on the minimization of the Kullback-Leibler (KL) divergence between the model and the posterior. For high-dimensional distributions, it is very challenging to construct an accurate density model due to the curse of dimensionality, where extra assumptions are often introduced for efficiency. For instance, the classical mean-field approach assumes mutual independence between dimensions, which often yields an underestimated variance due to oversimplification. To alleviate this issue, we include into the loss the maximization of the mutual information between the latent random variable and the original random variable, which helps keep more information from the region of low density such that the estimation of variance is improved. Numerical experiments have been presented to demonstrate the effectiveness of our model.

AMS subject classifications: 62C10, 62G07, 65C20, 65C60

Key words: Deep learning, variational Bayes, uncertainty quantification, Bayesian inverse problems, generative modeling.

*Corresponding author. *Email addresses:* xlwan@math.lsu.edu (X. Wan), swei@lsu.edu (S. Wei)

1 Introduction

The density estimation of high-dimensional data plays an important role in unsupervised learning, which is challenging due to the curse of dimensionality [27]. In the last decade, deep generative modeling has made a lot of progress by incorporating with deep neural networks. Deep generative models are usually with likelihood-based methods, such as the autoregressive models [14, 22–24], variational autoencoders (VAE) [16, 18, 21], and flow-based generative models [3, 6–9, 19, 25, 33]. One flexible model that does not need the likelihood is the generative adversarial network (GAN) [1, 13], which seeks a Nash equilibrium of a zero-sum game between the generator and the discriminator. Recently, the coupling of different modeling strategies has also been explored. The flow-based model was coupled with GAN in [15] to obtain a likelihood for GAN; The VAE, flow-based model and GAN were coupled in [34] for more flexibility and efficiency. The main goal of deep generative models is to generate new data that are consistent with the underlying distribution of the available data. To achieve this, a specific density model is not a necessity, e.g., GAN manages to focus on the mapping from a standard Gaussian to the desired data distribution without using the likelihood. Other than GANs, deep generative models usually provide a density model, e.g., the flow-based models actually define an invertible transport map between two random variables which yields an explicit push-forward measure. A common characteristic of deep generative models is that they employ neural networks to model the mapping between high-dimensional inputs and outputs whenever needed. Such a strategy is proved to be very effective for application problems although the models are usually not easy to analyze due to the strong nonlinearity induced by neural networks.

Classical density estimation techniques such as kernel density estimation and mixture of Gaussians, suffer severely from the curse of dimensionality, meaning that they are only effective for low-dimensional data. However, the approximation of high-dimensional distributions is often expected to alleviate the computational cost of sampling a complicated mathematical model. For example, a typical Uncertainty Quantification (UQ) model is a partial differential equation (PDE) subject to uncertainty. When studying rare events in such a system, we must have an effective strategy to reduce the number of samples since each sample corresponds to solving a PDE. One strategy is to use the reduced-order model to obtain the samples of the desired rare events followed by a density estimation step. The estimated distribution can then be coupled with the importance sampling technique for variance reduction [12, 26, 31]. Another important example is the variational Bayes [2]. Sampling strategies such as Markov Chain Monte Carlo (MCMC) become less effective as the number of dimensions increases. The variational Bayes approach, which seeks the optimal approximation of the distribution in a family of density models, may be more effective for high-dimensional problems than sampling strategies.

The available deep generative models usually focus on capturing the main features of the data instead of an accurate estimation of the density for that the dimensionality of the target data is often extremely high, e.g., high-resolution images that have mil-