Laboratório
Nacional de
Computação
Científica

COMOD
Coordenação de Modelagem Computacional

Relatório de Atividades

Servidor: Marcio Borges

Petrópolis-RJ
September 9, 2023

1 INTRODUTION

Variational auto-encoder (VAE) model is a stochastic inference and learning algorithm based on variational Bayes (VB) inference proposed by Kingma and Welling (2014). This is a generative technique whose central idea is the development of representations in a low-dimensional latent space that can be mapped back into a realistic-looking image.

Higgins et al. (2016) introduced the β-VAE, a modification of the original VAE, that introduces an adjustable hyperparameter β to balance latent channel capacity and independence constraints with reconstruction accuracy. They demonstrate that with tuned values of β ($\beta > 1$) the β-VAE outperforms VAE ($\beta = 1$).

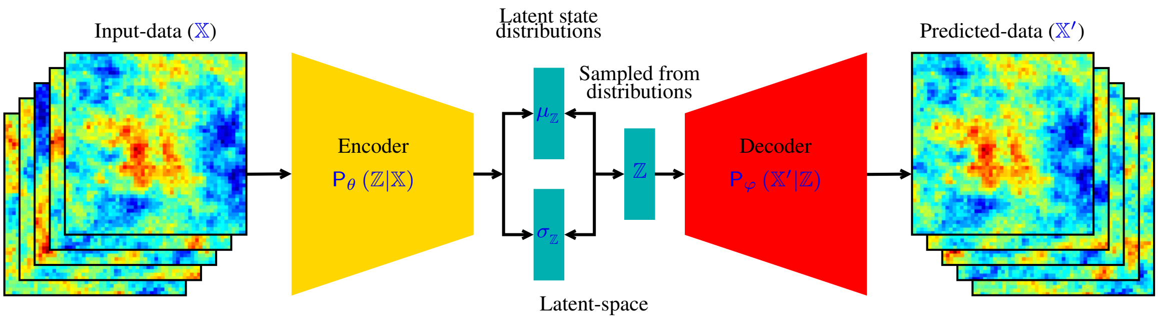
Makhzani et al. (2016); Louizos et al. (2017); Burda et al. (2016); Zheng et al. (2019); Vahdat and Kautz (2020)

2 ZHANG ET AL. (2022)

Variational auto-encoder (VAE) model is a generative network based on variational Bayes (VB) inference proposed by Kingma and Welling (2014).

Zhang et al. (2022) proposed a method to reconstruct porous media based on VAE and Fisher information with good quality and efficiency.

Consider the input data $\mathbb{X} = \left\{\boldsymbol{x}^{(i)}\right\}_{i=1}^{\mathsf{N}}$ consisting of N *i*ndependent and *i*dentically *d*istributed (*i.i.d.*) samples of the continuous (or discrete) variable $\boldsymbol{x}$.



$$\mathbb{Z} = \mu_{\mathbb{Z}} + \sigma_{\mathbb{Z}} \cdot \varepsilon, \quad \text{where} \quad \varepsilon \sim \mathbb{N}(0,1) \tag{1}$$

The Kullback–Leibler divergence (also called relative entropy and I-divergence) is a measure of divergence between two distributions (Kullback and Leibler, 1951; Csiszar, 1975):

$$\mathcal{D}_{\mathsf{KL}}\left(f_{\mathbb{P}} || f_{\mathbb{Q}}\right) = \int_{-\infty}^{\infty} \tag{2}$$

Zheng et al. (2019) proposed a Fisher autoencoder

3 Fisher Information

Let $f(x;\theta)$ be the probability density function of the random variable $\mathbb{X}$ conditioned on the parameter θ. The Fisher information measures the amount of information that an observation of $\mathbb{X}$ carries about the unknown parameter θ. The partial derivative of the natural logarithm of the likelihood function is called **score** (S):

$$\mathsf{S}(x,\theta) = \frac{\partial}{\partial\theta}\log\left[f(x;\theta)\right]. \tag{3}$$

Fisher information is defined as the variance of the score S:

$$\begin{aligned} \mathcal{I}(\theta) = \mathsf{E}[\mathsf{S}^2|\theta] &= \mathsf{E}\left[\left(\frac{\partial}{\partial\theta}\log\left[f(x;\theta)\right]\right)^2\middle|\theta\right] \\ &= \int_{\mathbb{R}}\left(\frac{\partial}{\partial\theta}\log\left[f(x;\theta)\right]\right)^2 f(x;\theta)\,\mathsf{d}x \end{aligned}. \tag{4}$$

If $\log\left[f(x;\theta)\right]$ is twice differentiable with respect to θ and under certain regularuty conditions, Fisher information can be written as

$$\mathcal{I}(\theta) = \mathsf{E}\left[-\frac{\partial^2}{\partial\theta^2}\log\left[f(x;\theta)\right]\middle|\theta\right]. \tag{5}$$

Let $\mathbb{X}$ be a scalar Gaussian random variable *i.e.* $\mathbb{X} \sim \mathbb{N}(\mu,\sigma^2)$. Then the probability density function is parameterized by the parameters μ and σ:

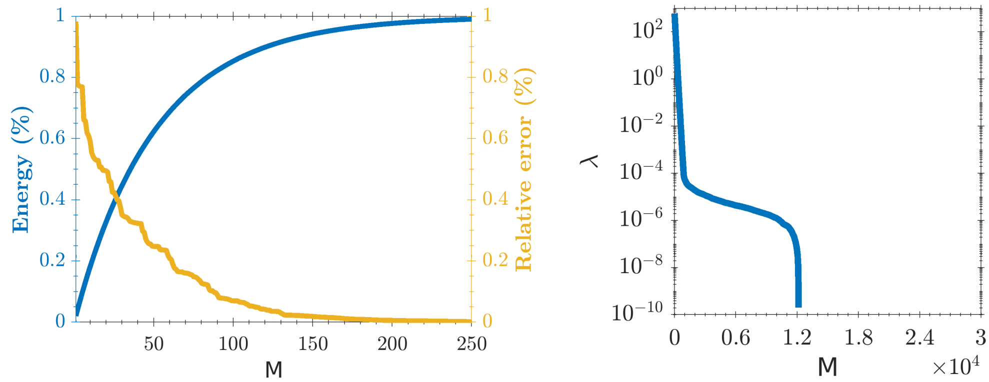
$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]. \tag{6}$$

Substituting Eq. (6) in Eq. (5) where $\theta = \mu$ or σ we can compute de Fisher information for a Gaussian variable as:

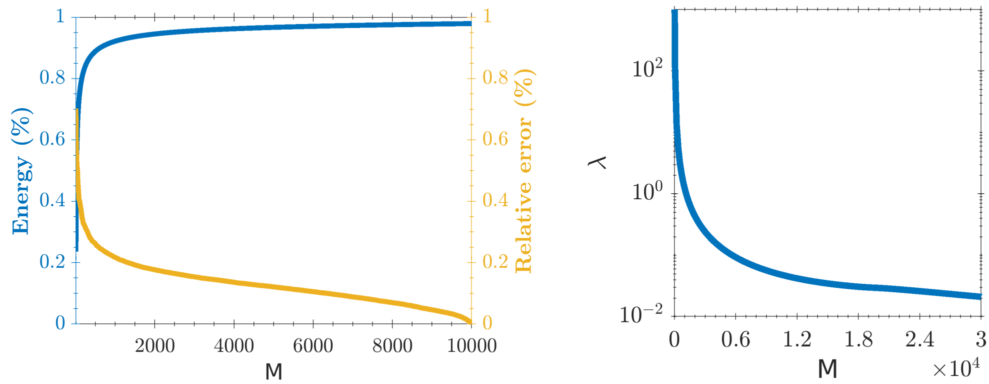
$$
\begin{aligned}
\mathcal{I}(\mu) &= \mathsf{E}\left[-\frac{\partial^2}{\partial\mu^2}\left[\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right)-\frac{1}{2}\left(\frac{\mathbb{X}-\mu}{\sigma}\right)^2\right]\middle|\mu\right] \\
&= \mathsf{E}\left[-\frac{\partial}{\partial\mu}\left(-\frac{\mathbb{X}-\mu}{\sigma^2}\right)\middle|\mu\right] \\
&= \mathsf{E}\left[\frac{1}{\sigma^2}\right]=\frac{1}{\sigma^2}
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
\mathcal{I}(\sigma) &= \mathsf{E}\left[-\frac{\partial}{\partial\sigma}\left[-\frac{1}{\sigma}+\frac{(\mathbb{X}-\mu)^2}{\sigma^3}\right]\middle|\sigma\right] \\
&= \mathsf{E}\left[-\frac{1}{\sigma^2}+\frac{3(\mathbb{X}-\mu)^2}{\sigma^4}\middle|\sigma\right] \\
&= -\frac{1}{\sigma^2}+\frac{3\sigma^2}{\sigma^4}=\frac{2}{\sigma^2}
\end{aligned} \tag{8}
$$

4 KLE



(a) Squared exponential covariance



(b) Exponential covariance

Table 1: Number of KL expansion terms needed to obtain an given energy level

Energy	M	
(%)	Exponential	Squared exp.
80	201	83
90	781	120
94	1928	148
96	3416	170
98	6111	208
100	∼30,000	∼

References

Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. Importance weighted autoencoders, 2016.

I. Csiszar. I-Divergence Geometry of Probability Distributions and Minimization Problems. *The Annals of Probability*, 3(1):146 – 158, 1975. doi: 10.1214/aop/1176996454. URL https://doi.org/10.1214/aop/1176996454.

Irina Higgins, Loïc Matthey, Arka Pal, Christopher P. Burgess, Xavier Glorot, Matthew M. Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. In *International Conference on Learning Representations*, 2016.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. In *International conference on learning representations*, pages 14–27, 2014.

S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79 – 86, 1951. doi: 10.1214/aoms/1177729694. URL https://doi.org/10.1214/aoms/1177729694.

Christos Louizos, Kevin Swersky, Yujia Li, Max Welling, and Richard Zemel. The variational fair autoencoder, 2017.

Alireza Makhzani, Jonathon Shlens, Navdeep Jaitly, and Ian Goodfellow. Adversarial autoencoders. In *International Conference on Learning Representations*, 2016. URL http://arxiv.org/abs/1511.05644.

Arash Vahdat and Jan Kautz. Nvae: a deep hierarchical variational autoencoder. In *NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, number 1650, page 19667–19679, Red Hook, NY, USA, 2020. Curran Associates Inc. ISBN 9781713829546.

Ting Zhang, Tu Hongyan, Xia Pengfei, and Du Yi. Reconstruction of porous media using an information variational auto-encoder. *Transport in Porous Media*, 143(2):271–295, 2022. doi: 10.1007/s11242-022-01769-5. URL https://doi.org/10.1007/s11242-022-01769-5.

Huangjie Zheng, Jiangchao Yao, Ya Zhang, Ivor Tsang, and Jia Wang. Understanding vaes in fisher-shannon plane. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33:5917–5924, 07 2019. doi: 10.1609/aaai.v33i01.33015917.

Index