

VARIATIONAL SPARSE CODING

简介

本文主要介绍了一种结合稀疏编码和变分自编码器（VAE）的方法，通过在VAE的潜在空间中显式地建模稀疏性，从而得到可解释性更好的潜在表示。作者提出了一种Spike and Slab先验分布，利用离散混合分布的识别函数推导出了证据下界，使近似后验推断的计算效率与标准VAE相同。作者在MNIST和Fashion-MNIST数据集上进行了实验，结果表明该方法能够得到更稀疏的表示，并且在分类任务中表现更好。

INTRODUCTION

1. Variational auto-encoders (VAEs)是在处理难以处理的生成模型时进行近似推断的一种可行方法。
2. 然而，标准VAE通常会产生分散并且缺乏可解释性的潜在编码，因此使得结果表示不适合辅助任务（例如分类）和人类解释。
3. 线性稀疏编码在解决这个问题方面提供了一个优雅的解决方案；表示空间被诱导成为稀疏的，从而鼓励编码函数使用最少数量的非零元素来描述每个观察结果并将信息凝聚在少量活动变量中，每个样本的变量不同。
4. 我们的目标是将线性稀疏编码的上述能力扩展到非线性概率生成模型中，从而在一般情况下允许有效、信息丰富和可解释的表示。
5. 为此，我们提出了一种基于Spike and Slab先验分布在VAE的潜在空间中显式建模稀疏性的新方法。
6. 我们使用离散混合识别函数推导了证据下界，从而使近似后验推断的计算效率与标准VAE情况下相当。
7. 通过新方法，我们能够推断出真正稀疏的表示，即使在一般情况下使用非线性概率模型也是如此。
8. 我们展示了这些稀疏表示在两个基准分类任务（MNIST和Fashion-MNIST）上的优越性，通过展示改进的分类准确性和对潜在维数数量的鲁棒性显著增加来证明。
9. 此外，我们定性地展示了稀疏元素捕捉到主观可理解的变化源的特性。

BACKGROUND AND RELATED WORK

1. 稀疏编码（Sparse coding）的目标是用少量的基向量加权线性组合来近似表示输入向量。这个问题可以被形式化为最小化一个目标函数，其中包括数据矩阵 X 、基向量矩阵 B 、稀疏编码矩阵 Z 、一个正实数 λ 和一个稀疏惩罚函数 $\phi(z)$ 。
2. 稀疏编码可以被概率化为一个生成模型，其中观察向量 x 是通过线性过程 Bz 和一个高斯噪声 ϵ 生成的。这个模型可以通过概率分布来描述。
3. 变分自编码器（VAE）通过引入一个识别模型 $q(z|x)$ 来近似真实的后验分布 $p(z|x)$ ，以及一个证据下限（ELBO）来代替真实的后验分布。ELBO由两部分组成：一个先验项，鼓励编码分布与先验分布之间KL散度的最小化，以及一个重构项，最大化在识别函数下数据似然的期望。这个ELBO通过梯度下降来优化VAE的参数。
4. 离散潜变量和稀疏VAE：离散潜变量和稀疏编码密切相关，因为精确的稀疏概率分布涉及从一些离散变量中采样。之前的一些工作已经在VAE中使用Stick-Breaking Process和Indian Buffet Process先验来实现离散潜变量。另外，一些模型通过在连续变量的条件下引入离散变量来捕获自然观测数据中的离散特征。本文提出了一种新的VAE变体，通过在潜空间中采用Spike and Slab分布来明确地建模稀疏性。
5. Epitomic VAE通过学习确定性的选择变量，来指示识别函数在潜空间中应该利用哪些潜变量。与Epitomic VAE不同的是，本文的目标是直接在连续潜空间中诱导稀疏性，并找到一个合适的ELBO来进行近似的变分推断。

VARIATIONAL SPARSE CODING

1. 本部分介绍了如何将稀疏编码的思想与变分自编码器（VAE）相结合，提出了一种新的模型——变分稀疏编码（VSC），旨在实现对非线性概率生成模型中潜变量的稀疏建模，从而获得更加有效和可解释的表示。
2. VSC采用Spike and Slab分布作为潜变量的稀疏先验分布，通过神经网络的离散混合识别函数将观测数据映射为稀疏潜变量向量。
3. VSC利用推导出的证据下界（ELBO）进行训练，通过优化梯度方法来恢复编码和解码函数，从而实现与标准VAE相当的计算效率。
4. 本文还通过在MNIST和Fashion-MNIST等基准数据集上进行分类任务的实验，展示了VSC相对于标准VAE的优越性，并证明了VSC的稀疏编码对主成分的控制能力更强。

5. 本文最后通过对CelebA人脸数据集的定性分析，进一步证实了VSC的潜变量具有更好的可解释性和可视化效果。

3.1讲述了VSC模型中ELBO先验项的推导过程。该项由两部分组成，第一部分是负的KL散度，衡量Slab变量分布与先验分布的差异，乘以Slab变量非零概率 $y_{i,j}$ ，这一部分与标准VAE的先验项类似，鼓励识别函数的高斯分量与先验匹配；第二部分是负的KL散度，衡量Spike变量分布与先验分布的差异，这一部分鼓励概率 $y_{i,j}$ 与先验Spike概率 α 匹配。

3.2讲述了VSC模型中ELBO重构项的推导过程。该项中样本 $z_{i,l}$ 从识别函数 $q_{\phi}(z|x_i)$ 中抽取。为了使其对编码参数 ϕ 可微分，我们采用了重新参数化技巧。为了参数化从 $q_{\phi}(z|x_i)$ 中抽取的二进制分量，我们使用了二进制变量的连续松弛方法。我们使用两个辅助噪声变量和 η ，分别服从正态分布和均匀分布。其中 η 用于从Slab分布中抽取样本，从而实现与标准VAE类似的重新参数化。 $y_{i,l}$ 用于通过非线性二元选择函数 $T(y_{i,l})$ 参数化Spike变量的抽样。最终，两个变量相乘得到从 $q_{\phi}(z|x_i)$ 中抽取的参数化样本。

3.3讲述了VSC模型的ELBO公式。该公式由3.1和3.2中的先验项和重构项组成。先验项鼓励编码分布与先验匹配，重构项最大化似然函数期望。为了训练VSC模型，我们通过梯度上升最大化该公式，得到编码和解码参数 ϕ 和 θ 。

需要注意的点：

- 在ELBO先验项的推导中，需要分别考虑高斯-离散混合分布和Dirac Delta-离散混合分布这两种情况。
- 在ELBO重构项的推导中，需要使用重新参数化技巧对离散分量进行参数化，以实现对编码参数 ϕ 的微分。
- 在训练VSC模型时，需要通过梯度上升最大化ELBO公式，同时注意避免过拟合。

Experiment

论文中的实验设计主要分为三个部分。

首先，作者通过比较不同的先验稀疏度和潜在空间维度，评估了新模型的ELBO性能。作者首先训练一个标准的VAE，然后在不同的潜在空间维度下测试其性能。然后，作者使用相同的设置测试不同先验稀疏度下的VSC的ELBO，以评估其性能。作者在MNIST和Fashion-MNIST数据集上进行了测试，结果显示，VSC在不同的先验稀疏度和潜在空间维度下都能产生稀疏、信息丰富和可解释的表示。

其次，作者使用潜在变量作为输入特征，通过标准分类实验比较VAE和VSC的表示效率。作者将先验Spike概率设置为0.01，以鼓励VSC生成信息丰富的代码。作者使用MNIST和Fashion-MNIST数据集进行测试，结果显示VSC表示比VAE更有效，并且在潜在空间维度的大变化下具有更强的鲁棒性。

最后，作者通过检查更改稀疏代码中的单个非零元素的效果，定性评估了潜在空间的解释能力。作者使用CelebA面部数据集进行了测试，并展示了许多稀疏元素控制生成的观察结果中的主观可识别特征的效果。

CONCLUSION AND FUTURE WORK

1. 介绍了本文提出的在VAE中引入稀疏性的方法，该方法可以在复杂的概率稀疏编码模型中进行近似变分推断。
2. 推导了一个清晰易懂且有效的ELBO下界，该下界与标准VAE的ELBO下界形式相同。
3. 通过实验表明，该方法可以在常规分类基准测试中获得最佳的学习输入，并且在许多非零分量中具有很好的解释性。
4. 认为在生成模型的潜在空间中引入稀疏性是获得有用代码、可解释的表示和可控数据合成的有前途的途径，这是VAE和表示学习中所有问题的未来挑战。
5. 未来工作将进一步研究稀疏潜在空间的解释性和特征分离能力，并期望该方法能够通过稀疏地填充大的潜在空间来建模各种各样的数据集，从而隔离控制相似对象的变异特征。

为什么标准VAE的潜在表示不够稀疏，有什么问题？

标准VAE的潜在表示不够稀疏，主要问题在于其潜在变量的编码往往是分散和缺乏可解释性的，这导致其生成的表示不适合辅助任务（例如分类）和人类解释。这使得VAE的表示缺乏有效性和可解释性，限制了其在实际应用中的应用。

作者提出的Spike and Slab先验分布是如何建模稀疏性的？

作者提出的Spike and Slab先验分布是通过将概率密度函数离散化来建模稀疏性的。该分布由两个变量组成：一个二元的spike变量和一个连续的slab变量。spike变量的取值为1或0，对应的概率分别为 α 和 $(1-\alpha)$ ，slab变量有两种分布，一种是以0为中心的delta函数，

另一种是以0为均值的高斯函数，根据spike变量的取值不同而定。该先验分布的特点是可以将未使用的维度置为零，从而在概率上建模稀疏性。

在MNIST和Fashion-MNIST数据集上的结果表明该方法的优势，这种方法是否可以推广到其他数据集上？

从该论文的实验结果来看，该方法在MNIST和Fashion-MNIST数据集上表现出色，能够得到具有稀疏性、信息量和可解释性的表示。虽然该论文没有在其他数据集上进行实验，但是基于该方法的理论基础和实验结果，我们可以推测该方法是否可以推广到其他数据集上。但是，不同的数据集具有不同的特点和难点，因此需要在具体的数据集上进行实验和针对性的调整。因此，我们需要在不同的数据集上进行验证和改进，以便更好地应用该方法。