



南 华 大 学
UNIVERSITY OF SOUTH CHINA

毕业论文/设计外文翻译

题 目	对深度学习的几何理解
学院名称	数理学院
指导老师	高有
职 称	讲师
班 级	信计 1802
学 号	20184390213
学生姓名	朱柳承

2022 年 1 月 15 日

对深度学习的几何理解

摘要：这篇论文详细地给出了在深度学习中如何运用最优传输理论的相关知识的方法与技巧。这项工作引入了生成式对抗网络 (GANs) 的最优运输 (OT) 视图。自然数据集具有内在的模式，可以概括为流形分布原理：一类数据的分布接近于一个低维流形。GANs 主要完成两个任务：流形学习和概率分布转换。后者可以用经典的 OT 方法来进行。从 OT 的角度来看，生成器计算 OT 图，而鉴别器计算生成的数据分布和真实数据分布之间的瓦瑟斯坦距离；两者都可以简化为凸几何优化过程。此外，OT 理论还发现了发生器和鉴别器之间存在内在的协作关系，而不是竞争关系，以及模式崩溃的根本原因。我们还提出了一种新的生成模型，该模型使用自编码器 (AE) 进行流形学习，并使用 OT 映射进行概率分布变换。该 AE-OT 模型提高了理论的严谨性和透明度，以及计算的稳定性和效率；特别是，它消除了模式崩溃。实验结果验证了我们的假设，并证明了我们所提出的模型的优点。

关键词：生成性；对抗性；深度学习；最优传输；模式崩溃

1 介绍

生成对抗网络 (GANs) 已成为无条件图像生成的主要方法之一。当在多个数据集上进行训练时, GANs 能够产生真实和视觉上吸引人的样本。GAN 方法训练了一个无条件的生成器, 从随机噪声中回归真实图像, 以及一个鉴别器, 测量生成的样本和真实图像之间的差异。GANs 已经得到了各种改进。生成对抗网络 (GANs) 已成为无条件图像生成的主要方法之一。当在多个数据集上进行训练时, GANs 能够产生真实和视觉上吸引人的样本。GAN 方法训练了一个无条件的生成器, 从随机噪声中回归真实图像, 以及一个鉴别器, 测量生成的样本和真实图像之间的差异。GANs 已经得到了各种改进。

1.1 流形分布假说

差距的巨大成功可以解释为差距有效地发现真实数据集的内在结构, 可以制定为流形分布假设: 一个特定类的自然数据集中在低维流形嵌入在高维背景空间 [2]。

图 1 显示了 MNIST 数据库的流形结构。每个手写数字图像具有尺寸 28×28 , 并被视为图像空间中的一个点。MNIST 数据库集中在一个低维流形附近。通过使用 t-SNE 流形嵌入算法 [3], MNIST 数据库被映射到一个平面域上, 每个图像被映射到一个点上。代表相同数字的图像被映射到一个簇上, 10 个簇被颜色编码。这表明 MNIST 数据库分布在嵌入单元立方体中的二维 (2D) 曲面附近。

1.2 GANs 的理论模型

图 2 为 GANs 的理论模型。真实的数据分布 ν 集中在嵌入在环境空间 χ 中的流形 Σ 上。(Σ, ν) 一起显示了真实数据集的内在结构。一个 GAN 模型计算一个从潜在空间 Z 到流形 Σ 的生成器映射 g_θ , 其中 θ 表示一个深度神经网络 (DNN) 的参数。 ζ 是潜在空间中的高斯分布, g_θ 将 ζ 向前推到 μ_θ 。鉴别器计算真实数据分布 ν 和生成的分布 μ_θ 之间的距离, 如瓦瑟斯坦距离 $W_c(\mu_\theta, \nu)$, 它相当于康

塔罗维奇的潜在 $\phi\xi$

尽管 GANs 有优势，但它们也有严重的缺点。在理论上，对深度学习的基本原则的理解仍然是原始的。在实践中，GANs 的训练是棘手的，对超参数是敏感的；GANs 遭受模式崩溃。最近，Meschede 等人。[4] 研究了 9 种不同的 GAN 模型和变体，表明基于梯度下降的 GAN 优化并不总是局部收敛的。

尽管 GANs 有优势，但它们也有严重的缺点。在理论上，对深度学习的基本原则的理解仍然是原始的。在实践中，GANs 的训练是棘手的，对超参数是敏感的；GANs 遭受模式崩溃。最近，Meschede 等人。[4] 研究了 9 种不同的 GAN 模型和变体，表明基于梯度下降的 GAN 优化并不总是局部收敛的。

图 3 为生成器图 $g_\theta = h \circ T$ 的分解，其中 h 为从潜在空间到 \mathbf{h} : $Z \rightarrow \Sigma$ 环境空间中的数据流形 Σ ，概率分布变换图 $T: Z \rightarrow Z$ 。解码图 h 用于流形学习，图 T 用于测量传输。

1.3 最优传输的观点

OT 理论 [5] 研究了以最经济的方式将一种概率分布转化为另一种概率分布的问题。OT 提供了严格而强大的方法来计算最优映射，将一个概率分布转换为另一个分布，并确定它们之间的距离为 [6]。

如前所述，GANs 完成了两个主要任务：流形学习和概率分布变换。后一种任务可以直接通过 OT 方法完全完成。具体地说，在图 3 中，可以用 OT 理论计算出概率分布变换图 T 。该鉴别器计算生成的数据分布与真实数据分布之间的瓦瑟斯坦距离 $W_c(\mu_\theta, \nu)$ ，可以用 OT 方法直接计算。

从理论的角度来看，天然气的解释使黑盒子透明的一部分，概率分布变换减少凸优化过程使用 OT 理论，解决方案的存在性和唯一性理论保证，收敛速度和近似精度充分分析。

OT 的解释也解释了模态崩溃的根本原因。根据蒙格-安培方程的正则性理论，运输映射在某些奇异集合上是不连续的。然而，DNN 只能建模连续的函数/映射。因此，目标交通映射是在 GANs 可表示的功能空间之外的。这种内在的冲突

使模式的崩溃不可避免。

OT 解释也揭示了发生器和鉴别器之间更为复杂的关系。在当前的 GAN 模型中，生成器和鉴别器相互竞争，而不共享中间的计算结果。OT 理论表明，在 12 代价函数下，生成器的最优解和鉴别器的最优解可以用封闭的形式相互表示。因此，应该用协作来代替发生器和鉴别器之间的竞争，并应该共享中间的计算结果，以提高效率。

1.4 自动编码器-最优运输模型

为了减少 GANs 的训练难度，特别是为了避免模式崩溃，我们提出了一个基于 OT 理论的更简单的生成模型：一个自动编码器 (AE)-OT 模型，如图 4 所示。

如前所述，生成模型的两个主要任务是流形学习和概率分布变换。AE 计算编码映射 $f_0: Z \rightarrow \Sigma$ ，和解码映射 $g_\xi: \Sigma \rightarrow Z$ ，用于流形学习。OT 图， $T: Z \rightarrow Z$ ，将白噪声 ζ 转换为由编码图推送转发的数据分布， $(f_0)_\# \nu$ 。

AE-OT 模型有很多优点。从理论的角度来看，OT 理论已经得到了很好的建立和充分的理解。通过对解码图和 OT 图进行解耦，可以提高生成模型的理论严谨性，使黑盒的一部分透明。在实际应用中，将 OT 映射简化为一个凸优化问题，保证了了解的存在性和唯一性，使训练过程不会陷入局部最优状态。与 OT 映射相关的凸能量具有显式的黑森矩阵；因此，优化可以采用二阶收敛的牛顿方法，或采用超线性收敛的拟牛顿方法进行优化。相比之下，目前的生成模型基于线性收敛的梯度下降法；未知数等于训练样本，以避免过参数化问题；蒙特卡罗方法中的采样密度可以完全控制 OT 映射的误差界；具有自适应的层次算法进一步提高了效率；并行 OT 映射算法可以使用图形处理单元 (GPU) 实现。最重要的是，AE-OT 模型可以消除模式崩溃。

1.5 贡献

本工作利用 OT 理论解释了 GAN 模型。GANs 可以完成两个主要任务：流形学习和概率分布变换。后一种任务可以使用 OT 方法来完成。生成器计算一个 OT 图，而鉴别器计算生成的数据分布和真实数据分布之间的瓦瑟斯坦距离。利用布雷尼尔的理论，发电机和鉴别器之间的竞争可以被协作所取代；根据蒙格-安培方程的正则性理论，运输图的不连续导致了模态崩溃。我们进一步提出了利用 AE-OT 模型来解耦流形学习和概率分布变换，使部分黑箱透明，提高了训练效率，防止了模式崩溃。实验结果证明了该方法的有效性和有效性。

本文的组织结构如下：第 2 节简要回顾了 OT 和 GANs 中最相关的工作；第 3 节简要介绍了 OT 的基本理论和蒙格-安培方程的正则性理论；第 4 节介绍了计算 OT 的变分框架；第 5 节从 OT 的角度分析了发生器与鉴别器之间的协作（而不是竞争）关系，揭示了模式崩溃的内在原因；第 6 节报告了实验结果；论文在第 7 节中总结。