深度整流幂次单元神经网络的逼近性质研究

唐珊珊

摘　要

随着电子设备的日渐普及，人类社会产生的数据量呈爆炸性增长，这使得基于大数据的机器学习在近十年内得到了快速发展。作为当今机器学习中最成功的数学模型，深度神经网络通常使用整流线性单元(ReLU)作为激活函数。深度 ReLU 网络不但易于训练而且被证明具有比浅层网络更好的表达能力，在很多应用中取得了巨大的成功。但并没有理论证明 ReLU 是深度网络的最优激活函数。在本文中，我们构造了一种以整流幂次单元(RePU)作为激活函数的深度神经网络，并证明在处理光滑函数时，其比深度 ReLU 网络有更好的逼近性质。我们的分析基于两点: (1)经典的多项式逼近理论和高维空间稀疏网格谱方法逼近理论；(2)用本文中提出的最优算法将多项式逼近精确转换成相应大小的深度 RePU 神经网络。同最近的文章[D.Yarotsky, Neural Network,94:103-114,2017] [H.Montanelli and Q.Du, SIAM J Math. Data Sci., 1:78-92,2019] 给出的 ReLU 逼近结果相比，当以误差 ε 逼近 Sobolev 空间和 Korobov 空间中的函数时，根据我们构造性的证明，需要的 RePU 网络的大小一般比相应的 ReLU 网络小 $O(\log_2(1/\varepsilon))$ 倍。我们还设计了将切比雪夫多项式逼近直接转换成深度 RePU 神经网络的算法，构造出的网络比一般的多项式网络表现出更好的稳定性。我们基于 RePU 的构造性方法清晰的揭示了深度网络的“深度”和谱方法的“阶数”之间的密切关系。RePU 网络还有一些其它好的性质,例如：高阶可微，激活函数求值时需要的计算更少。鉴于以上特点，我们提倡在处理高维光滑问题时，或者在处理损失函数中包含逼近函数导数的问题时，应用深度 RePU 网络。我们进行了一些相关的数值实验，验证了所提出算法的正确性和构造出的深度 RePU 网络的逼近性质。

关键词：深度神经网络，高维逼近，稀疏网格，整流线性单元(ReLU)，整流幂次单元(RePU)，整流二次单元(ReQU)，谱方法