Accepted Manuscript

Limitations of shallow nets approximation

Shao-Bo Lin

PII: S0893-6080(17)30152-1
DOI: http://dx.doi.org/10.1016/j.neunet.2017.06.016
Reference: NN 3781

To appear in: *Neural Networks*

Received date : 18 August 2016
Revised date : 29 March 2017
Accepted date : 30 June 2017

Please cite this article as: Lin, S., Limitations of shallow nets approximation. *Neural Networks* (2017), http://dx.doi.org/10.1016/j.neunet.2017.06.016

Limitations of shallow nets approximation ☆

Shao-Bo Lin*

College of Mathematics and Information Science, Wenzhou University, Wenzhou, 325035, P R China

Abstract

In this paper, we aim at analyzing the approximation abilities of shallow networks in reproducing kernel Hilbert spaces (RKHSs). We prove that there is a probabilistic measure such that the achievable lower bound for approximating by shallow nets can be realized for all functions in balls of reproducing kernel Hilbert space with high probability, which is different with the classical minimax approximation error estimates. This result together with the existing approximation results for deep nets shows the limitations for shallow nets and provides a theoretical explanation on why deep nets perform better than shallow nets.

Keywords: Shallow nets, deep nets, approximation, reproducing kernel Hilbert space

1. Introduction

Recent years have witnessed a tremendous growth of interest in deep nets, a.k.a., neural networks with more than one hidden layers. Applications include the image classification [21], speech recognition [23], manifold learning [4] and so on. All these applications show the excellent power of deep nets over shallow nets, i.e, neural networks with one hidden layer. We refer the readers to [18, 5, 24, 42, 12] and references therein for more applications and details of deep nets.

The comparison of performances between deep nets and shallow nets is a classical topic in approximation theory. Regardless of the computational burden, there are roughly two advantages of deep nets approximation. The first one, called as the expressivity [39], shows that there are various functions expressible by deep nets but cannot be approximated by any shallow nets with similar number of neurons. A typical example is that deep nets can provide localized approximation but shallow nets fail [10]. The other one, proposed in [11], is that deep nets can

☆The research was supported by the National Natural Science Foundation of China (Grant Nos. 61502342, 11401462).

*Corresponding author: sblin1983@gmail.com