Neural Network Approximation Theory

Introduction

This document delves into the theoretical capabilities of neural networks in approximating mathematical functions. The core idea is to explore whether a neural network can not only replicate a given function but also potentially enhance its performance.

Definitions

- Function f: A predefined mathematical function that we aim to approximate.
- 2. Function g: A hypothetical function that outperforms f based on a specific performance metric.
- 3. **Neural Network** N: A feedforward neural network with parameters θ that aims to approximate f.
- 4. **Performance Metric** L: A metric (e.g., Mean Squared Error) that measures the difference between the outputs of a function and the true values.

Assumptions

1. N perfectly replicates f, i.e., for all inputs x,

$$N(x;\theta) = f(x)$$

2. There exists a g such that for all x,

Objective

To demonstrate that there exists a set of parameters θ' for which N not only approximates f but also has the potential to approximate g, thereby outperforming f.

Proof

- $1. \ \, \textbf{Neural Network's Approximation Ability:} \\$
 - By the Universal Approximation Theorem, for any continuous function g and any $\epsilon > 0$, there exists a neural network N' and parameters θ' such that:

$$|N'(x;\theta') - g(x)| < \epsilon$$

for all x.

2. Constructing a Composite Network:

• Define a new neural network N'' as a combination of N and N'. Specifically, for any input x, the output of N'' is:

$$N''(x; \theta, \theta') = N(x; \theta) + N'(x; \theta')$$

3. Performance Analysis:

• Given the assumption that N perfectly replicates f, for all x:

$$N''(x; \theta, \theta') = f(x) + N'(x; \theta')$$

• Using the approximation ability of N', we deduce:

$$|f(x) + N'(x; \theta') - g(x)| < \epsilon$$

This implies that N'' can potentially approximate g and thereby outperform f.

4. Conclusion:

• Through the composite network N'', constructed from N and N', we've demonstrated the potential of neural networks to not only replicate but also enhance the performance of a given function f.