

Introduction to State Space Models (SSM)

Vinay, Subhadeep Sing, Vaibhav Mahore, and Snehal Biswas

Indian Institute of Science, Bangalore

{vinay2023 ,shubadeeps ,mvaibhav ,snehalbiswas}@iisc.ac.in

October 26, 2020

Abstract

This is a short abstract summarizing the paper.

1 Introduction

Sequential data modeling has emerged as a critical paradigm in modern machine learning and data analysis, addressing the fundamental challenge of processing and understanding information that unfolds over time or in a specific order. The inherent temporal dependencies and contextual relationships within sequential data present unique challenges that traditional static modeling approaches fail to adequately address. This limitation has driven significant research interest in developing specialized architectures and methodologies capable of capturing these temporal dynamics effectively. From natural language processing to financial forecasting, the applications of sequential data modeling span diverse domains, each presenting its own set of challenges and requirements. The evolution of sequential modeling techniques has witnessed several breakthrough developments, starting from simple Markov models to sophisticated neural architectures. Traditional feed-forward neural networks, while powerful for many tasks, proved inadequate for sequential data due to their inability to maintain state information across time steps. This limitation led to the development of recurrent neural networks (RNNs), which introduced the concept of memory through recurrent connections. However, RNNs themselves faced challenges with long-term dependencies, leading to the introduction of more sophisticated architectures like

Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs). These innovations marked significant progress in the field, enabling better handling of long-range dependencies and more stable training dynamics. Despite these advances, the quest for more efficient and effective sequential modeling approaches continues, driven by the ever-increasing complexity and scale of modern applications. The emergence of attention mechanisms and transformer architectures has further revolutionized the field, offering new perspectives on how to process sequential data more effectively.

2 Related Work

3 Method

4 Experiments

5 Conclusion