

双曲型偏微分方程式を利用した状態空間レイヤーについて

谷口隆晴・神戸大学大学院理学研究科数学専攻

研究目的：大規模言語モデルなどで利用されているTransformerは，入力系列の長さが長くなると計算量が大きくなりすぎるということが知られている．そこで，近年，状態空間モデルをニューラルネットワークの層として利用する方法が注目されている．本研究では，特に，そのような方法として，偏微分方程式を利用した方法を開発することを目的とする．

研究内容の概要：本研究では，深層学習分野で利用されている，既存の状態空間モデルに基づく層が，既存の状態空間モデルが双曲型偏微分方程式を離散化したものと解釈できることを示した．既存手法ではHiPPO行列と呼ばれる行列が状態空間モデルで利用されている．我々は，この行列が積分作用素と関係していることに着目し，モデル式を変形することで，双曲型偏微分方程式が導かれることを示した．

結果及び結果図：実際に，得られた双曲型偏微分方程式を離散化し直すことで，いくつかの機械学習モデルを構築した．ベンチマーク問題に適用したところ，確かに，これらのモデルが動作することが確認できた．

利用した計算機：SQUID GPUノード

