金融データサイエンス入門Ⅱシリーズ［第５回］

深層学習の数理

鈴　木　大　慈

目　　次

１．はじめに
２．深層ニューラルネットワークの表現能力
３．Transformerの理論
４．おわりに

深層学習の数理に関して、特に特徴学習が学習効率に与える影響を明らかにする理論を紹介する。深層構造を持つ基盤モデルは、中間層で問題に合わせた適切なデータの表現を得ることで、学習の効率を上げることができる。実際に、異方的滑らかさを持つ関数の推定問題において深層構造が学習効率向上に寄与することや、事前学習で適切な表現を獲得した基盤モデルは、テスト時推論においても最適な学習効率を達成することを紹介する。

１．はじめに

本稿では、深層学習の理論について、特にその特徴学習能力の観点からいくつかの重要な内容を解説する。深層学習は、現在のAIの発展に欠かすことのできない学習法であり、大規模言語モデルや拡散モデルといった基盤モデルは、いずれも深層構造を持っており、中間層においてデータの適切な表現（特徴量）を構築することで高度な情報処理を実現している。実際、今現在の基盤モデルは、数多くの知的情報処理を実現し、日常の文章作成からプログラミング、さらには数学に関しては数学オリンピックで金メダルを獲得できるスコアを達成している。拡散モデルは、画像の生成だけでなく、化合物や文章の生成といった様々な応用で高い性能を示している。これらの問題は文章・画像・グラフといった複雑なデータ構造を扱う必要があり、深層構造を持たないモデルでは高い精度を出すことは難しい。一方で、深層学習は、そのような複雑なデータを適切な特徴量に変換することで「解ける」問題へと変換している。この特徴学習が深層学習の本質であり、層を深くすることの恩恵にほかならない。特に、複雑データはどれも高次元データであり、そのまま扱えばいわゆる「次元の呪い」を受けてしまう。これまでの機械学習の歴史は、この次元の呪いにいかに立ち

鈴木　大慈（すずき　たいじ）
東京大学大学院情報理工学系研究科教授。2004年３月東京大学工学部計数工学科卒業。2009年３月東京大学大学院情報理工学系研究科数理情報学専攻博士課程修了。同年４月、同専攻にて助教に着任。2013年東京工業大学大学院情報理工研究科准教授、2017年東京大学大学院情報理工学系研究科准教授を経て、2024年４月より現職。主な著書に、『確率的最適化』（講談社、2015年）がある。

©日本証券アナリスト協会　2025