

代数的操作を行う神経回路網の構成に向けて

古澤 慎太郎[†] 山本 修身[†]
[†]名城大学理工学部情報工学科

1 はじめに

近年注目されているニューラルネットワーク (NN) による学習では、入力と出力の対を与えることで目的の計算を実行する機械を構成する。この場合、NN は完全にブラックボックスである。このような性質からどのようなアルゴリズムを構成して良いかわからないような問題についてもうまく解決することができる反面、論理的に構造のはっきりしている問題でも、その出力結果が保証されないという性質がある。

本稿では NN で扱いづらいと考えられるいくつかのビット演算 (16bit の加法および乗法, 32bit の後者関数) を取り上げ、NN に学習させることにより、その演算を NN がどのように学習するかについて調べた。

2 2つの予備的実験と考察

図 1 に示す順伝播型の BitNN(3) を用いて以下の 3 つのビット演算を学習させる:

- I. $z = (x|y) + 1$
- II. $z = x + y$
- III. $z = x \cdot y$

なお、学習用のデータは 10,000 個の一様な整数乱数であり、損失関数はクロスエントロピー関数である。また、エポック数は 500 であり、ドロップアウト率は 0.4 とした。こうして学習をさせた NN に、同じく一様な 16bit 整数の乱数である 1,000 個の検証用データを計算させ、結果の正答率を比較した。その結果を図 2 に示す。後者関数ではどの桁においても高い精度となったが、加法と乗法では特に中位の桁で低い精度となった。各出力は 0 か 1 の 2 通りであるため、正答率 50% は全く計算できていないことを意味する。II と III では端の出力は依存している入力が少ないため比較的正しい結果が得られたが、中位の桁はそれ以下の桁の繰り上がりの影響を受けやすく依存する入力が多いため、学習が困難になっていると考えられる。

また演算 III について、学習の条件は変えずに隠れ層を $n = 1, 3, 5, 10, 20$ と変化させた場合の BitNN(n) 検証データにおける正答率を比較した。その結果、隠れ

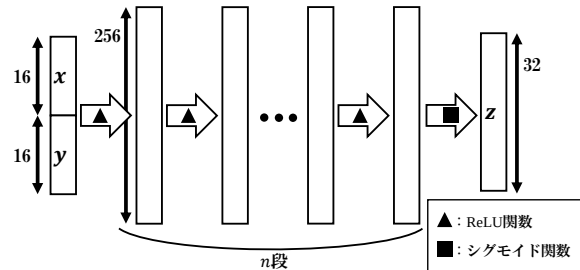


図 1: 実験で使用した隠れ層 n のネットワーク BitNN(n) の構造。

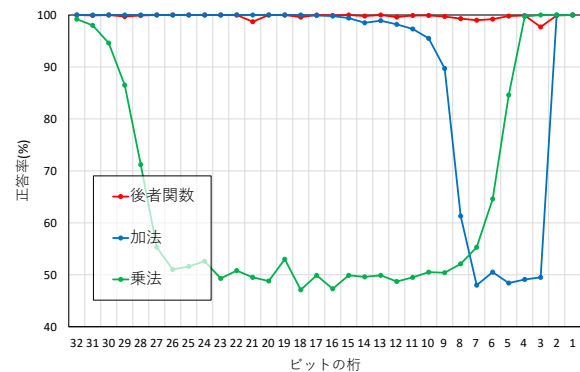


図 2: ビット演算 I~III のそれぞれのビットにおける正答率。

層を 3 程度にした時に最も精度が高くなり、以降は層を増やすにつれ正答率が低下していく傾向が見られた。通常の乗算では下位ビットの出力は上位ビットの入力に影響されないが、本実験で用いた NN では上位の入力と下位の出力の間にも経路が存在しているため、その影響を受ける。これにより本来考慮されるべきでない疑似相関が学習を妨害しており、層が増えることでその機会が増加したことが正答率が低下の原因だと考えられる。

参考文献

[1] 太田満, 須藤広大, 黒澤匠雅, 小田大輔: 現場で使える! TensorFlow 開発入門 Keras による深層学習モデル構築手法. 翔泳社 (2018).