線形の層は何枚重ねても一枚と厳密に同じで、折れ曲がりを一つ挟むと層ごとに区間が倍になる

— 線形の三層とその積の一層は、200 本の入力で出力の相対差が最大 5.192×10^{-15} で、深さは何の働きもしない — 分離子は、層のあいだに ReLU のような非線形を挟むかどうか — [論文745]

山岸 勇輝 (Yuuki YAMAGISHI)
Yamagishi Complexity Torus Lab, Japan｜山岸勇輝ラボ
2026年9月

— From thinking to syncing. —

要旨

層を深くすると表せるものが増える、と言われる。それは層のあいだに折れ曲がりがあるときだけ正しい。深さが無の効きになる場合と指数の効きになる場合を、同じ部品で並べる。新しい定理も法則も主張しない。 第一に、線形の層は何枚重ねても一枚である——三層とその積の一層で、出力の相対差は最大 5.192×10^{-15}（第2節）。 第二に、だから直線で作れないものは、深くしても作れない——$|x|$ を最良の直線で近似すると平均二乗誤差 **0.083342** が残る（第2節）。 第三に、折れ曲がりを二つ使えば、同じものが厳密に作れる——ReLU 二つで $|x|$ の誤差は 0（第2節）。 第四に、これが本稿の芯である。折れ曲がりがあると、深さは指数で効く——テント写像を L 回合成すると区間は 2^L、$L = 10$ で 1024（第3節）。 第五に、浅く作ると素子が桁で要る——同じ 1024 区間を、深さなら ReLU 20 個、一層なら 1023 個——**51.150000** 倍（第3節）。 第六に、分離子は、層のあいだに ReLU のような非線形を挟むかどうかである——深さの値打ちは 0 か指数かに分かれ、その間は無い（第4節）。 キーワード：活性化関数、ReLU、線形写像の合成、区分線形、深さと表現力、非線形の有無。

本稿の射程（射程注記）：新しい定理も法則も主張しない——線形写像の合成が線形であること、区分線形の網の区間の数、深さによる表現力の差は、いずれも既知である（テルガルスキー 2016、グッドフェロー 2016）。学習を扱わない——重みをどう見つけるかは扱わず、表せる関数だけを数える。一次元に限る——入力は一次元で、区間の数で表現力を数える。実在の網を主張しない——個別の網や課題は扱わない。浅い網の素子数の意味を限る——一層で要る素子の数は、一次元の区分線形関数を作る場合の数である。既刊との関係：論文330 は、正規方程式で条件数が二乗になるのは近似ではなく恒等式であることを示した——そこでは行列の積が条件数を悪くするのが恒等式だった。ここでは線形の層を重ねても一層になるのが恒等式であり、行列の積が深さを消す。行列の積が何を増やし、何を消すかになる。加えたのは、線形の三層と一層の一致を数値で確かめたこと、直線では $|x|$ に 1/12 の誤差が残ることを示したこと、テント写像の合成で区間が 2^L になることを $L = 10$ まで数えたこと、浅く作ると 51.150000 倍の素子が要ることを示したこと、分離子を非線形の有無に置いたことである。

1. 記法と問い

設定：層を行列 W_k とする。非線形が無ければ、層を重ねた写像は行列の積になる。ReLU は $\max(0,x)$ で、一つの折れ曲がりを作る。

$$W_3W_2W_1x = (W_3W_2W_1)\,x, \qquad |x| = \mathrm{ReLU}(x) + \mathrm{ReLU}(-x), \qquad t(x) = 2\,\mathrm{ReLU}(x) - 4\,\mathrm{ReLU}\big(x - \tfrac{1}{2}\big)$$

t は $[0,1]$ を $[0,1]$ に写すテント写像で、ReLU 二つで書ける。

問い：層を深くすることは、何を増やしているのか。

2. 線形なら、深さは何もしない

計算 2.1：6×6 の乱数の行列を三つ重ねた写像と、その積の一層に、200 本の入力を通した。

量	値
三層と一層の出力の最大の相対差	5.192×10^{-15}
$\lvert x\rvert$ を最良の直線で近似した平均二乗誤差	0.083342
$1/12$ との差	8.335×10^{-6}
$\lvert x\rvert$ を ReLU 二つで作った誤差	0

> **観察 2.2**
>
> 線形の三層は、一層と厳密に同じである。出力の差は計算機の丸めの大きさ 5.192×10^{-15} しか無い——近似で同じなのではなく、行列の積という恒等式で同じである。

> **観察 2.3**
>
> だから直線で作れないものは、深くしても作れない。$\lvert x\rvert$ を $[-1,1]$ で最良の直線（高さ 0.5）で近似すると、平均二乗誤差 0.083342 が残り、$1/12$ に一致する——何枚重ねても、この誤差は一つも減らない。

> **観察 2.4**
>
> 折れ曲がりを二つ使うと、誤差は 0 になる。$|x| = \mathrm{ReLU}(x) + \mathrm{ReLU}(-x)$ だからである——表せるものを増やしたのは、層の枚数ではなく折れ曲がりである。

3. 折れ曲がりがあれば、深さは指数で効く

計算 3.1：テント写像を L 回合成した関数の、$[0, 1]$ での直線の区間の数を、傾きの向きの変化で数えた。

L	深さで作る ReLU の数	区間の数	一層で作るなら要る ReLU
1	2	2	1
2	4	4	3
3	6	8	7
4	8	16	15
6	12	64	63
8	16	256	255
10	20	1024	1023

> **観察 3.2**
>
> これが本稿の芯である。区間の数はちょうど 2^L になる。一層ぶんの合成ごとに、すでにある区間のそれぞれが二つに折り返される——同じ ReLU 二つが、線形なら何も足さず、合成すれば区間を倍にする。

> **観察 3.3**
>
> 浅く作ると、素子が桁で要る。一次元で 1024 区間の関数を一層で作るには、折れ目の数だけ ReLU が要り 1023 個、深さで作れば 20 個で、51.150000 倍である——深さが稼いでいるのは、折れ目の数の指数である。

観察 3.4

深さの値打ちは、連続には増えない。非線形が無ければ 0、あれば指数で、その間の形は無い——「少し非線形なら少し得する」という中間の読みは、区間の数の上では成り立たない。

4. 分離子

観察 4.1

分離子は、層のあいだに ReLU のような非線形を挟むかどうかである。挟まなければ深さは 0 の効き、挟めば指数の効きになる——「深いほど強い」は、この条件を言わなければ真偽が決まらない。

観察 4.2

論文330 と並べる。330 は、正規方程式で条件数が二乗になるのが近似ではなく恒等式であることを示した——そこでは行列の積が条件数を悪くした。ここでは行列の積が深さを消す。どちらも近似でなく恒等式で、行列の積が何を増やし、何を消すかの両側である。

観察 4.3

言わないこと。学習も、多次元の入力も、実在の網も扱わず、区間の数だけを数えた——示したのは、深さの効きが非線形の有無で 0 か指数かに分かれることである。

5. 点呼と結び

1. 三層と一層の差 5.192×10^{-15}（計算 2.1）：恒等式。
2. 直線の誤差 0.083342（計算 2.1）：深くしても減らない。
3. $1/12$ との差 8.335×10^{-6}（観察 2.3）：閉じた値に一致。
4. ReLU 二つで誤差 0（観察 2.4）：折れ曲がりが作る。
5. 区間は 2^L（計算 3.1）：本稿の芯。
6. $L = 10$ で 1024 区間（計算 3.1）：ReLU 20 個。
7. 一層なら 1023 個（計算 3.1）：桁で要る。

8. **51.150000 倍**（観察 3.3）：深さが稼ぐもの。
9. **0 か指数か**（観察 3.4）：**中間は無い**。
10. **非線形を挟むか**（観察 4.1）：**分離子・論文330**。

結び：線形の層は何枚重ねても一枚と厳密に同じで、折れ曲がりを一つ挟むと層ごとに区間が倍になる。線形の三層とその積の一層は、200 本の入力で出力の相対差が最大 5.192×10^{-15} で、行列の積という恒等式によって同じである——だから $|x|$ を最良の直線で近似した平均二乗誤差 0.083342（$1/12$）は、何枚重ねても一つも減らない。ところが ReLU 二つを使えば $|x|$ は誤差 0 で作れ、表せるものを増やしたのは層の枚数ではなく折れ曲がりである。そしてテント写像を L 回合成すると、直線の区間はちょうど 2^L になり、$L = 10$ で 1024——深さで作れば ReLU 20 個、一層で作れば 1023 個が要り、51.150000 倍になる——深さの値打ちは非線形が無ければ 0、あれば指数で、その間の形は無い。分離子は、層のあいだに ReLU のような非線形を挟むかどうかである。既刊との位置——論文330 では行列の積が条件数を悪くする恒等式だった。ここでは行列の積が深さを消す恒等式である。正直に言えば——学習も多次元の入力も実在の網も扱わず、区間の数だけを数えた。

参考文献

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
2. Telgarsky, M. (2016). Benefits of depth in neural networks. *Conference on Learning Theory*.
3. 山岸勇輝 (2026). 『条件数が二乗になるのは、近似ではなく恒等式である』[論文330]. Zenodo. 10.5281/zenodo.22681423

— From thinking to syncing. —

山岸 勇輝 (Yuuki YAMAGISHI)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)