誤差には最小が在るが、その最小の位置は模型の性質ではなく、手持ちのデータ数で動く

— データ数 **20** では自由度 **5**、**2500** では自由度 **27** で、データを **125** 倍にすると最小を与える自由度は **5.400000** 倍になる — 分離子は、数えている誤差が当てはめたデータの上か、当てはめていないデータの上か — [論文733]

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

模型が複雑すぎても単純すぎても誤差は大きい。その間に最小が在ることは知られている。ここで数えるのは、その最小がどこに在るかは模型の性質ではないということである。最小の位置がデータ数で動くことを示す。新しい定理も法則も主張しない。 第一に、二つの項は逆向きに動く——自由度 1 で偏り 4.000000・分散 0.010000、自由度 32 で偏り 0.003906・分散 0.320000 である（第2節）。 第二に、だから内側に最小が在る——データ数 100 では自由度 9 で 1.139383 になり、両端の 5.010000 と 1.323906 より小さい（第2節）。 第三に、これが本稿の芯である。最小の位置はデータ数で動く——データ数 20 では自由度 5、2500 では自由度 27 で、5.400000 倍になる（第3節）。 第四に、だから「ちょうどよい複雑さ」は模型の性質ではない——同じ模型の族でも、手持ちのデータ数が変われば答が変わる（第3節）。 第五に、雑音の項は消えない——データ数を 10^6 にしても、誤差は雑音の 1.0 を下回れない（第3節）。 第六に、分離子は、数えている誤差が当てはめたデータの上か、当てはめていないデータの上かである——前者なら複雑にするほど小さくなり、最小は在らない（第4節）。 キーワード：偏りと分散、誤差の分解、最適な複雑さ、標本数、消えない雑音、どのデータで数えるか。

本稿の射程（射程注記）：新しい定理も法則も主張しない——偏り・分散・雑音への分解、複雑さに対する二つの項の向き、最適な複雑さが標本数とともに増えることは、いずれも標準的である（ゲマン 1992、ハスティー 2009）。形を前提として置く——偏りの二乗を $4/p^2$、分散を $\sigma^2 p/n$ と置いた。これは実在の学習器の振舞いを主張するものではなく、読むのは「逆向きに動く」「最小が動く」という構造だけである。手順を扱わない——交差検証、正則化、早期打ち切りの手順は扱わない。二重降下を扱わない——自由度をさらに増やすと誤差が再び下がる現象は、この形では出ない。実在の模型を主張しない——個別のデータや学習器は扱わない。既刊との関係：論文701 は、データを増やすほど罰則つきの規準が複雑なほうを選ぶようになることを示した——そこでは規準が何を選ぶかが主題で、交差点を数で出した。ここでは誤差そのものを三つに分け、最小を与える複雑さがデータ数とともに増えることを示す。規準の選択と、誤差の分解——どちらも「データが増えると複雑なほうへ動く」という同じ向きを指す。加えたのは、三項を同じ表に並べたこと、最小が内側に在ることを両端との比較で示したこと、最小の位置が 5.400000 倍動くことを出したこと、雑音の下限が残ることを示したこと、分離子を「どのデータの上で誤差を数えるか」に置いたことである。

1. 記法と問い

設定：当てていないデータの上での誤差は、三つに分かれる。偏りの二乗、分散、そして雑音である。

$$\text{誤差} = \text{偏り}^2(p) + \text{分散}(p, n) + \sigma^2, \qquad \text{偏り}^2 = \frac{4}{p^2}, \qquad \text{分散} = \frac{\sigma^2 p}{n}$$

p は模型の自由度、n はデータ数である。二つの型は前提として置いたもので、$\sigma^2 = 1.0$ とする。

問い：ちょうどよい複雑さは、何の性質なのか。

2. 二つの項が逆向きに動き、内側に最小が在る

計算 2.1：データ数 100 で、自由度を変えたときの三項を出した。

自由度	偏りの二乗	分散	合計
1	4.000000	0.010000	5.010000
2	1.000000	0.020000	2.020000
4	0.250000	0.040000	1.290000
8	0.062500	0.080000	1.142500
16	0.015625	0.160000	1.175625
32	0.003906	0.320000	1.323906

> **観察 2.2**
>
> 向きが逆である。偏りは自由度とともに減り、分散は増える——だから両端はどちらも悪く、内側に最小が在る。

> **観察 2.3**
>
> 最小は自由度 **9** で、**1.139383** である。自由度 **1** の **5.010000**、自由度 **32** の **1.323906** より小さい——単純すぎる側の悪さは偏りから、複雑すぎる側の悪さは分散から来ている。

3. 最小の位置は、データ数で動く

計算 **3.1**：データ数を変えて、最小を与える自由度と、そのときの三項を出した。

データ数	最小の自由度	偏りの二乗	分散	誤差
20	5	0.160000	0.250000	1.410000
100	9	0.049383	0.090000	1.139383
500	16	0.015625	0.032000	1.047625
2500	27	0.005487	0.010800	1.016287

観察 3.2

これが本稿の芯である。データを 125 倍にすると、最小を与える自由度は 5 から 27 へ 5.400000 倍になる。模型の族は一つも変えていない——「この模型はこれくらいの複雑さがちょうどよい」という言い方が、手持ちのデータ数を黙って前提にしている。

観察 3.3

向きは分散の側から来ている。分散はデータ数に反比例するので、データが増えると複雑さの代金が安くなる——偏りの側はデータ数に依らないので、安くなったぶんだけ複雑にできる。

観察 3.4

それでも雑音は消えない。データ数を 10^6 にしても、誤差は雑音の 1.0 を下回れない——誤差 0 を目指すという言い方は、この分解の中では成り立たない。

観察 3.5

三項は同じ単位だが、性質が違う。偏りは模型の選び方、分散はデータ数、雑音はどちらでもない——同じ表に並べられるのに、動かせるつまみが違う。

4. 分離子

観察 4.1

分離子は、数えている誤差が当てはめたデータの上か、当てはめていないデータの上かである。前者なら複雑にするほど小さくなり、最小は在らない——最小が在ることそのものが、外のデータを見ていることの帰結である。

> **観察 4.2**
>
> 論文701 と並べる。701 は、データを増やすほど罰則つきの規準が複雑なほうを選ぶようになることを示した——そこでは規準の選択が主題だった。ここでは誤差の分解で、最小の位置が同じ向きに動く。選ぶ側と、選ばれる側の誤差である。

> **観察 4.3**
>
> 言わないこと。交差検証も正則化も二重降下も扱わず、二つの型は前提として置いた——示したのは、最小が動くことと、何が動かしているかである。

5. 点呼と結び

1. 自由度 **1** で **5.010000**（計算 2.1）：偏りが大きい。
2. 自由度 **32** で **1.323906**（計算 2.1）：分散が大きい。
3. 最小は **1.139383**（計算 2.1）：自由度 **9**。
4. データ **20** なら自由度 **5**（計算 3.1）：最小の位置。
5. データ **2500** なら **27**（計算 3.1）：動いている。
6. **5.400000** 倍（観察 3.2）：本稿の芯。
7. 誤差は **1.016287** まで（計算 3.1）：データ **2500** で。
8. 雑音の **1.0** は残る（観察 3.4）：誤差 **0** は無い。
9. 分散は n に反比例（観察 3.3）：複雑さが安くなる。
10. どのデータの上で数えるか（観察 4.1）：分離子・論文701。

結び：誤差には最小が在るが、その最小の位置は模型の性質ではなく、手持ちのデータ数で動く。誤差を偏りの二乗・分散・雑音に分けると前の二つは逆向きに動き、データ数 100 では自由度 1 で偏り 4.000000＋分散 0.010000＝5.010000、自由度 32 で偏り 0.003906＋分散 0.320000＝1.323906、その間の自由度 9 で最小 1.139383 になる——単純すぎる側の悪さは偏りから、複雑すぎる側の悪さは分散から来ている。ところが最小の位置は動く——データ数 20 では自由度 5、2500 では自由度 27 で、データを 125 倍にすると 5.400000 倍になる——模型の族は一つも変えていないので、「この模型はこれくらいの複雑さがちょうどよい」という言い方は、手持ちのデータ数を黙って前提にしている。向きは分散の側から来ている——分散はデータ数に反比例するので、データが増えると複雑さの代金が安くなり、偏りの側はデータ数に依らないので、安くなったぶんだけ複雑にできる。それでも雑音は消えず、データ数を 10^6 にしても誤差は 1.0 を下回れない——誤差 0 を目指すという言い方は、この分解の中では成り立たない。分離子は、数えている誤差が当てはめたデータの上か、当てはめてい

ないデータの上かである——前者なら複雑にするほど小さくなるので、最小が在ることそのものが、外のデータを見ていることの帰結である。既刊との位置——論文701 は規準の選択を扱った。ここは選ばれる側の誤差で、同じ向きに動く。正直に言えば——二つの型は前提として置いたもので、交差検証も二重降下も扱わなかった。

参考文献

1. Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation*.
2. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.
3. 山岸勇輝 (2026).『データを増やすほど、罰則つきの規準は複雑なほうを選ぶようになる』[論文701]. Zenodo. 10.5281/zenodo.23014898

— 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)