チェザロ平均から学ぶ 数学の証明の基本と全て

February 20, 2026

チェザロ平均の定理

定理

正の実数列 $\langle a_i : i \in \mathbb{N} \setminus \{0\} \rangle$ が 0 に収束するとする．このとき，その始切片の平均からなる数列 $\langle b_i = \dfrac{a_1 + a_2 + \ldots + a_i}{i} : i \in \mathbb{N} \setminus \{0\} \rangle$ も 0 に収束する．

この定理でいうところの各 b_i のことをチェザロ平均と呼びます．この定理自体は有名ですし，証明もよく知られていると思います．

しかし，その証明を自力で構築し書き切るための「思考法」にスポットライトを当てた解説は多くないと思います．

証明技術の全てが詰まっている

チェザロ平均の定理を証明するためには, 高校以前の数学で目にする証明のセオリーはほとんど役に立ちません.

一方で, 量化子を含む命題の証明, およびその基本, 基本にして本質的な全ての注意事項を, この証明で学ぶことができます.

今回は, チェザロ平均の定理の証明を通して, 「証明問題にいかにして向き合うべきか」を解説していきます.

仮定を変形して結論を得ようとするな - 1

定理

正の実数列 $\langle a_i : i \in \mathbb{N} \setminus \{0\} \rangle$ が 0 に収束するとする．このとき，その始切片の平均からなる数列
$\langle b_i = \dfrac{a_1 + a_2 + \ldots + a_i}{i} : i \in \mathbb{N} \setminus \{0\} \rangle$ も 0 に収束する．

さて，この定理を証明しようというときに，a_i に関する条件を変形して b_i に関する条件を得ようと考えるのは，高校以前の数学ではよくある試みですが，
それは完全なる悪手です．
なぜなら，今問われているのは b_i に関する条件であり，
a_i の条件は利用できる仮定でしかなく，示すべき事柄に直接の関係がないからです．

仮定を変形して結論を得ようとするな - 2

今の注意の意味が分からないなら, 例えば次のことを考えてみてください.

「$b = 2$ であるとする. このとき, 111111 を b で割ると余りが 1 になることを示せ」

これを見て, $b = 2$ を変形して結論を出そうとする人はかなり筋が悪いことは理解できると思います.
実際は 111111 を b で割るときに, $b = 2$ であることを用いて代入し実際に割り算の筆算なりをして $111111 = 2k + 1$ のような結果を得て結論に至るはずですね. これと同じことなのです.

仮定というのは, 必要なタイミングで必要なら何度でも引用して用いるためのものであって,
証明すべき結論の言い換えなどではないのです.

基本に忠実に - 1

証明すべき命題を論理記号で表すと,
$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \in \mathbb{N} (n > N \longrightarrow b_n < \varepsilon)$
このように表現できるでしょう.

全称量化子が先頭に来ている命題の証明はまず, 何をするよりも最初に $\varepsilon > 0$ を任意に取って固定することです. これを証明の冒頭で「$\varepsilon > 0$ とする」と表現します. これはおよそどんな全称命題の証明でも同様の書き出しになります. これが基本というもので, 誰でもできます.

何が言いたいかというと, 前述のことと関係しますが, 数列 a_i に関する条件はいろいろ出ますし, 使わなければならないのですが, それが証明の冒頭にくることはありません. a_i の条件の変形や分析から書き出さないでください.

基本に忠実に - 2

また, 示すべき命題が「任意の $\varepsilon > 0$ について何々」だからといって,「任意の $\varepsilon > 0$ について」という文章のままいじったり,「ε は任意のものだったので」などと任意性を引きずるのも悪手です. ε を固定することによって任意性を文面からさっさと除去しないと目標を見失いやすいです.

この節でした注意は「誰でも出来る機械的な基本事項には忠実に従いましょう」ということです.

全称/存在命題の証明の実践における基本を, より詳細に解説した文献には, 嘉田勝先生による『論理と集合から始める数学の基礎』などがあります.

いきなり証明を書き始めるな

さて, それでは証明をしていきましょう. $\varepsilon > 0$ とします.

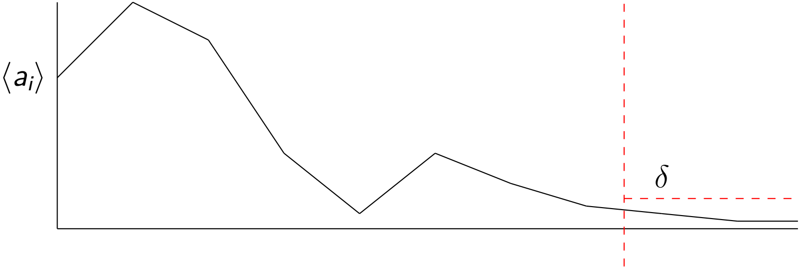
... とするのは気が早すぎます.
ここまでの注意は, 証明を書き始める段階になってから適用する話であって, 証明で最も重要なのは
「証明に手を付ける前に問題を観察しまくること」です.

1 に観察 2 に観察, 3,4 も観察, 5 も観察.

往年のファンクバンド風に言えば「観察, 観察 & 観察」
数学の四大元素があるならそのうち三つは観察です.

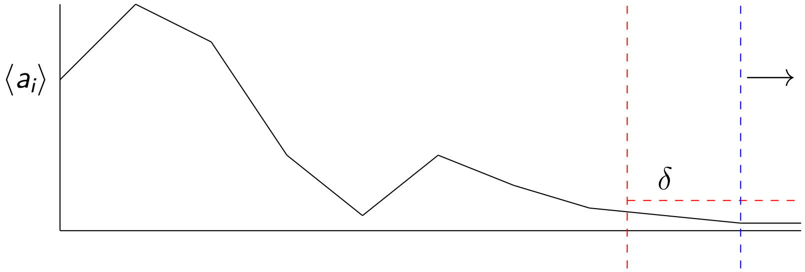
観察 A

それでは, 問題を観察していきましょう.
数列 $\langle a_i \rangle$ は 0 に収束する列とします. 0 に収束する列というのは, 「どんなに小さい数 δ を指定しても, あるところから先はその数未満になる」ということです. ここで $\langle a_i \rangle$ を前後半に分けると, 後半だけで平均を考えれば δ を越えることはないことが分かります.



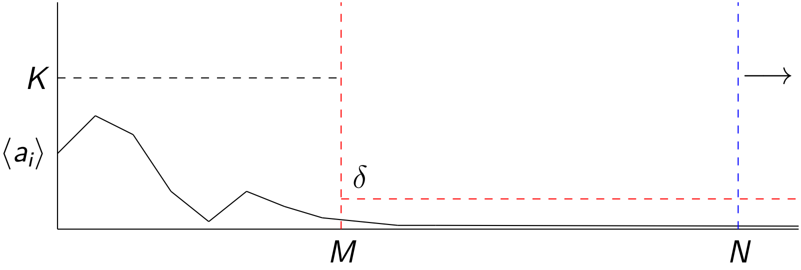
観察 B

$\langle a_i \rangle$ の前半部分は暴れ散らかしている可能性がありますが, $\langle a_i \rangle$ を十分に伸ばしていけば, 追加される数は全て, 平均を δ 以上にする方向には寄与しないため, いずれ全体の平均も δ を下回ってくれるでしょう. ただし, どれぐらい伸ばすべきかは δ に依存し, δ が小さいほど長く伸ばさねばなりません.



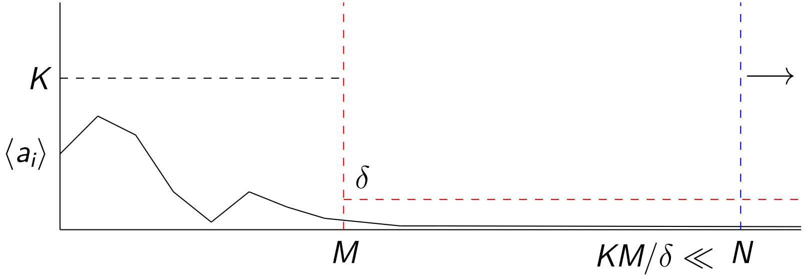
観察 C

ここは観察の本質からは蛇足になりますが, 証明を実際に構築する上では, 考える必要のある $\langle a_i \rangle$ の長さ N はどれぐらいか軽く見積もる必要があります. $\langle a_i \rangle$ 前半の長さを M, 上界を K とします.



観察 C

すると, 前半部分の総和は高々 KM, これを高さ δ に均すのなら, 必要な長さは KM/δ. N はそれより十分に大きく取ればよいはずです. (後半で追加される部分を平均に考慮していないので, KM/δ ギリギリはよろしくない.)
このぐらい観察すれば具体的な証明を書けそうです.



証明 - 1

定理

正の実数列 $\langle a_i : i \in \mathbb{N} \setminus \{0\} \rangle$ が 0 に収束するとする．このとき，その始切片の平均からなる数列 $\langle b_i = \dfrac{a_1 + a_2 + \ldots + a_i}{i} : i \in \mathbb{N} \setminus \{0\} \rangle$ も 0 に収束する．

証明: $\varepsilon > 0$ とする．証明すべき論理式は
$\exists N \in \mathbb{N} \forall n \in \mathbb{N}(n > N \longrightarrow b_n < \varepsilon)$ である．
$\varepsilon > 0$ は任意のものであったが，証明すべき論理式の性質上，一般性を失わずに $\varepsilon < 1$ としてよい．というのも $\varepsilon < 1$ である 1 つの ε に対して $\forall n \in \mathbb{N}(n > N \longrightarrow b_n < \varepsilon)$ である $N \in \mathbb{N}$ が取れたのなら，この $N \in \mathbb{N}$ は $\varepsilon \geq 1$ である全ての ε に対しても $\forall n \in \mathbb{N}(n > N \longrightarrow b_n < \varepsilon)$ を満たすからである．

証明 - 2

本当は ε にこのような制限をかけなくてもよいのですが，収束性を示す際のこの $\varepsilon > 0$ は非常に小さい任意の値を想定されているものです．ですから本質的には $x < x/\varepsilon$ のような不等式が使えるはずの場面でも，その式は $\varepsilon < 1$ のときのみ有効であるため利用できません．実際は $\varepsilon \geq 1$ の場合は相対的にトリビアルなので，$\varepsilon < 1$ を仮定しておいて $x < x/\varepsilon$ を自然に利用した方が感覚にマッチしていてスマートな場合があります．
数学では似たような理由から「一般性を失わずに追加の仮定をする」ことが多々あります．本当に一般性が失われていないかのチェックはちゃんとせねばなりません．

証明 - 3

本題に戻って，ここで，$\langle a_i \rangle$ が 0 に収束することから，$M \in \mathbb{N} \setminus \{0\}$ を $\forall i \in \mathbb{N}(i > M \longrightarrow a_i < \delta)$ であるように取れる．ただし，$\delta = \dfrac{\varepsilon}{10^{10}}$ である．

本当は $\delta = \varepsilon/2$ ぐらいでも良いのですが，δ は仮定から自由に定めていいものなので，証明の都合が良くなるようにガッツリ小さく取ります．証明を書いていくにあたって重要なポイントを外さないようにする工夫の一つです．もし，途中で上手く行かなくなったらそれは観察不足ということになります．「小さく取り過ぎてはいけない理由を見落としている」あるいは「本質的に小さく取れていない」ということになります．

証明 - 4

ここで, $K = \max\{100, a_1, a_2, \ldots, a_M\}$ とおく.
このとき, $K > 1$ である.

$N = \lceil \frac{KM \times 10^{100}}{\varepsilon} \rceil$ で定める.

$K > 1$ かつ $\varepsilon < 1$ であるから, $N > M$ である.

この N も過剰すぎるガッツリ大きい場所に取っています. 予め $\varepsilon < 1$ にしておいたので N を簡単に定義できています. K の定義に 100 を混ぜてるのは $K > 1$ を保証するため. a_M までの a_i を全て押さえて十分デカいものならなんでも良いです. ここは観察 B, C に対応しています.

証明 - 5

後は, $\forall n \in \mathbb{N}(n > N \longrightarrow b_n < \varepsilon)$ を示せばよい. $n > N$ とする. $M < N$ なので,

$$
\begin{aligned}
b_n &= \frac{a_1 + a_2 + \ldots a_M + a_{M+1} + \ldots + a_n}{n} \\
&\leq \frac{KM}{n} + \frac{a_{M+1} + \ldots + a_n}{n} \leq \frac{KM}{N} + \frac{(n-M)\frac{\varepsilon}{10^{10}}}{n} \\
&\leq \frac{KM \times \varepsilon}{KM \times 10^{100}} + \frac{\varepsilon}{10^{10}} \leq \frac{\varepsilon}{10^{100}} + \frac{\varepsilon}{10^{10}} < \varepsilon
\end{aligned}
$$

以上より, 定理は示された.

上述の式変形で, a_i の後半部分の変形は観察 A に基づいていて, 前半部分の変形は観察 B, C に基づいていることに注意してください.

補足 - 1

数列の収束性を示す際に，最初に固定した ε の定数倍で数列を押さえられればよいとする証明も見かけます．チェザロ平均の証明では

$\varepsilon > 0$ とする．$\exists N \in \mathbb{N} \forall n \in \mathbb{N}(n > N \longrightarrow b_n < \varepsilon)$ を示す．
$\langle a_i \rangle$ が 0 に収束することから，
$M \in \mathbb{N} \setminus \{0\}$ を $\forall i \in \mathbb{N}(i > M \longrightarrow a_i < \varepsilon)$ であるように取る．
$\cdots$ （中略） $\cdots$
このとき，$b_n < 2\varepsilon$ なのでこれでよい．

と結ぶようなパターンです．
これは私は全く推奨しません．

補足 - 2

$M \in \mathbb{N} \setminus \{0\}$ を $\forall i \in \mathbb{N}(i > M \longrightarrow a_i < \varepsilon)$ であるように取る.

この証明の良くない点は $\langle a_i \rangle$ を押さえるのに ε を無批判にそのまま用いていることです. ε は $\langle b_n \rangle$ を押さえる目標であって, $\langle a_i \rangle$ には直接関係ありません. $\langle a_i \rangle$ についての観察が十分に足りておらず, ガチャガチャやった結果として偶然 2ε で押さえられただけのような証明に見えます.

そもそも文面上 ε で押さえることが当初の目的で進めているのに, それに失敗しているのですから, 証明としては不満が残ります.

まとめ

今回説明した内容を簡単に振り返りましょう.

- 仮定を変形して結論を得ようとしてはいけない.
- 証明の書き方は基本に忠実に.
- 証明を書き始める前に, 必要な論証の本質部分がはっきりするまで十分な観察を行う.
- 十分な観察を行った後, 本質的でないノイズを一般性を失わないように排除して議論を簡素化できるか考察する.
- あえて汚い証明を書く必要はないが, 教科書のマネのような綺麗な証明を書こうとしなくてよい.

最後に追加で言いたい

証明問題はあくまで問題ですが，証明というのは本来，未知の命題の正否を決定するための未知な論証です．
であれば，むしろ最初から教科書のような綺麗な証明が書けるほうがおかしいのです．

また，既に書かれた証明の正当性を追うことだけなら，多少訓練された学部生にもできます．しかし，それだけで「研究」ができるようにはなりません．

重ねて言いますが，数学において最重要なのは「観察」です．膨大な経験と観察によって数学的現象の本質をつかみ，論理を組み合わせて真偽不明の命題に瑕疵の無い厳密な論証を与える．これが「研究」です．

学問と向き合う学生諸君の健闘を祈ります．