

読書ノート：地力をつける微分積分

tomixy

2025 年 3 月 8 日

* * *

微分と積分は何を捉えているか

微分は、「微小な変化でどのような変動が起こるか」を分析する数学の手法

数学では、**極限**という概念を用いて「無限小レベルの変化」として厳密に微分を定義する

* * *

現実の事象を解明しようとする、その事象に関わる要因は1つではなく、複数の要因が絡み合っていることが多い

複数の要因が絡み合っている状況を数量的に表すのが**多変数関数**

* * *

変数が2個以上あると、「変数を少し動かす」といってもいろいろな動かし方がある

その動かし方を精密に扱うのが**偏微分**

* * *

積分は「そこにある量を算出する道具」

「そこにある量」を小分けして合算して求める考え方（**区分求積法**）が積分論の主軸

本書の内容

数学では、相対的に無視できるものを切り捨てるという操作を論理的に積み重ねて、**無限**という概念に向き合う

近似をしたつもりなのに大きな誤差が生じているとすると、その兆候はどこかに現れているはず
局所近似の誤差の兆候は**高階の微分**に現れる
これを逆に突き詰めると**テイラー展開**という概念に到達する

大きな数を感覚的に捉える方法

分子も分母も無限大に近づく中で、分母に比べて分子を無視できるという等式

$$\lim_{n \rightarrow \infty} \frac{n^2}{n^3} = 0$$

分子を1辺 n の正方形の面積（2次元）、分母を1辺 n の立方体の体積（3次元）とみなせば、「次数が高くなると巨大な数が現れやすい」という性質の1つの姿と思うこともできる

収束や発散の速さ

a_1, a_2, a_3, \dots という数列があったときに、その初めの N 個の和を次のように表す

$$\sum_{k=1}^N a_k = a_1 + a_2 + a_3 + \dots + a_N$$

2個や3個の和なら $a_1 + a_2$ や $a_1 + a_2 + a_3$ と書けるが、たとえば N が 10^{16} というようなときには、

すべての項を書き出せないため、このように記法を決めておく

* * *

等比級数 A_N の収束の速さを考える

$$A_N = \sum_{k=1}^N \frac{1}{2^k} = \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^N}$$

この和に $\frac{1}{2^N}$ を足すと、1 になる

たとえば、

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} = 1$$

なので、 N 個の総和は、

$$1 - \frac{1}{2^N}$$

$\frac{1}{2^N}$ がどれくらい小さいかがわかれば、 A_N の和が 1 にどれくらい近いかわかる

では、 $N = 10^{16}$ のとき、 $\frac{1}{2^N}$ はどれくらいの大きさか？

まず、 $N = 10$ の場合を考えると、

$$2^{10} = 1024 \doteq 1000$$

より、

$$\begin{aligned} \frac{1}{2^{10}} &\doteq 0.001 \\ 1 - \frac{1}{2^{10}} &\doteq 0.999 \end{aligned}$$

というふうに、小数点以下に 9 が 3 つ連続して並ぶ $N = 10^2$ のときは、

$$2^{100} = (2^{10})^{10} = (1024)^{10} \doteq (1000)^{10} = 10^{30}$$

から、9 がおよそ 30 個並ぶ ($1024 > 1000$ なので、「少なくとも」30 個並ぶ)

$N = 10^{16}$ のときは、

$$2^{10^{16}} = (2^{10})^{10^{15}} \doteq 10^{30 \times 10^{15}}$$

から、 $3 \times 10^{15} = 3000$ 兆以上の 9 が並ぶ

このように、 N を大きくしていくと、 2^N はきわめて 1 に近い数になる

$N \rightarrow \infty$ としたときに、等比級数 A_N は 1 に収束する

誤差評価

測定や推定で何らかの値を得たとき、**誤差**は次のように定義される

$$\text{誤差} = |\text{測定値 or 推定値} - \text{真の値}|$$

誤差が論理的に小さいと言えれば、その測定値や推定値は一定の安心感を持って**近似値**として使うことができる

誤差がある数 ε より小さくなるという不等式

$$|\text{測定値 or 推定値} - \text{真の値}| < \varepsilon$$

を**誤差評価**という

N が 10^{16} のとき、右辺は小数点以下に少なくとも 3×10^{15} 個の 0 が並ぶ

$$|A_N - 1| < 0.00 \dots 0 \dots$$

この不等式は、 $N = 10^{16}$ のときの A_N が、極限值である 1 をどの程度の精度で近似しているかを表す誤差評価と考えることもでき、この誤差評価は、相対的な収束の速さを数値的に表すものでもある (右辺の小数点以下に並ぶ 0 の数で速さがわかる)

関数の全体を見る

$y = x^4$ と $y = x^2$ の差異を考えてみる

x が大きい場合の例として、たとえば $x = 10$ のと

きは、

$$y = x^4 = 10^4 = 10000$$
$$y = x^2 = 10^2 = 100$$

x をもっと大きくすると、 x^2 も x^4 も大きな数になるが、この2つだけを比較すると、 x^4 の方がはるかに大きいので x^2 は相対的に無視できそうだとはいえる

x が0に近い場合の例として、たとえば $x = 0.1$ のときは、

$$y = x^4 = 0.1^4 = 0.0001$$
$$y = x^2 = 0.1^2 = 0.01$$

$x = 0$ の近くでグラフを書くと、 $y = x^4$ の方は x 軸すれすれになる
このように、 x が0に近いときは x^2 も x^4 も小さな数だが、この2つだけを比較すると、 x^2 の方が相対的に大きいので x^4 は無視できるくらい小さそうだといえる

3つ以上のものを比較するときも、**圧倒的に大きいものが1つだけあれば残りを無視しよう**という考え方ができる
微分や積分における極限にも、この考え方が用いられる

* * *

厳密な論理体系である数学では、どういう基準で何が無視するかというルールを明確に決めてから議論を積み重ねることになる
その一方で、論理的な証明を導く際には、**効くものと無視できるものを区別する**という直観が役立つ

二項係数の4つの側面

パスカルの三角形を帰納的に定義する

- 1段目（最上段）は1とする
- n 段目に n 個の数を定めたとして、 $n + 1$ 段目は線で繋がっているすぐ上の数を足し合わせて定めることにする

このとき、次の4つの数が一致する

1. $(a + b)^n$ の展開式における $a^{n-k}b^k$ の係数
2. パスカルの三角形の $n + 1$ 段目、左から $k + 1$ 番目の数
3. パスカルの三角形で $n + 1$ 段目、左から $k + 1$ 番目の地点から最上段に登る最短経路の個数
4. ${}_nC_k = \binom{n}{k} = \frac{n!}{k!(n - k)!}$

* * *

1と4の一致 次のように $(a + b)^n$ を n 個の積を並べたものとして表すことで示す

$$(a + b)^n = \underbrace{(a + b) \times (a + b) \times \cdots \times (a + b)}_n$$

この展開式において、たとえば $a^{n-k}b^k$ という項がどこから生じるかを考える

右辺を展開するとき、 $a + b$ のそれぞれの項で a か b の2通りの選択肢があるため、展開すると全部で 2^n 通りの単項式の和になる

展開したときに $a^{n-k}b^k$ という項に寄与するのは、 n 個の $(a + b)$ の中から b を k 個選ぶ場合の数なので、 ${}_nC_k$ 通りある

こうして、 ${}_nC_k$ 通りの $a^{n-k}b^k$ が現れることから、 $a^{n-k}b^k$ の係数は ${}_nC_k$ となる

* * *

2と3の一致 ある地点から最上段に行くには、その地点の左上か右上に移動するしかない

そのため、この地点から最上段に登る最短経路の個数は、そのすぐ左上の地点を経由して最上段に登る最短経路の個数と、そのすぐ右上の地点を経由して最上段に登る最短経路の個数の和になる（和の法則）

これはまさにパスカルの三角形を定義したときのルールであり、パスカルの三角形の $n+1$ 段目、左から $k+1$ 番目の数は、 n 段目かつ左から k 番目の数と、 n 段目かつ左から $k+1$ 番目の数の和になる

* * *

1と2の一致 次のように書くことで見えてくる

$$\begin{aligned}(a+b)^n &= (a+b)^{n-1} \times (a+b) \\ &= (a+b)^{n-1} \times a + (a+b)^{n-1} \times b\end{aligned}$$

これは、 $(a+b)^{n-1}$ の各項に、 a と b をそれぞれかけて足し合わせる計算になっている

$(a+b)^{n-1}$ を展開すると、それぞれの項は、 a を $(n-1)-k$ 回、 b を k 回かけた形になる

$$(a+b)^{n-1} = \sum_{k=0}^{n-1} {}_{n-1}C_k \cdot a^{n-k-1} b^k$$

$(a+b)^{n-1}$ の各項に a をかけると、 a の指数が1増えるので、 a^{n-k} の項が現れる

$$\begin{aligned}(a+b)^{n-1} \times a &= \sum_{k=0}^{n-1} {}_{n-1}C_k a^{n-k} b^k \\ &= a^n + \sum_{k=1}^{n-1} {}_{n-1}C_{k-1} a^{n-k} b^k\end{aligned}$$

同様に、 $(a+b)^{n-1}$ の各項に b をかけると、 b の指数が1増えるので、 b^{k+1} の項が現れる

$$\begin{aligned}(a+b)^{n-1} \times b &= \sum_{k=0}^{n-1} {}_{n-1}C_k a^{n-k-1} b^{k+1} \\ &= \sum_{k=1}^n {}_{n-1}C_{k-1} a^{n-k} b^k + b^n\end{aligned}$$

よって、これらを足し合わせると、

$$(a+b)^n = a^n + \sum_{k=1}^{n-1} ({}_{n-1}C_{k-1} + {}_{n-1}C_k) a^{n-k} b^k + b^n$$

${}_{n-1}C_{k-1} + {}_{n-1}C_k$ の部分は、それぞれ

- ${}_{n-1}C_{k-1}$ はパスカルの三角形の n 段目、左から k 番目の数
- ${}_{n-1}C_k$ はパスカルの三角形の n 段目、左から $k+1$ 番目の数

を表すので、これらの和 ($a^{n-k}b^k$ の係数) がパスカルの三角形の $n+1$ 段目、左から $k+1$ 番目の数になることがいえる

(注: 段数は1から始まるので、 $n+1$ 段目が $(a+b)^n$ の展開に対応する。同様に、左から数える番号も1始まりなので、 $k+1$ 番目が $a^{n-k}b^k$ の係数に対応する)

* * *

一般の展開公式 1と4の一致から、一般の展開公式は次のように書ける

$$(a+b)^n = \sum_{k=0}^n \frac{n!}{k!(n-k)!} a^{n-k} b^k$$

この展開公式を二項展開という

そして、 $a^{n-k}b^k$ の係数

$${}_nC_k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1}$$

は二項係数という

* * *

二項展開において、 $a = 1$ の場合を考えると、

$$(1+b)^n = 1 + nb + \frac{n(n-1)}{2!}b^2 + \frac{n(n-1)(n-2)}{3!}b^3 + \dots$$

ここで、二項係数はいつでも正なので、 $b > 0$ ならば、上の二項展開に現れる項はすべて正となる

特に、 $n \geq 2$ ならば、次の不等式が成り立つ

$$(1+b)^n > 1 + nb$$

$100!$ と 10^{100} はどちらが大きいのか？

$n = 100$ のとき、 10^n はとても大きい数だが、 $n!$ と比べたら取るに足らないことを示す ($n!$ を概算するスターリングの公式は使わずに)

* * *

$100!$ において、10 から先をすべて 10 に置き換える
10 から 100 までの数は 91 個あるので、

$$100! > 10^{91}$$

がわかる

より精密に評価するために、10 から 99 までの 90 個の数の積を 10 個ずつまとめてみると、

$$10^{10} < 10 \cdot 11 \cdots 19 < 10^{20}$$

$$10^{20} < 20 \cdot 21 \cdots 29 < 10^{30}$$

\vdots

$$10^{90} < 90 \cdot 91 \cdots 99 < 10^{100}$$

さらに、これを縦にかけ合わせて、

$$10^{10} \cdots 90^{10} < 10 \cdots 99 < 20^{10} \cdots 100^{10}$$

左辺は、

$$\begin{aligned} 10^{10} \cdots 90^{10} &= (10 \cdots 90)^{10} \\ &= ((10 \cdot 1) \cdot (10 \cdot 2) \cdots (10 \cdot 9))^{10} \\ &= (10^9 \cdot (1 \cdot 2 \cdots 9))^{10} \\ &= (10^9 \cdot 9!)^{10} \\ &= 10^{90} \cdot (9!)^{10} \end{aligned}$$

右辺は、

$$\begin{aligned} 20^{10} \cdots 100^{10} &= (20 \cdots 100)^{10} \\ &= ((10 \cdot 2) \cdot (10 \cdot 3) \cdots (10 \cdot 10))^{10} \\ &= (10^9 \cdot (1 \cdot 2 \cdots 9 \cdot 10))^{10} \\ &= (10^9 \cdot 9! \cdot 10)^{10} \\ &= (10^{10} \cdot 9!)^{10} \\ &= 10^{100} \cdot (9!)^{10} \end{aligned}$$

なので、

$$10^{90} \cdot (9!)^{10} < 10 \cdots 99 < 10^{100} \cdot (9!)^{10}$$

ここで、

$$100! = 9! \cdot 10 \cdots 99 \cdot 100$$

と表し、後回しにしていた $9! \cdot 100$ を不等式の各項にかけると、

$$10^{90} \cdot (9!)^{10} \cdot 10^2 \cdot 9! < 100! < 10^{100} \cdot (9!)^{10} \cdot 10^2 \cdot 9!$$

$$10^{92} \cdot (9!)^{11} < 100! < 10^{102} \cdot (9!)^{11}$$

ところで、 $9! = 362880 \approx 3.6 \times 10^5$ から、

$$3 \times 10^5 < 9! < 4 \times 10^5$$

と粗く評価しておき、この式の両辺を 11 乗すると、

$$3^{11} \times 10^{55} < (9!)^{11} < 4^{11} \times 10^{55}$$

ここで、次のように考えると、 $10^5 < 3^{11}$ という大まかな不等式が成り立つ

$$\begin{aligned} 10^5 &= 3^5 \times \frac{10^5}{3^5} \\ &= 3^5 \times \left(\frac{10}{3}\right)^5 \\ &\doteq 3^5 \times (3.3)^5 \\ &= 3^5 \times 3^{0.3 \times 5} \\ &= 3^{10} \times 3^{0.3} \\ &< 3^{11} \end{aligned}$$

また、次のように考えることで、 $4^{11} < 10^7$ という大まかな不等式も成り立つ

$$\begin{aligned} 4^{11} &= 2^{22} \\ &= 2^{10} \times 2^{10} \times 2^2 \\ &= 1024^2 \times 4 \\ &= \left(1000 \times \frac{1024}{1000}\right)^2 \times 4 \\ &= \left(10^3 \times 1.024\right)^2 \times 4 \\ &= 10^6 \times 1.024^2 \times 4 \\ &< 10^6 \times 1.1^2 \times 4 \\ &< 10^7 \end{aligned}$$

これらを使うと、 $(9!)^{11}$ に関する不等式の左辺と右辺は、

$$\begin{aligned} 3^{11} \times 10^{55} &< (9!)^{11} < 4^{11} \times 10^{55} \\ 10^5 \times 10^{55} &< (9!)^{11} < 10^7 \times 10^{55} \\ 10^{60} &< (9!)^{11} < 10^{62} \end{aligned}$$

したがって、 $100!$ に関する不等式の左辺と右辺は、

$$\begin{aligned} 10^{92} \times (9!)^{11} &< 100! < 10^{102} \times (9!)^{11} \\ 10^{92} \times 10^{60} &< 100! < 10^{102} \times 10^{62} \\ 10^{152} &< 100! < 10^{164} \end{aligned}$$

これで、大まかな手計算で $100!$ の大きさを評価できた

特に、 $10^{152} < 100!$ から、 $100!$ が 10^{100} よりもはるかに大きいことがわかる

cos x のテイラー展開

単項式におまじないの係数をつけて足したり引いたりした関数を考える

$$y = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \dots$$

この無限級数を途中で打ち切らずに、ずっと続けることを考えてみる

x を固定しておくとし、 n が大きいとき、分子のべき乗 x^n と分母の階乗 $n!$ では $n!$ の方が圧倒的に大きくなり、この無限級数は収束する

この無限級数をプロットしたグラフは、 $|x|$ が小さいところでは $y = \cos x$ のグラフとほぼ重なり合う。原点から少し離れたところでも、多項式の項の個数を増やすとよく近似できる

このように、調べたい関数を多項式で近似し、局所的に近似の精度を上げるときは多項式の項を増やすという形で定理を形式化したものが **テイラー展開** である

関数の局所的な様子を見る

簡単な関数のグラフは拡大していくと急に様子が変わったりせず、むしろ、だんだん安定したものになると考えられる

局所的な部分を拡大すると安定した姿になるとき、その様子を数学的にとらえる概念が **微分**

ものによっては、拡大するとどんどん見え方が変

わるものもある

拡大を何度繰り返しても同じ複雑さを保つ数学的
構造（フラクタル）も自然界には現れる

拡大すれば何でも簡単になるわけではないが、微
分では、拡大したとき安定していく「素直」なもの
を主な対象とする

つまり、**微分は局所を分析するのに強力な手法だ
が、万能ではない**

微分の定義

関数は変化の法則性をとらえる数学的言語

数 x に対して数 $f(x)$ が定まるとき、 $f(x)$ を変数 x
の関数という

* * *

座標 $(x, f(x))$ を xy 平面でプロットした曲線を関数
 $f(x)$ の**グラフ**という

これは、 x 座標の点 x における高さが $f(x)$ となる
曲線

* * *

この曲線の局所的な様子を見るのに、変数 x を
 $x + h$ に動かしてみる

そうすると、関数の値は $f(x)$ から $f(x + h)$ に変
わる

「素直」な関数のグラフをどんどん拡大すると、拡
大部分はだんだん直線のように見えるだろう、と
考えられる

h が小さいとき、斜めの曲線がほぼ一定の傾き
の直線に見えるというのは、関数の値の変化量
 $f(x + h) - f(x)$ が h にほぼ正比例するということ

式で表すと、 x から $x + h$ の区間のグラフを直線と
みなしたときの勾配

$$\frac{f(x + h) - f(x)}{h}$$

は、 h が 0 に近づくとある 1 つの数に近づくと、すな
わち、収束するはずである

* * *

■定義 h を 0 に近づけると、 $\frac{f(x + h) - f(x)}{h}$ が
ある数に収束するとき、 $f(x)$ は x において**微分可
能**であるという

このとき、極限値を

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h}$$

と書き、 $f(x)$ の**微分**または**微分係数（微係数）**と
いう

* * *

定数関数の微分 「収束する」ことを「限りなく近
づく」と言うこともある

日常的な言葉だと「限りなく近づく」には「その
値に達していない」というニュアンスを感じるが、
数学では、最初からずっと同じ値のときも「収束
する」場合を含める

$f(x)$ が x の値によらないとき、 $f(x)$ を**定数関数**と
いう

このときは h がどんな数でも $f(x + h) - f(x) = 0$
となるので、定数関数の微分は 0 である

* * *

微分係数が定まらない例

$$\lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h}$$

が収束しない状況の例として、 $y = |x|$ を考える

$f(x) = |x|$ の場合、 $x = 0$ で

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

を計算しようとする、

$h > 0$ のときは

$$\frac{f(h) - f(0)}{h} = \frac{h}{h} = 1$$

$h < 0$ のときは

$$\frac{f(h) - f(0)}{h} = \frac{-h}{h} = -1$$

となり、 h を正から 0 に近づけると、負から 0 に近づけると、 $\frac{f(h) - f(0)}{h}$ の極限の値が異なってしまうので、微分係数 $f'(0)$ が定まらない

* * *

■定理 $a < x < b$ で定義された、微分可能な関数 $f(x)$ が $x = c$ で最大値または最小値をとるならば、 $f'(c) = 0$ である

* * *

$f'(c)$ が最大値となる場合の証明

$$f'(c) = \lim_{h \rightarrow 0} \frac{f(c+h) - f(c)}{h}$$

において、 $f(c)$ が最大値であることから、

$$f(c) \geq f(c+h)$$

$$f(c+h) - f(c) \leq 0$$

したがって、 $h > 0$ のときは、

$$\frac{f(c+h) - f(c)}{h} \leq 0$$

となり、 h を正の側から 0 に近づけた極限值として $f'(c) \leq 0$ が成り立つ

一方、 $h < 0$ のときは、

$$\frac{f(c+h) - f(c)}{h} \geq 0$$

となり、 h を負の側から 0 に近づけた極限值として $f'(c) \geq 0$ が成り立つ

$f'(c) \leq 0$ かつ $f'(c) \geq 0$ なので、 $f'(c) = 0$ が導かれた □

* * *

$f'(c)$ が最小値となる場合の証明 $f'(c)$ が最大値となる場合と同様に示される □

導関数

x を止めて考えると、 $f(x)$ の微分は 1 つの数

$$\frac{df}{dx}(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

また別の視点として、

- x に数を与えると、何か 1 個、数が出てくる
- また別の x に対しては、別の数が出る

そう思うと、 x から $\frac{df}{dx}(x)$ への対応は 1 つの関数を与えていると考えることができる

このように、 $\frac{df}{dx}(x)$ を x の関数と見たとき、それを $f(x)$ の導関数という

* * *

「微分」と「導関数」は視点の違いで使い分けられる言葉

- x を止めて $\frac{df}{dx}(x)$ という 1 個の数（微分係数）に注目するのか
- x を変数と思って $\frac{df}{dx}(x)$ を関数とみなす（導関数として扱う）のか

後者の立場に立って、 $\frac{df}{dx}(x)$ を関数だと思えば、さらに微分を考えることができる

微分できないからといってそこで終わりではない

たとえば、関数概念を拡張した超関数の理論は、
極限

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

が存在しない場合にも、より広く「微分」という概念をとらえる枠組みを与えるもの

単項式 x^n の微分

$f(x+h) = (x+h)^n$ の二項展開

$$f(x+h) = x^n + nx^{n-1}h + {}_nC_2x^{n-2}h^2 + \cdots + h^n$$

を用いると、

$$\begin{aligned} f(x+h) - f(x) &= (x+h)^n - x^n \\ &= x^n + nx^{n-1}h + {}_nC_2x^{n-2}h^2 + \cdots + h^n - x^n \\ &= nx^{n-1}h + {}_nC_2x^{n-2}h^2 + \cdots + h^{n-1} \end{aligned}$$

上の式変形で、最初の x^n は最後の $-x^n$ と相殺されている

両辺を h で割ると、

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{nx^{n-1}h + {}_nC_2x^{n-2}h^2 + \cdots + h^{n-1}}{h} \\ &= nx^{n-1} + {}_nC_2x^{n-2}h + \cdots + h^{n-1} \end{aligned}$$

h が 0 に近づくと、

- h に無関係な最初の項 nx^{n-1} はそのまま残る
- 次の h の項は 0 に近づく
- その後の h^2, h^3, \dots, h^{n-1} の項はさらに速く 0 に近づく

というわけで、 h を 0 に近づけると nx^{n-1} に収束し、

$$\frac{d}{dx}x^n = nx^{n-1}$$