Yaminabe -備忘録-

栗原 雅

最終更新日：2025/10/6

README

- この PDF を第三者へ譲渡される場合には、必ず私に許可を取ってください。
- 思いついたことをとりあえずまとめる闇鍋みたいな PDF です。
- 実際の闇鍋みたいにそれぞれの具がコンフリクトして絶望的なまずさになるのかは書いてみないと分からない。
- 誤りや誤植を見つけた際には適宜指摘して頂けると助かります。

目次

1　e^x の定義について

ネイピア数の定義：

$$e \equiv \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \qquad (n \in \mathbb{N}) \tag{1.1}$$

から出発してはさみうちの定理を用いることで実数 x に関する指数関数

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \tag{1.2}$$

を得る。

e^x のべき級数展開を行う際に一見困ったことが生じる。

$$\begin{aligned} e^x &= \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \\ &= \lim_{n\to\infty}\sum_{k=0}^{n} {}_n\mathrm{C}_k \left(\frac{x}{n}\right)^k \\ &= \lim_{n\to\infty}\sum_{k=0}^{n} 1\cdot\left(1-\frac{1}{n}\right)\cdot\cdots\cdot\left(1-\frac{k-1}{n}\right)\frac{x^k}{k!} \end{aligned} \tag{1.3}$$

ここまでは問題ない。ここで、見知っている形：

$$e^x = \sum_{k=0}^{\infty}\frac{x^k}{k!} \tag{1.4}$$

と比較する（導出の観点からは比較する必要はないが）と、各 x のべき毎に展開して考えて

$$\lim_{n\to\infty} 1\cdot\left(1-\frac{1}{n}\right)\cdot\cdots\cdot\left(1-\frac{k-1}{n}\right)\frac{x^k}{k!} = \frac{x^k}{k!} \tag{1.5}$$

となるので

$$e^x = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{x^k}{k!} = \sum_{k=0}^{\infty}\frac{x^k}{k!} \tag{1.6}$$

を得る（k を固定したまま、n の極限を取ってから k について和を取った）。

ここで生じる疑問点は式 (1.3) では和を取ってから極限を取ったが、式 (1.5)、(1.6) ではそれぞれの項で極限をとってから和をとっていて、この操作は正当化されるのか、ということである。一般に和と極限の操作は交換しない。

例 1.1. [和と極限が交換しない例]

以下の数列を考える：

$$a_{n,k} = \begin{cases} \frac{1}{\sqrt{n}} & \text{if } k < n, \\ 0 & \text{if } k \geq n. \end{cases}$$

和を取ってから極限を取ると

$$\lim_{n\to\infty} \sum_{k=0}^{n} a_{n,k} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{\sqrt{n}} = \lim_{n\to\infty} \sqrt{n} = \infty$$

となり、正の無限大に発散する。一方、k を固定したまま、n の極限を取ってから k について和を取ると

$$\lim_{n\to\infty} a_{n,k} = 0$$

$$\therefore \sum_{k=0}^{\infty} a_{n,k} = \sum_{k=0}^{\infty} 0 = 0$$

となり、0 に収束する。 ◁

このように、安易に和と極限を交換すると怖いことが起こりうる。確認したいことを提示する。

命題 1.2. [先ほどの操作は正当化できるのか？]

$f_{n,k}(x) = 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \cdots \cdot \left(1 - \frac{k-1}{n}\right) \frac{x^k}{k!}$ について各 k を固定したまま n に関して極限を取ると

$$\lim_{n\to\infty} \sum_{k=0}^{n} f_{n,k}(x) = \sum_{k=0}^{\infty} \lim_{n\to\infty} f_{n,k}(x)$$

が成り立つ。 ◁

実は、数学科の学部 1 年の後期で学習する事実（もはや飛び道具）を用いることで、先に挙げた操作は正当化できる。

定理 1.3. [**Weierstrass の M 判定法**]

I 上の実数値関数列 $\{f_n\}$ に対し、ある数列 $\{M_n\}$ が存在して、$|f_n(x)| \leq M_n$ $(x \in I)$ かつ $\sum_{n=1}^{\infty} M_n < \infty$ ならば、$\sum_{n=1}^{\infty} f_n < \infty$ は I 上絶対一様収束（絶対収束かつ一様収束）する。 ◁

一様収束の強力さは、物理学科が通常お祈りする強力な性質である項別収束が成り立つ（逆に項別収束が成り立つが、一様収束が成り立たない例もある）ところにある。

定理 1.4. [**d'Alembert の収束判定法**]

数列 $\{a_n\}$ に対し、$\left|\frac{a_{n+1}}{a_n}\right|$ の極限：

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = r$$

が存在するとする。このとき

- $0 \le r < 1 \Rightarrow \sum_{n=1}^{\infty} a_n$ は絶対収束
  （$\sum_{n=1}^{\infty} |a_n|$ が収束することを $\sum_{n=1}^{\infty} a_n$ が絶対収束するといい、このとき $\sum_{n=1}^{\infty} a_n$ は収束する。）
- $r > 1 \Rightarrow \sum_{n=1}^{\infty} a_n$ は発散

が成り立つ。 ◁

証明. [命題 1.2 について]

Weierstrass の M 判定法を用いる。

$$f_{n,k}(x) = 1\cdot\left(1-\frac{1}{n}\right)\cdot\cdots\cdot\left(1-\frac{k-1}{n}\right)\frac{x^k}{k!},\quad x\in[-R,R]\quad (R>0)$$

とすると、和を取る範囲から $k \le n$ となり

$$|f_{n,k}(x)| < \frac{R^k}{k!}(\equiv M_k). \tag{1}$$

ここで、$\sum_{k=0}^{\infty} M_k$ に関して d'Alembert の収束判定法を用いると

$$\left|\frac{M_{n+1}}{M_n}\right| = \frac{\frac{R^{k+1}}{(k+1)!}}{\frac{R^k}{k!}} = \frac{R}{k+1} < \infty \quad (\text{as } k\to\infty)$$

となるので

$$\sum_{k=0}^{\infty} M_k < \infty. \tag{2}$$

1、2から Weierstrass の M 判定法より、$\sum_{k=0}^{\infty} f_{n,k}(x)$ は $[-R,R]$ 上一様収束する。
(d'Alembert の収束判定法を用いた際に、$R\in\mathbb{R}$ で収束したので、収束半径は $\mathbb{R}$、すなわち $\mathbb{R}$ 上一様収束する。)

以上より $\sum_{k=0}^{\infty} 1 \cdot \left(1-\frac{1}{n}\right) \cdot \cdots \cdot \left(1-\frac{k-1}{n}\right) \frac{x^k}{k!}$ は項別収束するので

$$1 \cdot \left(1-\frac{1}{n}\right) \cdot \cdots \cdot \left(1-\frac{k-1}{n}\right) \frac{x^k}{k!} \to \frac{x^k}{k!} \quad (\text{as } n \to \infty)$$

としてよい。 □

余談だが、$(1+\frac{1}{n})^n$ が収束することも示せる。

証明.

$b_n = (1+\frac{1}{n})^n$ とおくと

$$\begin{aligned} b_n &= \sum_{k=0}^{n} 1 \cdot \left(1-\frac{1}{n}\right) \cdot \cdots \cdot \left(1-\frac{k-1}{n}\right) \frac{1^k}{k!} \\ &< \sum_{k=0}^{n} \frac{1}{k!} \\ &< 1+1+\sum_{k=2}^{n} \frac{1}{k!} \\ &< 1+1+\sum_{k=1}^{n-1} \frac{1}{2^k} \\ &< 3. \end{aligned}$$

$\{b_n\}$ は単調増大かつ上に有界な数列なので収束する。この値をネイピア数 e と定義した（収束すれば一意であることに注意)。 □

参考文献

[1] 埼玉大学理学部物理学科、物理学演習 IA 演習問題

2　多変数関数の極値と Hesse 行列

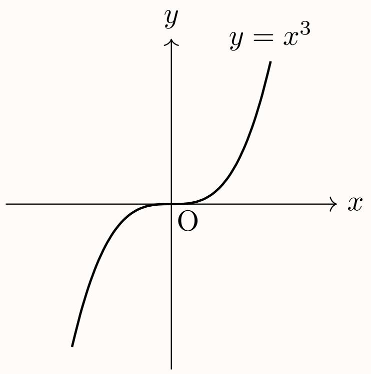
多変数関数の極値判定には Hesse 行列が有用である。しかしながら、Hesse 行列のいいたいことが明示的には分かりにくいため、理解に苦しむだろう[*1]。ここでは、2 変数関数の場合に限定して、Hesse 行列が示していることを理解したい。なお、ここでは学部 1 年前期の解析学で学習する Taylor 展開、線形代数学で学習する対角化の知識を前提とする。

2.1　1 変数のときは？

そもそも極値をとるとはどういうことだろうか？

例 2.1. [微分して **0** になるが極値をとらない例]

$f(x) = x^3$ は $f'(0) = 0$ だが、$x = 0$ で極値をとらない。



◁

微分して 0 になることは、極値をとるための必要条件にすぎない。ある点で極大／極小であるとは、その点の近傍では値が最も大きい／小さいことを意味する。

2 次関数は極大／極小の概念について直感的に良い例である。

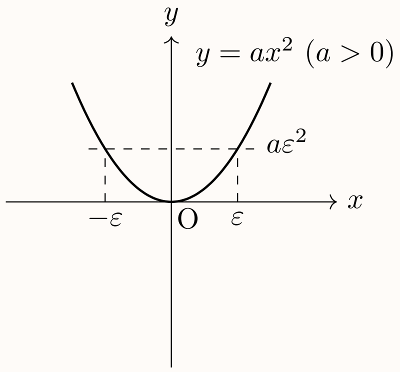
例 2.2. [$f(x) = ax^2$]

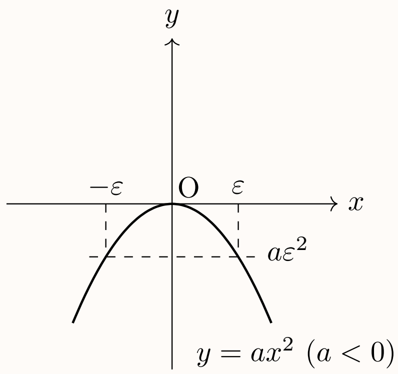
$f(x) = ax^2$ において $x \in (-\varepsilon, \varepsilon) \quad (\varepsilon > 0)$ を考える。

- $a > 0$ のとき、$f(-\varepsilon) = f(\varepsilon) = a\varepsilon^2 > f(0) = 0$ であり、$x = 0$ で極小をとる。

[*1] 学部 1 年のときは理解していなかったし、今 TA の立場として演習問題を解くにあたって、Hesse 行列が気になってこの記事を書くに至った。

- $a<0$ のとき、$f(-\varepsilon)=f(\varepsilon)=a\varepsilon^2<f(0)=0$ であり、$x=0$ で極大をとる。





微分して 0 の点の近くを考えた際に、常にそこより小さくなってしまう／大きくなってしまうという状況、もっとラフに言えば山／谷の形が極大／極小に対するもっとも簡単なイメージである。 ◁

ここでは 2 次の Taylor 展開まで考えることで極大／極小の判定を行いたい[*2]。$x=a$ 周りの展開は

$$f(a+x)\simeq f(a)+f'(a)\,x+\frac{1}{2}f''(a)\,x^2 \tag{2.1}$$

とすればよい。特に、$x=a$ で極大／極小をとるときには $f'(a)=0$ より

$$f(a+x)\simeq f(a)+\frac{1}{2}f''(a)\,x^2 \tag{2.2}$$

が成り立つ。このとき、$f''(a)<0$ ならば $f(a+x)\leq f(a)$、$f''(a)>0$ ならば $f(a+x)\geq f(a)$ がそれぞれ成り立ち、これこそが極大／極小をそれぞれ意味する。

2.2　2 変数のときは？

例 2.3.

$f_k(x,y)=x^2+kxy+y^2$ を考える。

$$\begin{cases} f_k(0,0)=0, \\ f_0(x,y)=x^2+y^2\geq 0, \quad f_{\pm 2}(x,y)=(x\pm y)^2\geq 0 \end{cases} \tag{2.3}$$

がそれぞれ成り立つ。それでは任意の k に対して $f_k(x,y)$ は $(0,0)$ で極小だろうか？ 答えは No である。

[*2] 2 次で 0 になって、4 次の依存性を考えなくてはいけないときもあるがここでは扱わない。

$\varepsilon > 0$ として $f_k(\varepsilon,\varepsilon) = (k+2)\varepsilon^2$ を考えると $k < -2$ で $f_k(\varepsilon,\varepsilon) < 0$ となり、そもそも $f_k(x,y) \geq 0$ すら保証されていない。 ◁

2 次関数（平方完成）が極大／極小の判定に役立つという感覚なので $f(x_1,x_2)$ が $(x_1,x_2) = (a_1,a_2)$ 周りで

$$f(a_1+x_1, a_2+x_2) \simeq A + Bx_1 + Cx_2 + Dx_1^2 + Ex_1x_2 + Fx_2^2 \tag{2.4}$$

という形に展開できると仮定したとき、係数 A,B,C,D,E,F がどうなるのか考えてみる。

復習 2.4. [**1 変数関数の Taylor 展開の係数決定**]

$f(x)$ が $x = a$ 周りで

$$f(a+x) \simeq A + Bx + Cx^2$$

という形に展開できると仮定する。このとき $x = 0$ を代入することで $A = f(a)$、x で微分してから $x = 0$ を代入することで $B = f'(a)$、x で 2 回微分してから $x = 0$ を代入することで $C = \frac{1}{2}f''(a)$ がそれぞれ得られる。以上より

$$f(a+x) \simeq f(a) + f'(a)x + \frac{1}{2}f''(a)x^2$$

となる。 ◁

$(x_1,x_2) = (0,0)$ を代入することで $A = f(a_1,a_2)$、x_1 で偏微分してから $(x_1,x_2) = (0,0)$ を代入することで $B = \frac{\partial f(a_1,a_2)}{\partial x_1}$、$x_2$ で偏微分してから $(x_1,x_2) = (0,0)$ を代入することで $C = \frac{\partial f(a_1,a_2)}{\partial x_2}$ がそれぞれ得られる。同様に偏微分してから $(x_1,x_2) = (0,0)$ を代入することで $D = \frac{1}{2}\frac{\partial^2 f(a_1,a_2)}{\partial x_1^2}$、$E = \frac{\partial^2 f(a_1,a_2)}{\partial x_1 \partial x_2}$、$F = \frac{1}{2}\frac{\partial^2 f(a_1,a_2)}{\partial x_2^2}$ がそれぞれ得られる。以上より

$$\begin{aligned} f(a_1+x_1, a_2+x_2) \simeq f(a_1,a_2) &+ \frac{\partial f(a_1,a_2)}{\partial x_1}x_1 + \frac{\partial f(a_1,a_2)}{\partial x_2}x_2 \\ &+ \frac{1}{2}\frac{\partial^2 f(a_1,a_2)}{\partial x_1^2}x_1^2 + \frac{\partial^2 f(a_1,a_2)}{\partial x_1 \partial x_2}x_1x_2 + \frac{1}{2}\frac{\partial^2 f(a_1,a_2)}{\partial x_2^2}x_2^2 \end{aligned} \tag{2.5}$$

となる。

ここで、$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$、$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ とおくと Hesse 行列：

$$H(\boldsymbol{a}) \equiv \begin{pmatrix} \frac{\partial^2 f(a_1,a_2)}{\partial x_1^2} & \frac{\partial^2 f(a_1,a_2)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(a_1,a_2)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(a_1,a_2)}{\partial x_2^2} \end{pmatrix} \tag{2.6}$$

を用いて

$$f(\boldsymbol{a}+\boldsymbol{x}) \simeq f(\boldsymbol{a}) + \boldsymbol{\nabla} f(\boldsymbol{a}) \cdot \boldsymbol{x} + \frac{1}{2}\boldsymbol{x}^\mathsf{T} H(\boldsymbol{a})\boldsymbol{x} \tag{2.7}$$

と表せる[*3]。特に、$\boldsymbol{a}$ で極大／極小をとるときには $\boldsymbol{\nabla} f(\boldsymbol{a}) = \boldsymbol{0}$ より

$$f(\boldsymbol{a}+\boldsymbol{x}) \simeq f(\boldsymbol{a}) + \frac{1}{2}\boldsymbol{x}^\mathsf{T} H(\boldsymbol{a})\boldsymbol{x} \tag{2.8}$$

が成り立つ。

それでは、唐突に表れた Hesse 行列は何を教えてくれるのだろうか。(2.2) と (2.8) を比較すると $\boldsymbol{x}^\mathsf{T} H(\boldsymbol{a})\boldsymbol{x}$ の正負が極大／極小を決定することが期待される。Hesse 行列は実対称行列で定義されているので対角化を考えて[*4]見通しの良いものにしたい。ここでは固有値を対角に並べた行列を $\Lambda(\boldsymbol{a})$、直交行列を P として

$$\boldsymbol{x}^\mathsf{T} H(\boldsymbol{a})\boldsymbol{x} = \boldsymbol{y}^\mathsf{T} \Lambda(\boldsymbol{a})\boldsymbol{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 \qquad \left(\boldsymbol{y} = P\boldsymbol{x} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \Lambda(\boldsymbol{a}) = PH(\boldsymbol{a})P^\mathsf{T} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\right) \tag{2.9}$$

と表すことにすると、λ_1, λ_2 の符号によって以下のように場合分けできる。

- $\lambda_1 > 0, \lambda_2 > 0$ のとき、$(y_1, y_2) \neq (0,0)$ に対して $\boldsymbol{y}^\mathsf{T} \Lambda(\boldsymbol{a})\boldsymbol{y} > 0$ となり、$\boldsymbol{a}$ で極小をとる。
- $\lambda_1 < 0, \lambda_2 < 0$ のとき、$(y_1, y_2) \neq (0,0)$ に対して $\boldsymbol{y}^\mathsf{T} \Lambda(\boldsymbol{a})\boldsymbol{y} < 0$ となり、$\boldsymbol{a}$ で極大をとる。
- $\lambda_1 \lambda_2 < 0$ のとき、$\boldsymbol{y}^\mathsf{T} \Lambda(\boldsymbol{a})\boldsymbol{y}$ は正にも負にもなりうる[*5]ので、$\boldsymbol{a}$ では極値をとらない。

いきなり $\boldsymbol{x}$ が $\boldsymbol{y}$ に置き換わっているが、直交行列による変換は見通しの良い座標で考え直していると思えば、大した違いはない。

要は、Hesse 行列の固有値の符号を調べることで極大／極小／鞍点を判定できることが Hesse 行列の存在意義であったわけだ。実際に多変数関数の極値判定条件を調べると正定値行列だとか、首座小行列だとか、色々な言葉を目にするが、それは Hesse 行列の固有値の符号判定条件を言い換えたものに他ならない。

[*3] 実はこれ以外の表し方も考えられるが、2 次の部分を実対称行列で表そうとするとこの形に限定される。

[*4] 実対称行列は直交行列で対角化可能である。

[*5] 例えば $\lambda_1 > 0$ かつ $\lambda_2 < 0$ のときを考えると $y_2 = 0$ のまま y_1 方向に動かすと $\boldsymbol{y}^\mathsf{T} \Lambda(\boldsymbol{a})\boldsymbol{y} > 0$ だが、$y_1 = 0$ のまま y_2 方向に動かすと $\boldsymbol{y}^\mathsf{T} \Lambda(\boldsymbol{a})\boldsymbol{y} < 0$ になる。このようにある変数が変化するときに関数値が増加するが、もう一方の変数が変化するときに関数値が減少する場合にはグラフの形状から鞍点（saddle point）を持つという。