微分積分・線形代数の次に学ぶ数学

川西　琢也

2025 年 4 月 21 日

目次

1　多変数関数の極大・極小

1.1　はじめに

最大・最小問題, あるいは極大・極小問題は多くの分野に現れる. 多くの産業で, 費用やエネルギーを最小にするための試みは日常的に行われているし, また, 機械学習では計算の大半が何らかの極小値を求めるのに費やされている.

1 変数の連続関数の極大値・極小値については, 高校数学で習うように, その関数の導関数を調べれば分かる. これに比べて, 多変数関数の極大・極小問題は少し難しくなる. しかし, これは, 大学の 1 年生で学んだ微分積分学と線形代数学の知識を使うことで, 見通しよく解けるのである.

本章では, 2 変数関数の極大・極小問題を例として多変数関数の極大・極小問題の解き方を学ぶ. 本章で学ぶ事項は, 広く使われている「最適化手法」の基礎をなすものである.

なお, この章では, 実行列のみを扱う. 応用上重要な極大・極小問題のほとんどは実行列で表現できると考えられるためである.

1.2　極大・極小問題とは

1.2.1　1 変数関数の極大極小

まずは, 1 変数関数の場合の復習から始めよう. 図 1.1 は, 関数

$$y = f(x) = x^3 - 3x$$

をプロットしたものである. この関数は $x = -1$ で極大値, $x = 1$ で極小値をもつ. つまり $x = -1$ と $x = 1$ はそれぞれこの関数の極大点, 極小点だが, この 2 つの点は $f'(x) = 3x^2 - 3 = 0$ の解になっており, $f''(-1) = 6x|_{x=-1} = -6 < 0$ より前者が極大値, $f''(1) = 6x|_{x=1} = 6 > 0$ より, 後者が極小値を与えることが分かる. す

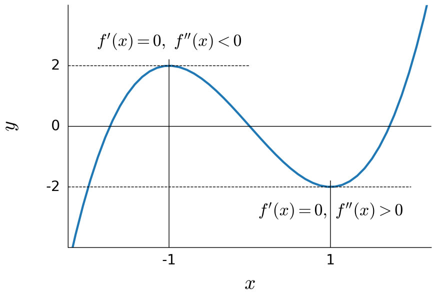


図 1.1　1 変数関数の極大値と極小値

なわち, 極値を与えるかもしれない x は, $f'(x)=0$ の解（停留点という）であり, その点で $f''(x)$ の値が正であればその点が極小点, 負であれば極大点である. もし $f''(x)=0$ ならば, 極大極小を調べるには, さらに $f^{(3)}(x)$ などを調べる必要があるが, 1 変数関数の場合, その導関数を調べることにより, 極大点・極小点を求めることができるわけである.

1.2.2　2 変数関数の場合の例

では, 2 変数関数の場合はどうだろうか. 例として,

$$z=f(x,y)=\frac{1}{4}x^2+\frac{1}{4}y^2+xy$$

を考えてみよう. この関数は, $x=0$ 平面, $y=0$ 平面上ではそれぞれ

$$z=f(0,y)=\frac{1}{4}x^2,\quad z=f(x,0)=\frac{1}{4}x^2$$

である. この 2 つの平面での値を 3D の線図に示すと 図 1.2 のようになる. つまり, $x=0$ 平面上と $y=0$ 平面上の値に限っては, この関数は $x=y=0$ で極小値を持っている. では, $(x,y)=(0,0)$ は極小点なのだろうか.

答えを先に言ってしまえば, $(x,y)=(0,0)$ はこの関数 $z=f(x,y)$ の極小値ではない. 3D グラフで見ると, $z=\frac{1}{4}x^2+\frac{1}{4}y^2+xy$ は図 1.2 のようになっている. 先ほど見たように, $x=0$ 平面, $y=0$ 平面では, 関数の値は $(x,y)=0$ で極小となっている. しかし, $y=-x$ 平面では, 図から分かるように, 関数の値は上に凸になっており, $(x,y)=(0,0)$ で極大値となっている. よって, (x,y) はこの関数の極小点ではないのである. 因みにこの点 $(x,y)=(0,0)$ のように, ある方向に沿った場合は極小値で, ある方向に沿った場合は極大値であるような点を**鞍点 (saddle point)** という.

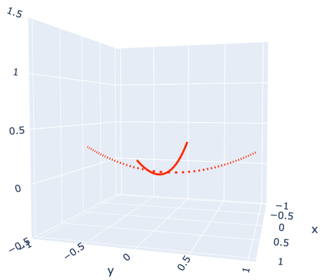


図 1.2　2 変数関数 $z = \frac{1}{4}x^2 + \frac{1}{4}y^2 + xy$ の $x = 0$ 平面, $y = 0$ 平面上の値

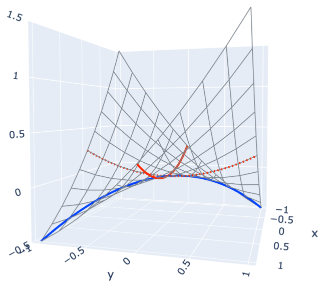


図 1.3　2 変数関数 $z = \frac{1}{4}x^2 + \frac{1}{4}y^2 + xy$: ワイヤーフレーム表示

1.2.3　多変数関数の極大・極小問題の難しさ

あらためて 2 変数の場合の難しさを考えてみよう. 先ほどの例では, z 軸に平行な面のうち, 面 $x = 0$, 面 $y = 0$ では, $z = f(x, y)$ のグラフは $(x, y) = (0, 0)$ で極小値をとるが, 面 $y = -x$ では $(x, y) = (0, 0)$ で極大値をとった.

関数 $z = f(x, y)$ が $(x, y) = (0, 0)$ で極小値をとるためには, $x = 0$, $y = 0$, $y = -x$

を含む, z 軸を含むあらゆる面で（方向としては, 360°（180°？）あらゆる方向で）その面の中の関数のグラフが極小値をとらなければならない．これが, 1 変数の場合との違いである．いわば, 2 変数の場合は, 問題の質が違っているのである．では, どうすれば良いのだろうか？実は, 2 変数を含め, n 変数 ($n \geq 2$ の場合の極大・極小問題は, これから述べる方法によって, 全て同じ手順で解けるのである．次節では, まず, この方法を説明し, そのあとに, なぜその方法で解けるのかを解説しよう．

1.3　2 変数関数の極大・極小問題

1.3.1　極点, 極値を求める手順

手順 1.1 (2 変数関数 $z = f(x, y)$ の極点, 極値を求める手順)**.**　2 変数関数 $f : \mathbb{R}^2 \to \mathbb{R}$ は 2 回連続微分可能であると仮定する[1)]．

1. $z = f(x, y)$ の**停留点 (stationary point)** をすべて求める．
2. 各停留点において, **Hesse 行列 (Hessian matrix)** を求める．
3. 上で求めた Hesse 行列の固有値を求める．
4. Hesse 行列の固有値がすべて正ならば, その点は極小値, すべて負ならば, その点は極大値, もし正と負であれば, その点は鞍点である．固有値が 0 を含む場合, さらに高階の導関数を調べる必要がある．

1.3.2　2 変数関数の Taylor 展開と極小点

2 変数あるいはそれ以上の多変数関数の極大・極小問題では, Hesse 行列（とその固有値）を用いる．これが 1 変数の場合との大きな違いとなる．ここでは, 極大・極小問題と Hesse 行列をつなぐために, Taylor 展開を用いる．Taylor 展開は, 大学の 1

---

1)　連続微分可能とは continuously differentiable の訳で, 関数が微分可能でかつ得られる導関数が連続であることをいう．**「何回も連続して微分できる」という意味ではない**ので間違わないように．英語で考えたほうが意味がとりやすいかもしれない．2 回連続微分可能は, 2 回微分可能で, 2 階の偏導関数が全て連続であることをいう．連続微分可能性は, まず, 領域内の点について定義され, 例えば点 (a, b) で連続微分可能, と表現される．さらに, 領域内の全ての点で連続微分可能なとき, 関数はその領域で連続微分可能であるという．読者は, 大学 1 年生の教科書で, 連続微分可能がどう扱われていたか復習してみてほしい（記述がない場合もあり得る）．

年生で学ぶが, 高次項や剰余項の取扱など面倒に思ったかもしれない. しかし, 応用では, たいてい 1 次までの近似で十分（まれに 2 次まで）であり, 剰余項はたいてい無視するので実はさほど面倒ではない. なので, 積極的に手を動かして計算をしてみて欲しい. まずは, **2 回連続微分可能な** 2 変数関数 $z = f(x, y)$ の Taylor 展開を下に示そう. 剰余項は R で示す.

(1) 1 次までの展開

$$\begin{aligned} & f(x+\Delta x, y+\Delta y) \\ & = f(x,y) + \frac{\partial f}{\partial x}(x,y)\Delta x + \frac{\partial f}{\partial y}(x,y)\Delta y + R \end{aligned}$$

(2) 2 次までの展開

$$\begin{aligned} f(x+\Delta x, y+\Delta y) &= f(x,y) + \frac{\partial f}{\partial x}(x,y)\Delta x + \frac{\partial f}{\partial y}(x,y)\Delta y \\ &+ \frac{1}{2}\left\{\frac{\partial^2 f}{\partial x^2}(x,y)(\Delta x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x,y)\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(x,y)(\Delta y)^2\right\} + R \end{aligned} \tag{1.1}$$

ここに示した展開では, $\partial^2 f/\partial x \partial y = \partial^2 f/\partial y \partial x$ を用いている. これを用いることは, 次の定理 1.2 と例題 1.1 で正当化される.

定理 1.2. 次式が成立するためには, $\partial^2 f/\partial x \partial y$ と $\partial^2 f/\partial y \partial x$ が存在し, かつ, ともに $(x, y) = (a, b)$ において連続であることが十分である.

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) = \frac{\partial^2 f}{\partial y \partial x}(a,b)$$ ┘

証明は微積分学の教科書などを参照していただきたい.

例題 1.1. 関数 $z = f(x, y)$ が点 (a, b) を含むある開領域で 2 回連続微分可能であれば, 式 (**??**) が成立することを示せ.

解答 1.1. 関数が $z = f(x, y)$ が 2 回連続微分可能ということは, f の 2 階導関数, すなわち次の 4 つの関数がすべてが存在して, かつ連続であるということである.

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial y^2}$$

したがって, 定理 1.2 より, (**??**) が成り立つ.

備考 1.3. 例題 1.1 の (a, b) を含む**開領域**で, という条件だが, もし領域が閉領域だと, その境界上の点での微分の定義が少々面倒になる. 例えば, 1 次元の閉区間の端点では, 右連続, 左連続の概念が必要になる. 開領域では境界上の点がなく, 特別なことを考えなくてすむのである.

定義 1.4. 点 (a, b) が関数 $f(x, y)$ の停留点であるとは,

$$\frac{\partial f}{\partial x}(a, b) = 0, \quad \frac{\partial f}{\partial y}(a, b) = 0$$

が成り立つことである.

$\boldsymbol{a} = [a_1, \ldots, a_n]^{\mathrm{T}}$ が n 変数関数 $f(\boldsymbol{x}) = f(x_1, \ldots, x_n)$ の停留点であるとは,

$$\frac{\partial f}{\partial x_i}(\boldsymbol{a}) = 0, \quad (i = 1, \ldots, n)$$

が（すべての i について）成り立つことである.

定理 1.5 (極小値をとるための条件)**.** 2 回連続微分可能な関数 $z = f(x, y)$ が $(x, y) = (a, b)$ で極小値をとるためには, 次の 2 つの条件が成立することが十分である.

条件 1（停留点）

$$\frac{\partial f}{\partial x}(a, b) = 0, \quad \frac{\partial f}{\partial y}(a, b) = 0$$

条件 2（正値性） $\Delta x \neq 0$ または $\Delta y \neq 0$ に対して

$$\frac{\partial^2 f}{\partial x^2}(a, b)(\Delta x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(a, b)\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(a, b)(\Delta y)^2 > 0. \text{」}$$

証明.$f(x, y)$ の 2 次までの展開である (1.1) において, $(x, y) = (a, b)$ とおく.

$$\begin{aligned} f(a + \Delta x, b + \Delta y) &= f(a, b) + \underline{\frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y} \\ &+ \underline{\underline{\frac{1}{2}\left\{\frac{\partial^2 f}{\partial x^2}(a, b)(\Delta x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(a, b)\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(a, b)(\Delta y)^2\right\}}} + R \end{aligned}$$

条件 1 により, 右辺第 2 項, 第 3 項（RHS_{23} とおく）は

$$\underline{\mathrm{RHS}_{23} = \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y = 0}$$

となる．また, 条件 2 により, 右辺第 4 項（RHS_4 とおく）は $\Delta x \neq 0$ または $\Delta y \neq 0$ のもとで

$$\underline{\underline{\mathrm{RHS}_4 = \frac{1}{2}\left\{\frac{\partial^2 f}{\partial x^2}(a,b)(\Delta x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(a,b)\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(x,y)(\Delta y)^2\right\}}} > 0.$$

さらに, Δx, Δy の絶対値が十分小さければ, 剰余項の絶対値は RHS_4 の絶対値よりも小さくなる．つまり, $\mathrm{RHS}_4 + R > 0$ である．以上より

$$\underline{\mathrm{RHS}_{23}} + \underline{\underline{\mathrm{RHS}_4}} + R > 0$$

すなわち

$$f(a+\Delta x, b+\Delta y) - f(a,b) = \underline{\mathrm{RHS}_{23}} + \underline{\underline{\mathrm{RHS}_4}} + R > 0.$$

これは, (x,y) が (a,b) からどのような方向に変化 $(\Delta x, \Delta y)$ したとしても,（$\Delta x \neq 0$ または $\Delta y \neq 0$ である限り）$f(a+\Delta x, b+\Delta y)$ の値は $f(a,b)$ の値よりも大きくなること, すなわち (a,b) が極小点であることを示している． □

備考 1.6. 2 変数の場合は, 2 次関数の判別式を用いても解ける．しかし, この方法は 3 変数以上の場合には使えないのでここでは扱わない．

1.3.3　2 次形式と Hesse 行列 (Hessian matrix)

さて, 定理 1.5 の条件 2 は, Hesse 行列 H（あとで定義する）を用いると次のようになる．

任意の $\boldsymbol{x} \in \mathbb{R}^2$, $\boldsymbol{x} \neq \boldsymbol{0}$ に対して

$$\boldsymbol{x}^{\mathrm{T}} H \boldsymbol{x} > 0 \tag{1.2}$$

ここで $\boldsymbol{x}^{\mathrm{T}}$ はベクトル $\boldsymbol{x}$ の転置 (transpose) を表す[2]．式 (1.2) の左辺は $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}$ の形をしているが, 一般に, 対称行列 A に対し, $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}$ を 2 次形式（あるいは 2 次形式の行列表現）という．

2)　行列の転置を表す記号には, ${}^t A$, A', など様々なものが使われるが, 本書では A^{T} を用いる．

定義 1.7. 2 変数関数の 2 次形式とは

$$Q(x,y) = ax^2 + bxy + cy^2 \tag{1.3}$$

で, a, b, c のうち少なくともひとつは 0 でないものをいう.

備考 1.8. n 変数の 2 次形式は次のように定義できる.

$$Q(x_1, \dots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j \quad \text{少なくともひと組の } (i,j) \text{ について } a_{ij} \neq 0$$

命題 1.9. 2 次形式は対称行列により表現できる. ⌟

証明. ここでは 2 変数の場合のみについて示す（3 変数以上も同様に示すことができる）. $\boldsymbol{x} = [x, y]^{\mathrm{T}}$, 対称行列を

$$S = \begin{bmatrix} a & \frac{1}{2}b \\ \frac{1}{2}b & c \end{bmatrix}$$

とおくと

$$\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = [x, y] \begin{bmatrix} a & \frac{1}{2}b \\ \frac{1}{2}b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + bxy + cy^2$$

つまり, 任意の 2 次形式 (1.3) について, 上記の対称行列 S を取れば, $Q(x) = \boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$ とできる. □

定義 1.10 (2 変数関数の Hesse 行列, Hessian matrix)**.** 2 回偏微分可能な関数 $f(x,y)$ に対し次の行列を Hesse 行列という

$$H(x,y) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

備考 1.11. n 変数関数の Hesse 行列は

$$H(x_1, \dots, x_n) = \begin{bmatrix} f_{x_1 x_1} & \cdots & f_{x_n x_1} \\ \vdots & \ddots & \vdots \\ f_{x_1 x_n} & \cdots & f_{x_n x_n} \end{bmatrix}.$$

備考 1.12. 英語で単に Hessian といえば, Hesse 行列の行列式になる. Hesse 行列は Hessian matrix という.

命題 1.13. 関数 $f(x,y)$ が 2 回**連続微分可能**な場合, Hesse 行列は対称行列になる. ⌟

証明.2 回連続微分可能な関数では

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$$

が成立するので $H(x,y)$ は対称行列となる. □

命題 1.13 は一般の n 変数でも成立する. さて, 定理 1.5 を Hesse 行列を用いて言い換えておこう.

定理 1.14. 2 回連続微分可能な関数 $z = f(x,y)$ が $(x,y) = (a,b)$ で極小値をとるためには, 次の 2 つの条件が成立することが十分である.

条件 1（停留点）

$$\frac{\partial f}{\partial x}(a,b) = 0, \quad \frac{\partial f}{\partial y}(a,b) = 0$$

条件 2（正値性）

$(x,y) = (a,b)$ における Hesse 行列

$$H(a,b) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2}(a,b) & \dfrac{\partial^2 f}{\partial x \partial y}(a,b) \\ \dfrac{\partial^2 f}{\partial x \partial y}(a,b) & \dfrac{\partial^2 f}{\partial y^2}(a,b) \end{bmatrix}$$

について, 任意の $\boldsymbol{x} \in \mathbb{R}^2$, $\boldsymbol{x} \neq 0$ に対し

$$Q_H(\boldsymbol{x}) := \boldsymbol{x}^{\mathrm{T}} H(a,b) \boldsymbol{x} > 0.$$ ⌟

1.3.4 対称行列と転置行列の性質

さて, $f(x,y)$ が 2 回連続微分可能であれば, Hesse 行列は対称（$\partial^2/\partial x \partial y = \partial^2/\partial y \partial x$）だった. 対称行列は以下に述べるような性質を持つ. 1 年生の線形代数で学ぶ内容だが, 応用上きわめて重要なので, ぜひここで覚えてしまってほしい.

定理 1.15. 実対称行列は必ず直交行列によって対角化できる. すなわち, 任意の実対称行列 S に対しある直交行列 U が存在し, 次のようにできる.

$$J = U^{-1}SU \quad J \text{ は対角行列}$$ ⌟

定理 1.16. 実対称行列の固有値はすべて実数である. ⌟

定理 1.17. 直交行列の逆行列はその転置行列に等しい.

$$U^{\mathrm{T}} = U^{-1}$$ ⌟

定理 1.15, 1.16, 1.17 から, 次が言える.

系 1.18. 任意の実対称行列 S に対してある直交行列 U が存在し, 次のようにできる.

$$J = U^{\mathrm{T}}SU \quad J \text{ は対角行列}$$

このとき, 対角行列 J の成分は全て実数となる. ⌟

備考 1.19. 実対称行列はこのほかにも, 固有ベクトルが直交するなど, 応用上ありがたい性質をもつ. 詳しくは線形代数の教科書を見てほしい.

さて, 転置行列も応用でよく出てくる. 特に次の定理は重要である.

定理 1.20. 転置行列については, 次が成立する.

$$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}} \tag{1.4}$$

⌟

備考 1.21. 定理 1.20 は, A, B が正方行列でなくても成立する. 証明は線形代数の教科書に譲るとして, 行列とベクトルの大きさだけ確認しておこう. A が $l \times m$ 行

列, B が $m \times n$ 行列のとき AB は $l \times n$ 行列. B^{T} は $n \times m$ 行列, A^{T} は $m \times l$ 行列だから, $B^{\mathrm{T}}A^{\mathrm{T}} = (AB)^{\mathrm{T}}$ は $n \times l$ 行列.

n 行の列ベクトルは $n \times 1$ 行列とみなせるから, 式 (1.4) の特別な場合として, 行列とベクトルの積の転置について次が成立する.

系 1.22. 行列 A とベクトル b の積の転置については次が成立する.

$$(A\boldsymbol{b})^{\mathrm{T}} = \boldsymbol{b}^{\mathrm{T}}A^{\mathrm{T}} \tag{1.5}$$

⌟

例題 1.2. 次の場合に, 式 (1.5) を確かめよ.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} x \\ y \end{bmatrix}$$

解答 1.2.

$$\begin{aligned} A\boldsymbol{b} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax+by \\ cx+dy \end{bmatrix} \\ (A\boldsymbol{b})^{\mathrm{T}} &= \begin{bmatrix} ax+by, & cx+dy \end{bmatrix} \end{aligned}$$

一方

$$\begin{aligned} \boldsymbol{b}^{\mathrm{T}}A^{\mathrm{T}} &= \begin{bmatrix} x, & y \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} xa+yb, & xc+yd \end{bmatrix} \\ &= \begin{bmatrix} ax+by, & cx+dy \end{bmatrix} \end{aligned}$$

となり, $(A\boldsymbol{b})^{\mathrm{T}} = \boldsymbol{b}^{\mathrm{T}}A^{\mathrm{T}}$ であることが分かる.

1.3.5　対称行列の対角化と 2 次形式の標準形

実対称行列は直交行列を用いて系 1.18 のように対角化できるが, これを用いて 2 次形式を標準形に変換できる. この変形は, 極大・極小問題にとって極めて重要である. まずは 2 次形式を対角化行列を用いて表すことから始めよう。

例題 1.3. 実対称行列 S が直交行列 U により $J = U^{\mathrm{T}} S U$ のように対角化できるとき, 2 次形式

$$Q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$$

を対角化行列 J を用いて表せ.

解答 1.3 式 $J = U^{\mathrm{T}} S U$ の両辺に左から U, 右から U^{T} を作用させると, $UU^{\mathrm{T}} = U^{\mathrm{T}} U = E$（$E$ は単位行列）だから

$$UJU^{\mathrm{T}} = U\left(U^{\mathrm{T}} S U\right) U^{\mathrm{T}} = S$$

これを $Q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$ に代入して

$$Q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} UJU^{\mathrm{T}} \boldsymbol{x} \tag{1.6}$$

解答はここまでとして, さらに変形を続ける. $\boldsymbol{y} = U^{\mathrm{T}} \boldsymbol{x}$ とおくと系 1.22 より

$$\boldsymbol{y}^{\mathrm{T}} = \left(U^{\mathrm{T}} \boldsymbol{x}\right)^{\mathrm{T}} = \boldsymbol{x}^{\mathrm{T}} U$$

これを用いて式 (1.6) を書き換えると, 次のようになる.

$$Q(\boldsymbol{x}) = \boldsymbol{y}^{\mathrm{T}} J \boldsymbol{y}$$

2 変数の場合で考えよう.

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad J = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

とおくと,

$$Q(\boldsymbol{x}) = \boldsymbol{y}^{\mathrm{T}} J \boldsymbol{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \lambda_1 y_1^2 + \lambda_2 y_2^2 \tag{1.7}$$

式 (1.7) の形にした 2 次形式を（2 次形式の）**標準形 (canonical form)** という. 標準形の y_1^2, y_2^2 の係数は, 2 次形式の表現行列の固有値となる. この標準形から分か

るように, $\lambda_1 > 0$ かつ $\lambda_2 > 0$ のとき, $\boldsymbol{y} \neq \mathbf{0}$ であれば $Q(\boldsymbol{x}) > 0$ が成立する. さらに, U は直交行列なので $\|\boldsymbol{y}\| = \left\|U^{\mathrm{T}}\boldsymbol{x}\right\| = \|\boldsymbol{x}\|$ だから[3), $\boldsymbol{y} \neq \mathbf{0} \Leftrightarrow \boldsymbol{x} \neq \mathbf{0}$ となり, 結局, 次の定理が成り立つことになる.

定理 1.23. 実対称行列 S の固有値が全て正であれば, $\boldsymbol{x} \neq \mathbf{0}$ なる任意の $\boldsymbol{x}$ に対して次が成立する.

$$Q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} > 0$$

┘

備考 1.24. 定理 1.23 のように固有値が全て正の行列を正定値行列 (positive definite matrix) という. せっかくなので, 関連の用語をまとめておく.

定義 1.25 (正定値, 半正定値, 負定値, 半負定値)**.**

1. $n \times n$ 実対称行列 A が**正定値 (positive definite)** であるとは, $\boldsymbol{x} \neq \mathbf{0}$ なる任意の n 次元ベクトル $\boldsymbol{x}$ に対して　$\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} > 0$ が成立することである.
2. $n \times n$ 実対称行列 A が**負定値 (negative definite)** であるとは, $\boldsymbol{x} \neq \mathbf{0}$ なる任意の n 次元ベクトル $\boldsymbol{x}$ に対して　$\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} < 0$ が成立することである.
3. $n \times n$ 実対称行列 A が**半正定値 (positive semidefinite)** あるいは**非負定値 (nonnegative definite)** であるとは, $\boldsymbol{x} \neq \mathbf{0}$ なる任意の n 次元ベクトル $\boldsymbol{x}$ に対して　$\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} \geq 0$ が成立することである.
4. $n \times n$ 実対称行列 A が**半負定値 (negative semidefinite)** あるいは**非正定値 (nonpositive definite)** であるとは, $\boldsymbol{x} \neq \mathbf{0}$ なる任意の n 次元ベクトル $\boldsymbol{x}$ に対して　$\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} \leq 0$ が成立することである.

1.3.6　2 変数関数の極大点・極小点のまとめ

以上の議論をまとめよう. 2 変数関数 $f(x, y)$ が 2 回連続微分可能であるとき, その Hesse 行列 $H(x, y)$ は（すべての (x, y) において）対称行列となる. さて, $(x, y) = (a, b)$ が 2 回連続微分可能な関数 $f(x, y)$ の停留点であるとする. この時, $H(a, b)$（対称行列である）の固有値が全て正であれば, 定理 1.23 により, $\boldsymbol{x} \neq \mathbf{0}$ なる $\boldsymbol{x}$ について $Q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} H(a, b) \boldsymbol{x} > 0$ が成立する, 従って (a, b) は極小点となる. これを定理にまとめよう.

3）　直交行列による変換＝直交変換は, ベクトルの大きさを変えない

定理 1.26. 2 回連続微分可能な関数 $z = f(x, y)$ が $(x, y) = (a, b)$ で極小値をとるためには, 次の 2 つの条件が成立することが十分である.

条件 1（停留点）

$$\frac{\partial f}{\partial x}(a, b) = 0, \quad \frac{\partial f}{\partial y}(a, b) = 0$$

条件 2（正値性）点 (a, b) における Hesse 行列の固有値が全て正である. 」

備考 1.27. 極大点についても全く同様の議論ができる. $z = f(x, y)$ が点 (a, b) で極大値をとるためには, (a, b) が停留点で, かつ, (a, b) における Hesse 行列 $H(a, b)$ の固有値が全て負であれば十分である. 停留点の Hesse 行列の固有値が正と負ならばその点は鞍点となる. また, もし Hesse 行列の固有値が 0 を含むならば, 極値の判定のためには, さらに高次の導関数を調べる必要がある.

1.4 3 変数以上への拡張

前節までは, 基本的に 2 変数の場合のみを扱った. しかし, 第 1.3.4 節の対称行列と転置行列に関する部分は全ての次元の行列に対して言えるものであるし, Hesse 行列も 2 次形式も 3 変数以上の関数に対しても（備考に示したように）定義でき, 2 次形式の標準化とその表現行列である対称行列の対角化が対応していることは 3 変数以上でも同じである. つまり, 3 変数以上でも停留点における Hesse 行列の固有値を調べれば, その停留点が極大点か極小点か鞍点かいずれでもないか判定できる.

この章のまとめとして, n 変数の場合の Hesse 行列と極大・極小について記す.

定理 1.28. 2 回連続微分可能な実関数 $z = f(\boldsymbol{x})$, $\boldsymbol{x} = [x_1, \ldots, x_n]^{\mathrm{T}}$ の停留点 $\boldsymbol{a} = [a_1, \ldots, a_n]$

$$\frac{\partial f}{\partial x_1}(\boldsymbol{a}) = \cdots = \frac{\partial f}{\partial x_n}(\boldsymbol{a}) = 0$$

における Hesse 行列

$$H(\boldsymbol{a}) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(\boldsymbol{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\boldsymbol{a}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\boldsymbol{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(\boldsymbol{a}) \end{bmatrix}$$

の固有値が (1) 全て正の実数であればこの停留点は極小値であり, (2) 全て負の実数であればこの点は極大値である. また, 固有値が (3) 正と負の数を含み 0 を含まない場合この停留点は鞍点であり, (4) 固有値が 0 を含む場合, 極値判定のためには, さらに高次の微分係数を調べる必要がある. ⌟

1.5 Python による固有値の計算例

例 1.29. 関数

$$z = f(x, y) = \cos(\pi x)\sin(\pi y)$$

に対して, $[1, 1/2]^{\mathrm{T}}$ が 1) 停留点である（すなわち 1 階の導関数が 0 となる）ことを確認し, 2) この点における Hesse 行列の固有値からこの停留点が極大点か極小点かどちらでもないかを判定する.

```
import numpy as np


def func(x):
    return np.cos(np.pi * x[0]) * np.sin(np.pi * x[1])


def grad(x):
    return np.array([
        - np.pi * np.sin(np.pi * x[0]) * np.sin(np.pi * x[1]),
        np.pi * np.cos(np.pi * x[0]) * np.cos(np.pi * x[1])
    ])


def hess(x):
    return np.array([
        [- np.pi ** 2 * np.cos(np.pi * x[0]) * np.sin(np.pi * x[1]),
          - np.pi ** 2 * np.sin(np.pi * x[0]) * np.cos(np.pi * x[1])],
        [- np.pi ** 2 * np.sin(np.pi * x[0]) * np.cos(np.pi * x[1]),
          - np.pi ** 2 * np.cos(np.pi * x[0]) * np.sin(np.pi * x[1])]
    ])


def main():
    x = np.array([1, 0.5])
    g = grad(x)
    h = hess(x)
```

```
    eigvals = np.linalg.eigvals(h)
    print("x = ", x)
    print("gradient of f at x = ")
    print(g)
    print("Eigenvalues of Hessian matrix at x = ")
    print(eigvals)


if __name__ == '__main__':
    main()
```

出力は以下のとおり.

```
x =  [1.   0.5]
gradient of f at x =
[-3.84734139e-16 -1.92367069e-16]
Eigenvalues of Hessian matrix at x =
[9.8696044 9.8696044]
```

1) $\boldsymbol{x} = [1, 1/2]^{\mathrm{T}}$ における勾配（gradient）の値は

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

であるから, この点は停留点である. 2) $\boldsymbol{x} = [1, 1/2]^{\mathrm{T}}$ における Hesse 行列の値は, ともに 9.869 で, 正である. よってこの点 $\boldsymbol{x} = [1, 1/2]^{\mathrm{T}}$ は極小点である.

1.6　学習の参考に

線形代数を復習する場合, 応用上で出てきたトピックスに対して, 関連する部分を復習するのがよい. 大学 1 年生のときには, 線形代数を学ぶ意味など全く分からなかっただろうが, 応用上では, 行列式, 固有値・固有ベクトル, 2 次形式, 線形空間と, 1 年生で学ぶ線形代数の全ての分野が「役にたつ」ことが分かるだろう. 1 年生で使った教科書が 2 次形式までしっかりと解説してあればよいが, そうでない場合は, もう少し高度なテキストが必要であろう. 線形代数については, 優れた教科書は山ほどあるが, 筆者が使ったものの中で特に印象に残ったものとして, 以下を挙げておく.

1. 児玉慎三, 須田信英, システム制御のためのマトリクス理論, コロナ社, 1978.
   https://www.coronasha.co.jp/np/isbn/9784339083309/
2. 笠原皓司, 新装版改訂増補, 線形代数と固有値問題–スペクトル問題を中心に,

現代数学社, 2019. https://www.gensu.co.jp/book_print.cgi?isbn=978-4-7687-0519-3

洋書では,

3. Strang, G, Introduction to linear algebra, 5th Ed., Cambridge University Press, 2016.

邦訳も出ている. この本の著者のビデオ講義は世界的に人気である. 通読するのもよいと思う.

1.7 演習問題

問題 1.1. 次の関数 $f(x)$ の (a) 1, 2, 3 階の導関数 $f'(x)$, $f''(x)$, $f^{(3)}(x)$ を求め, (b) それぞれの $f(0)$, $f'(0)$, $f''(0)$, $f^{(3)}(0)$ の値を計算し, (c) 3 次の項までの Taylor 展開を求めよ. ただし, ここで log は自然対数である.

例: $f(x)=x^3$ の場合,

$$
\begin{aligned}
&\text{(a) } f'(x)=3x^2, \quad f''(x)=6x, \quad f^{(3)}(x)=6\\
&\text{(b) } f(0)=0, \quad f'(0)=0, \quad f''(0)=0, \quad f^{(3)}(0)=6\\
&\text{(c) } f(x)=0+0x+\frac{1}{2!}x^2+\frac{6}{3!}x^3+R=x^3+R
\end{aligned}
$$

(1) $f(x)=e^x$, (2) $\log(1+x)$, (3) $\sin x$, (4) $\cos x$ (5) $\sinh x$, (6) $\tan x$

問題 1.2. 次の関数 f の 2 階導関数 (a) $\partial^2 f/\partial x^2$, (b) $\partial^2 f/\partial x\partial y$, (c) $\partial^2 f/\partial y^2$ を求めよ.

(1) $f(x,y)=x^2+y^2+2x^2y^2$, (2) $f(x,y)=x^2+2x+y^2-1$, (3) $f(x,y)=\sin x\cos y$,
(4) $f(x,y)=\sin(xy)$, (5) $f(x,y)=\tan(x+y)$, (6) $f(x,y)=e^{x+y}$
(7) $f(x,y)=e^{-\frac{x^2+y^2}{2}}$, (8) $f(x,y)=\log(xy)$, $\quad (x>0, y>0)$

問題 1.3. (1) $f(x)=\cos x\cos y$ の, $(x,y)=(0,0)$ における次の導関数の値を求めよ.

(a) $\frac{\partial f}{\partial x}(0,0)$, (b) $\frac{\partial f}{\partial y}(0,0)$, (c) $\frac{\partial^2 f}{\partial x^2}(0,0)$, (d) $\frac{\partial^2 f}{\partial x\partial y}(0,0)$, (e) $\frac{\partial^2 f}{\partial y^2}(0,0)$

(2) $f(x)=e^{-\frac{x^2+y^2}{2}}$ の, $(x,y)=(0,0)$ における次の導関数の値を求めよ. (a) $\frac{\partial f}{\partial x}(0,0)$, (b) $\frac{\partial f}{\partial y}(0,0)$, (c) $\frac{\partial^2 f}{\partial x^2}(0,0)$, (d) $\frac{\partial^2 f}{\partial x\partial y}(0,0)$, (e) $\frac{\partial^2 f}{\partial y^2}(0,0)$

問題 1.4. $(x, y) = (0, 0)$ は (a) 次の関数の停留点であるか否か, (b) 停留点の場合, 極大点か極小点か鞍点かそれ以外かを答えよ.

(1) $f(x, y) = x^2 + y^2 + 2x^2y^2$, (2) $f(x, y) = x^2 + 2x + y^2 - 1$

(3) $f(x, y) = \sin x \cos y$, (4) $f(x, y) = \sin(xy)$

(5) $f(x, y) = \tan(x + y)$, (6) $f(x, y) = e^{x+y}$

(7) $f(x, y) = e^{-\frac{x^2+y^2}{2}}$, (8) $f(x, y) = \log(xy), \quad (x > 0, y > 0)$

問題 1.5. 次の各行列に対して二次形式 $Q(\boldsymbol{x})$ を求めよ. ただし $\boldsymbol{x} = [x, y]^{\mathrm{T}}$ とする.

$$(1)\ A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (2)\ B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad (3)\ C = \begin{bmatrix} \pi/2 & -\pi/2 \\ \pi/2 & \pi/2 \end{bmatrix},$$

$$(4)\ D = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

問題 1.6. 次のようなベクトルと行列を考える.

$$\boldsymbol{a} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad A = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

これらのベクトル, 行列について次の計算をせよ. (1) $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}$, (2) $\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}$, (3) AG, (4) GA, (5) AA^{T}.

問題 1.7. 次の行列の (1) 固有値を求め, (2) 正定値行列か, 負定値行列か, どちらでもないか答えよ.

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} -3 & 1 \\ 1 & -3 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$$

問題 1.8. 次の対称行列が正定値か負定値か, どちらでもないか示せ.

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 1 \\ 1 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 3 \\ 3 & -3 \end{bmatrix}, \quad D = \begin{bmatrix} -2 & -1 \\ -1 & 2 \end{bmatrix}$$

問題 1.9. (1) 次の関数について, $(x, y) = (1, 1)$ が停留点であることを確かめ, (2) これが極小点か極大点か鞍点かいずれでもないか答えよ.

$$f(x, y) = \log(xy) - \frac{1}{3}\left(x^3 + y^3\right)$$

問題 1.10. (1) 次の関数について, $(x, y, z) = (0, 0, 0)$ が停留点であることを確かめよ. (2) また, $(x, y, z) = (0, 0, 0)$ が極小点か極大点か鞍点かいずれでもないか答えよ.

固有値の計算には Python の Numpy, Wolfram Alpha のツールを用いよ.

$$f(x, y, z) = 2x^2 + 2y^2 + 2z^2 - xy - yz - zx$$

問題 1.11. 次の関数の (1) 停留点を求め, (2) その停留点が極小点か極大点か鞍点かいずれでもないか答えよ.

$$f(x, y) = (x + y - 3)^2 + 2(x - 1)^2 + (y - 1)^2$$

2　2×2 行列は何をやっているのか

2.1　この章の目的

線形代数をより深く理解するために, 実 2×2 行列が図形にどのように作用をするのかを調べる.

2.2　固有値がすべて実数で行列が対角化できる場合

2.2.1　対角化できる行列の作用

命題 2.1 (対角化できる行列の作用)**.** 実 2×2 行列が対角化できて固有値がすべて実数の場合, 行列は, 図形 (ベクトル) を**固有ベクトルの方向に**（その固有ベクトルに対応する）**固有値倍**する. 」

2.2.2　対角行列

まずは対角行列から始める.

例題 2.1. 次の行列 A の固有値と固有ベクトルを求めよ.

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \tag{2.1}$$

解答 2.1 A はすでに対角行列なので, 特に変換などの操作は不要である. A の固有値はその対角成分 $\lambda_1 = 3$, $\lambda_2 = 2$, であり, これに対応する固有ベクトルはそれぞれ x 方向, y 方向のベクトルである. 基準化すると以下のとおりである.

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

解答は以上として, 例えば, ベクトル $\boldsymbol{a} = [2,1]^{\mathrm{T}}$ に上記の行列 A を作用させると次のようになる.

$$\boldsymbol{b} = A\boldsymbol{a} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

この例では, ベクトル $\boldsymbol{b}$ は, 固有ベクトル $\boldsymbol{e}_1 = [1,0]^{\mathrm{T}}$ の方向に固有値 $\lambda_1 = 3$ 倍,

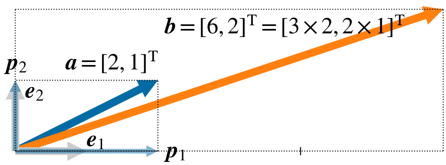


図 2.1　対角行列 A （式 (2.1)）のベクトル $\boldsymbol{a} = [2,1]^{\mathrm{T}}$ への作用

固有ベクトル $\boldsymbol{e}_2 = [0,1]^{\mathrm{T}}$ の方向に固有値 $\lambda_2 = 2$ 倍されている.

2.2.3　一般の対角化可能行列

例 2.2 (対角化可能行列で固有値が 2 つの実数の場合)**.** 例えば次の行列 B について考える.

$$B = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \tag{2.2}$$

この行列の固有値は $\lambda_1 = -1$, $\lambda_2 = -3$ であり, ともに実数である. λ_1, λ_2 に対応する固有ベクトル（基準化せず）は例えば $\boldsymbol{p}_1 = [1,1]^{\mathrm{T}}$ と $\boldsymbol{p}_2 = [1,-1]^{\mathrm{T}}$ である. よって, 行列 B は $P = [\boldsymbol{p}_1, \boldsymbol{p}_2]^{\mathrm{T}}$ を使って対角化できる.

$$\begin{aligned} J_B = P^{-1}BP &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix} \end{aligned}$$

例えば, ベクトル $\boldsymbol{a} = [3, 0]^{\mathbb{T}}$ に B が作用する場合を考えよう. $\boldsymbol{b} = B\boldsymbol{a}$ とする. 行列 B の作用を固有ベクトルの基底で考える. 基底変換の概念を図示すると次の図 2.2 のようになる. すなわち, ベクトル $\boldsymbol{a}$ に行列 B を作用させることは, ベクトル $\boldsymbol{a}$ に, P^{-1}, J_B, P を順次作用させることと同じであり, $B\boldsymbol{a} = PJ_BP^{-1}\boldsymbol{a}$ と書ける. ここで, $\boldsymbol{\xi}_{\boldsymbol{a}} = P^{-1}\boldsymbol{a}$ は, ベクトル a を元の基底（座標系）から固有ベクトルの基底に写したものであり, その $\boldsymbol{\xi}_{\boldsymbol{a}}$ に固有ベクトル基底でのこの行列の作用を表す対角行列 J_B を作用させる, すなわち, 固有ベクトルの座標でそれぞれの成分を対応する固有値倍する作用となる. 最後に P は固有ベクトルの基底から元の基底に戻す操作を担っている.

$$
\begin{tikzcd}
\boldsymbol{\xi}_{\boldsymbol{a}} \arrow[r, "J_B"] & \boldsymbol{\xi}_{\boldsymbol{b}} \arrow[d, "P"] \\
\boldsymbol{a} \arrow[u, "P^{-1}"] \arrow[r, "B"] & \boldsymbol{b}
\end{tikzcd}
$$

図 2.2　基底変換

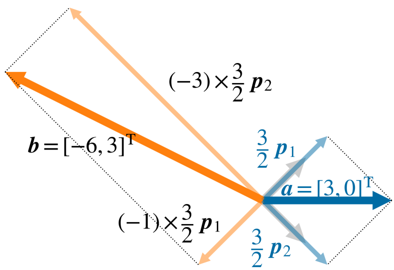


図 2.3　対角化可能で固有値がすべて実数の行列 B（式 (2.2)）のベクトル $\boldsymbol{a} = [1, 2]^{\mathrm{T}}$ への作用

また, $\boldsymbol{a}$, $\boldsymbol{b} = B\boldsymbol{a}$ と固有ベクトル, 固有値との関係を図示すると次の図 2.3 のようになる. 固有ベクトルの座標では, $\boldsymbol{a}$ は次のように表される.

$$
\boldsymbol{\xi}_{\boldsymbol{a}} = P^{-1}\boldsymbol{a} = \begin{bmatrix} 3/2 \\ 3/2 \end{bmatrix}
$$

つまり, $\boldsymbol{a}$ は固有ベクトル $\boldsymbol{p}_1 = [1, 1]^{\mathrm{T}}$ と $\boldsymbol{p}_2 = [1, -1]^{\mathrm{T}}$ の次の線型結合となる.

$$\boldsymbol{a} = \boldsymbol{\xi_a}[0]\boldsymbol{p}_1 + \boldsymbol{\xi_a}[1]\boldsymbol{p}_2 = \frac{3}{2}p_1 + \frac{3}{2}\boldsymbol{p}_2$$

これに対角行列 J_B が作用し, 固有ベクトル座標系での $\boldsymbol{b}$ の座標 $\boldsymbol{\xi_b}$ を得る.

$$\boldsymbol{\xi_b} = J_B\boldsymbol{\xi_a} = \begin{bmatrix} -3/2 \\ -9/2 \end{bmatrix}$$

J_B が作用することは, $\boldsymbol{\xi_a}$ の第 1 成分を λ_1 倍, 第 2 成分を λ_2 倍すること, すなわち, 固有ベクトルの方向に固有値倍する操作に他ならない. 元の座標系に戻すと

$$\boldsymbol{b} = P\boldsymbol{\xi_b} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$$

となる. 以上, P と P^{-1} は基底変換（座標変換）を行っており, J_B が固有ベクトルの方向への変形という役割をになっている. この合成として, 固有ベクトルの方向に（対応する）固有値倍, という結果が得られるのである.

2.3 固有値が全て実数だが縮退している場合

2.3.1 縮退行列と Jordan 細胞

2×2 行列の固有値が縮退している場合, 適当な基底変換 P, P^{-1} により 次のような 2 次の Jordan 細胞 J_2 の形になる. 2 次以上の場合の Jordan 細胞については線形代数の教科書を参照するか, 検索するかしてほしい. Jordan 細胞を対角に並べた形の行列を **Jordan 標準形**というのだった. また, 対角行列は 1 次の Jordan 細胞を対角に並べたものなので, Jordan 標準形の特別な場合である.

$$J_2 = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \tag{2.3}$$

まずは, 行列 J_2 がベクトルにどのように作用するかについて見る.

2.3.2　2 次の Jordan 細胞

例 2.3 (2 次の Jordan 細胞)**.**

$$C = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \tag{2.4}$$

この C は, 式 2.3 の J_2 で $\lambda = 2$ とおいた行列である. C がベクトルに作用すると, x 軸, y 軸の方向にそれぞれ 2 倍して, さらに x 軸の計算結果に 元の y 軸の値を加えることになる. 図に示すと図 2.4 のようになる.

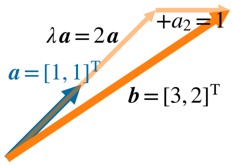


図 2.4　2 次の Jordan 細胞 C（式 (2.4)）のベクトル $\boldsymbol{a} = [1, 1]^{\mathrm{T}}$ への作用

参考までに, この行列の固有方程式は次のとおりで, 固有値は $\lambda = 2$ （重解）.

$$\det(\lambda E - A) = (\lambda - 2)(\lambda - 2) = 0$$

固有, 広義固有ベクトルはそれぞれ $\boldsymbol{p}_1 = [1, 0]^{\mathrm{T}}$, $\boldsymbol{p}_2' = [0, 1]^{\mathrm{T}}$ である.

例 2.4. 次の行列 D を考える.

$$D = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix} \tag{2.5}$$

固有値は $\lambda = 2$（重解）. 固有ベクトルは $\boldsymbol{p}_1 = [1, 1]^{\mathrm{T}}$, 広義固有ベクトルは $\boldsymbol{p}_2 = [1/2, -1/2]^{\mathrm{T}}$ よって, $P = [\boldsymbol{p}_1, \boldsymbol{p}_2]$ を使って標準化して

$$J = P^{-1}DP = \begin{bmatrix} 1/2 & 1/2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 1 & -1/2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

$\boldsymbol{a} = [1, 0]^{\mathrm{T}}$ にこの行列 D を作用させた場合を考える.

次式に計算過程を, 図 2.5 にこの $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{p}_1$, $\boldsymbol{p}_2'$ の関係を示す. $\boldsymbol{b}$ はベクトル $\boldsymbol{a}$ に固有値 $\lambda = 2$ を掛けたものに, $(P^{-1}\boldsymbol{a})_2 \times \boldsymbol{p}_1$ を足したものになっている.

$$\boldsymbol{b} = PDP^{-1}a = \begin{bmatrix} 1 & 1/2 \\ 1 & -1/2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 1/2 \\ 1 & -1/2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 1 & -1/2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

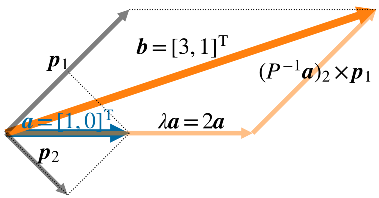


図 2.5　退化行列 D (2.5) のベクトル $\boldsymbol{a} = [1, 0]^{\mathrm{T}}$ への作用

2.4　固有値が共役複素数を含む場合（回転・回転と拡大）

2.4.1　基礎知識

命題 2.5. 実行列が虚数の固有値をもつ場合, その共役複素数も固有値となる.　⌟

命題 2.6. 実行列が共役複素数の固有値をもつ場合, 行列は図形を回転する作用を含む.　⌟

命題 2.7 (Euler の公式, Euler's formula)**.**

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{2.6}$$

⌟

2.4.2 回転を表す行列

原点まわりの角度 θ だけの回転を表す行列は,

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{2.7}$$

行列 R の固有方程式は

$$\det(R - \lambda E) = \lambda^2 - 2\cos\theta\lambda + 1 = 0$$

これより,

$$\lambda_1, \lambda_2 = \cos\theta \pm i\ \sin\theta$$

となり, 固有値が互いに共役複素数であることが分かる. また, Euler の公式 (2.6) より

$$\lambda_1, \lambda_2 = e^{\pm i\theta}$$

例 2.8.

$$F = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

この行列の固有方程式は

$$\det(F - \lambda E) = \lambda^2 - \lambda + 1 = 0$$

よって, 固有値は

$$\lambda_1 = \frac{1+\sqrt{3}i}{2} = \cos\frac{\pi}{3} + i\sin\frac{\pi}{3} = e^{i\frac{\pi}{3}},$$
$$\lambda_2 = \frac{1-\sqrt{3}i}{2} = \cos\frac{\pi}{3} - i\sin\frac{\pi}{3} = e^{-i\frac{\pi}{3}}$$

で, この行列 F は原点まわりの角度 $\pi/3$ あるいは $-\pi/3$ の回転を表すことが分かる. 実際

$$F = \begin{bmatrix} \cos\pi/3 & -\sin\pi/3 \\ \sin\pi/3 & \cos\pi/3 \end{bmatrix}$$

である. 固有値のみでは, 回転角が $+\pi/3$ か $-\pi/3$ かは分からない.

2.4.3　回転と拡大・縮小を表す行列

次の形の行列は, 回転と拡大・縮小を行う

$$G = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad a, b \in \mathbb{R}$$

なぜなら

$$\begin{aligned} G &= \sqrt{a^2+b^2} \begin{bmatrix} \frac{a}{\sqrt{a^2+b^2}} & -\frac{b}{\sqrt{a^2+b^2}} \\ \frac{b}{\sqrt{a^2+b^2}} & \frac{a}{\sqrt{a^2+b^2}} \end{bmatrix} \\ &= \sqrt{a^2+b^2} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \end{aligned} \tag{2.8}$$

ここで $\cos\theta = \frac{a}{\sqrt{a^2+b^2}}$, $\sin\theta = \frac{b}{\sqrt{a^2+b^2}}$

例 2.9.

$$H = \begin{bmatrix} 2 & -4 \\ 4 & 2 \end{bmatrix}$$

固有値は $\lambda_1 = 2+4i$, $\lambda_2 = 2-4i$. 上記の式 (2.8) に倣って書き換えると

$$H = 2\sqrt{5} \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

H は $\cos\theta = 1/\sqrt{5}$, $\sin\theta = 2/\sqrt{5}$ となる角度 θ だけ回転して大きさを $2\sqrt{5}$ 倍に拡大する作用をもつ.

2.5　固有値が共役複素数を含む場合（一般の場合・極分解）

2.5.1　半正定値行列の平方根 *

固有値に共役複素数を含む実行列が図形に作用するとき, 回転の作用を含む. このとき, もとの行列は, 回転行列と半正定値行列の積で表すことができる. このような, 回転行列と半正定値行列の積への分解は, 極分解と言われるものの一部である. 極分解は入門レベルの教科書には扱われない場合が多いので, ここで解説しておく.

極分解について学ぶ前に, 行列の平方根について知る必要がある. この項では, 半正定値行列の平方根と主平方根について学ぶ. 半正定値（実）行列とは対称行列でその固有値がすべて非負のもののことであった. 実対称行列は必ず対角化できるので, 半正定値行列 A は必ず対角化できる. A が固有ベクトルを並べた行列 P によって $J = P^{-1}AP$ と対角化できるとしよう. このとき, $A = PJP^{-1}$ だが, J は対角行列でその成分は仮定から非負なので, $K^2 = J$ となるような行列 K を考えることができる. 具体的に 2×2 行列で考えると $a \geq 0$, $b \geq 0$ として,

$$J = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad K = \begin{bmatrix} \sqrt{a} & 0 \\ 0 & \sqrt{b} \end{bmatrix}$$

が $K^2 = J$ をみたしている. ここで $B = PKP^{-1}$ なる行列を考えよう. 簡単な計算で

$$\begin{aligned} B^2 &= \left(PKP^{-1}\right)^2 = \left(PK\underbrace{P^{-1}\right)\left(P}_{=E}KP^{-1}\right) \\ &= PK^2P^{-1} = PJP^{-1} = A \end{aligned}$$

すなわち $B = PKP^{-1}$ が A の平方根であることが分かる. さて, あきらかに, 次の行列

$$K_1 = \begin{bmatrix} \sqrt{a} & 0 \\ 0 & -\sqrt{b} \end{bmatrix}, \quad K_2 = \begin{bmatrix} -\sqrt{a} & 0 \\ 0 & \sqrt{b} \end{bmatrix}, \quad K_3 = \begin{bmatrix} -\sqrt{a} & 0 \\ 0 & -\sqrt{b} \end{bmatrix}$$

も, $K_i^2 = J$ $(i = 1, 2, 3)$ を満たしている. 実数 $a \geq 0$ の平方根が $\pm\sqrt{a}$ と符号の不定性をもつのと同じである. 実数の場合に $b^2 = a \geq 0$ のうち, $b \geq 0$ を満たすものを $\sqrt{a}$ と記したのと同じように, 行列の平方根においても, 対角化行列の成分がすべて非負であるもの（上の例では $B = PKP^{-1}$）を選び, 主平方根とよび, $\sqrt{A}$ などと記す. 平方根のうち, 対角化行列の成分がすべて非負, すなわち固有値が全て非負のものはひとつしかないから, これは一意に定まる.

$A = PK^2P^{-1}$ は, 対角行列 $J = K^2$ の左右に P と P^{-1} を作用させたものだが, これが対称だった. このことから, その主平方根 $B = PKP^{-1}$ も対角行列 K の左右に P と P^{-1} を作用させたものだから対称だと分かる. B は固有値がすべて非負の対称行列なので半正定値である. 以上のことを定理にまとめる.

定理 2.10. A を半正定値行列とすると, $B^2 = A$ となる半正定値行列 B が存在し一意に定まる. このような行列 B を A の主平方根（principal square root）とよび, $\sqrt{A}$ などと表す. ⌟

系 2.11. 1. 成分が非負（$j_i \geq 0,\ i = 1, \ldots, n$）の対角行列

$$J = \begin{bmatrix} j_1 & & & \\ & j_2 & & \\ & & \ddots & \\ & & & j_n \end{bmatrix} \tag{2.9}$$

の主平方根は

$$\sqrt{J} = \begin{bmatrix} \sqrt{j_1} & & & \\ & \sqrt{j_2} & & \\ & & \ddots & \\ & & & \sqrt{j_n} \end{bmatrix} \tag{2.10}$$

である．

2. 半正定値行列 A が $A = PJP^{-1}$ と対角化できるとき，その主平方根は $P\sqrt{J}P^{-1}$ で与えられる．

┘

備考 2.12. 表記 $\sqrt{A}$ は常に主平方根を表すわけでもないようである．文脈に注意する必要がある．

例題 2.2. 次の行列の主平方根を求めよ．

$$A = \begin{bmatrix} 19 & -15 \\ 10 & -6 \end{bmatrix}$$

解答 2.2

A の固有値は $\lambda_1 = 4,\ \lambda_2 = 9$ 対応する固有ベクトルは $\boldsymbol{p}_1 = [1, 1]^{\mathrm{T}}$, $\boldsymbol{p}_2 = [3, 2]^{\mathrm{T}}$, これより，$P = [\boldsymbol{p}_1, \boldsymbol{p}_2]$ を使って

$$\begin{aligned} A &= PJP^{-1} \\ &= \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix} \end{aligned}$$

と表すことができる．A の主平方根 B はつぎのように計算される．

$$\begin{aligned} B = P\sqrt{J}P^{-1} &= \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 5 & -3 \\ 2 & 0 \end{bmatrix} \end{aligned}$$

2.5.2 極分解 *

定理 2.13 (極分解（右極分解）)**.** 任意の実正方行列 A は, 直交行列 U と半正定値対称行列 S によって

$$A = US$$

と分解できる．この分解は一意である． ⌟

証明の前に補題をふたつ記しておく．

補題 2.14. 任意の実行列 A について（正方行列以外でもよい）$A^{\mathrm{T}}A$ は対称行列である． ⌟

証明.（補題 2.14）

$$(A^{\mathrm{T}}A)^{\mathrm{T}} = A^{\mathrm{T}}\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A^{\mathrm{T}}A.$$

□

補題 2.15. 任意の実行列 A に対して, $A^{\mathrm{T}}A$ は半正定値である． ⌟

証明.（補題 2.15）任意の実ベクトル $\boldsymbol{x}$ に対し

$$\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x} = (A\boldsymbol{x})^{\mathrm{T}}(A\boldsymbol{x}) \geq 0.$$

最後の不等号であるが, $A\boldsymbol{x}$ は実ベクトルで, $(A\boldsymbol{x})^{\mathrm{T}}(A\boldsymbol{x})$ は, $A\boldsymbol{x}$ の各要素を 2 乗したものの和なので, 非負である． □

極分解の証明の考え方は次のとおり．行列 A が直交行列 U と半正定値行列 S に

よって $A = US$ と分解できたとすると

$$\begin{aligned} A^{\mathrm{T}}A &= (US)^{\mathrm{T}}(US) = S^{\mathrm{T}}U^{\mathrm{T}}US \\ &= S(U^{\mathrm{T}}U)S = S^2. \end{aligned}$$

ここで, 直交行列の次の性質を使っている $U^{\mathrm{T}}U = U^{-1}U = E$. S は半正定値行列と仮定したので対称だから $S^{\mathrm{T}} = S$. さて, この S は $A^{\mathrm{T}}A$ の平方根であるが, S は半正定値行列, という条件があるので, 主平方根として一意に定まる. では, 証明に移るが, ここでは簡単のため, 正則行列についてのみ証明を記す. 一般の場合は逆行列の存在が保証されないので, 少々面倒になる.

証明.（定理 2.13, A が正則行列の場合）$A^{\mathrm{T}}A$ は, 補題 2.15 により半正定値行列となるので, その主平方根 （principal square root）$S = \sqrt{A^{\mathrm{T}}A}$ は半正定値行列として一意に定まる. A は正則と仮定したから $A^{\mathrm{T}}A$ も正則, したがって S も正則. よって S^{-1} が存在する. $U = AS^{-1}$ とおく. この U については,

$$U^{\mathrm{T}}U = (AS^{-1})^{\mathrm{T}}AS^{-1} = S^{-1}A^{\mathrm{T}}AS^{-1} = S^{-1}S^2S^{-1} = E$$

となり, U が直交行列であることが分かる. □

2.5.3　固有値が共役複素数の正則行列・回転と半正定値行列への分解 *

定理 2.16. 2×2 実正方行列 A が, 共役複素数固有値をもつ場合, その行列の作用は, 回転を表す行列 R と半正定値行列 S に $A = RS$ と分解できる. ⌟

証明.任意の実行列は, 直交行列と半正定値行列の積に分解できる（極分解）が, 2×2 実正方行列が共役複素数固有値をもつ場合, この行列には回転の作用が含まれる. 2×2 直交行列は, 原点まわりの回転か, 原点を通る直線に対する線対称移動を表すが, 元の行列の固有値が共役複素数の場合, 直交行列は回転を表す行列になる. □

備考 2.17. 2×2 直交行列の固有値が共役複素数の場合, その直交行列が回転行列であることの証明は省略する. 気になる人は, 回転の行列, 原点を通る直線に対する線対称移動の行列を作って, 固有値を計算してみるとよい.

例題 2.3. 次の行列 M を回転の行列 R と 半正定値行列 S の積に $M = RS$ と極分解せよ.

$$M = \begin{bmatrix} 0.9 & -0.5 \\ 1.3 & 1.5 \end{bmatrix}$$

解答 2.3

$$M^{\mathrm{T}}M = \begin{bmatrix} 2.5 & 1.5 \\ 1.5 & 2.5 \end{bmatrix}$$

$M^{\mathrm{T}}M$ は実対称行列なので対角化できる. $M^{\mathrm{T}}M$ の固有値は $\lambda = 4, 1$ で, 対応する固有ベクトルはそれぞれ例えば, $\boldsymbol{p}_1 = [1, 1]^{\mathrm{T}}$, $\boldsymbol{p}_2 = [1, -1]^{\mathrm{T}}$. これを並べた行列を $P = [\boldsymbol{p}_1, \boldsymbol{p}_2]$ として

$$M^{\mathrm{T}}M = PJP^{-1}, \quad J = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$

$M^{\mathrm{T}}M$ の行列の平方根は, これより $S = P\sqrt{J}P^{-1}$ だから

$$S = \sqrt{M^{\mathrm{T}}M} = P \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix}$$

$$U = MS^{-1} = \begin{bmatrix} -0.9 & 0.5 \\ -1.3 & 1.5 \end{bmatrix} \begin{bmatrix} 3/4 & -1/4 \\ -1/4 & 3/4 \end{bmatrix} = \begin{bmatrix} 4/5 & -3/5 \\ 3/5 & 4/5 \end{bmatrix}$$

これは $\cos\theta = 4/5$, $\sin\theta = 3/5$ なる角度 θ だけの回転を表す行列である. よって

$$M = US = \begin{bmatrix} 4/5 & -3/5 \\ 3/5 & 4/5 \end{bmatrix} \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix}$$

と, 回転を表す行列 U と半正定値行列 S に分解できた.

備考 2.18. 計算手順から, この分解は, S として $\pm\sqrt{J}$ がとれることを除けば一意であるから, S が半正定値であるとの条件から $S = \sqrt{J}$ と一意に定まる.

2.6 演習問題

問題 2.1. 次の行列が図形に作用した時, (a) 各固有ベクトルの方向にそれぞれの固有ベクトル倍, (b) 回転, (c) 回転と拡大・縮小 (d) いずれでもない, のいずれの結果となるか示せ.

$$A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad a \neq 0,\ b \neq 0,$$
$$B = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 4 & -1 \\ -2 & 5 \end{bmatrix},$$
$$D = \begin{bmatrix} 4 & 1 \\ -2 & 5 \end{bmatrix}, \quad G = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$$

問題 2.2. 次の行列 A を考える.

$$A = \begin{bmatrix} \frac{11}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{9}{4} \end{bmatrix}$$

この行列 A は対称行列であるから, 対角化できる. この行列の固有値は $\lambda_1 = 3$, $\lambda_2 = 2$ で, それぞれに対応する固有ベクトルは (正規化して) 次の通りである.

$$\boldsymbol{p}_1 = \left[\sqrt{3}/2, 1/2\right], \quad \boldsymbol{p}_2 = \left[-1/2, \sqrt{3}/2\right]$$

次の問に答えよ.

1. 対角化行列 J を求めよ.
2. 固有ベクトルの座標で $\boldsymbol{a} = [1, 1]^{\mathrm{T}}$ を表せ. すなわち, 次を満たすような $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]^{\mathrm{T}}$ を求めよ.
$$\boldsymbol{a} = \alpha_1 \boldsymbol{p}_1 + \alpha_2 \boldsymbol{p}_2$$
3. 対角化行列を $\boldsymbol{\alpha}$ に作用させた結果のベクトル $\boldsymbol{\beta}$ を求めよ.
4. $[\boldsymbol{p}_1, \boldsymbol{p}_2]\boldsymbol{\beta}$ の結果を求めよ. ただし, $[\boldsymbol{p}_1, \boldsymbol{p}_2]$ は固有ベクトルを並べて作られる行列とする.
5. $A\boldsymbol{a}$ が 4. の結果と一致することを確かめよ.

問題 2.3. 次の行列の主平方根を求めよ.

$$A = \begin{bmatrix} -1 & 5 \\ -10 & 14 \end{bmatrix}, \quad B = \begin{bmatrix} -8 & 18 \\ -6 & 13 \end{bmatrix}$$

問題 2.4. つぎの行列 C は回転と拡大 (縮小) を表す行列である.

1. 回転角の sin と cos を求めよ. 半時計回りに回転するとする.
2. 図形の拡大を拡大倍数 (縮小の場合はこの数が 1 未満) を求めよ.

$$C = \begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}$$

問題 2.5. 次の行列を考える.

$$D = \begin{bmatrix} 4 & -1 \\ 4 & 8 \end{bmatrix}, \quad P = \begin{bmatrix} 2 & -1 \\ -4 & 0 \end{bmatrix}$$

1. $P^{-1}DP$ を計算せよ.
2. D の固有値と固有ベクトルを求めよ (1 組だけで良い. ただし, 一般化固有ベクトルが分かる人はもう 1 組みとしてそれも求められればなおよい).

問題 2.6. 次の行列を回転の行列と半正定値行列の積に右極分解せよ.

$$A = \begin{bmatrix} 5 & -2 \\ 10 & 11 \end{bmatrix}$$

3　連立線形微分方程式

3.1　この章で学ぶこと

- 連立線形微分方程式（線形微分方程式系）とその解
- 行列の指数関数
- 線形微分方程式系の解の挙動
    - 固有値が実数で対角化可能な場合
    - 固有値が実数で縮退している場合
    - 固有値が共役複素数の場合
- 力学系の入門

3.2　連立線形微分方程式

3.2.1　この章で考える連立線形微分方程式

この章では, 次のような連立微分方程式を考える. $\boldsymbol{x} = [x_1, \ldots, x_n]^{\mathrm{T}}$ を n 項列ベクトルとする[4]. $\boldsymbol{x}$ に関する 1 階の微分方程式と初期条件を次のように表す.

$$\frac{d\boldsymbol{x}}{dt} = A\boldsymbol{x}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \quad t \in [0, \infty),\ \boldsymbol{x} \in (-\infty, \infty)^n \tag{3.1}$$

ここで

$$\frac{d\boldsymbol{x}}{dt} = \begin{bmatrix} \dfrac{dx_1}{dt} \\ \vdots \\ \dfrac{dx_n}{dt} \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0 = \begin{bmatrix} x_{1,0} \\ \vdots \\ x_{n,0} \end{bmatrix}$$

4）　$\boldsymbol{x}$ は行ベクトル $[x_1, \ldots, x_n]$ の転置だから列ベクトルとなる.

ただし a_{ij} は実定数とする．この微分方程式系は, 定係数同次 1 階連立微分方程式とよばれる．$\boldsymbol{x}_0$ が与えられているので, 初期値問題である．

英語では, 連立線形微分方程式を system of linear differential equations という．しばしば, 線形微分方程式系と訳される．

3.2.2　1 変数の場合の解と連立微分方程式の解の比較

まずは, 1 変数の場合, すなわち $dx/dt = ax$, $x(0) = x_0$ の場合の解と n 変数の場合（式 (3.1)）の解を定理として提示し, その後に解説をすることにしよう．ちなみに, 1 変数の場合は変数分離法でも解けるし, 1 階線形微分方程式の解の公式を用いても解ける．解法が気になる人は微分方程式の教科書を参照して欲しい．ただ, 理工系を志す人ならば, 次の式 (3.2) を見ただけで, 解 (**??**) がすぐ出てくるようになってほしい.

定理 3.1. 1 変数 1 階線形微分方程式の初期値問題 （$x, x_0, t \in [0, \infty)$, a は定数）

$$\frac{dx}{dt} = ax, \quad x(0) = x_0 \tag{3.2}$$

の解は

$$x = e^{at}x_0$$

⌟

定理 3.2. 線形微分方程式系の初期値問題（$\boldsymbol{x}, \boldsymbol{x}_0 \in \mathbb{R}^n$, $t \in [0, \infty)$, $A \in \mathbb{R}^{n\times n}$ は実定数行列）

$$\frac{d\boldsymbol{x}}{dt} = A\boldsymbol{x}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$$

の解は

$$\boldsymbol{x} = e^{At}\boldsymbol{x}_0$$

⌟

解 (**??**) と解 (**??**) とはほとんど同じ形をしている．違いは, 指数関数が式 (**??**) では, e^{at} なのに対して, 式 (**??**) では e^{At} となっていることである．このように, 多変数になっても, 解の形が 1 変数と似たような形になることは, 線形微分方程式系の挙動にある見通しを与えることになる．しかし, そもそも e^{At} は何を表しているのだろうか.

3.3 行列の指数関数

3.3.1 行列の指数関数の定義

指数関数については, いくつか異なった定義のしかたがあるが, 級数による定義は複素数 a に対しても成り立つ. 行列の指数関数は, この定義の a を行列に置き換えたもので定義される.

定義 3.3.

$$e^a (= \exp a) = \sum_{k=0}^{\infty} \frac{a^k}{k!} \;= 1 + a + \frac{1}{2}a^2 + \frac{1}{6}a^3 + \cdots$$

定義 3.4 (行列の指数関数, matrix exponential)**.** A を $n \times n$ 行列とするとき（単位行列を E として）

$$e^A (= \exp A) = \sum_{k=0}^{\infty} \frac{A^k}{k!} \;= E + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \cdots$$

例 3.5 (対角行列の指数関数)**.** 対角行列

$$J = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix}$$

に対し e^{tJ} は

$$e^{tJ} = \begin{bmatrix} e^{td_1} & & 0 \\ & \ddots & \\ 0 & & e^{td_n} \end{bmatrix} \tag{3.3}$$

例題 3.1. 次に与える A, $\boldsymbol{x}_0$ の定める初期値問題 $\frac{d\boldsymbol{x}}{dt} = A\boldsymbol{x}$, $\boldsymbol{x}(0) = \boldsymbol{x}_0$ の解を求めよ.

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}, \quad \boldsymbol{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

解答 3.1 初期値問題の解は式 (**??**) で与えられる．A は対角行列だから, 式 (3.3) を使って

$$\boldsymbol{x}(t) = e^{tA}\boldsymbol{x}_0 = \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{2t} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} e^{-t} \\ e^{2t} \end{bmatrix}$$

3.3.2　実固有値の正方行列は標準形にできる

さて, 前項で行列の指数関数を定義し対角行列の指数関数を例として示した．対角行列の場合は行列の指数関数の計算も, それを用いた解の挙動の解析も簡単である．一般の行列も対角化できるならば基底変換（座標変換）により指数行列の部分は簡単になる．まずは線形代数から次の定理を引こう．

定理 3.6. $n \times n$ 実数行列 A の固有値がすべて実数の場合, この行列はある行列 P によって, 必ず Jordan 標準形 $J_{\mathrm{c}} = P^{-1}AP$ に変形できる． ⌟

備考 3.7. Jordan 標準形は, Jordan 細胞を対角に並べたものである．例として, 1, 2, 3 次の Jordan 細胞 J_1, J_2, J_3 を次に示そう．

$$J_1 = \begin{bmatrix} \lambda \end{bmatrix}, \quad J_2 = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad J_3 = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

このように, r 次の Jordan 細胞は, 対角に r 個の固有値 λ が並び, その 1 つ上の行に 1 が入る行列となる．$n \times n$ 行列の Jordan 標準形 J_{c} は

$$J_{\mathrm{c}} = \begin{bmatrix} J_{k_1} & & & \\ & J_{k_2} & & \\ & & \ddots & \\ & & & J_{k_l} \end{bmatrix}$$

ここで

$$k_1 + k_2 + \cdots + k_l = n$$

また, 各 Jordan 細胞 J_{k_i} の対角成分には, それぞれの固有値が入る.

備考 3.8. 対角行列は, 1 次の Jordan 細胞を対角に並べたものだから, Jordan 標準形である. つまり, Jordan 標準形は対角行列を含む.

3.3.3 対角化可能行列の指数関数の計算法

ここでは, 固有値がすべて実数の行列について, 行列が対角化できる場合と, 対角化できない, つまり Jordan 細胞を含む形にしか変形できない場合に分けて, 行列の指数関数の求め方を示そう. まずは, 行列が対角化できる場合について述べる.

命題 3.9 (正方行列の対角化)**.** 行列 $A \in \mathbb{R}^{n\times n}$ が対角化できるとき, 対角化行列 J は, 列固有ベクトルを並べた行列 P を用いて次のように表される.

$$J = P^{-1}AP$$ ⌟

例題 3.2 (対角行列のべき)**.** $A \in \mathbb{R}^{n\times n}$ が $J = P^{-1}AP$ と対角化されるとき, 対角化行列 J のべき J^n について以下が成立することを示せ.

$$J^n = P^{-1}A^nP \tag{3.4}$$

解答 3.2

$$J^n = (P^{-1}AP)^n = \underbrace{(P^{-1}A\underbrace{P)(P^{-1}}_{E}AP)\cdots(P^{-1}AP)}_{n} = P^{-1}A^nP$$

式 (3.4) の両辺に左から P を, 右から P^{-1} をかけることによって, A^n に関する次の結果が得られる.

定理 3.10. $A \in \mathbb{R}^{n\times n}$ が J に対角化されるとき,

$$\boxed{A^n = PJ^nP^{-1}} \tag{3.5}$$

例題 3.3. 次の A について, A^5 を計算せよ.

$$A = \begin{bmatrix} \frac{5}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{5}{2} \end{bmatrix}$$

解答 3.3 まず A を対角化する. 固有方程式 $\det(A - \lambda E) = 0$ より

$$\lambda^2 - 5\lambda - 6 = 0$$

固有値は $\lambda_1 = 3$, $\lambda_2 = 2$. まず, $\lambda_1 = 3$ に対応する固有ベクトルは, 行列の第 1 行と $[x, y]^{\mathrm{T}}$ の積が $\lambda_1 x$ に等しいことから

$$\frac{5}{2}x + \frac{1}{2}y = 3x$$

より, 例えば $[1, 1]^{\mathrm{T}}$, $\lambda_2 = 2$ に対応する固有ベクトルは

$$\frac{5}{2}x + \frac{1}{2}y = 2x$$

より, 例えば $[1, -1]^{\mathrm{T}}$. これより, P として次の行列をとることができる.

$$P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

この逆行列は

$$P^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

よって, 対角化行列[5)]は

$$J = P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

5) 単に対角化行列を求めるだけなら, 固有ベクトルを規準化（長さを 1 にすること）する必要はない.

行列のべき A^5 は式 (3.5) に上記 J と $n=5$ を代入して

$$A^5 = PJ^5P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3^5 & 0 \\ 0 & 2^5 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{275}{2} & \frac{211}{2} \\ \frac{211}{2} & \frac{275}{2} \end{bmatrix}$$

また, 行列の指数関数の定義より, 次の重要な結果が得られる.

定理 3.11. $n \times n$ 実数行列 A （$A \in \mathbb{R}^{n\times n}$）が $J = P^{-1}AP$ と対角化可能なとき,

$$\boxed{e^{tA} = Pe^{tJ}P^{-1}}\lrcorner$$

備考 3.12. t がスカラー（scalar) のとき, $e^{tA} = e^{At}$ である. どちらの表記も可.

証明.

$$\begin{aligned} e^{tJ} &= e^{(P^{-1}AP)t} \\ &= E + (P^{-1}AP)t + \frac{(P^{-1}AP)^2t^2}{2!} + \cdots + \frac{(P^{-1}AP)^kt^k}{k!} + \cdots \\ &= P^{-1}EP + P^{-1}(At)P + \frac{P^{-1}A^2t^2P}{2!} + \cdots + \frac{P^{-1}A^kt^kP}{k!} + \cdots \\ &= P^{-1}\left(E + At + \frac{A^2t^2}{2!} + \cdots + \frac{A^kt^k}{k!} + \cdots\right)P = P^{-1}e^{tA}P \end{aligned}$$

この両辺に左から P, 右から P^{-1} をかけて

$$Pe^{tJ}P^{-1} = e^{tA}$$

□

上記の計算では

$$(P^{-1}AP)^2 = P^{-1}A\underbrace{PP^{-1}}_{=E}AP = P^{-1}A^2P$$

等を使っている. 第 1 章と同じである.

3.3.4　対角化できない行列の指数関数の計算法

固有値がすべて実数の正方行列は, 対角化できない場合には, Jordan 細胞を含む Jordan 標準形に変形できる（定理 3.6）. この場合も, 対角化できる場合と全く同じ議論で, e^{tA} が計算できる.

定理 3.13. $n \times n$ 実数行列 A ($A \in \mathbb{R}^{n\times n}$ が $J_{\mathrm{c}} = P^{-1}AP$ と標準化されるとき

$$e^{tA} = Pe^{tJ_{\mathrm{c}}}P^{-1} \quad J_{\mathrm{c}} \text{ は Jordan 標準形」}$$

例えば, 上記の 2 次の Jordan 細胞の場合, e^{tJ_2} は次のようになる.

$$e^{tJ_2} = \begin{bmatrix} e^{ta} & te^{ta} \\ 0 & e^{ta} \end{bmatrix}$$

3.4 固有値が実数の場合の解の挙動

3.4.1 解の時間変化は e^{tJ} で決まる

行列 A が $J = P^{-1}AP$ と対角化できる場合, 定理 3.2 と定理 3.11 から, 連立微分方程式の初期値問題の解については以下が成立する.

$$\boldsymbol{x}(t) = e^{tA}\boldsymbol{x}_0 = Pe^{tJ}P^{-1}\boldsymbol{x}_0 \tag{3.6}$$

$$e^{tJ} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix} \tag{3.7}$$

さて, この解の形から分かるように, 時間 t は, e^{tJ} の部分のみ影響する.

$$\boldsymbol{x}(t) = P \underbrace{e^{tJ}}_{\text{時間依存はここだけ}} P^{-1}\boldsymbol{x}_0$$

すなわち, 解の時間変化については, この部分に集約されている. さらに, J は A の対角化行列なので, A の固有値を対角に並べたものになっている. よく, **線形システムの挙動はその固有値によって決まる**, と言われるが, それはこのような意味である[6].

6) ここで述べたのは連続な線形システムの場合である. 連続でない（離散）線形システムも, 似たようなやり方で, 時間変化（離散の場合は A が作用する回数）は対角化行列（のべき）によってのみ決まることが示される.

実正方行列が対角化できない場合も, 定理 3.13 により, $J_{\mathrm{c}} = P^{-1}AP$ と変形できて, $\boldsymbol{x}(t) = Pe^{tJ_{\mathrm{c}}}P^{-1}\boldsymbol{x}_0$ とできる. この場合も時間変化は $e^{tJ_{\mathrm{c}}}$ の部分のみで定まる.

3.4.2 指数関数の挙動（1 変数）

1 変数の関数としての指数関数 $y = e^{at}$ の挙動は, $a > 0$ の場合, $x \to -\infty$ で $e^{at} \to 0$, $x \to \infty$ で $e^{at} \to \infty$. $a < 0$ の場合 $x \to -\infty$ で $e^{at} \to \infty$, $x \to \infty$ で $e^{at} \to 0$ となる.

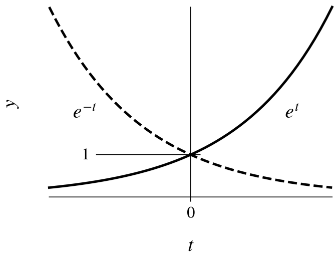


図 3.1 指数関数 e^t と e^{-t}の挙動

言い換えれば, $a < 0$ のとき $x \to \infty$ で e^{ax} は 0 に収束するが, $a > 0$ のときは, $x \to \infty$ で e^{ax} は無限大に発散する. これを用いて, 連立線形微分方程式の解が, 時間とともに（$t \to \infty$）$\mathbf{0}$ に収束するか発散するか, さらにはあるベクトルの方向に収束するか発散するかを判定できる.

3.4.3 行列 A が対角行列の場合

行列の指数関数 e^{tA} の 行列 A が対角行列の場合, 式 (3.7) がそのまま使える. e^{tJ} の各対角項は, $t \to \infty$ のとき, $d_i > 0$ ならば無限大に発散し, $d_i < 0$ ならば 0 に収束する.

例題 3.4. 例題 3.1 の $\boldsymbol{x}(t)$ は次のようなベクトルだった. $\boldsymbol{x}(t)$ の $t \to \infty$ での挙動はどうなるか.

$$\boldsymbol{x}(t) = \begin{bmatrix} e^{-t} \\ e^{2t} \end{bmatrix}$$

解答 3.4 $\boldsymbol{x} = [x, y]^{\mathrm{T}}$ とすると, $t \to \infty$ のとき, $x \to 0$ であり, y 軸の初期値 y_0 が正 $(y_0 > 0)$ のとき $y \to \infty$ で, 負 ($y_0 < 0$ のとき, $y \to -\infty$ である. $y_0 = 0$ のとき, すなわち初期値が x 軸上にある時, $y \equiv 0$ であり, （$t \to \infty$ に対して） $(x, 0) \to (0, 0)$ となる.

連立微分方程式の解をみるとき, 軌跡を図示すると挙動を把握しやすい. 図 3.2 に, この例題 3.4 の様々な初期値に対する軌跡を示す. 解答にみるように, この微分方程式の解は y_0 の正負に対応して, $y \to \pm\infty$ に発散し, x 方向では 0 に近づく. $y_0 = 0$ すなわち初期値が x 軸上にある場合は, $(x, y) \to (0, 0)$ に収束する.

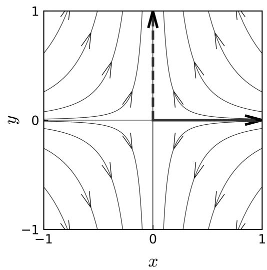


図 3.2　様々な初期値に対する軌跡（3.4）

さて, 例題 3.1, 例題 3.4 の微分方程式の行列 A の a_{11} 項の符号を変えた場合

$$\frac{d\boldsymbol{x}}{dt} = A\boldsymbol{x}, \quad A = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \tag{3.8}$$

の様々な初期値に対する軌跡を図 3.3 に示す. a_{11} 項の符号が変わるだけで, 軌跡の形が大きく変わることが分かる. 初期値問題 (3.8) の場合, 初期値にかかわらず $t \to \infty$ に対して, $(x, y) \to (0, 0)$ に収束する.

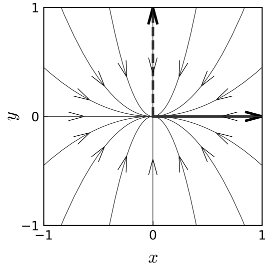


図 3.3　初期値問題（式 (3.8)）の様々な初期値に対する軌跡

上のふたつの例では, 初めから行列が対角になっている場合を示した. 次に, 一般の対角化行列で固有値がすべて実数の場合の例を示そう.

3.4.4　行列 A が対角化可能な場合

行列の指数関数 e^{tA} の 行列 A の固有値を $\lambda_1, \ldots, \lambda_n$ とすると, e^{tA} の対角化行列 e^{tJ} は, 第 i 成分が $e^{t\lambda_i}$ の対角行列になる（式 (3.7)）.　この各対角項は, $t \to \infty$ のとき, $\lambda_i > 0$ ならば無限大に発散し, $\lambda_i < 0$ ならば 0 に収束する.

例題 3.5. 次の A, $\boldsymbol{x}_0$ に対して, A を対角化した行列を J とするとき, (1) e^{tJ} を求めよ. (2) $t \to \infty$ で $\boldsymbol{x} = e^{tA}\boldsymbol{x}_0$ はどうなるか.

$$A = \begin{bmatrix} -1.4 & 0.8 \\ 0.3 & -1.6 \end{bmatrix}, \quad \boldsymbol{x}_0 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

ただし, この行列 A の固有値（対角化行列 J の対角成分）は $\lambda_1 = -1$, $\lambda_2 = -2$ となる.

解答 3.5 (1) 固有値が -1, -2 なので, 次式のようになる. (2) $\boldsymbol{x} \to \mathbf{0} = [0, 0]^{\mathrm{T}}$ に収

束する.

$$e^{tJ} = \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-2t} \end{bmatrix}$$

解答はここまでとして, 参考までに e^{tA} を計算しておこう. 固有値 λ_1, λ_2 にそれぞれに対応する固有ベクトル（規準化せず）は $\boldsymbol{p}_1 = [2,1]^{\mathrm{T}}$, $\boldsymbol{p}_2 = [4,-3]^{\mathrm{T}}$ だから, 固有値を並べた行列 $P = [\boldsymbol{p}_1, \boldsymbol{p}_2]$ とその逆行列 P^{-1} を用いて.

$$\begin{aligned} e^{tA} = Pe^{tJ}P^{-1} &= \begin{bmatrix} 2 & 4 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-2t} \end{bmatrix} \begin{bmatrix} 0.3 & 0.4 \\ 0.1 & -0.2 \end{bmatrix} \\ &= \frac{1}{10} \begin{bmatrix} 6e^{-t} + 4e^{-2t} & 8e^{-t} - 8e^{-2t} \\ 3e^{-t} - 3e^{-2t} & 4e^{-t} + 6e^{-2t} \end{bmatrix} \end{aligned}$$

となる.

計算が面倒なわりに, 出てきた答えを見ても, この解がどのような挙動をするかいまいちよく分からない. この式よりも,

$$\boldsymbol{x}(t) = P \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-2t} \end{bmatrix} P^{-1}\boldsymbol{x}_0$$

の形の式のほうが, 解の挙動の見通しが良い. この式で, P^{-1} は, 元の基底 $\boldsymbol{e}_1 = [1,0]^{\mathrm{T}}$, $\boldsymbol{e}_2 = [0,1]^{\mathrm{T}}$ を, 固有ベクトルの基底 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ に変換し, P は固有ベクトルの基底の基底を元の基底 $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ に変換している. P^{-1} と P は基底変換（座標変換）を行っているだけで, 先ほども述べたとおり, 系の時間変化はすべて e^{tJ} に含まれている.

これを図にすると, 図 3.4 のようになる. 図の一番下の矢印で, 初期値（初期条件）$\boldsymbol{x}_0$ に対して, 微分方程式の演算子 e^{tA} が作用して, $\boldsymbol{x}(t)$ になると考える. 次に, $\boldsymbol{x}_0$ の固有ベクトル $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ の基底のもとでの座標を $\boldsymbol{\xi}_0$ とする. これは, $\boldsymbol{x}_0$ に行列（線形変換）P^{-1} を作用させることで, $\boldsymbol{\xi}_0 = P^{-1}\boldsymbol{x}_0$ として求まる. 固有ベクトルの座標では, 時間 t 経過後の値 $\boldsymbol{\xi}(t)$ は, この $\boldsymbol{\xi}_0$ に e^{tJ} が作用して, $\boldsymbol{\xi}(t) = e^{tJ}\boldsymbol{\xi}_0$ として求まる. これを P で元の座標系に戻して, $\boldsymbol{x}(t) = P\boldsymbol{\xi}(t)$ を得る.

このようにすると, 座標変換, 時間による変化, 座標変換（元に戻す）, という操作で, 解の挙動の見通しが良くなることが分かる. 図 3.5 に, xy 平面でのこの微分方程式の様々な初期値に対する解の軌跡を示す. また, 図 3.6 に, 固有ベクトルの座標系での解の軌跡を示す.

$$
\begin{tikzcd}
\boldsymbol{\xi}_0 \arrow[r, "e^{tJ}"] & \boldsymbol{\xi}(t) \arrow[d, "P"] \\
\boldsymbol{x}_0 \arrow[u, "P^{-1}"] \arrow[r, "e^{tA}"] & \boldsymbol{x}(t)
\end{tikzcd}
$$

図 3.4　行列の指数関数の座標変換

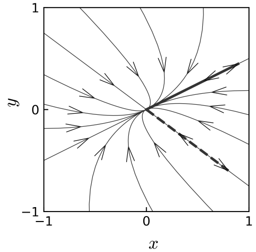


図 3.5　例題 3.5 の 解のさまざまな初期値に対する軌跡, 太い線の矢印は, 固有ベクトル

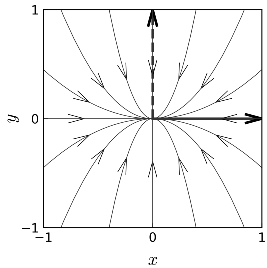


図 3.6　例題 3.5 の $e^{tJ}\boldsymbol{\xi}_0$ の固有ベクトル基底における座標 (ξ_1, ξ_2) での様々な $\boldsymbol{\xi}_0$ に対する軌跡, 太い線の矢印は, この座標系における単位ベクトル $\boldsymbol{e}_{\xi 1}$（実線）, $\boldsymbol{e}_{\xi 2}$（破線）

図 3.6 の軌跡は, 例題 3.4 の場合の図 3.3 と, 座標が違う（前者が固有ベクトルの座標, 後者は xy 座標）だけで, 全く同じになる. すなわち, 例題 3.5 において, 解の時間変動の本質的な部分は, 対角化された J を用いた e^{tJ} に含まれており, P と P^{-1} は座標変換のみを行っていることが, 図からも了解される.

解の挙動のうち, $\boldsymbol{x} \to \boldsymbol{0}$ となる場合は, システムの安定性等において重要な役割を果たす. これを命題としてまとめておく.

命題 3.14. 連立微分方程式の初期値問題 (3.1) において, 行列 A が対角化可能な場合, 固有値が全て負である場合, $t \to \infty$ のとき $\boldsymbol{x} \to \boldsymbol{0}$ に収束する. 」

3.4.5 行列が対角化できない場合

この場合, Jordan 細胞を含む形にしか変形できない. まずは 行列が jordan 細胞の形をしている場合を考えよう.

例 3.15.

$$A = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$$

この 2 次の Jordan 細胞 に対する e^{tA} の様々な初期値に対する軌跡を, 図 3.7 に示す.

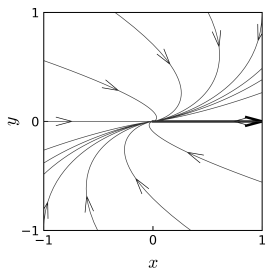


図 3.7 2 次の Jordan 細胞, 固有ベクトルは $\boldsymbol{e}_1$ ひとつ

例 3.16. 一般の縮退行列の例として, 次の行列 A を考え, e^{tA} の軌跡をみる.

$$A = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$$

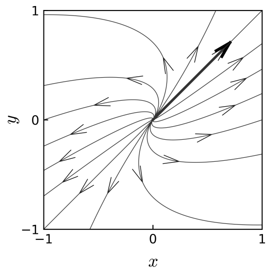


図 3.8 縮退行列の例, 固有ベクトルはひとつ

備考 3.17. 以上 2 つの例では, 固有ベクトルがそれぞれひとつしかない（例 3.15 では $\boldsymbol{p} = [1, 0]^{\mathrm{T}}$, 例 3.16 では $\boldsymbol{p} = \left[1/\sqrt{2}, 1/\sqrt{2}\right]^{\mathrm{T}}$）. 例 3.15, 例 3.16 の固有値はそれぞれ -1, 2 なので, $t \to \infty$ で軌跡は 前者では $\boldsymbol{x} = \boldsymbol{0}$ に収束し, 後者では発散する. 収束するときに, 最終的には固有ベクトル方向に漸近しながら $\boldsymbol{x} = \boldsymbol{0}$ に近づく. この場合も収束・発散は固有値によってのみ決まる.

命題 3.18. 連立微分方程式の初期値問題 (3.1) において, 行列 A の固有値がすべて実数であるが, 対角化できずに, Jordan 細胞を含む形にしか変形できない場合においても, 固有値が全て負である場合, $t \to \infty$ のとき $\boldsymbol{x} \to \boldsymbol{0}$ に収束する. ⌟

3.5 固有値が共役複素数の場合の解の挙動

3.5.1 基礎知識

固有値が共役複素数を表す行列の図形への作用は回転を含む. よって, このような行列 A に対する微分方程式の初期値問題 (3.1) の解は何らかの回転を含む. この

とき, 固有値が純虚数であれば, 解の軌跡は円となる. それ以外の場合は, 固有値の実部が正であれば, $t \to \infty$ の時 $\boldsymbol{x}$ は発散, 固有値の実部が負であれば $\boldsymbol{x} = \boldsymbol{0}$ に収束する.

3.5.2 回転を表す行列・一般形

前章で論じたように, 行列

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{3.9}$$

は, 図形に対して, 反時計回りの角度 θ だけの回転の作用をする. さて, この固有値であるが, 固有方程式が

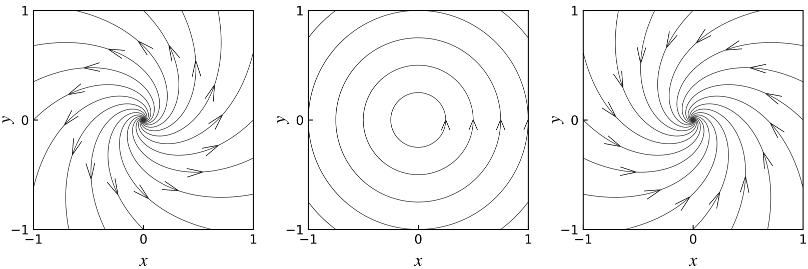
$$\lambda - 2\cos\theta + 1 = 0$$

であるから, 固有値が

$$\lambda = \cos\theta \pm \sqrt{\cos^2\theta - 1} = \cos\theta \pm i\sin\theta$$

である. 初期値問題 (3.1) の解の挙動は, この固有値によって決まる.

図 3.9 に, θ が $\pi/4$, $\pi/2$, $3\pi/4$ の場合の解の挙動を示す. それぞれの場合に実部は $1/\sqrt{2}$, 0, $-1/\sqrt{2}$ だが, この実部の正, 0, 負に対応して, 解の軌跡は発散, 円, 収束, となっている. いずれの場合も回転角は $0 < \theta < \pi$ なので, 反時計回りの回転が図に見られる.



(a) $\theta=\pi/4$　(b) $\theta=\pi/2$　(c) $\theta=3\pi/4$

図 3.9　回転を表す行列の場合の様々な初期値に対する解の軌跡

3.5.3 回転と拡大, 縮小を表す行列の場合

$$A = a\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad a \neq 0, 1 \tag{3.10}$$

行列がこの式 (3.10) の形をしている場合, 初期値問題 (3.1) の解は, 式 (3.9) における t が ta になるだけで, まったく同じ軌跡となる.

3.5.4 一般の共役複素数固有値

例 3.19.

$$A = \begin{bmatrix} 7/8 & -5/8 \\ 1/8 & 5/8 \end{bmatrix}$$

の固有値は $\lambda = 3/4 \pm i/4$ である. 解の挙動は次のようになる.

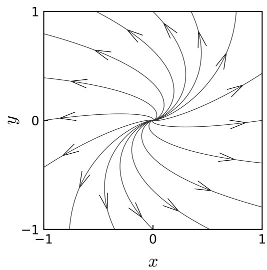


図 3.10　共役複素数の固有値をもつ行列による連立微分方程式の様々な初期値に対する軌跡

この場合の軌跡は図 3.10 のように, 図 3.9 (a) のグラフ少し変形した形になる. 固有値の実部が 3/4 で正なので, $t \to \infty$ で発散する. 固有値が共役複素数の場合は, このように, 実部の政府で発散・収束が決まってくる.

命題 3.20. 連立微分方程式の初期値問題 (3.1) において, 行列 A が共役複素数をもつ場合, その実部が全て負であれば, $t \to \infty$ のとき $\boldsymbol{x} \to \mathbf{0}$ に収束する. 」

3.6 解の挙動のまとめ

3.6.1 収束する場合

前節で様々な場合における収束の条件をみてきたが, これらをまとめて次が言える.

定理 3.21. 連立微分方程式の初期値問題 (3.1) において, 行列 A の固有値の実部が全て負であれば, $t \to \infty$ のとき $\boldsymbol{x} \to \mathbf{0}$ に収束する. 」

3.6.2 場合分け

さて, 2×2 実正方行列は, 固有値が実数の場合と共役複素数のペアの場合に分けられる. さらに, 前者は, 対角化できる場合と, 2 次の Jordan 細胞 の形に標準化できる場合に分けられる. よって, ここでは, (1) 対角化できて固有値が実数の場合, (2) 2 次の Jordan 細胞 に標準化できる場合, (3) 固有値が共役複素数の場合, に分けて整理する.

3.6.3 行列が対角化可能で固有値が実数の場合

この場合行列は適当な座標変換で対角行列になる. 対角化行列の挙動の分類のために, 次の形の行列を考える.

$$A = \begin{bmatrix} a & 0 \\ 0 & -1 \end{bmatrix} \tag{3.11}$$

この行列の挙動が a の値によってどのように変化するかを見てみよう（図 3.11）.

(a) $a < -1$ の場合: $|a|$ は, 右下の対角要素 -1 よりも大きいので, x 方向のほうが y 方向よりも速く小さくなる. よって, 軌跡は x 軸を軸とする放物線状になる.
(b) $a = -1$ の場合: 対角成分がともに -1 で等しいので, x 方向と y 方向で同じ速さで変化する. よって, 軌跡は初期値からまっすぐに原点をめざす形になる.
(c) $-1 < a < 0$ の場合: $|a| < |-1|$ なので, y 方向のほうが x 方向より速く小さくなる.
(d) $a = 0$ の場合: x 方向の変化はなく, y 方向にのみ変化する. $t \to \infty$ で, $y \to 0$

となる.

(e) $a > 0$ の場合: もう一方の対角成分は -1 なので, $t \to \infty$ で x 方向は無限大に発散, y 方向は 0 に収束する.

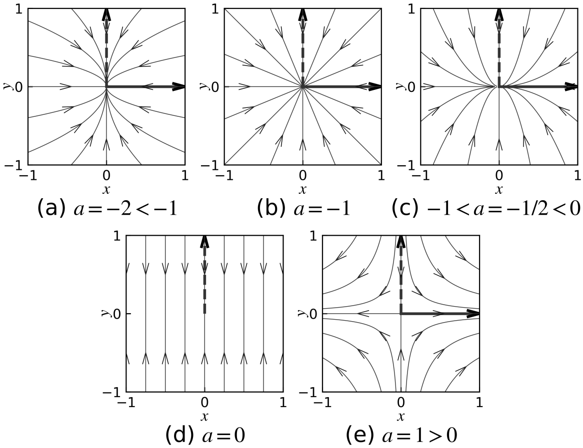


図 3.11　対角化可能行列の e^{tJ} の挙動

さて, 式 (3.11) の右下の項 $a_{22} = -1$ の符号を変えた行列

$$B = \begin{bmatrix} b & 0 \\ 0 & 1 \end{bmatrix}$$

について考える. この場合の e^{tB} については, (a) $b > 1$ のとき, 図 3.11 (a) の変化の向きが逆になった場合, (b) $b = 1$ のとき, 図 3.11 (b) の逆向き, (c) $0 < b < 1$ のとき, 図 3.11 (c) の逆向き, (d) $b = 0$ のとき, 図 3.11 (d) の逆向き, (e) $b < 0$ のとき, 図 3.11 (e) の逆向き, となる.

3.6.4 固有値が実数で行列が対角化できない場合の解の挙動

2×2 行列の固有値がすべて実数で対角化できない場合, すなわち, 固有値が縮退している場合, 適当な基底変換により, 次の形の行列となる.

$$A = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}$$

このとき, e^{tA} は, a の正負により発散か収束かが決まる.

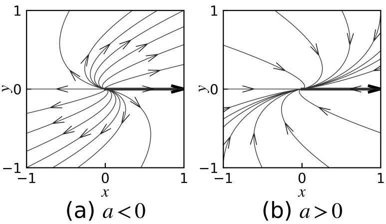


図 3.12 縮退行列の解の挙動

3.6.5 固有値が共役複素数の場合の解の挙動

固有値が共役複素数の場合は, 固有値の実部の正負によって発散か収束かが決まる.

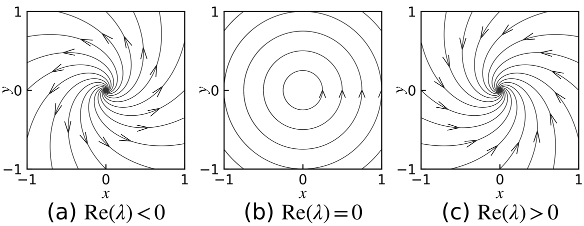


図 3.13 固有値が共役複素数となる場合の解の挙動

3.6.6　まとめの例題

例題 3.6. 次の行列の e^{tA} の軌跡は, 座標変換の作用を除けば, 図 3.11 の (a) から (e) のうちのどれに最も近くなるか. ただし, 解の向き（収束か発散か）は問わず, (a) と (c) は区別しないとする.

$$(1) \quad A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad (2) \quad A = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}$$

解答 3.6 (1) 固有方程式は $\lambda^2 - 5\lambda - 2 = 0$ より, $\lambda = 5/2 \pm \sqrt{33}/2$ より, $\lambda = -0.3723, 5.3723$. 一方が正でもう一方が負なので, 図 3.11 の (e) に最も近い.
(2) 固有方程式は $\lambda^2 - 3\lambda + 2 = 0$ より $\lambda = -1, -2$ ともに負で異なる値なので, 図 3.11 の (a) （あるいは (c) ）に最も近い.

3.7　参考書

講義内容の理解のためには, 次の小寺の 4 章が良い.

- 小寺平治, なっとくする微分方程式, 講談社, 2000,
  ISBN=978-4-06-154521-2 https://bookclub.kodansha.co.jp/product?item=0000148737

線形微分方程式系から力学系, カオスへは例えば,

- Strogatz, S. H., Nonlinear Dynamics and Chaos, 3rd Ed., Chapman and Hall (2024).
  https://www.taylorfrancis.com/books/mono/10.1201/9780429492563/nonlinear-dynamics-chaos-steven-strogatz

3.8　演習問題

問題 3.1. $x = f(t) = e^{-kt}$ を $t = 0$ の周りで t の 3 次の項まで Taylor 展開せよ. ただし, 剰余項はただ R と示すのみでよい.

問題 3.2. 次の微分方程式を初期値問題を解け.

$$\frac{dx}{dt} = -kx, \quad x(0) = 1$$

問題 3.3. 次の微分方程式を, 下に示す行列と初期条件のもとで解け.

$$\frac{d\boldsymbol{x}}{dt} = A\boldsymbol{x}, \quad A = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}, \quad \boldsymbol{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

問題 3.4. 行列 A は次のように対角化される.

$$\begin{aligned} A &= \begin{bmatrix} -\frac{3}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{-3}{8} \end{bmatrix} = PJP^{-1} \\ &= \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \end{aligned}$$

e^{tA} を求めよ.

問題 3.5. 連立微分方程式 (3.1) の行列が次であるとき, 適当な座標変換後の軌跡の形はそれぞれ図 3.11 の (a)〜(e) のどれに最も近くなるか. ただし, 軌跡の向き（収束か発散か）は問わないとし, (a) と (c) は区別しないとする.

$$(1) \quad A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad (2) \quad A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

$$(3) \quad A = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}, \quad (4) \quad A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

問題 3.6. 連立微分方程式 (3.1) の行列が次であるとき, 適当な座標変換後の軌跡の形は図 3.13 の (a)〜(c) のどれに最も近くなるか. ただし, 渦巻きが左回りか右回りかは問わないとする.

$$(1) \quad A = \begin{bmatrix} 4 & -6 \\ 6 & 4 \end{bmatrix}, \quad (2) \quad A = \begin{bmatrix} -2 & -4 \\ 3 & -1 \end{bmatrix}$$

$$(3) \quad A = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}, \quad (4) \quad A = \begin{bmatrix} 0 & 2 \\ -1 & 0 \end{bmatrix}$$

問題 3.7. 連立微分方程式 (3.1) の行列が次であるとき, 適当な座標変換後の軌跡の形は, 図 3.11 (a)〜(e), 図 3.12 (a), (b), 図 3.13 (a) 〜 (c) のどれに最も近くなるか. ただし, 図 3.11 では, 軌跡の向き（収束か発散か）は問わず, (a) と (c) は区別しないとし, 図 3.13 では, 渦巻きが左回りか右回りかは区別しないとする.

$$(1)\quad A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \quad (2)\quad A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

4 極大点・極小点を求めるアルゴリズム

4.1 この章で学ぶこと

現在, 最適化で最もよく用いられているのは, おそらく, Newton 法と勾配降下法だろう. 前者は効率よく最適値を求めるのに適しており, 後者は 1 階微分のみを使うという計算負荷上の利点から機械学習で用いられる（実際には派生法である確率的勾配降下法がよく用いられる）.

最適化の難しさは, 過学習や, 初期値の設定方法, などアルゴリズムとは違ったところにあることが多いが, それでも, **アルゴリズムを理解**していれば, **問題に対してより適切な対処ができる**可能性が高まる. そのような観点から, この章では, Newton 法と勾配降下法を学ぶ.

4.2 最適化のアルゴリズム（1 変数）

4.2.1 Newton 法

Newton 法（Newton Raphson 法）は, 極小値を求めるというよりは, 方程式 $f(\boldsymbol{x}) = 0$ の解を求める方法である. 極小値を求めるためには, 1 階微分のゼロ点 $df/dx = 0$ つまり停留点を求める. 図 4.1 に $y = f(x) = 0$ の解を求める例を示す. まず, 適当な初期値 x_0 を決める. 次に, その初期値における関数の値 $f(x_0)$ を求め

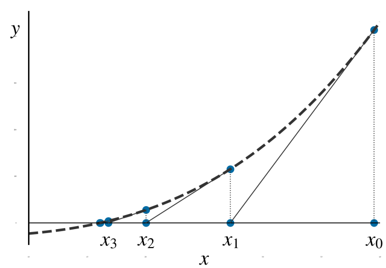


図 4.1 1 変数の場合の Newton 法による方程式 $y = f(x) = 0$ の解の求め方

る．点 $(x_0, f(x_0))$ から, 接線（傾き $f'(x)$）を引く．接線と x 軸（$y=0$） との交点を求め, x_1 とする．以下, $(x_i, f(x_i))$ から接線を引いて x 軸との交点 x_{i+1} を求めることを繰り返す．$|f(x_{i+1})| < \varepsilon_0$ となったら, x_{i+1} を解の近似値として採用し, 計算を終える．

Newton 法で極大・極小点を求めるには, 停留点を求める．求めた停留点が極大点か極小点かについては, 第 1 章に述べた方法で別に調べる必要がある．図 4.2 に停留点を Newton 法で求めるときの x の更新過程の例を示す．$f'(x) = 0$ を解くことで, 関数 $f(x)$ の極小点が求まっている．図 4.3 に, Newton 法で停留点を求めるアルゴリズムを示しておく．

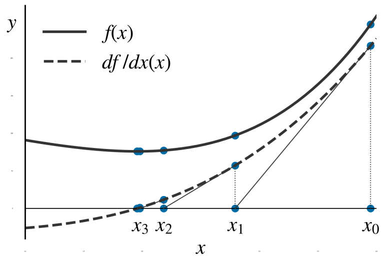


図 4.2　1 変数の場合の Newton 法による停留点の求め方の例, 実線 $f(x)$, 破線 $f'(x)$

x_0 を設定

$$x_{i+1} = x_i + \Delta x_i, \quad \boxed{\Delta x_i = -\frac{f'(x_i)}{f''(x_i)}} \qquad \text{(a.1)}$$

if $|f'(x_{i+1})| < \varepsilon_0$:

　x_{i+1} を近似値として採用

else :

　(a.1) にもどる．

図 4.3　Newton 法により停留点を求めるアルゴリズム

4.2.2 勾配降下法

勾配降下法では関数の勾配（が最大になる）方向に勾配に比例するだけ移動する. 関数の谷を降りていき, 谷底（極小点）に到達すれば, 勾配が 0 になるので, 移動が止まる. 計算には, 学習率（learning rate）と呼ばれるパラメーターを用いるが, この値により収束速度が大きく異なり, 場合によっては収束しないこともある. 図 4.4 に勾配降下法での極小点を求めるまでの x_i の更新の様子の例を示す.

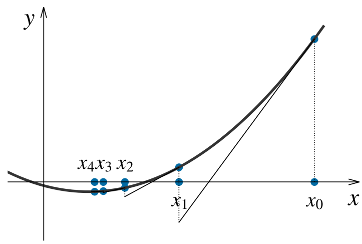


図 4.4　勾配降下法による極小点の求め方の例

図 4.5 に, 勾配降下法のアルゴリズムを示す.

x_0 を設定

$$x_{i+1} = x_i + \Delta x_i, \quad \boxed{\Delta x_i = -\eta f'(x_i)} \qquad \text{(a.2)}$$

if $|f'(x_{i+1})| < \varepsilon_0$:

　　x_{i+1} を近似値として採用

else :

　　(a.2) にもどる

図 4.5　勾配降下法のアルゴリズム

4.2.3 Newton 法と勾配降下法の比較

Newton 法では, f' の接線と $y=0$ との交点を更新点として用いるが, 勾配降下法では, f の接線方向に, その傾きに比例した分だけ x を変化させる. Newton 法で停留点を求めるためには, アルゴリズムからも分かるように, $f'(x_i)$ と $f''(x_i)$ が必要である. 勾配降下法では, $f'(x_i)$ だけが必要である.

備考 4.1. アルゴリズムから分かるが, Newton 法は停留点に収束する. つまり極大値にも収束する. 勾配降下法は極小値のみに収束し, 極大値に収束することはない.

例題 4.1. 次の関数の極小値を, 初期値 $x_0 = 2$ からの 1) Newton 法, 2) 勾配降下法で, ともに収束条件を $|f'(x_i)| < \varepsilon_0$, $\varepsilon = 1.0 \times 10^{-4}$ として求めよ. ただし, 勾配降下法の学習率は $\eta = 0.1111$ とする.

$$y = f(x) = x^3 - 3x$$

この問題はできれば関数電卓で解いてみよう.

解答 4.1

$y = f(x)$ の 1 階, 2 階の導関数は以下のとおりである.

$$f'(x) = 3x^2 - 3, \quad f'' = 6x$$

$x_0 = 2$ として, それぞれのアルゴリズムに従って, $f'(x_i)$, $f''(x_i)$ を計算し, それらをもとに Newton 法では $\Delta x_i = f'(x_i)/f''^{(x_i)}$, 勾配降下法では $\Delta x_i = -\eta f'(x_i)$ を算出. そこから $x_{i+1} = x_i + \Delta x_i$ を算出した. 結果を表 4.1 にまとめた.

備考 4.2. 以上の操作で, Newton 法と勾配降下法で x_i が近似解に近づいていく様子がよく分かる. この問題の場合, 勾配降下法のほうが, Newton 法より速く収束した. 勾配法でも, 学習率をうまく選べば, 収束を早めることができる. ただし, ここでは収束条件を $\varepsilon_0 = 1.0 \times 10^{-4}$ とかなり緩いものにした. 例えばこれを 1.0×10^{-6} に変えると Newton 法のほうが速くなる.

一般に, Newton 法は近似解への収束が速いことが知られている. これに対し, 勾配降下法は収束速度が遅いが, 1 階微分を計算するだけでよいのが大きなメリット

表 4.1　例題 4.1 の計算

	Newton				Gradient descent		
i	x_i	$f'(x_i)$	$f''(x_i)$	Δx_i	x_i	$f'(x_i)$	Δx_i
0	2.00000	9.00000	12.00000	-0.75000	2.00000	9.00000	-0.99990
1	1.25000	1.68750	7.50000	-0.22500	1.00010	0.00060	-0.00007
2	1.02500	0.15187	6.15000	-0.02470	1.00003	0.00020	-0.00002
3	1.00030	0.00183	6.00183	-0.00030	1.00001		
4	1.00000						

となる. 特に機械学習など変数の数 n が大きい場合には, 2 解微分である Hessian の計算に n^2 の計算が必要となることを考えると, 勾配降下法が選ばれる理由が納得できる.

4.2.4　初期値依存性

非線形の最適化問題は, 初期値の選び方が重要である. 領域内の「最小点」を選ぶ問題の場合, 初期値の選び方によっては, 最小点とは別の極小点が解として得られてしまう場合もあれば, 全く解が得られない（発散する）場合もある.

大まかな傾向とすれば, 目的とする極小値（例えば最小値）の近くに初期値を持ってくると, その点に収束する場合が多い. しかし, 勾配降下法などでは適切な学習率を設定しないと, 初期値が解の近くにあっても x_i が解に収束しない場合がある（後述）.

また, そもそもの問題として, 解が未知なのだから, 初期値をその近くに設定するというのは無理がある. このため, 往々にして様々な初期値を試すことになる. このあたりに非線形最適化問題の難しさがある.

まずは Newton 法で初期値依存性をみよう.

例題 4.2. 例題 4.1 の次の関数の停留点を求めるのに Newton 法を用いる.

$$y = f(x) = x^3 - 3x$$

初期値を $x_0 = -0.1, 0.1$ としたときに, それぞれ, どのような停留点に収束するか.

解答 4.2 実際に Newton 法で計算してみると（表 4.2）$x_0 = -0.1$ のときには, $x = -1$ に, $x_0 = 0.1$ のときには 1 に収束することが分かる.

備考 4.3. Newton 法は, 停留点 $df/dx = 0$ を求める方法なので, 極大点にも収束する. 表 4.2 に, $x_0 = -0.5, -0.1, 0.1, 0.6$ のときに, x_6 までを計算した結果を示す. 図を描いてみれば分かるが, この関数の変曲点 $x = 0$ より初期値 x_0 が小さければ極大点 $x = -1$ に収束し, 大きければ極小点 $x = 1$ に収束する. $x_0 = -0.5$ のほうが, $x_0 = -0.1$ よりも $x = -1$ に近いので, 前者の場合のほうが収束が速い. また, この関数は原点まわりで点対称なので, 初期値 x_0 の絶対値が同じであれば, 符号以外の結果は同じになる. これを踏まえて, $x_0 = -0.5$ と $x_0 = 0.6$ を比較すると, $|-1-(-0.5)| > |1-0.6|$ なので, 後者のほうが速く収束する.

表 4.2　Newton 法の初期値依存性

$i \backslash x_0$	-0.5	-0.1	0.1	0.6
0	-0.5000	-0.1000	0.1000	0.6000
1	-1.2500	-5.0500	5.0500	1.1333
2	-1.0250	-2.6240	2.6240	1.0078
3	-1.0003	-1.5026	1.5026	1.0000
4	-1.0000	-1.0840	1.0840	1.0000
5	-1.0000	-1.0033	1.0033	1.0000
6	-1.0000	-1.0000	1.0000	1.0000

例 4.4. 例題 4.1 の関数

$$y = f(x) = x^3 - 3x$$

に勾配降下法を適用する. 様々な初期値と学習率の組で計算を行った場合の $i = 8$ までの x_i を表 4.3 にまとめた.

表 4.3　勾配降下法における初期値と学習率の影響の例

x_0	-1.1	-0.9	0.99	0.99	0.99	0.99
$i \setminus \eta$	0.1	0.167	0.167	0.1	0.3	0.6
0	-1.1000	-0.9000	0.9900	0.9900	0.9900	0.9900
1	-1.1630	-0.8048	1.0000	0.9960	1.0079	1.0258
2	-1.2688	-0.6283	1.0000	0.9984	0.9936	0.9317
3	-1.4517	-0.3251	1.0000	0.9994	1.0051	1.1693
4	-1.7839	0.1229	1.0000	0.9997	0.9959	0.5084
5	-2.4387	0.6164	1.0000	0.9999	1.0032	1.8432
6	-3.9228	0.9270	1.0000	1.0000	0.9974	-2.4720
7	-8.2393	0.9975	1.0000	1.0000	1.0021	-11.6716
8	-28.3052	1.0000	1.0000	1.0000	0.9983	-255.0771

勾配降下法では, 初期値 x_0 と学習率 η の両方によって収束が違ってくる. 左から見ていこう. 初期値と学習率が $x_0 = -1.1$, $\eta = 0.1$ の場合, 初期値が極大点 -1 より小さいので, 勾配を下ると, x はどんどん負の方向に移動する. よって発散してしまう. $x_0 = -0.9$, $\eta = 0.167$ のときは, 初期値は極大値より右なので収束する. 学習率はこれが最適だが, 収束は遅い. $x_0 = 0.99$ の場合, 学習率を最適, 0.1, 0.3, 0.6 と変化させてその影響をみた. 最適学習率 $\eta = 0.167$ の場合, x_1 ですでに誤差が 5×10^5 以下になっている. $\eta = 0.1$ の時には, 同じようなレベルになるのに 6 回かかっている. $\eta = 0.3$ の場合, x_i は, $x_i > 1$, $x_i < 1$ を繰り返してながら $x = 1$ に近づいている. この場合はいずれ収束するが, このように x_i が振動する場合は学習率が大きすぎると考えてよい. $\eta = 0.6$ の場合, x_i は, $x_i > 1$, $x_i < 1$ を繰り返してながら発散し, $x_i < -1$ となったところから勾配を降ってどんどん負の方向に移動することになる.

4.3 多変数関数の場合

4.3.1 2 変数関数のゼロ点と停留点

第 1 章で見たように, 2 変数以上の多変数関数の極大・極小問題はすべて同じように扱える. よって, ここでは 2 変数関数のみ考える. 次の関数を考えよう.

$$y = f(x, y) = x^2 + y^2 - 1$$

この関数のゼロ点は $x^2 + y^2 - 1 = 0$ の解だから, 原点を中心として半径を 1 とする円となる. つまり, この関数のゼロ点は 1 点には定まらない. 一方, 停留点は

$$\frac{\partial f}{\partial x} = 2x = 0, \quad \frac{\partial f}{\partial y} = 2y = 0$$

の解なので, $(x, y) = (0, 0)$ と 1 点に定まる. 停留点が極大値を与えるか極小値を与えるか等については, 第 1 章で解説した方法で調べればよい. 以上を考えて, この節では Newton 法でもゼロ点はとりあげず, 停留点と極小点のみを議論する.

4.3.2 Newton 法による停留点の求め方（2 変数）

1 変数の場合と, アルゴリズムは同じになる. ただし, f が多変数関数なので, x_i がベクトル $\boldsymbol{x}_i = [x_i, y_i]^{\mathrm{T}}$ に変わるほか, x_i の更新の部分（1 変数の場合は $\Delta x_i = -f'(x_i)/f''(x_i)$）が次のように変わる.

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \Delta\boldsymbol{x}_i, \quad \boxed{\Delta x_i = -H_f(\boldsymbol{x}_i)^{-1}\boldsymbol{\nabla} f(\boldsymbol{x}_i)} \tag{4.1}$$

ここで, H_f は第 1 章でもとりあげた Hesse 行列である. 関数 f の Hesse 行列であることを強調したいために, 添字として f を付けてある. $H(\boldsymbol{x}_i)^{-1}$ は $\boldsymbol{x}_i$ における Hesse 行列の逆行列である. $\nabla f(\boldsymbol{x}_i)$ は, 次に示す**勾配 (gradient)** である. 勾配は ∇f とか $\mathrm{grad}\, f$ で表される.

$$\nabla f(\boldsymbol{x}_i) = \mathrm{grad}\, f(\boldsymbol{x}_i) = \begin{bmatrix} \dfrac{\partial f}{\partial x}(\boldsymbol{x}_i) \\[2em] \dfrac{\partial f}{\partial y}(\boldsymbol{x}_i) \end{bmatrix}$$

式 (4.1) の $\Delta \boldsymbol{x}_i$ は, 2×2 行列 $H(\boldsymbol{x}_i)^{-1}$ と 2 項列ベクトル $\nabla f(\boldsymbol{x}_i)$ の積だから, 2 項列ベクトルとなる.

これだけでは分かりにくいだろうから, 簡単な計算例をみていこう.

例 4.5. 関数

$$y = f(x, y) = x^2 + y^2 + x^2 y^2$$

の停留点は $\boldsymbol{x} = [0, 0]^{\mathrm{T}}$ だが, 初期値を $\boldsymbol{x}_0 = [2, 1]^{\mathrm{T}}$ として Newton 法で求める最初のステップを調べる.

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} 2x + 2xy^2 \\ 2y + 2x^2 y \end{bmatrix}, \quad H_f(\boldsymbol{x}) = \begin{bmatrix} 2 + 2y^2 & 4xy \\ 4xy & 2 + 2x^2 \end{bmatrix}$$

さて, $\boldsymbol{x}_0 = [2, 1]^{\mathrm{T}}$ のときに

$$\nabla f(\boldsymbol{x}_0) = \begin{bmatrix} 8 \\ 10 \end{bmatrix}, \quad H_f(\boldsymbol{x}_0) = \begin{bmatrix} 4 & 8 \\ 8 & 10 \end{bmatrix}, \quad H_f(\boldsymbol{x}_0)^{-1} = \begin{bmatrix} -5/12 & 1/3 \\ 1/3 & -1/6 \end{bmatrix}$$

$$\Delta \boldsymbol{x}_0 = -H_f(\boldsymbol{x}_0)^{-1} \nabla f(\boldsymbol{x}_0) = -\begin{bmatrix} -5/12 & 1/3 \\ 1/3 & -1/6 \end{bmatrix} \begin{bmatrix} 8 \\ 10 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 + \Delta \boldsymbol{x}_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

あとは, x_i の更新を繰り返せばよい.

4.3.3 勾配降下法による極小点の求め方（2 変数）

勾配降下法の場合も, $f'(x_i)$ が $\nabla f(x_i)$ となるだけで, アルゴリズムは同じである. 学習率はスカラーのままでよい.

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \Delta \boldsymbol{x}_i, \quad \boxed{\Delta \boldsymbol{x}_i = -\eta \nabla f(\boldsymbol{x}_i)} \tag{4.2}$$

例 4.6. 例 4.5 と同じ関数

$$y = f(x, y) = x^2 + y^2 + x^2 y^2$$

の関数の極小点を勾配降下法で求る最初のステップを示そう.

ただし学習率は $\eta = 0.05$ とする.

例 4.5 でみたように, 勾配は, $\nabla f = \left[2x + 2xy^2, 2y + 2x^2y\right]^{\mathrm{T}}$ だから, 初期値 $\boldsymbol{x}_0 = [2, 1]^{\mathrm{T}}$ に対して

$$\nabla f(\boldsymbol{x}_0) = \begin{bmatrix} 8 \\ 10 \end{bmatrix}, \quad \Delta x = -\eta \nabla f(\boldsymbol{x}_0) = \begin{bmatrix} -0.4 \\ -0.5 \end{bmatrix}$$

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \Delta \boldsymbol{x}_i = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} -0.4 \\ -0.5 \end{bmatrix} = \begin{bmatrix} 1.6 \\ 0.5 \end{bmatrix}$$

あとは, $\boldsymbol{x}_i$ の更新を繰り返すことになる.

4.3.4　収束判定（多変数）

多変数の場合の収束判定は, 1 変数の場合と違って, ベクトル $\nabla(\boldsymbol{x}_i)$ の大きさを調べることになる. これにはいくつかの方法があるが, $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_n]^{\mathrm{T}}$ のユークリッドノルム

$$\|\boldsymbol{\xi}\|_2 = \sqrt{\xi_1^2 + \cdots + \xi_n^2}$$

は, 変数の数が多くなると, 計算コストが高い. そのためだろうが, 多くの場合, 次の最大値ノルムが用いられる.

$$\|\boldsymbol{\xi}\|_\infty = \max |\xi_i|$$

4.3.5　アルゴリズム

整理の意味もこめて, 多変数の Newton 法と勾配降下法のアルゴリズムを載せる.

4.4　Python による計算例

4.4.1　はじめに

Newton 法にせよ勾配降下法にせよ, 1 変数の場合はともかく, 多変数の場合はコンピューターを使って計算すべきものである. このとき, 勾配などを計算する必要

$\boldsymbol{x}_0$ を設定

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \Delta\boldsymbol{x}_i, \quad \boxed{\Delta x_i = -H_f(\boldsymbol{x}_i)^{-1}\boldsymbol{\nabla} f(\boldsymbol{x}_i)} \qquad \text{(a.3)}$$

if $\|\nabla(\boldsymbol{x}_{i+1})\|_\infty < \varepsilon_0$:

x_{i+1} を近似値として採用

else :

(a.3) にもどる

図 4.6 多変数関数で Newton 法により停留点を求めるアルゴリズム

$\boldsymbol{x}_0$ を設定

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \Delta\boldsymbol{x}_i, \quad \boxed{\Delta x_i = -\eta\nabla f(\boldsymbol{x}_i)} \qquad \text{(a.4)}$$

if $\|\nabla f(\boldsymbol{x}_{i+1})\|_\infty < \varepsilon_0$:

x_{i+1} を近似値として採用

else :

(a.4) にもどる

図 4.7 多変数関数で勾配降下法により極小点を求めるアルゴリズム

があるが, 特に Newton 法の場合, Hesse 行列, 及びその逆行列を計算する必要があり, これを一々入力するのは大変である. このようなとき, 自動微分関連のパッケージを用いればよい.

ここでは, Google 製の JAX（https://github.com/google/jax）を使う. JAX は自動微分と高速化線形代数をひとつにしたパッケージであり, 色々なことができるのだが, ここでは, numpy の機能を代替する `jnp.numpy` と, 自動微分の機能のみを使う. 自動微分を使うだけで, 最適化計算にかかる労力の大半を削減することができる.

4.4.2 Newton 法の計算例（2 変数）

ここでは, 次の関数の停留点を初期値 $\boldsymbol{x}_0 = [0.2, 2.2]^{\mathrm{T}}$ として Newton 法で解く.

$$z = f(x, y) = -\cos(2x)\sin(y)$$

```python
import jax  # jax パッケージを使う
import jax.numpy as jnp  # numpy の代わりに jax.numpy を使う.
from jax import config  # この行と次の行は jax での 64-bit 浮動小数点計算のため
config.update("jax_enable_x64", True)


def func(x):
    return - jnp.cos(2 * x[0]) * jnp.sin(x[1])


# jax では, Hesse 行列を計算する関数がないので作る.
def jax_hess(f):
    return jax.jacfwd(jax.jacrev(f))


def main():
    eps_0 = 1.e-15  # 収束条件の設定
    x = jnp.array([.2, 2.2])  # 初期値の設定
    # 最大値ノルムの計算
    nrm_grad = jnp.linalg.norm(jax.grad(func)(x), ord=jnp.inf)
    while nrm_grad > eps_0:
        grad = jax.grad(func)(x)
        hess = jax_hess(func)(x)
        hess_inv = jnp.linalg.inv(hess)  # Hesse 行列の逆行列
        dltx = jnp.matmul(hess_inv, grad) # H_f^-1 と grad f の積
        x = x - dltx
        print(x)
        nrm_grad = jnp.linalg.norm(grad, ord=jnp.inf)
    # 極大・極小判定のため Hessian matrix の固有値を出力
    print(jnp.linalg.eigvals(hess))


if __name__ == '__main__':
    main()
```

出力は以下の通り.

```
[-0.15723497  1.25222653]
[0.02518525 1.62116685]
[-8.55469635e-05  1.57062523e+00]
[3.33897023e-12 1.57079633e+00]
[0.         1.57079633]
```

```
[0.          1.57079633]
[4.+0.j  1.+0.j]
```

出力から, 停留点として $[0, 1.57]^{\mathrm{T}}$ が得られ, Hesse 行列の固有値が 4, 1, で共に正だから, これが極小点であることが分かる.

4.5 演習問題

問題 4.1. 次の関数の極小点を $x \in (0, 2)$ の範囲で求めるために, 例題 1 と同様のやり方で x_4 までの各値を求めて表にせよ. ただし, 初期値は $x_0 = 2$ とし, 勾配降下法の学習率は $\eta = 0.033$ とせよ.

$$y = f(x) = x^4 - 6x^2$$

問題 4.2. 次の関数の停留点を Newton 法で求める. 1) 関数 $y = f(x)$ のグラフを $x \in (-1.5, 1.5)$ の範囲で描け. 2) 初期値 $x_0 = -3, -1, 0.3, 2$ から初めたとき, 解はどの停留点に収束するか.

$$y = f(x) = x^4 - 2x^2 \tag{4.3}$$

ちなみに, 停留点は,

$$y = f'(x) = 4x^3 - 4x = 0$$

の解なので, $x = -1, 0, 1$ の 3 つである. また, 変曲点は

$$y = f''(x) = 12x^3 - 4 = 0$$

の解であるから, $x = \pm 1/\sqrt{3}$ である.

問題 4.3.（この問題は Python を使うほうがよい.）上記の式 (4.3) の関数の極小点を勾配降下法で求めることにする. グラフの形から, 初期値を正の数にとると, 極小値 $x = 1$ に, 初期値を負の値にとると, 極小値 $x = -1$ に収束するはずである. 1) 初期値を負にとり, その値で $x = -1$ に収束するように学習率 η を定めよ. 2) 初期値を正にとり, その値で $x = 1$ に収束するように学習率 η を求めよ. 3) 上記 1), 2) のいずれかにおいて, 学習率 η を変えたときに, 収束速度がどのように変化するか計算例を示せ.

演習問題解答

演習問題 1 解答

例 (c) には剰余が入っていないが, 等号が成立するには, 剰余 R が必要である (20231227 問題文訂正済）.

1.1.

(1) $f(x) = e^x$

$$
\begin{aligned}
&\text{(a)} \quad f'(x) = e^x, \quad f''(x) = e^x, \quad f^{(3)}(x) = e^x \\
&\text{(b)} \quad f'(0) = 1, \quad f''(0) = 1, \quad f^{(3)}(0) = 1 \\
&\text{(c)} \quad f(x) = 1 + \frac{1}{2}x + \frac{1}{6}x^2 + R
\end{aligned}
$$

(2) $f(x) = \log(1+x)$

$$
\begin{aligned}
&\text{(a)} \quad f'(x) = \frac{1}{1+x}, \quad f''(x) = -\frac{1}{(1+x)^2}, \quad f^{(3)} = \frac{2}{(1+x)^3} \\
&\text{(b)} \quad f'(0) = 1, \quad f''(0) = -1, \quad f^{(3)}(0) = 2 \\
&\text{(c)} \quad f(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + R
\end{aligned}
$$

(3) $f(x) = \sin x$

$$
\begin{aligned}
&\text{(a)} \quad f'(x) = \cos x, \quad f''(x) = -\sin x, \quad f^{(3)}(x) = -\cos x \\
&\text{(b)} \quad f'(0) = 1, \quad f''(0) = 0, \quad f^{(3)}(0) = -1 \\
&\text{(c)} \quad f(x) = x - \frac{1}{6}x^3 + R
\end{aligned}
$$

(4) $f(x) = \cos x$

$$
\begin{aligned}
&\text{(a)} \quad f'(x) = -\sin x, \quad f''(x) = -\cos x, \quad f^{(3)}(x) = \sin x \\
&\text{(b)} \quad f'(0) = 0, \quad f''(0) = -1, \quad f^{(3)}(0) = -0 \\
&\text{(c)} \quad f(x) = 1 - \frac{1}{2}x^2 + R
\end{aligned}
$$

(5) $f(x) = \sinh x$

$$
\begin{aligned}
&\text{(a)} \quad f'(x) = \cosh x, \quad f''(x) = \sinh x, \quad f^{(3)}(x) = \cosh x \\
&\text{(b)} \quad f'(0) = 1, \quad f''(0) = 0, \quad f^{(3)}(0) = 1 \\
&\text{(c)} \quad f(x) = x + \frac{1}{6}x^3 + R
\end{aligned}
$$

(6) $f(x) = \tan x$

$$\text{(a)}\quad f'(x) = \sec^2 x, \quad f''(x) = 2\sec^2 \tan x,$$
$$f^{(3)}(x) = 2\sec^4 x + 4\sec^2\tan^2 x$$
$$\text{(b)}\quad f'(0) = 1, \quad f''(0) = 0, \quad f^{(3)}(0) = 2$$
$$\text{(c)}\quad f(x) = 1 + \frac{1}{3}x^3 + R$$

1.2.

$\frac{\partial^2 f}{\partial x^2}(x,y)$ 等の (x,y) は省略可.

(1) $f(x,y) = x^2 + y^2 + 2x^2y^2$

$$\left(\frac{\partial f}{\partial x}(x,y) = 2x + 4xy^2, \quad \frac{\partial f}{\partial y}(x,y) = 2y + 4x^2y\right)$$

(a) $\dfrac{\partial^2 f}{\partial x^2}(x,y) = 4y^2 + 2$ (b) $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = 8xy$ (c) $\dfrac{\partial^2 f}{\partial y^2}(x,y) = 4x^2 + 2$

(2) $f(x,y) = x^2 + 2x + y^2 - 1$

$$\left(\frac{\partial f}{\partial x}(x,y) = 2x + 2, \quad \frac{\partial f}{\partial y}(x,y) = 2y\right)$$

(a) $\dfrac{\partial^2 f}{\partial x^2}(x,y) = 2$ (b) $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = 0$ (c) $\dfrac{\partial^2 f}{\partial y^2}(x,y) = 2$

(3) $f(x,y) = \sin x \cos y$

$$\left(\frac{\partial f}{\partial x} = \cos x \cos y, \quad \frac{\partial f}{\partial y} = -\sin x \sin y\right)$$

(a) $\dfrac{\partial^2 f}{\partial x^2}(x,y) = -\sin x \cos y$ (b) $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = -\cos x \sin y$

(c) $\dfrac{\partial^2 f}{\partial y^2}(x,y) = -\sin x \cos y$

(4) $f(x,y) = \sin xy$

$$\left(\frac{\partial f}{\partial x} = y\cos(xy), \quad \frac{\partial f}{\partial y} = x\cos(xy)\right)$$

(a) $\dfrac{\partial^2 f}{\partial x^2}(x,y) = -y^2 \sin(xy)$ (b) $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = \cos(xy) - xy\sin(xy)$

(c) $\dfrac{\partial^2 f}{\partial y^2}(x,y) = -x^2 \sin(xy)$

(5) $f(x, y) = \tan(x, y)$

$$\left(\frac{\partial f}{\partial x} = \sec^2(x+y), \quad \frac{\partial f}{\partial y} = \sec^2(x+y)\right)$$

$$\text{(a), (b), (c)}\ \frac{\partial^2 f}{\partial x^2}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y^2}(x, y) = \tan(x+y)\sec^2(x+y)$$

(6) $f(x, y) = e^{x+y}$

$$\left(\frac{\partial f}{\partial x} = e^{x+y}, \quad \frac{\partial f}{\partial y} = e^{x+y}\right)$$

$$\text{(a), (b), (c)}\ \frac{\partial^2 f}{\partial x^2}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y^2}(x, y) = e^{x+y}$$

(7) $f(x) = e^{-\frac{x^2+y^2}{2}}$

$$\left(\frac{\partial f}{\partial x}(x, y) = -xe^{-\frac{x^2+y^2}{2}}, \quad \frac{\partial f}{\partial y}(x, y) = -ye^{-\frac{x^2+y^2}{2}}\right)$$

$$\text{(a)}\ \frac{\partial^2 f}{\partial x^2}(x, y) = -e^{-\frac{x^2+y^2}{2}} + (-x)^2 e^{-x^2-y^2} = \left(x^2 - 1\right) e^{-\frac{x^2+y^2}{2}}$$

$$\text{(b)}\ \frac{\partial^2 f}{\partial x \partial y}(x, y) = xye^{-\frac{x^2+y^2}{2}} \qquad \text{(c)}\ \frac{\partial^2 f}{\partial y^2}(x, y) = \left(y^2 - 1\right) e^{-\frac{x^2+y^2}{2}}$$

(8) $f(x, y) = \log(xy)$

$$\left(\frac{\partial f}{\partial x}(x, y) = y\frac{1}{xy} = \frac{1}{x}, \quad \frac{\partial f}{\partial y}(x, y) = \frac{1}{y}\right)$$

$$\text{(a)}\ \frac{\partial^2 f}{\partial x^2}(x, y) = -\frac{1}{x^2} \quad \text{(b)}\ \frac{\partial^2 f}{\partial x \partial y}(x, y) = 0 \quad \text{(c)}\ \frac{\partial^2 f}{\partial y^2}(x, y) = -\frac{1}{y^2}$$

1.3.

(1) $f(x, y) = \cos x \cos y$

$$\text{(a)}\ \frac{\partial f}{\partial x}(0, 0) = \left.\frac{\partial f}{\partial x}\right|_{(0,0)} = \left.-\sin x \cos y\right|_{(0,0)} = 0$$

$$\text{(b)}\ \frac{\partial f}{\partial y}(0, 0) = \left.\frac{\partial f}{\partial y}\right|_{(0,0)} = \left.-\cos x \sin y\right|_{(0,0)} = 0$$

$$\text{(c)}\ \frac{\partial^2 f}{\partial x^2}(0, 0) = \left.\frac{\partial^2 f}{\partial x^2}\right|_{(0,0)} = \left.-\cos x \cos y\right|_{(0,0)} = -1$$

$$\text{(d)}\ \frac{\partial^2 f}{\partial x \partial y}(0, 0) = \left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(0,0)} = \left.\sin x \sin y\right|_{(0,0)} = 0$$

$$\text{(e)}\ \frac{\partial^2 f}{\partial y^2}(0, 0) = \left.\frac{\partial^2 f}{\partial y^2}\right|_{(0,0)} = \left.\cos x \cos y\right|_{(0,0)} = -1$$

(2) $f(x) = e^{-\frac{x^2+y^2}{2}}$

$$\text{(a)}\ \frac{\partial f}{\partial x}(0,0) = -xe^{-\frac{x^2+y^2}{2}}\bigg|_{(0,0)} = 0$$

$$\text{(b)}\ \frac{\partial f}{\partial y}(0,0) = -ye^{-\frac{x^2+y^2}{2}}\bigg|_{(0,0)} = 0$$

$$\text{(c)}\ \frac{\partial^2 f}{\partial x^2}(0,0) = \left(x^2-1\right)e^{-\frac{x^2+y^2}{2}}\bigg|_{(0,0)} = -1$$

$$\text{(d)}\ \frac{\partial^2 f}{\partial x \partial y}(0,0) = xye^{-\frac{x^2+y^2}{2}}\bigg|_{(0,0)} = 0$$

$$\text{(c)}\ \frac{\partial^2 f}{\partial y^2}(0,0) = \left(y^2-1\right)e^{-\frac{x^2+y^2}{2}}\bigg|_{(0,0)} = -1$$

1.4.

(1) $f(x,y) = x^2 + y^2 + 2x^2y^2$

(a) 次の 2 式より $(0,0)$ は停留点である.

$$\frac{\partial f}{\partial x}(0,0) = 2x + 4xy^2\bigg|_{(0,0)} = 0, \quad \frac{\partial f}{\partial y}(0,0) = 2y + 4x^2y\bigg|_{(0,0)} = 0$$

(b) 次の結果より, (0, 0) は極小点である.

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 4y^2 + 2\bigg|_{(0,0)} = 2, \quad \frac{\partial^2 f}{\partial x \partial y}(0,0) = 8xy\bigg|_{(0,0)} = 0$$

$$\frac{\partial^2 f}{\partial y^2}(0,0) = 4x^2 + 2\bigg|_{(0,0)} = 2$$

(2) $f(x,y) = x^2 + 2x + y^2 - 1$

(a) 次の式により, $(0,0)$ は停留点ではない.

$$\frac{\partial f}{\partial x}(0,0) = 2x + 2\bigg|_{(0,0)} = 2 \neq 0$$

(3) $f(x,y) = \sin x \cos y$

(a) 次の式により, $(0,0)$ は停留点ではない.

$$\frac{\partial f}{\partial x}(0,0) = \cos x \cos y\bigg|_{(0,0)} = 1 \neq 0$$

(4) $f(x,y) = \sin xy$

(a) 次の結果より, $(0,0)$ は停留点である.

$$\frac{\partial f}{\partial x}(0,0) = y\cos(xy)\Big|_{(0,0)} = 0, \quad \frac{\partial f}{\partial y}(0,0) = x\cos(xy)\Big|_{(0,0)} = 0$$

(b)

$$\frac{\partial^2 f}{\partial x^2}(0,0) = -y^2\sin(xy)\Big|_{(0,0)} = 0, \quad \frac{\partial^2 f}{\partial x \partial y}(0,0) = \cos(xy) - xy\sin(xy) = 1$$

$$\frac{\partial^2 f}{\partial y^2}(0,0) = -x^2\sin(xy)\Big|_{(0,0)} = 0$$

$$\therefore \frac{\partial^2 f}{\partial x^2}(0,0)(\Delta x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(0,0)\Delta x\Delta y + \frac{\partial^2 f}{\partial y^2}(0,0)(\Delta y)^2$$

$$= 2\Delta x\Delta y$$

これは $\Delta x\Delta y > 0$ の時に正, $\Delta x\Delta y < 0$ の時に負. すなわち, 鞍点である.

鞍点は, Hesse 行列の固有値を用いると明確に判定できる.

(5) $f(x,y) = \tan(x+y)$

(a) 次の結果より, $(0,0)$ は停留点ではない.

$$\frac{\partial f}{\partial x}(0,0) = \sec^2(x+y)\Big|_{(0,0)} = 1 \neq 0$$

(6) $f(x,y) = e^{x+y}$

(a) 次の結果より, $(0,0)$ は停留点ではない.

$$\frac{\partial f}{\partial x}(0,0) = e^{x+y}\Big|_{(0,0)} = 1 \neq 0$$

(7) $f(x,y) = e^{-\frac{x^2+y^2}{2}}$

(a)

$$\frac{\partial f}{\partial x}(0,0) = -xe^{-\frac{x^2+y^2}{2}}\Big|_{(0,0)} = 0, \quad \frac{\partial f}{\partial y}(0,0) = -ye^{-\frac{x^2+y^2}{2}}\Big|_{(0,0)} = 0$$

従って $(0,0)$ は停留点である.

(b)

$$\frac{\partial^2 f}{\partial x^2}(0,0) = \left(x^2 - 1\right) e^{-\frac{x^2+y^2}{2}}\bigg|_{(0,0)} = -1, \quad \frac{\partial^2 f}{\partial x \partial y}(0,0) = xye^{-\frac{x^2+y^2}{2}}\bigg|_{(0,0)} = 0$$

$$\frac{\partial^2 f}{\partial y^2}(0,0) = \left(y^2 - 1\right) e^{-\frac{x^2+y^2}{2}}\bigg|_{(0,0)} = -1$$

$$\therefore \frac{\partial^2 f}{\partial x^2}(0,0)(\Delta x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(0,0)\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(0,0)(\Delta y^2$$

$$= -(\Delta x)^2 - (\Delta y)^2 < 0$$

従って, (0, 0) は極大点である.

(8) $f(x, y) = \log(xy)$

(a) そもそも $f(x, y)$ が $(0, 0)$ で定義されていないので, $(0, 0)$ は停留点ではない.

1.5.

(1) $\quad \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} = [x,\ y] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [x,\ y] \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + y^2$

(2) $\quad \boldsymbol{x}^{\mathrm{T}} B \boldsymbol{x} = [x,\ y] \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [x,\ y] \begin{bmatrix} x + y \\ y \end{bmatrix} = x^2 + xy + y^2$

(3) $\quad \boldsymbol{x}^{\mathrm{T}} C \boldsymbol{x} = [x,\ y] \begin{bmatrix} \frac{\pi}{2} & -\frac{\pi}{2} \\ \frac{\pi}{2} & \frac{\pi}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [x,\ y] \begin{bmatrix} \frac{\pi}{2}x - \frac{\pi}{2}y \\ \frac{\pi}{2}x + \frac{\pi}{2}y \end{bmatrix} = \frac{\pi}{2}x^2 + \frac{\pi}{2}y^2$

(4) $\quad \boldsymbol{x}^{\mathrm{T}} D \boldsymbol{x} = [x,\ y] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [x,\ y] \begin{bmatrix} y \\ x \end{bmatrix} = 2xy$

1.6

$$(1)\quad \boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = \begin{bmatrix} a & b \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = a^2 + b^2$$

$$(2)\quad \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}} = \begin{bmatrix} a \\ b \end{bmatrix}\begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}$$

$$(3)\quad AG = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} a+4b & 2a+5b & 3a+6b \\ c+4d & 2c+5d & 3c+6d \\ e+4f & 2e+5f & 3e+6f \end{bmatrix}$$

$$(4)\quad GA = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a+2c+3e & b+2d+3f \\ 4a+5c+6e & 4b+5d+6f \end{bmatrix}$$

$$(5)\quad AA^{\mathrm{T}} = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} a & c & e \\ b & d & f \end{bmatrix} = \begin{bmatrix} a^2+b^2 & ac+bd & ae+bf \\ ca+db & c^2+d^2 & ce+df \\ ea+fb & ec+fd & e^2+d^2 \end{bmatrix}$$

1.7.

A の固有値は $4, 2$, よって A は正定値. B の固有値は $4, 2$, よって B は正定値. C の固有値は $-2, -4$, よって, C は負定値. D の固有値は $4, -2$, よって, どちらでもない.

1.8.

A: $\lambda^2 - 10\lambda + 21 = 0, \quad \lambda = 5 \pm 2$, 固有値が共に正なので正定値.
B: $\lambda^2 - 2\lambda - 9 = 0, \quad \lambda = -1 \pm \sqrt{10}$, 固有値が正と負なので, どちらでもない.
C: $\lambda^2 + 2\lambda - 12 = 0, \quad \lambda = -1 \pm \sqrt{13}$, 固有値が正と負なので, どちらでもない.
D: $\lambda^2 - 5 = 0, \quad \lambda = \pm\sqrt{5}$, 固有値が正と負なので, どちらでもない.

1.9.

$$f(x, y) = \log(xy) - \frac{1}{3}\left(x^3 + y^3\right)$$

$$\frac{\partial f}{\partial x}(1,1) = \frac{1}{x} - x^2\bigg|_{(1,1)} = 0, \quad \frac{\partial f}{\partial y}(1,1) = \frac{1}{y} - y^2\bigg|_{(1,1)} = 0$$

$$\frac{\partial^2 f}{\partial x^2}(1,1) = -\frac{1}{x^2} - 2x\Big|_{(1,1)} = -1,$$
$$\frac{\partial^2 f}{\partial y^2}(1,1) = -\frac{1}{y^2} - 2y\Big|_{(1,1)} = -1$$
$$\frac{\partial^2 f}{\partial x \partial y}(1,1) = 0$$

$$H(1,1) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

固有値はともに -1 (重複) であるから, $(x,y) = (1,1)$ は極大点である.

1.10.

$$\frac{\partial f}{\partial x}(0,0,0) = 4x - y - z\Big|_{(0,0,0)} = 0, \quad \frac{\partial f}{\partial y}(0,0,0) = 4y - z - x\Big|_{(0,0,0)} = 0$$
$$\frac{\partial f}{\partial z}(0,0,0) = 4z - x - y\Big|_{(0,0,0)} = 0$$

従って, 停留点である.

$$\frac{\partial^2 f}{\partial x^2}(0,0,0) = 4, \quad \frac{\partial^2 f}{\partial y^2}(0,0,0) = 4, \quad \frac{\partial^2 f}{\partial z^2}(0,0,0) = 4$$
$$\frac{\partial^2 f}{\partial x \partial y}(0,0,0) = -1, \quad \frac{\partial^2 f}{\partial y \partial z}(0,0,0) = -1, \quad \frac{\partial^2 f}{\partial z \partial x}(0,0,0) = -1$$

ヘッセ行列は

$$H(0,0,0) = \begin{bmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{bmatrix}$$

ヘッセ行列の固有値は, $5, 5, 2$ で全て正, 従って $(x,y,z) = (0,0,0)$ は極小点である.

1.11.

$$f(x,y) = (x + y - 3)^2 + 2(x-1)^2 + (y-1)^2$$

$$\frac{\partial f}{\partial x}(x,y) = 6x + 2y - 10 = 0$$
$$\frac{\partial f}{\partial y}(x,y) = 2x + 4y - 8 = 0$$

この 2 つの方程式を連立に解いて, 停留点は $(x,y)=\left(\frac{6}{5},\frac{7}{5}\right)$.

$$\frac{\partial^2 f}{\partial x^2}\left(\frac{3}{5},\frac{7}{10}\right)=6,\quad \frac{\partial^2 f}{\partial y^2}\left(\frac{3}{5},\frac{7}{10}\right)=4$$
$$\frac{\partial^2 f}{\partial x\partial y}\left(\frac{3}{5},\frac{7}{10}\right)=2$$

$$H\left(\frac{6}{5},\frac{7}{5}\right)=\begin{bmatrix}6 & 2\\ 2 & 4\end{bmatrix}$$

Hesse 行列の固有値は $5\pm\sqrt{5}$. 固有値がともに正であるから, 点 $\left(\frac{6}{5},\frac{7}{5}\right)$ は極小点である.

演習問題 2 解答

2.1 A は明らかに (a).

B: 固有方程式は $\lambda^2-4\lambda+5=0$, $\lambda=2\pm i=\sqrt{5}\left(2/\sqrt{5}\pm i/\sqrt{5}\right)$, 固有値は

$$\begin{aligned}\lambda=2\pm i&=\sqrt{5}\left(\frac{2}{\sqrt{5}}\pm\frac{i}{\sqrt{5}}\right)\\ &=\sqrt{5}\,(\cos\theta\pm i\sin\theta)\end{aligned}$$

この情報を元に変形すると

$$B=\begin{bmatrix}\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\\ \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\end{bmatrix}\begin{bmatrix}\sqrt{5} & 0\\ 0 & \sqrt{5}\end{bmatrix}=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\sqrt{5} & 0\\ 0 & \sqrt{5}\end{bmatrix}$$

ここで $\cos\theta=\frac{2}{\sqrt{5}}$ $\sin\theta=-\frac{1}{\sqrt{5}}$, これは, 回転と $\sqrt{5}$ 倍の組み合わせなので (c)

C の固有値は 3, 6. 従って, (a)

D の固有方程式は, $\lambda^2-9\lambda+22$, 固有値は $9/2\pm i\sqrt{7}/2$, 従って, D は回転を含んでいる. しかし, 形が回転行列の定数倍にはなっていないので, (b) でも (c) でもない. よって (d).

G は 2 次の Jordan セルなので (d).

2.2

対角化ベクトルは固有値を対角に並べたものだから

$$J=\begin{bmatrix}3 & 0\\ 0 & 2\end{bmatrix}$$

対角化に際して, 固有ベクトルを並べた行列を $P = [\boldsymbol{p}_1, \boldsymbol{p}_2]$ とおくと,

$$A = PJP^{-1}$$

固有ベクトルの座標での表現は

$$\boldsymbol{\alpha} = P^{-1}\boldsymbol{a} = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{3}}{2} \\ \frac{-1+\sqrt{3}}{2} \end{bmatrix}$$

$$\boldsymbol{\beta} = J\boldsymbol{\alpha} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{1+\sqrt{3}}{2} \\ \frac{-1+\sqrt{3}}{2} \end{bmatrix} = \begin{bmatrix} \frac{3+3\sqrt{3}}{2} \\ -1+\sqrt{3} \end{bmatrix}$$

$$[\boldsymbol{p}_1, \boldsymbol{p}_2]\,\boldsymbol{\beta} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} \frac{3+3\sqrt{3}}{2} \\ -1+\sqrt{3} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}+11}{4} \\ \frac{\sqrt{3}+9}{4} \end{bmatrix}$$

$$\boldsymbol{b} = A\boldsymbol{a} = \begin{bmatrix} \frac{11}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{9}{4} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{11+\sqrt{3}}{4} \\ \frac{\sqrt{3}+9}{4} \end{bmatrix}$$

2.3

A: 固有値は $\lambda^2 - 13\lambda + 36 = 0$ の根だから $\lambda_1 = 4$, $\lambda_2 = 9$ である. 対応する固有ベクトル（基準化しない）は, $\boldsymbol{p}_1 = [1, 1]^{\mathrm{T}}$, $\boldsymbol{p}_2 = [1, 2]^{\mathrm{T}}$, $P = [\boldsymbol{p}_1, \boldsymbol{p}_2]$ を用いて PJ_AP^{-1} と対角化し, 対角化行列の平方根をとって, $P\sqrt{J_A}P^{-1}$ で平方根が計算できる.

$$A = PJ_AP^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

$$\sqrt{A} = P\sqrt{J_A}P^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$$

B: 固有値は $\lambda^2 - 5\lambda + 4 = 0$ の根だから $\lambda_1 = 1$, $\lambda_2 = 4$ である. 対応する固有ベクトル（基準化しない）は, $\boldsymbol{p}_1 = [2, 1]^{\mathrm{T}}$, $\boldsymbol{p}_2 = [3, 2]^{\mathrm{T}}$, $P = [\boldsymbol{p}_1, \boldsymbol{p}_2]$ を用いて PJ_BP^{-1} と対角化し, 対角化行列の平方根をとって, $P\sqrt{J_B}P^{-1}$ で平方根が計算

できる.

$$
\begin{aligned}
B = PJ_BP^{-1} &= \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \\
\sqrt{B} = P\sqrt{J_B}P^{-1} & \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} -2 & 6 \\ -2 & 5 \end{bmatrix}
\end{aligned}
$$

2.4

回転を表す行列は $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ の形をしているので, $\cos^2\theta + \sin^2\theta = 1$ を使って変形する.

$$3^2 + 4^2 = 25$$

よって,

$$C = 5 \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

$\cos\theta = \frac{3}{5}$, $\sin\theta = -\frac{4}{5}$ 拡大率は 5.

2.5

$$P^{-1} = \frac{1}{(-4)} \begin{bmatrix} 0 & 1 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{4} \\ -1 & -\frac{1}{2} \end{bmatrix}$$

$$
\begin{aligned}
P^{-1}DP &= \begin{bmatrix} 0 & -\frac{1}{4} \\ -1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 & -1 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -4 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & -\frac{1}{4} \\ -1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 12 & -4 \\ -24 & -4 \end{bmatrix} \\
&= \begin{bmatrix} 6 & 1 \\ 0 & 6 \end{bmatrix}
\end{aligned}
$$

固有方程式は

$$\lambda^2 - 12\lambda + 36 = (\lambda - 6)^2 = 0$$

従って, 固有値は $\lambda_1 = \lambda_2 = 6$ (重解). 固有ベクトルは,

$$A\boldsymbol{p}_1 = \lambda \boldsymbol{p}_1$$

$4x - y = 6x$ より, $2x = -y$, これより, 例えば

$$\boldsymbol{p}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

よって, 固有値は 6, 固有ベクトルは $[1, -2]^{\mathrm{T}}$

(以上で十分だが, 以下, 参考までに一般化固有ベクトルの求め方)

$$(A - \lambda I)\boldsymbol{p}_2 = \boldsymbol{p}_1$$

ここで, λ は固有値, I は単位ベクトル. $\boldsymbol{p}_2 = [p_{21}, p_{22}]^{\mathrm{T}}$ として

$$\begin{bmatrix} -2 & -1 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} p_{21} \\ p_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$-2p_{21} - p_{22} = 1, \quad 4p_{21} + 2p_{22} = -2$$

これを満たす p_{21}, p_{22} として例えば $p_{21} = -1/2$, $p_{22} = 0$ (この場合, $\boldsymbol{p}_1$ との関係があるので, ベクトルの大きさを自由に選ぶことはできない.) よって, $\boldsymbol{p}_1 = [1, -2]^{-1}$ のとき, (もうひとつの) 一般化固有ベクトルは

$$\boldsymbol{p}_2 = \begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix}$$

問題で与えられている P は $P = 2\,[\boldsymbol{p}_1, \boldsymbol{p}_2]$ は, (一般化) 固有ベクトルを並べた行列である. これと, これの逆行列を使って, $P^{-1}DP$ を計算すると, Jordan 標準形が得られる.

2.6

次の行列を右極分解せよ.

$$A = \begin{bmatrix} 5 & -2 \\ 10 & 11 \end{bmatrix} \tag{4.4}$$

$$M := A^{\mathrm{T}}A = \begin{bmatrix} 5 & 10 \\ -2 & 11 \end{bmatrix} \begin{bmatrix} 5 & -2 \\ 10 & 11 \end{bmatrix} = \begin{bmatrix} 125 & 100 \\ 100 & 125 \end{bmatrix}$$

M の固有方程式は

$$\lambda^2 - 250\lambda + 5625$$

より, 固有値は $\lambda_1 = 25$, $\lambda_2 3 = 225$, 対応する固有ベクトルは, $\boldsymbol{p}_1 = [1, -1]^{\mathrm{T}}$, $\boldsymbol{p}_2 = [1, 1]^{\mathrm{T}}$, これより M は以下の様に対角化できる.

$$M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 25 & 0 \\ 0 & 225 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

$S := M^{1/2}$ は次のように計算できる.

$$\begin{aligned} S := M^{1/2} &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 25 & 0 \\ 0 & 225 \end{bmatrix}^{1/2} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 15 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 10 & 5 \\ 5 & 10 \end{bmatrix} \end{aligned}$$

さらに

$$\begin{aligned} S^{-1} &= \begin{bmatrix} 2/15 & -1/15 \\ -1/15 & 2/15 \end{bmatrix} \\ U = AS^{-1} &= \begin{bmatrix} -2 & 5 \\ 11 & 10 \end{bmatrix} \begin{bmatrix} 2/15 & -1/15 \\ -1/15 & 2/15 \end{bmatrix} \\ &= \begin{bmatrix} 4/5 & -3/5 \\ 3/5 & 4/5 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \end{aligned}$$

ここで, $\cos\theta = 4/5$, $\sin\theta = 3/5$

演習問題 3 解答

3.1

$$f(t) = 1 - (kt) + \frac{1}{2}(kt)^2 - \frac{1}{6}(kt)^3 + R$$

3.2 1 階線形微分方程式の公式より

$$x = e^{-kt}$$

3.3 式 (3.9) を使って

$$\boldsymbol{x}(t) = e^{tA}\boldsymbol{x}_0 = \begin{bmatrix} e^{3t} & 0 \\ 0 & e^{-t} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} e^{3t} \\ e^{-t} \end{bmatrix}$$

3.4

$$\begin{aligned} e^{tA} &= Pe^{tJ}P^{-1} \\ &= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{-\frac{1}{2}t} & 0 \\ 0 & e^{-\frac{1}{4}t} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{-(1/2)t} & e^{-(1/2)t} \\ e^{-(1/4)t} & -e^{-(1/4)t} \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} e^{-(1/2)t} + e^{-(1/4)t} & e^{-(1/2)t} - e^{-(1/4)t} \\ e^{-(1/2)t} - e^{-(1/4)t} & e^{-(1/2)t} - e^{-(1/4)t} \end{bmatrix} \end{aligned}$$

3.5

(1)固有値は　2, 1 なので, (a)（または (c)）となる.
(2)固有値は 2, 0 で, 一つの固有値が 0 なので, 一番近いのは (d) となる.
(3)固有値は　1, 4 なので, (a)（または (c)）となる.
(4)固有値は $\pm\sqrt{2}$ で片方が正, もう片方が負なので, 一番近いのは (e) となる.

3.6

(1)固有値は　$4 \pm 6i$ で, 実部が正であるから, (a) が近い.
(2)固有値は $-1.5 \pm 3.42i$ で, 実部が負なので, (c) が近い.
(3)固有値は $\pm 2i$ で, 実部が 0 なので, (b) が近い.
(4)固有値は $\pm\sqrt{2}i$ で, 実部が 0 なので, (b) が近い.

3.7

(1)固有値は　1, -1　（正と負）なので, 図 3.11 の (e) が最も近い.
(2)固有値は　$\pm i$ なので, 図 3.13 の (b) が最も近い.

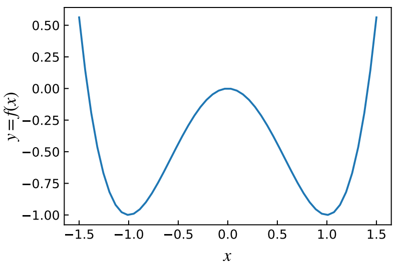
演習問題 4 解答

4.1 結果は次のとおり

表 4.4　問題 4.1 の解答

i	Newton				Gradient descent		
	x_i	$f'(x_i)$	$f''(x_i)$	Δx_i	x_i	$f'(x_i)$	Δx_i
0	2.0000	8.0000	36.0000	-0.2222	2.0000	8.0000	-0.2640
1	1.7778	1.1413	25.9259	-0.0440	1.7360	0.0951	-0.0031
2	1.7338	0.0410	24.0709	-0.0017	1.7329	0.0195	-0.0006
3	1.7321	0.0001	24.0001	-0.0000	1.7322	0.0040	-0.0001
4	1.7321				1.7321		

4.2 まず, 図は以下の通り



Newton 法であるから, 変曲点を境に収束点が決まる. 勾配降下法のように発散することはない. よって, 次のようになる.

- $x < -1/\sqrt{3}$ の場合, 一番左の停留点 $x = -1$ に収束.
- $-1/\sqrt{3} < x < 1/\sqrt{3}$ の場合, 停留点 $x = 0$ に収束.
- $1/\sqrt{3} < x$ の場合, 停留点 $x = 1$ に収束