

第 5 章

ベクトルと行列の数学的性質

前章で確認したように、ベクトルや行列は複数の要素から構成されるものであるため、その「大きさ」を表現するには、実数や複素数における絶対値と同じ性質を持つ関数、即ちノルムが必要となる。本章では今後使用する具体的なノルムを解説する。また、絶対値やノルムに基づいて定義される「誤差」の定義も合わせて述べる。

5.1 IEEE754-1985 単精度・倍精度浮動小数点数のデータ構造

現在の PC や WS で良く用いられているのは IEEE で規格化されているもので、一般には IEEE754 standard と呼ばれているものである。

この IEEE754 規格の浮動小数点数の基数は 2 であり、図 5.1 に示すような有限桁の浮動小数点表現が定められている。単精度・倍精度の仮数部が実際の bit 列より 1bit 分長いのは、最初の bit が必ず 1 になるように正規化されるため、実際にはその分は削ってあるからである。これをケチ表現と呼ぶ。C や

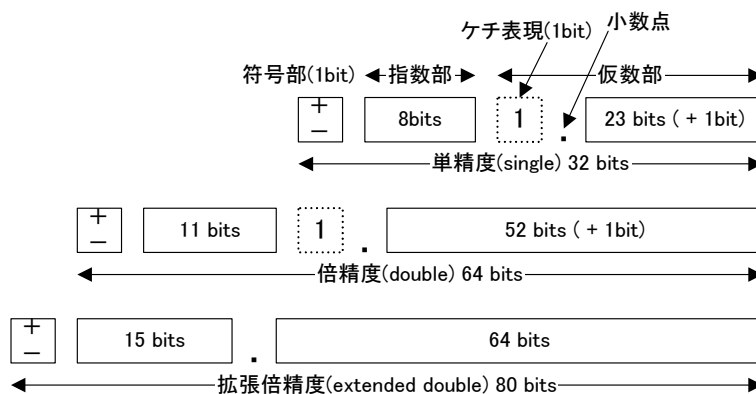


図 5.1 IEEE754 standard

Fortran 等のプログラミング言語で単精度・倍精度の値を 10 進数の文字列として出力すると、例えば 10.65387 は

10.65387

0.1065387e+02
 1.065387e+01
 1.065387e+1
 1.065387e1
 1.065387E+01
 1.065387D+01
 1.065387Q+01

のようになる。e や E, D, Q の前が 10 進数に変換後の仮数部，後ろが指数部を示している。

有限桁の浮動小数点数は離散的に存在するため，表現可能な数の間には隙間ができる。この隙間は指数部が大きくなるにつれて，絶対的な距離は大きくなるが相対的にはほぼ同じという特徴がある。この間隔をマシンイプシロン (machine epsilon) と呼ぶ。これにも幾つかの定義方法があるが，本書では以下のよう

定義 5.1 (マシンイプシロン)

1 の次に大きい最小の正の有限桁の浮動小数点数を $1 + \varepsilon_M$ と表現する時，この ε_M をマシンイプシロン (machine epsilon) と呼ぶ。特に基数 m で仮数部が p 桁の場合，

$$\varepsilon_M = m^{-(p-1)} \quad (5.1)$$

となる。

従って，IEEE754 standard の場合のマシンイプシロンは

単精度 $2^{-23} \approx 1.1920928955078125 \times 10^{-7}$

倍精度 $2^{-52} \approx 2.2204460492503130808472633361816 \times 10^{-16}$

となる。

この IEEE754 単精度，倍精度浮動小数点表現を用いてコンピュータで実行される計算には，マシンイプシロン程度の丸め誤差 (round-off error) が含まれているものと想定しておく，ゴミのような数値が計算結果の下桁に現れることも納得できる筈である。

5.2 絶対値と丸め誤差

絶対値の数学的性質については既に述べた。

本書で扱うコンピュータ上の線型計算は，前述したように有限桁の浮動小数点数に基づくものである。従って，計算過程において，短い桁に「丸め」られることが多い。そのため，実際に得られた数値 \tilde{a} と，真の値 a とのズレが生じる。このズレを誤差 (error) と呼び，誤差を伴う数値 \tilde{a} を真値 a に対する近似値 (approximation) と呼ぶ。

例えば，円周率 $\pi = 3.14159265\dots$ という無限小数で表現するほかない実数は， π が持つ数学的性質を持った特殊なデータとして記号的に扱うか，有限桁の小数として表現する他ない。前述したように，本書では後者の取り扱い方法のみを扱うことにしているので，以降，コンピュータ上で扱う実数を要素として

持つデータ型（実数型 (float, double), 複素数型 (complex), ベクトル型, 行列型）の要素は有限桁の小数として表現されるものとする。

長い桁数の小数を短い桁の小数に変換することを丸める (rounding, round-off) と呼ぶ。例えば π を有効数字 5 桁の浮動小数点数 $\tilde{\pi}$ に丸める処理は、6 桁目の値を四捨五入して行うことが標準的である。

$$\tilde{\pi} = 3.1416 \approx \pi$$

この場合、真値 (真の値) は π であり、近似値は $\tilde{\pi}$ である。

ここで発生する真値と近似値のずれを誤差 (error) と呼ぶ。ここでは近似値 \tilde{a} に含まれる誤差を

$$E(a) = a - \tilde{a}$$

と表現することにする。

通常、誤差は符号より大きさのみを調べるために用いられる。そこで誤差の絶対値

$$aE(\tilde{a}) = |E(\tilde{a})| = |a - \tilde{a}|$$

を絶対誤差 (absolute error) と呼ぶ。また、近似値に含まれる誤差の割合を調べる際には、相対誤差 (relative error) と呼ばれる次の量 rE が使用される。

$$rE(\tilde{a}) = \begin{cases} \frac{aE(\tilde{a})}{|a|} = \left| \frac{a - \tilde{a}}{a} \right| & (a \neq 0) \\ aE(\tilde{a}) = |a - \tilde{a}| & (a = 0) \end{cases}$$

数学的にはこの定義が良いが、実際には真値 a が不明なことが多い。従って、その場合は \tilde{a} より真値に近い (と保証できる) 近似値 \tilde{a}' を用いて、絶対誤差も相対誤差も、その近似値

$$aE(\tilde{a}) \approx |\tilde{a}' - \tilde{a}|$$

$$rE(\tilde{a}) \approx \begin{cases} \left| \frac{\tilde{a}' - \tilde{a}}{\tilde{a}'} \right| & (\tilde{a}' \neq 0) \\ |\tilde{a}' - \tilde{a}| & (\tilde{a}' \approx 0) \end{cases}$$

で代用する。

この相対誤差 (あるいは相対誤差の近似値) を用いて、近似値の有効桁数 (有効数字, valid digits) を次のように定義できる。10 進数で表現される近似値 \tilde{a} の有効桁数は、相対誤差 $E_{\text{rel}}(\tilde{a})$ を用いて

$$d = -\log_{10} E_{\text{rel}}(\tilde{a})$$

と表現される。例えば、 $\tilde{\pi} = 3.14159$ の場合、真値 π の代わりに $\tilde{\pi}' = 3.1415926535$ を用いると、絶対誤差、相対誤差、(10 進表現の) 有効桁数は

$$aE(\tilde{\pi}) \approx 0.0000026535 = 2.6535 \cdot 10^{-6}$$

$$rE(\tilde{\pi}) \approx E_{\text{abs}}(\tilde{\pi})/\tilde{\pi}' \approx 8.44635 \cdot 10^{-7}$$

$$d = -\log_{10} rE(\tilde{\pi}) \approx 6.073$$

となる。大体 10 進 6 桁ぴったり真値と一致する、と考えてよい。

問題 5.1

1. 自然対数の底 $e = 2.71828\dots$ を真値とする時、その近似値 $\tilde{e} = 2.7183$ と与えられる時、絶対誤差、相対誤差、10 進有効桁数を求めよ。
2. 真値 $a = \sqrt{2000}$ に対して、10 進 10 桁に丸めた値を \tilde{a} とする。この時、絶対誤差、相対誤差、10 進有効桁数を求めよ。

5.3 ベクトルノルム

ベクトルや行列のように、多数の要素のカタマリをそのまま扱うことは難しく、手間のかかることである。この「手間」については後述する。ここではこの手間を省くため、実数や複素数における絶対値と同じ性質を持った量、即ちノルム (norm) について述べる。

ノルムの定義は次のようになる。

定義 5.2 (ノルム (norm))

a はベクトル、あるいは行列とする。この時、次のような性質を持つ正の実数への写像 $\|\cdot\|$ をノルムと呼ぶ。

1. 任意の a に対して $\|a\| \geq 0$ である。特に $a = 0$ の時のみ $\|a\| = 0$ となる。
2. 任意の複素数 (実数) c に対して

$$\|ca\| = |c| \cdot \|a\|$$

が成立する。

3. 任意の a, b に対して三角不等式が成立する。

$$\|a + b\| \leq \|a\| + \|b\|$$

このノルムはベクトルや行列の「大きさ」を表わす指標として使用される。実用上便利なように、さまざまなノルムが定義されているが、ベクトルノルムとして良く使用されるものとしては次のユークリッドノルム (2 ノルム) がある。

定義 5.3 (ユークリッドノルム (euclid norm), 2 ノルム (2-norm))

$\mathbf{v} \in \mathbb{C}^n$ とする時、ユークリッドノルム $\|\mathbf{v}\|_2$ は次のように定義する。

$$\|\mathbf{v}\|_2 = \sqrt{\sum_{i=1}^n |v_i|^2} \quad (5.2)$$

内積を用いると

$$\|\mathbf{v}\|_2 = \sqrt{(\mathbf{v}, \mathbf{v})}$$

と表現することもできる。

例えば $\mathbf{v} = [-1 \ -2 \ 3]^T \in \mathbb{R}^3$ の時は

$$\|\mathbf{v}\|_2 = \sqrt{(-1)^2 + (-2)^2 + 3^2} = \sqrt{14}$$

となる。

しかし、ユークリッドノルムは演算の回数が多いため、より軽い計算で済む次の二つのベクトルノルムもよく用いられる。

定義 5.4 (1 ノルム (1-norm) と無限大ノルム (infinity norm))

$\mathbf{v} \in \mathbb{C}^n$ とする時, 1 ノルム $\|\mathbf{v}\|_1$ 及び, 無限大ノルム $\|\mathbf{v}\|_\infty$ は次のように定義する。

$$\|\mathbf{v}\|_1 = \sum_{i=1}^n |v_i| \quad (5.3)$$

$$\|\mathbf{v}\|_\infty = \max_{1 \leq i \leq n} |v_i| \quad (5.4)$$

先ほどと同じ $\mathbf{v} = [-1 \ -2 \ 3]^T \in \mathbb{R}^3$ の時は

$$\|\mathbf{v}\|_1 = |-1| + |-2| + |3| = 6$$

$$\|\mathbf{v}\|_\infty = \max(|-1|, |-2|, |3|) = 3$$

となる。複素ベクトルの時も複素数の絶対値を使い, 例えば $\mathbf{a} = [3 + 2i \ 1 - 3i]^T \in \mathbb{C}^2$ の時は

$$\|\mathbf{a}\|_1 = |3 + 2i| + |1 - 3i| = \sqrt{(3^2 + 2^2)} + \sqrt{(1^2 + 3^2)} = \sqrt{13} + \sqrt{10}$$

$$\|\mathbf{a}\|_2 = \sqrt{|3 + 2i|^2 + |1 - 3i|^2} = \sqrt{23}$$

$$\|\mathbf{a}\|_\infty = \max(|3 + 2i|, |1 - 3i|) = \sqrt{13}$$

となる。

これら 3 つのベクトルノルムに対しては, 任意の $\mathbf{v} \in \mathbb{C}^n$ に対してそれぞれある正定数 α_{pq}, β_{pq} が存在し

$$\|\mathbf{x}\|_p \leq \alpha_{pq} \|\mathbf{x}\|_q$$

という不等式が成立する。これを表にすると表 5.1 のようになる。従って, どのノルムを使うにしても,

表 5.1 1 ノルム, ユークリッドノルム, 無限大ノルム間における α_{pq}

$p \rightarrow$ $q \downarrow$	1	2	∞
1	1	\sqrt{n}	n
2	1	1	\sqrt{n}
∞	1	1	1

その違いは高々 $\sqrt{n} \sim n$ 倍程度ということになる。

なお \mathbb{R}^2 空間における「単位円」を各ノルムを用いて描画すると図 5.2 のようになる。

問題 5.2

1. $\mathbf{a} = [-5 \ 2 \ -3 \ -4]^T \in \mathbb{R}^4$ の時, $\|\mathbf{a}\|_1, \|\mathbf{a}\|_2, \|\mathbf{a}\|_\infty$ をそれぞれ求めよ。
2. $\mathbf{b} = [2 + i \ 3i]^T \in \mathbb{C}^2$ の時, $\|\mathbf{b}\|_1, \|\mathbf{b}\|_2, \|\mathbf{b}\|_\infty$ をそれぞれ求めよ。
3. 1 ノルム, ユークリッドノルム, 無限大ノルムがそれぞれノルムの性質 (定義 5.2) を満足することを示せ。

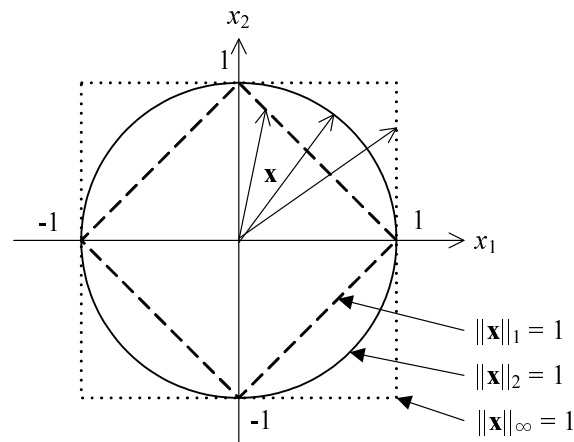


図 5.2 $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$ ノルムを用いた時の単位円

5.4 Scilab におけるベクトルと行列のノルム，絶対誤差，相対誤差の計算

ノルムの計算は `norm` 関数を用いる。ベクトルの場合，今まで学習してきたユークリッドノルム，1 ノルム，無限大ノルムは

```
// ノルム
norm(vec_v)
norm(vec_v, 1)
norm(vec_v, 'inf')
```

と計算する。

行列ノルムに関しても同様に，ユークリッドノルム，1 ノルム，無限大ノルム，フロベニウスノルム (定義は後述) は

```
norm(mat_a)
norm(mat_a, 1)
norm(mat_a, 'inf')
norm(mat_a, 'fro')
```

として計算する

このノルム関数を用いると，ベクトル，行列の誤差の計算ができる。

5.5 ベクトルのユークリッドノルムと「直交」の概念

n 次元ベクトル $\mathbf{v} \in \mathbb{C}^n$ において、ユークリッドノルム $\|\mathbf{v}\|$ は内積を用いて

$$\|\mathbf{v}\| = \sqrt{(\mathbf{v}, \mathbf{v})}$$

と記述することもできる。これをベクトルの長さ (length) と呼ぶことにする。

従って、二つの実ベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ が形成する三角形を考えると、この二つのベクトルがなす角を θ (ラジアン^{*1}) とすると、余弦定理より

$$\frac{1}{2} (-\|\mathbf{u} - \mathbf{v}\|_2^2 + \|\mathbf{u}\|_2^2 + \|\mathbf{v}\|_2^2) = \|\mathbf{u}\|_2 \|\mathbf{v}\|_2 \cos \theta$$

となる。

ここで左辺は

$$\begin{aligned} & \frac{1}{2} (-\|\mathbf{u} + \mathbf{v}\|_2^2 + \|\mathbf{u}\|_2^2 + \|\mathbf{v}\|_2^2) \\ &= \frac{1}{2} \left(-\sum_{i=1}^n |u_i - v_i|^2 + \sum_{i=1}^n |u_i|^2 + \sum_{i=1}^n |v_i|^2 \right) \\ &= \frac{1}{2} \left(\sum_{i=1}^n 2u_i v_i \right) \\ &= \sum_{i=1}^n u_i v_i = (\mathbf{u}, \mathbf{v}) \end{aligned}$$

となり、内積 (\mathbf{u}, \mathbf{v}) である。従って次の定理が成立する。

定理 5.1 (内積とノルムの関係式)

任意の $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ に対して、この二つのベクトルがなす角を θ とする時

$$\cos \theta = \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{u}\|_2 \|\mathbf{v}\|_2} \quad (5.5)$$

が成立する。特に内積が 0 になる時、この二つのベクトルは直交している (orthogonal) と言い、 $\theta = \pi/2$ である。

(5.5) 式を利用すると、二つのベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ のなす角 θ を求めることができる。即ち、逆三角関数 (アークコサイン) $\cos^{-1} x = \theta$ を用いて

$$\theta = \cos^{-1} \left(\frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{u}\|_2 \|\mathbf{v}\|_2} \right) \quad (5.6)$$

を計算すればよい。

問題 5.3

$\mathbf{u} = [1 \ 2]^T, \mathbf{v} = [2 \ 1]^T \in \mathbb{R}^2$ の時、 \mathbf{u} と \mathbf{v} のなす角 θ を Scilab を用いて求めよ。なおアークコサインは $\text{acos}(x)$ 関数を使用すること。

^{*1} 本書では特に断らない限り角度の単位はラジアン (弧度法) を用いる。

5.6 座標系という考え方と正規直交基底

全ての n 次元ベクトルは n 個の要素，即ち座標値を用いて表現されることは既に示した。これは n 個の標準単位ベクトル

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

というベクトルを用いて

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^n v_i \mathbf{e}_i$$

というように，標準単位ベクトルの和として表現できる。

例えば， $\mathbf{v} = [1 \ 2 \ 3]^T \in \mathbb{R}^3$ というベクトルは，

$$\mathbf{v} = \mathbf{e}_1 + 2\mathbf{e}_2 + 3\mathbf{e}_3$$

となる。

この標準単位ベクトルの組 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ は明らかに次の二つの性質，正規性と直交性を持つ。

正規性 長さ（ノルム）が常に 1 である。

$$\|\mathbf{e}_i\|_2 = 1 \quad (i = 1, 2, \dots, n)$$

直交性 異なる単位ベクトルとは必ず直交している。

$$(\mathbf{e}_i, \mathbf{e}_j) = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}$$

このような性質を持つベクトルの組を正規直交基底 (orthonormal bases) と呼ぶ。

なお，一次独立なベクトルは正規直交基底に変換することが出来る。例えば二つの一次独立なベクトル $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ に対して

$$\begin{aligned} \text{(正規化)} \quad \mathbf{u} &= \frac{\mathbf{a}}{\|\mathbf{a}\|_2} = \frac{1}{\|\mathbf{a}\|_2} \mathbf{a} \\ \text{(直交化)} \quad \mathbf{v}' &= \mathbf{b} - (\mathbf{b}, \mathbf{u})\mathbf{u} \\ \text{(正規化)} \quad \mathbf{v} &= \frac{\mathbf{v}'}{\|\mathbf{v}'\|_2} \end{aligned} \tag{5.7}$$

という，正規化（ノルムを 1 にする）と直交化（直交ベクトル成分を抜き出す）を行うことで得られるベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ は正規直交基底をなす。

このように正規化と直交化を繰り返して n 本の一次独立なベクトルを正規直交基底に変換するアルゴリズムについては QR 分解の所で解説する。

問題 5.4

- (5.7) によって生成される \mathbf{u}, \mathbf{v} は正規直交であることを示せ。
- $\mathbf{a} = [1 \ 2]^T, \mathbf{b} = [2 \ 1]^T \in \mathbb{R}^2$ である時, (5.7) の手順に従って正規直交ベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ を生成せよ。

5.7 行列のナチュラルノルム (p ノルム) とフロベニウスノルム

行列 $A \in \mathbb{C}^{m \times n}$ に対してもノルム (定義 5.2) を定義できる。よく使用されるのは, ベクトルノルム ($\|\cdot\|_p, p = 1, 2, \infty$) を用いて定義されるナチュラルノルム (natural norm) である。

定義 5.5 (行列のナチュラルノルム)

ベクトルのノルムが $\|\cdot\|_p$ であるとき, 行列 $A \in \mathbb{C}^{m \times n}$ のナチュラルノルム $\|A\|_p$ を次のように定義する。

$$\|A\|_p = \max_{\mathbf{x} \in \mathbb{C}^n} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} \quad (5.8)$$

$\|\mathbf{x}\|_p = 1$ に限定すると

$$\|A\|_p = \max_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p$$

と定義できる。

このナチュラルノルムは, 次のようにして計算できる。

定理 5.2 (行列のナチュラルノルムの計算法)

行列 $A \in \mathbb{C}^{m \times n}$ に対して, $\|A\|_1, \|A\|_\infty$ は次のように計算できる。

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|$$

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$$

また, $A \in \mathbb{R}^{n \times n}$ の時, A の固有値を $\lambda_i(A)$ とすると, $\|A\|_2$ は

$$\|A\|_2 = \sqrt{\max_{1 \leq i \leq n} |\lambda_i(AA^T)|}$$

となる。特に $A^T = A$ の時は

$$\|A\|_2 = \max_{1 \leq i \leq n} |\lambda_i(A)|$$

となる。

行列要素をベクトル同様に考えてユークリッドノルムのように計算したのも行列ノルムとなる。これをフロベニウス (Frobenius) ノルムと呼ぶ。

定義 5.6 (フロベニウスノルム)

$A \in \mathbb{C}^{m \times n}$ に対して, フロベニウスノルム $\|A\|_F$ を次のように定義する。

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}$$

例えば $A \in \mathbb{R}^{2 \times 2}$ が

$$A = \begin{bmatrix} -1 & -3 \\ -5 & -3 \end{bmatrix}$$

であるとき,

$$\|A\|_1 = \max(|-1| + |-5|, |-3| + |-3|) = \max(6, 6) = 6$$

$$\|A\|_\infty = \max(|-1| + |-3|, |-5| + |-3|) = \max(4, 8) = 8$$

となる。また $\lambda_1(AA^T) = 22 - 2\sqrt{85}$, $\lambda_2(AA^T) = 22 + 2\sqrt{85}$ であるから

$$\|A\|_2 = \sqrt{\max(|\lambda_1(AA^T)|, |\lambda_2(AA^T)|)} = \sqrt{22 + 2\sqrt{85}}$$

となる。また, フロベニウスノルム $\|A\|_F$ は

$$\|A\|_F = \sqrt{|-1|^2 + |-3|^2 + |-5|^2 + |-3|^2} = \sqrt{44}$$

である。

問題 5.5

1. 行列 A が

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

で与えられる時, $\|A\|_1, \|A\|_\infty, \|A\|_F$ をそれぞれ求めよ。

2. 行列 B が

$$B = \begin{bmatrix} i & -i \\ 2+i & 1-3i \end{bmatrix}$$

で与えられる時, $\|B\|_1, \|B\|_\infty, \|B\|_F$ をそれぞれ求めよ。

5.8 ノルム誤差と行列の条件数

実数や複素数同様, ベクトルや行列の誤差, そして絶対誤差, 相対誤差は次のように定義できる。

定義 5.7 (ベクトルの誤差)

$\mathbf{a} \in \mathbb{C}^n$ を真値, $\tilde{\mathbf{a}} \in \mathbb{C}^n$ をその近似値とする。このときノルム $\|\cdot\|$ を用いた時の $\tilde{\mathbf{a}}$ の絶対誤差は

$$aE(\tilde{\mathbf{a}}) = \|\mathbf{a} - \tilde{\mathbf{a}}\| \quad (5.9)$$

とする。更に $\tilde{\mathbf{a}}$ の相対誤差は

$$rE(\tilde{\mathbf{a}}) = \begin{cases} \frac{\|\mathbf{a}-\tilde{\mathbf{a}}\|}{\|\mathbf{a}\|} = \frac{aE(\tilde{\mathbf{a}})}{\|\mathbf{a}\|} & (\mathbf{a} \neq 0) \\ \|\mathbf{a}-\tilde{\mathbf{a}}\| = aE(\tilde{\mathbf{a}}) & (\mathbf{a} = 0) \end{cases} \quad (5.10)$$

とする。

正方行列についても同様に、真値 $A \in \mathbb{C}^{n \times n}$ とその近似値 $\tilde{A} \in \mathbb{C}^{n \times n}$ に対して、 $aE(\tilde{A})$, $rE(\tilde{A})$ が定義できる。

本章の最後に、連立一次方程式の誤差解析において重要なファクターである、行列の条件数を定義する。

定義 5.8 (行列の条件数)

正則行列 $A \in \mathbb{R}^{n \times n}$ の条件数 $\kappa(A)$ を

$$\kappa(A) = \|A\| \cdot \|A^{-1}\|$$

と定義する。1, ユークリッド, 無限大ノルムを用いたものをそれぞれ $\kappa_1(A)$, $\kappa_2(A)$, $\kappa_\infty(A)$ と書く。

Scilab では `cond(行列, ノルムの種類)` 関数で計算ができる。

結論から先に言うと、この $\kappa(A)$ が非常に大きい行列 A を悪条件である (ill-conditioned) と呼び、行列成分、定数ベクトル成分に含まれる初期誤差や、解法の過程で発生した丸め誤差を条件数倍して、数値解に忍び込む可能性が高い、といわれている。これについては後述する。

問題 5.6

次の行列 $A \in \mathbb{R}^{2 \times 2}$ の条件数 $\kappa_1(A)$ 及び $\kappa_\infty(A)$ を求めよ。

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$$