連立１次方程式 III

— 自分が使うための線型代数 —

桂田 祐史

2002 年 11 月 3 日～ 2016 年 10 月 13 日

つぶやき

かって、

私は線形代数を三回勉強した

と言っていたことがあった。大学の講義で一回、それが終わったあとに復習で一回、さらに大学院入試に備えて一回。その過程で、講義や文献に載っていた項目を自分なりにまとめたノートを作った。

そして今、必要を感じて四回目の勉強である。ちょっと驚くのは、今回必要に応じてまとめ始めた (この) ノートと、三回目の勉強で作ったノートの内容がほぼ完全に disjoint であることである。

このノートは講義ノートではないので、必ずしも証明はつけない (ただし証明が何に載っているかはなるべく書くことにする)。また簡潔性もめざさない。使えそうな命題はなるべく沢山収録する。

注意: この文書は書き初めてからしばらく放っておいたので、かなり古くなってしまった。山本 [25] が改訂されたのは大きい。また他のノートでより完全な記述を書いた事項もある。

(本当はいつから書き始めたのか良く覚えていない。2002 年にほぼ現在 (2011/11/6) と同様の内容があったのは分かっている。[25] の改訂は 2003 年で、そのとき「かなり古くなってしまった」とあるので、20 世紀中かなあ。)

この文書の大部分を書いてしまってから、山本 [44], 杉原・室田 [27] など、何冊か注意を払うべき本が出版されている。

目 次

第1章　行列の固有値

1.1　「線形代数」の復習

1.2　Gerschgorin の定理

定理 1.2.1 (Gershgorin, 1931)　$A = (a_{ij}) \in M(n; \mathbf{C})$ に対し

$$\Lambda_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad (i = 1, 2, \cdots, n)$$

とおくと、A の固有値は、円盤の合併

$$\bigcup_{i=1}^{n} \{z \in \mathbf{C}; |z - a_{ii}| \leq \Lambda_i\}$$

に含まれる。

証明　(伊理 [16], 仁木-河野 [26] などを見よ。) ■

系 1.2.2　$A = (a_{ij}) \in M(n; \mathbf{C})$ に対して

$$A \text{ のスペクトル半径} \leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.$$

証明　任意の i について

$$\{z \in \mathbf{C}; |z - a_{ii}| \leq \Lambda_i\} \subset \{z \in \mathbf{C}; |z| \leq |a_{ii}| + \Lambda_i\} = \{z \in \mathbf{C}; |z| \leq \sum_{j=1}^{n} |a_{ij}|\}$$

が成り立つことに注意すれば明らか。■

行列のノルムに関する理論を用いた別証　証明すべき不等式の右辺が $\|A\|_\infty$ で、これはベクトルのノルム $\|\cdot\|_\infty$ と両立しているから、スペクトル半径より大きいか等しい (系 4.3.14 を見よ)。 ■

定理 1.2.3 (Gershgorin の定理 (強い形)) $A=(a_{ij})\in M(n;\mathbf{C})$ の n 個の固有値 $\lambda_1, \cdots, \lambda_n$ (重複固有値は重複度だけ繰り返して数える) は、

$$D_i=\{z\in\mathbf{C}; |z-a_{ii}|\le\Lambda_i\},\quad \Lambda_i=\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|\quad (i=1,2,\cdots,n)$$

で定義される n 個の円盤 $D_1, \cdots, D_n$ の合併集合 $\cup_{i=1}^{n}D_i$ の各連結成分の中に、それを構成する円盤の個数に等しい個数ずつ含まれる。

証明 (伊理 [16] などを見よ。) ■

1.3 行列の関数の固有値

1.3.1 Frobenius の定理

線型代数の教科書 (例えば佐武 [31]) に次の定理がある。

定理 1.3.1 (Frobenius の定理) A を n 次正方行列、A の固有値を (重複度までいれて) $\alpha_1, \alpha_2, \cdots, \alpha_n$, また $f(x)\in\mathbf{C}[x]$ ($\mathbf{C}[x]$ は複素係数多項式全体を表わす) とするとき、$f(A)$ の固有値は (重複度までいれて) $f(\alpha_1), \cdots, f(\alpha_n)$ である。

注意 1.3.2 べき級数 $f(x)=\displaystyle\sum_{i=0}^{\infty}a_nx^n$ の収束半径が r で、$|\alpha_i|<r$ $(i=1,2,\cdots,n)$ ならば、$f(A)$ の固有値は $f(\alpha_1), \cdots, f(\alpha_n)$ である。

例 1.3.3 (巡回行列の固有値)

$$A\stackrel{\text{def.}}{=}\begin{pmatrix} 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ & & \ddots & \ddots & \\ & & & 0 & 1\\ 1 & 0 & \cdots & 0 & 0\end{pmatrix},\quad C\stackrel{\text{def.}}{=}\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1}\\ a_{n-1} & a_0 & a_1 & & a_{n-2}\\ & \ddots & \ddots & \ddots & \\ a_2 & & \ddots & a_0 & a_1\\ a_1 & a_2 & & a_{n-1} & a_0\end{pmatrix}$$

とおけば

$$C=a_0I+a_1A+a_2A^2+\cdots+a_{n-1}A^{n-1}$$

となる。A の固有値は

$$\omega^j\quad (j=0,1,2,\cdots,n-1),\quad \omega=\exp\left(\frac{2\pi\sqrt{-1}}{n}\right)$$

であるから、C の固有値は

$$a_0+a_1\omega^j+a_2\omega^{2j}+\cdots+a_{n-1}\omega^{(n-1)j}\quad (j=0,1,2,\cdots,n-1).■$$

1.3.2　スペクトル写像定理

まず、岩波数学辞典の線形作用素の「作用素解析」の項から引用する。

> T の関数 $f(T)$ を合理的に定義して、複素数値関数 f の代数演算と作用素 $f(T)$ の代数演算との間のある種の同形関係を考察する理論を、一般に**作用素解析** (operational calculus) という。T と f の範囲の選び方により種々の理論があるが、以下代表的なもの 2 つを述べる。1) T を Hilbert 空間 $\mathcal{X}$ の中の自己随伴作用素、$T = \int_{-\infty}^{\infty} \lambda\, dE(\lambda)$ をそのスペクトル分解とする。$f(\lambda)$ を複素数値可測関数とするとき、$\int_{-\infty}^{\infty} |f(\lambda)|^2 d(E(\lambda)x, x) < \infty$ となる $x \in \mathcal{X}$ に対して、$(f(T)x, y) = \int_{-\infty}^{\infty} f(\lambda) d(E(\lambda)x, y)$ によって $f(T)x$ を定義すれば、$f(T)$ は正規作用素になり、次の関係が成り立つ。i) $(\alpha f + \beta g)(T) \supset \alpha f(T) + \beta g(T)$; ii) $(fg)(T) \supset f(T)g(T)$; iii) $f(T)^* = \overline{f}(T)$. とくに例えば g が有界ならば、 i), ii) で等号が成り立つ。 2) $\mathcal{X}$ を複素 Banach 空間、$T \in B(\mathcal{X})$ とし、$\sigma(T)$ の近傍で正則な関数の全体を $\mathcal{F}(T)$ と書く。$f \in \mathcal{F}(T)$ に対して、$\sigma(T)$ を内部に含む閉曲線 C をとって、
>
> $$f(T) = \frac{1}{2\pi i} \int_C f(t) R(t; T) dt$$
>
> とすれば、$f(T) \in B(\mathcal{X})$ である。$f, g \in \mathcal{F}(T)$ であるとき、上記 i), ii) は等号をもって成り立つ。また、iv) $f \in \mathcal{F}(T')$ で $f(T') = f(T)'$. この形の積分はしばしば **Dunford 積分** (Dunford integral) とよばれる。$T \in B(\mathcal{X})$ が自己随伴のときには、この 2 つの定義は一致する。1), 2) いずれの場合においても**スペクトル写像定理** (spectral mapping theorem) $\sigma(f(T)) = f(\sigma(T))$ が成り立つ。
>
> 他の種類の作用素解析として、T が作用素の半群の生成作用素である場合などがある (→ 作用素の半群、発展方程式)。

1.4　Rayleigh の原理、min-max 原理

定義 1.4.1 (Rayleigh 商) A を Hermite 行列、$x \neq 0$ とするとき、

$$R_A(x) \overset{\text{def.}}{=} \frac{(Ax, x)}{(x, x)} = \frac{x^* A x}{x^* x}$$

を **Rayleigh 商** (**Rayleigh quotient**) と呼ぶ。ここで $x^* = \overline{x}^T$ (共役ベクトルを転置した横ベクトル).

定理 1.4.2 (Rayleigh の原理) A を Hermit 行列とするとき、Rayleigh 商 $R_A(x) = (Ax, x)/(x, x)$ は、最小固有値 λ_1 に属する固有ベクトル x_1 によって最小となり、その最小値は最小の固有値 λ_1 である:

$$\min_{x \neq 0} R_A(x) = R_A(x_1) = A \text{ の最小固有値 } \lambda_1.$$

もちろん、大小を入れ換えた命題「$R_A(x)$ は最大固有値 λ_n に属する固有ベクトル x_n によって最大となり、その最大値は最大の固有値 λ_n である」が成り立つ。

定理 1.4.3 (Fischer-Poincaré のミニ・マックス原理) A を n 次 Hermite 行列、$\lambda_1 \le \cdots \le \lambda_n$ をその固有値とするとき、$\forall k \in \{1, 2, \cdot, n\}$ に対して

$$\lambda_k = \min_{V_k \text{ は } \mathbf{C}^n \text{ の } k \text{ 次元部分空間}} \max_{x \in V_k \setminus \{0\}} \frac{(Ax, x)}{(x, x)}.$$

特に

$$\lambda_1 \le \frac{(Ax, x)}{(x, x)} \le \lambda_n \quad (x \in \mathbf{C}^n \setminus \{0\}).$$

この命題を Courant-Fischer のミニ・マックス定理と呼ぶ人も。

定理 1.4.4 (Courant-Weyl のマックス・ミニ原理) A を n 次 Hermite 行列、$\lambda_1 \le \cdots \le \lambda_n$ をその固有値とするとき、$\forall k \in \{1, 2, \cdot, n\}$ に対して

$$\lambda_k = \max_{S \text{ は } \mathbf{C}^n \text{ の } k-1 \text{ 次元部分空間}} \min_{u \in S^\perp \setminus \{0\}} \frac{(Au, u)}{(u, u)}.$$

1.5 レゾルベントによる特徴づけ — 解析的な取り扱い —

(準備中 — 例えば、杉浦 [43] などを見よ。)

第2章　対角化に関するメモ

(この章の命題について、証明のありかを明らかにしておきたい。)

2.1　正規行列

定義 2.1.1 (正規行列) $A \in M(n; \mathbf{C})$ が条件

$$AA^* = A^*A$$

を満すとき、A を**正規行列 (normal matrix)** であるという。

系 2.1.2 Hermite 行列、unitary 行列は正規行列である。

証明　A が Hermite 行列 $\Leftrightarrow A^* = A$, A が unitary 行列 $\Leftrightarrow A^*A = AA^* = I$ であるから明らか。 ■

系 2.1.3 $A \in M(n; \mathbf{C})$ が正規行列であるための必要十分条件は

$$\|Ax\| = \|A^*x\| \quad (x \in \mathbf{C}^n).$$

(ただし、$\|y\| \stackrel{\text{def.}}{=} \sqrt{y^*y}$ である。)

証明　(略) ■

定理 2.1.4 (Toeplitz) $A \in M(n; \mathbf{C})$ が unitary 行列によって対角化できるための必要十分条件は A が正規行列であることである。

証明　(略) ■

命題 2.1.5 $A_1 \in M(n; \mathbf{C})$, $A_2 \in M(n; \mathbf{C})$, $\cdots$, $A_m \in M(n; \mathbf{C})$ が、一つの unitary 行列によって同時に対角化されるための必要十分条件は、$2m$ 個の行列

$$A_1, A_1^*, A_2, A_2^*, \cdots A_m, A_m^*$$

が互いに交換可能なことである。

証明　(略) ■

系 2.1.6 正規行列 A について、

$$A \text{ が Hermite} \Leftrightarrow A \text{ の固有値はすべて実数。}$$

$$A \text{ が unitary} \Leftrightarrow A \text{ の固有値はすべて絶対値が 1 の複素数。}$$

注意 2.1.7 複素数と行列の間には次のアナロジーがある。

複素正方行列	$\longleftrightarrow$	複素数
Hermite 行列 (実対称行列)	$\longleftrightarrow$	実数
正値 Hermite 行列 (正値対称行列)	$\longleftrightarrow$	正数
歪 Hermite 行列 (実交代行列)	$\longleftrightarrow$	純虚数
unitary 行列 (実直交行列)	$\longleftrightarrow$	絶対値 1 の複素数

命題 2.1.8 $A = B + iC$ (B, C は Hermite 行列) とすれば、

$$A \text{ が正規行列} \Leftrightarrow BC = CB.$$

命題 2.1.9 任意の $A \in GL(n; \mathbf{C})$ は

$$A = HU, \quad H \text{ は正値 Hermite 行列}, U \text{ は unitary 行列}$$

という形に一意的に表わされる。

命題 2.1.10 $A_1, \cdots, A_m$ が Hermite 行列とすると、

$$\sum_{j=1}^{m} A_i^2 = O \quad \Leftrightarrow \quad A_j = O \quad (j = 1, 2, \cdots, m).$$

命題 2.1.11 正規行列 A のスペクトル分解を

$$A = \sum_{j=1}^{r} \lambda_j P_j$$

とするとき (各 P_j は λ_j に関する固有空間への直交射影作用素[a]で、

$$P_1 + P_2 + \cdots + P_r = I, \quad P_j P_k = O \quad (j \neq k)$$

を満す)、A^* のスペクトル分解は

$$A^* = \sum_{j=1}^{r} \overline{\lambda}_j P_j$$

で与えられる。

また $f(x) \in C[x]$ に対して、$f(A)$ のスペクトル分解は

$$f(A) = \sum_{j=1}^{r} f(\lambda_j) P_j.$$

[a] P が直交射影とは $P^2 = P,\ P^* = P$ を満すこと。

第3章　行列の特異値分解

3.1　特異値分解の定義

$A \in M(n,m;\mathbf{R})$ の階数を r とすると、n 次 unitary 行列 U と m 次 unitary 行列 V が存在して、

$$\begin{aligned} A &= U\Sigma V^*, \\ \Sigma &= \begin{pmatrix} D & O_{r,m-r} \\ O_{n-r,r} & O_{n-r,m-r} \end{pmatrix}, \quad D = \mathrm{diag}(\sigma_1,\cdots,\sigma_r), \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \end{aligned} \tag{3.1}$$

となる。

(3.1) を A の**特異値分解**, σ_i $(i=1,2,\cdots,r)$ を A の**特異値**と呼ぶ。

$$A^*A = (U\Sigma V^*)^*(U\Sigma V^*) = V\Sigma^*\Sigma V^*$$

であるから、

$$V^*A^*AV = \Sigma^*\Sigma = \begin{pmatrix} D^2 & O_{r,m-r} \\ O_{n-r,r} & O_{n-r,m-r} \end{pmatrix}, \quad D^2 = \mathrm{diag}((\sigma_1)^2,\cdots,(\sigma_r)^2).$$

ゆえに

A の特異値は A^*A の非零固有値の $\sqrt{\cdot}$ に等しく、V の列ベクトルは A^*A の固有ベクトル。

同様に

$$AA^* = (U\Sigma V^*)(U\Sigma V^*)^* = U\Sigma\Sigma^*U^*,$$
$$U^*AA^*U = \Sigma\Sigma^* = \begin{pmatrix} D^2 & O_{r,m-r} \\ O_{n-r,r} & O_{n-r,m-r} \end{pmatrix},$$

であるから、

A の特異値は AA^* の非零固有値の $\sqrt{\cdot}$ に等しく、U の列ベクトルは AA^* の固有ベクトル。

3.2　特異値の公式

この項は伊理 [16] による。

命題 3.2.1 (1) $A \in M(m,n;\mathbf{C})$ の特異値を $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ $(r = \operatorname{rank} A \geq 1)$ とするとき、

$$\|A\|_2 = \sigma_1 = \text{最大特異値}, \quad \|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}.$$

($\|\cdot\|_2$ はベクトルのノルムに 2 乗ノルムを採用したときの行列の作用素ノルム、$\|\cdot\|_F$ はフロベニウス・ノルム[a]とする。)

(2) $A \in GL(n;\mathbf{C})$ の特異値を $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$ とするとき、

$$|\det A| = \sigma_1 \cdots \sigma_n, \quad \|A^{-1}\|_2 = \frac{1}{\sigma_n}, \quad \operatorname{cond}_2(A) = \frac{\sigma_1}{\sigma_n}.$$

[a]別名ユークリッド・ノルム。平方和の平方根。

証明 まず

$$(A^*A)^* = A^*(A^*)^* = A^*A,$$
$$(A^*Ax, x) = (Ax, Ax) = \|Ax\|^2 \geq 0.$$

であるから、A^*A は半正値である。

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sup_{x \neq 0} \sqrt{\frac{(Ax, Ax)}{(x,x)}} = \sup_{x \neq 0} \sqrt{\frac{(A^*Ax, x)}{(x,x)}} = \sqrt{A^*A \text{ の最大固有値}} = \sqrt{(\sigma_1)^2} = \sigma_1.$$

同様に

$$\begin{aligned}\|A^{-1}\|_2 &= \sup_{x \neq 0} \frac{\|A^{-1}x\|_2}{\|x\|_2} = \sup_{y \neq 0} \frac{\|y\|_2}{\|Ay\|_2} = \frac{1}{\inf\limits_{y \neq 0} \dfrac{\|Ay\|_2}{\|y\|_2}} \\ &= \left(\sqrt{\inf_{y \neq 0} \frac{(A^*Ay, y)}{(y,y)}}\right)^{-1} = \left(\sqrt{A^*A \text{ の最小固有値}}\right)^{-1} = \left(\sqrt{(\sigma_n)^2}\right)^{-1} = \frac{1}{\sigma_n}.\end{aligned}$$

ゆえに

$$\operatorname{cond}_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_1}{\sigma_n}.$$

さて、$A \in GL(n;\mathbf{C})$ とすると、$\operatorname{rank} A = n$ であるから、特異値の定義から、二つの n 次 unitary 行列 Q_1, Q_2 が存在して、

$$Q_2^* A Q_1 = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}.$$

これから

$$\det(Q_2^* A Q_1) = \prod_{j=1}^n \sigma_j.$$

ゆえに、

$$\overline{\det Q_2} \det Q_1 \det A = \prod_{j=1}^n \sigma_j.$$

unitary 行列の行列式は、絶対値 1 の複素数であるから、

$$|\det A| = \prod_{j=1}^n \sigma_j. \quad \blacksquare$$

第4章　有限次元ベクトルと行列のノルム

4.1　ベクトルのノルム

以下 $K=\mathbf{R}$ または $K=\mathbf{C}$ として、$V=K^n$ とおくと、V は自然に体 K 上の線型空間 (ベクトル空間, linear space, vector space) の構造を持つ。

定義 4.1.1 (ノルム, ノルム空間) 体 K 上の線型空間 V で定義された関数 $\|\cdot\|\colon V\ni x\mapsto\|x\|\in\mathbf{R}$ が、次の三条件を満たすとき、関数 $\|\cdot\|$ は V のノルムであるという。(このとき、V と $\|\cdot\|$ を組にして考えた $(V,\|\cdot\|)$ は K 上のノルム空間 (ノルム線型空間) であるという。

(i) $\forall x\in V$ に対して $\|x\|\ge 0$ で、$x=0\Leftrightarrow\|x\|=0$.

(ii) $\forall c\in K$, $\forall x\in V$ に対して $\|cx\|=|c|\|x\|$.

(iii) $\forall x\in V$, $\forall y\in V$ に対して $\|x+y\|\le\|x\|+\|y\|$.

定義 4.1.2 (最大ノルム, p 乗ノルム, ユークリッド・ノルム) ベクトル $x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\in K^n$, $p\in[1,\infty)\cup\{\infty\}$ に対して、x の p 乗ノルム (あるいは単に p ノルム) $\|x\|_p$ を

$$\|x\|_\infty\overset{\text{def.}}{=}\begin{cases}\left(\displaystyle\sum_{i=1}^n|x_i|^p\right)^{1/p} & (1\le p<\infty)\\ \displaystyle\max_{1\le i\le n}|x_i| & (p=\infty)\end{cases}$$

で定める。特に $p=2$ のときユークリッド・ノルム、$p=\infty$ のとき最大ノルムとも呼ぶ。

命題 4.1.3 p ノルム $\|\cdot\|_p$ は K^n のノルムである $(1\le p\le\infty)$。

証明　略 (常識であろう) ■

問　$\mathbf{R}^2$ において $\|\cdot\|_p$ $(1\le p\le\infty)$ に関する単位球

$$\{x\in\mathbf{R}^2;\|x\|_p<1\}$$

を図示せよ (少なくとも $p=1,\ 2,\ \infty$ の場合の図を描け)。

例 4.1.4 W を n 次正則行列, $\|\cdot\|$ を $\mathbf{R}^n$ のノルムとするとき、

$$\|x\|_W\overset{\text{def.}}{=}\|Wx\|\quad(x\in\mathbf{R}^n)$$

で $\|\cdot\|_W$ を定めると、これは $\mathbf{R}^n$ のノルムとなる。特に W が正の対角成分を持つ対角行列である:

$$W = \mathrm{diag}[d_1, d_2, \cdots, d_n], \quad d_i > 0 \quad (i = 1, 2, \cdots, n)$$

の場合が良く現れる。■

4.2 有限次元ノルムの同値性

定義 4.2.1 (ノルムの同値性) 体 K 上のベクトル空間 V 上の二つのノルム $\|\cdot\|, \|\cdot\|'$ が条件

$$(\exists M > 0)\ (\exists m > 0)\ (\forall x \in K^n)\ \text{s.t.}\ m\|x\| \le \|x\|' \le M\|x\|.$$

を満すとき、この二つのノルムは互いに**同値**であるという。

ここで定義したノルムの同値性はいわゆる同値関係である。

例 4.2.2 K^n のノルム $\|\cdot\|_\infty, \|\cdot\|_2, \|\cdot\|_1$ について、簡単に

$$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty,$$

$$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$$

が示せるから、$\|\cdot\|_\infty, \|\cdot\|_2, \|\cdot\|_1$ は同値である。■

実は、有限次元ベクトル空間の任意の二つのノルムは互いに同値であることが証明できる。

補題 4.2.3 K^n 上の任意のノルム $\|\cdot\|$ に対して、次式を満す正数 C が存在する。

$$\|x\| \le C\|x\|_2 \quad (x \in K^n).$$

証明 第 i 成分のみが 1 で、他の成分は 0 である (要するに第 j 成分が Kronecker のデルタ δ_{ij} である) ベクトルを $e^{(i)}$ とすると、

$$\|x\| = \left\|\sum_{j=1}^n x_j e^{(j)}\right\| \le \sum_{j=1}^n |x_j|\|e^{(j)}\| \le \max_{1\le j\le n}|x_j|\sum_{j=1}^n \|e^{(j)}\| = \|x\|_\infty \sum_{j=1}^n \|e^{(j)}\| \le \left(\sum_{j=1}^n \|e^{(j)}\|\right)\|x\|_2.$$

ゆえに

$$C = \sum_{j=1}^n \|e^{(j)}\|$$

とおけばよい。■

補題 4.2.4 K^n の任意のノルム $\|\cdot\|$ に対して、

$$m\|x\|_2 \le \|x\| \le M\|x\|_2 \quad (x \in K^n)$$

が成り立つような正数 M, m が存在する。

証明　K^n にノルム $\|\cdot\|_2$ から導かれる距離を導入して考える。単位球面 $S \overset{\text{def.}}{=} \{x \in K^n; \|x\|_2 = 1\}$ はコンパクトである (解析学の常識)。上の補題の C を取ると、

$$\Big| \|x\| - \|y\| \Big| \le \|x - y\| \le C\|x - y\|_2$$

であるから、$\|\cdot\|$ は ($\|\cdot\|_2$ から定まる距離に関して) 連続関数である。ゆえに、

$$m = \min_{x \in S} \|x\|, \quad M = \max_{x \in S} \|x\|$$

が存在する。こうして、

$$m \le \|x\| \le M \quad (\|x\|_2 = 1 \text{ なる任意の } x \in K^n)$$

が導かれる。これから容易に

$$m\|x\|_2 \le \|x\| \le M\|x\|_2 \quad (x \in K^n)$$

が成り立つことが分かる。■

定理 4.2.5 (有限次元ベクトル空間の任意のノルムの同値性) K^n の任意の二つのノルムは互いに同値である。

証明　上の補題によって明らか。■

4.3　行列のノルム

K の要素を成分とする n 次正方行列全体の集合 $M(n;K)$ は K 上の多元環 (代数) の構造を持つ。

定義 4.3.1 (行列のノルム) $M(n;K)$ 上で定義された関数 $\|\cdot\|\colon M(n;K) \ni A \mapsto \|A\| \in \mathbf{R}$ が、次の四条件を満たすとき、関数 $\|\cdot\|$ は $M(n;K)$ のノルムであるという。

(i) $\forall A \in M(n;K)$ に対して $\|A\| \ge 0$ で、$A = 0 \Leftrightarrow \|A\| = 0$.

(ii) $\forall c \in K, \forall A \in M(n;K)$ に対して $\|cA\| = |c|\|A\|$.

(iii) $\forall A \in M(n;K), \forall B \in M(n;K)$ に対して

$$\|A + B\| \le \|A\| + \|B\|.$$

(iv) $\forall A \in M(n;K), \forall B \in M(n;K)$ に対して

$$\|AB\| \le \|A\|\|B\|.$$

注意 4.3.2 ベクトル空間のノルムに対しては (i), (ii), (iii) だけを要求することが多いが、$M(n;K)$ では積も定義されいるので、積 $(A,B) \mapsto AB$ の連続性を保証するために (iv) も要求しておくのである。

系 4.3.3 $\|\cdot\|$ が $M(n;K)$ のノルムであるならば、単位行列 I のノルムは 1 以上である:

$$\|I\| \ge 1.$$

証明

$$\|I\| = \|II\| \le \|I\|\|I\|$$

であるが、$I \neq O$ なので $\|I\| > 0$ に注意すると

$$\|I\| \ge 1$$

が得られる。■

定義 4.3.4 (ベクトルのノルムと両立する行列のノルム) K^n のノルム $\|\cdot\|$, $M(n;K)$ のノルム $\|\cdot\|'$ について

$$\|Ax\| \le \|A\|'\|x\| \quad (A \in M(n;K),\ x \in K^n)$$

が成り立つとき、$\|\cdot\|$ と $\|\cdot\|'$ は互いに両立するという。

例 4.3.5 (フロベニウス・ノルム)

$$\|A\|_{\mathrm{F}} \stackrel{\text{def.}}{=} \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2} = \sqrt{\operatorname{trace}(A^*A)}$$

で定義される $\|\cdot\|_{\mathrm{F}}$ は $M(n;K)$ のノルムとなる。これをフロベニウス・ノルム (Frobenius norm) と呼ぶ (ユークリッド・ノルムと呼ぶ流儀もある[1])。K^n のユークリッド・ノルム (2 乗ノルム) $\|\cdot\|_2$ と行列のフロベニウス・ノルム $\|\cdot\|_{\mathrm{F}}$ は互いに両立する。すなわち

$$\|Ax\|_2 \le \|A\|_{\mathrm{F}}\|x\|_2 \quad (x \in K^n)$$

が成り立つ (これは Schwarz の不等式から証明できる)。■

Frobenius ノルムは計算するのが簡単だが、$\|I\| = \sqrt{n}$ であり、あまり扱いやすくない (後で見るように作用素ノルムでは $\|I\| = 1$ となる)。

定義 4.3.6 (スペクトル半径) 行列 A の固有値の絶対値の最大値を A の**スペクトル半径** (**spectrum radius**) と呼び、$r(A)$ で表わす。

注意 4.3.7 (スペクトル半径を表わす記号について) A のスペクトル半径のことを $\rho(A)$ という記号で表わしている本も多いが、線型作用素のスペクトル論において、非常にしばしば、$\rho(A)$ は線型作用素 A のレゾルベント・セット

$$\rho(A) \stackrel{\text{def.}}{=} \{z \in \mathbf{C};\ \text{有界な}\ (zI-A)^{-1}\ \text{が存在する}\ \}$$

を表わすために使われる (ちなみにその補集合 $\sigma(A) \stackrel{\text{def.}}{=} \mathbf{C}\setminus\rho(A)$ は A の**スペクトル** (**spectrum**) と呼ばれる) ので、ここではスペクトル半径は $r(A)$ という記号で表わすことにした (あまり一般的な記号ではないと思う)。他の本を読むときは混同しないこと。

例 4.3.8 任意のベクトル x_0 から出発する反復法

$$x_{j+1} = Ax_j$$

が 0 に収束するための必要十分条件は $r(A) < 1$. (これは A の Jordan 標準形を考えればすぐ分かる。例えば杉浦 [43] を見よ。) ■

[1] これは $M(n;K)$ を K^{n^2} と自然に同一視するとき、$\|\cdot\|_{\mathrm{F}}$ が K^{n^2} のユークリッド・ノルム $\|\cdot\|_2$ に対応することによるのであろう。

命題 4.3.9 (ベクトルのノルムと両立する行列ノルムはスペクトル半径以上である) 少なくとも一つの K^n のノルム $\|\cdot\|$ と両立する $M(n;K)$ のノルム $\|\cdot\|'$ について

$$r(A) \le \|A\|' \quad (A \in M(n;K)).$$

証明 λ を A の任意の固有値とするとき、対応する A の固有ベクトル $x \neq 0$ を取ると、

$$|\lambda|\|x\| = \|\lambda x\| = \|Ax\| \le \|A\|'\|x\|$$

より

$$|\lambda| \le \|A\|'.$$

これから $r(A) \le \|A\|'$. ■

(上の命題は $r(A)$ はベクトル・ノルムと両立する行列ノルムの下界を与えることを意味するが、実は下限を与える。森 [20] にある。)

注意 4.3.10 上の命題は行列の固有値の存在範囲をごく大まかに知る目的に利用できる: A の任意の固有値 λ について

$$|\lambda| \le \|A\|' \quad \text{i.e.} \quad \lambda \in \{z \in \mathbf{C}; |z| \le \|A\|'\}.$$

言い換えると

$$\sigma(A) \subset \{z \in \mathbf{C}; |z| \le \|A\|'\}.$$

(もちろん、Gershgorin の定理を使った方がより精密に調べられるであろう。)■

補題 4.3.11 K^n のノルム $\|\cdot\|$ を用いて、

$$\|A\|_{\mathrm{op}} \stackrel{\text{def.}}{=} \sup_{\|x\| \le 1} \|Ax\| \quad (A \in M(n;K))$$

とおくと、$\|\cdot\|_{\mathrm{op}}$ は $M(n;K)$ のノルムになる。また、$\|\cdot\|$ と $\|\cdot\|_{\mathrm{op}}$ は互いに両立する、つまり

$$\|Ax\| \le \|A\|_{\mathrm{op}}\|x\|.$$

証明 略 (大抵の関数解析の教科書に書いてある)。■

定義 4.3.12 (行列の作用素ノルム) K^n のノルム $\|\cdot\|$ に対して、

$$\|A\|_{\mathrm{op}} \stackrel{\text{def.}}{=} \sup_{\|x\| \le 1} \|Ax\| \quad (A \in M(n;K))$$

で定まる $M(n;K)$ のノルム $\|\cdot\|_{\mathrm{op}}$ を**作用素ノルム** (**operator norm**) または**自然ノルム** (**natural norm**) と呼ぶ。

系 4.3.13 (恒等作用素の作用素ノルムは 1 に等しい) $M(n;K)$ の任意の作用素ノルム $\|\cdot\|$ に対して

$$\|I\| = 1.$$

証明　$\|Ix\| = \|x\|$ より $\|I\| = \sup_{\|x\|\le 1} \|Ix\| = \sup_{\|x\|\le 1} \|x\| = 1$. ■

系 4.3.14 (作用素ノルムはスペクトル半径より大きい) $M(n;K)$ の任意の作用素ノルム $\|\cdot\|$ は、スペクトル半径よりも大きい:

$$r(A) \le \|A\| \quad (A \in M(n;K)).$$

証明　命題 4.3.9 の系である。■

注意 4.3.15 作用素ノルムのことを「スペクトル・ノルム」と呼ぶこともあるが、この文書では、行列の 2 乗ノルム $\|\cdot\|_2$ のことを「スペクトル・ノルム」という言葉で呼ぶことにする[2]。

定義 4.3.16 ベクトルの p 乗ノルム $\|\cdot\|_p$ に対応する作用素ノルムも記号 $\|\cdot\|_p$ で表わす。

$$\|A\|_p \overset{\text{def.}}{=} \sup_{\|x\|_p \le 1} \|Ax\|_p.$$

特に $p = 2$ のとき、$\|A\|_2$ を A のスペクトル・ノルムと呼ぶ。

行列の空間 $M(n;K)$ のノルムについても (ベクトル空間 K^n のノルムと同様に) 同値性を定義することが出来るが、すべてのノルムは互いに同値になってしまう ($M(n;K)$ と K^{n^2} を自然に同一視することができるが、その同一視で、行列のノルムはベクトルのノルムになるから)。例えば

$$\frac{1}{\sqrt{n}}\|A\|_\infty \le \|A\|_2 \le \sqrt{n}\|A\|_\infty,$$

$$\frac{1}{\sqrt{n}}\|A\|_F \le \|A\|_2 \le \|A\|_F.$$

次の命題は $\|\cdot\|_p$ ($p = 1, 2, \infty$) の評価法に関するものである。評価の簡単さから言うと $p = 1, \infty$ の場合が簡単だが、応用上 $p = 2$ の場合が重要なので、それについて色々と苦労をするわけである。

[2] この辺りの用語の混乱は大変面倒である。

命題 4.3.17 以下では A^T, A^* はそれぞれ A の転置 (a_{ji}), Hermite 共役 $(\overline{a_{ji}})$ を表わす。

(1) K^n の任意のノルム $\|\cdot\|$, $\forall A \in M(n;K)$ に対して、

$$\sup_{\|x\|\le 1}\|Ax\| = \sup_{\|x\|=1}\|Ax\| = \sup_{x\ne 0}\frac{\|Ax\|}{\|x\|}.$$

(K^n が有限次元空間であることから、いずれの sup も実は max である。)

(2) $\forall A = (a_{ij}) \in M(n;K)$ に対して、

$$\|A\|_1 = \max_{1\le j\le n}\sum_{i=1}^n |a_{ij}|, \quad \|A\|_\infty = \max_{1\le i\le n}\sum_{j=1}^n |a_{ij}|.$$

これから

$$\|A^T\|_1 = \|A^*\|_1 = \|A\|_\infty, \quad \|A^T\|_\infty = \|A^*\|_\infty = \|A\|_1.$$

が成り立つ。特に A が対称行列または Hermite 行列ならば

$$\|A\|_1 = \|A\|_\infty.$$

(3)

$$\|A\|_2 = \sqrt{r(AA^*)} = \sqrt{r(A^*A)} = \sqrt{\text{半正値 Hermite 行列 } A^*A \text{ の最大固有値}}.$$

A が実行列である場合、A^* は A の転置行列 $A^T = (a_{ji})$ に他ならない。

(4)

$$\|A\|_2 \le \sqrt{\|A\|_1\|A\|_\infty}.$$

特に A が対称行列または Hermite 行列ならば

$$\|A\|_2 \le \|A\|_1 = \|A\|_\infty.$$

(5) A が Hermite 行列であるならば、

$$\|A\|_2 = r(A).$$

(6) Q, Q' を unitary 行列とするとき、

$$\|QAQ'\|_2 = \|A\|_2, \quad \|QAQ'\|_F = \|A\|_F.$$

証明

(1) 本当に明らかなので証明略。

(2) (1 ノルムの場合) 任意の $x \in K^n$ に対して、

$$\|Ax\|_1 = \left\|\sum_{j=1}^n x_j a_j\right\|_1 \le \sum_{j=1}^n |x_j|\,\|a_j\|_1 \le \max_{1\le j\le n}\|a_j\|_1 \sum_{j=1}^n |x_j| = \max_{1\le j\le n}\|a_j\|\cdot\|x\|_1$$

であるから

$$\|A\|_1 \le \max_{1\le j\le n}\|a_j\|_1. \tag{4.1}$$

一方

$$\max_{1\le j\le n} \|a_j\|_1 = \|a_J\|_1$$

とすると、$x = e_J$ に対して

$$\|Ax\|_1 = \|Ae_J\|_1 = \|a_J\|_1 = \max_{1\le j\le n} \|a_j\|_1$$

であるから、実は (4.1) は等号で成立する:

$$\|A\|_1 = \max_{1\le j\le n} \|a_j\|_1 = \max_{1\le j\le n} \sum_{i=1}^n |a_{ij}|.$$

(∞ ノルムの場合)

$$\|Ax\|_\infty = \max_{1\le i\le n} \left|\sum_{j=1}^n a_{ij}x_j\right| \le \max_{1\le i\le n} \left(\sum_{j=1}^n |a_{ij}|\,|x_j|\right) \le \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}|\,\|x\|_\infty$$

であるから、

$$\|A\|_\infty \le \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}|\,.$$

一方、番号 I を

$$\max_{1\le i\le n} \sum_{j=1}^n |a_{ij}| = \sum_{j=1}^n |a_{Ij}|$$

で定めるとき、

$$x_j = \begin{cases} 1 & (a_{Ij} \ge 0) \\ -1 & (a_{Ij} < 0) \end{cases}$$

で $x = (x_1, x_2, \cdots, x_n)^T$ を定めると、$\|x\|_\infty = 1$ で、

$$\|A\|_\infty \ge \|Ax\|_\infty = \max_{1\le i\le n} \left|\sum_{j=1}^n a_{ij}x_j\right| \ge \left|\sum_{j=1}^n a_{Ij}x_j\right| = \left|\sum_{j=1}^n |a_{Ij}|\right| = \sum_{j=1}^n |a_{Ij}| = \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}|\,.$$

ゆえに

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}|\,.$$

(3) まず A^*A は半正定値 Hermite 行列であることを注意しておく (Hermite 行列であることは明らかで、$(A^*Ax, x) = (Ax, Ax) \ge 0$)。

$$\|A\|_2 = \sup_{x\ne 0} \frac{\|Ax\|_2}{\|x\|_2} = \sup_{x\ne 0} \sqrt{\frac{(Ax, Ax)}{(x, x)}} = \sup_{x\ne 0} \sqrt{\frac{(A^*Ax, x)}{(x, x)}} = \sqrt{A^*A \text{ の最大固有値}} = r(A^*A).$$

最後から二つ目の等号は、いわゆる Rayleigh の原理である (定理 1.4.2 を見よ)。

(4)

$$\|A\|_2 = \sqrt{r(A^*A)} \le \sqrt{\|A^*A\|_1} \le \sqrt{\|A^*\|_1\|A\|_1} \le \sqrt{\|A\|_\infty\|A\|_1}.$$

(5) A の固有値を $\lambda_1, \cdots, \lambda_n$ とするとき、$A^*A = AA = A^2$ の固有値は $(\lambda_1)^2, \cdots, (\lambda_n)^2$ となる。

$$\|A\|_2 = \sqrt{\max_{1\le j\le n} (\lambda_j)^2} = \max_{1\le j\le n} |\lambda_j|. \quad \blacksquare$$

命題 4.3.18 (正規行列のスペクトル・ノルムはスペクトル半径に等しい) A が正規行列 (i.e. $A^*A = AA^*$) であるならば

$$\|A\|_2 = r(A).$$

証明　A はある unitary 行列 Q によって対角化できる:

$$Q^* A Q = \mathrm{diag}(\lambda_1, \cdots, \lambda_n) \overset{\text{def.}}{=} D.$$

これから

$$Q^* A^* Q = D^*.$$

ゆえに

$$Q^* A^* A Q = (Q^* A^* Q)(Q^* A Q) = D^* D = \mathrm{diag}(|\lambda_1|^2, \cdots, |\lambda_n|^2).$$

これから $A^* A$ は DD^* と同じ固有値 $|\lambda_1|^2, \cdots, |\lambda_n|^2$ を持つ。ゆえに

$$r(A^* A) = \max_{1 \le j \le n} |\lambda_j|^2 = \left(\max_{1 \le j \le n} |\lambda_j| \right)^2 = r(A)^2.$$

ゆえに

$$\|A\|_2 = \sqrt{r(A^* A)} = r(A). \quad \blacksquare$$

第5章　連立1次方程式の誤差解析

5.1　消去法の丸め誤差解析について歴史覚え書き

(一松 [15] (1995) を見よ。)

消去法の際に生ずる丸め誤差を解析すると、実際には大変うまく行くことが多いのだが、およそ現実とはかけはなれた結果、我々は悲観的にならざるを得ない、と主張する研究結果が出た。

> von Neumann and Goldstine, 『高位の行列の数値的逆転』[38] (1947) — 素朴な (前進) 誤差解析で Gauss の消去法を解析した[1]。これを信じると 100 元程度の方程式を解くのは絶望的と考えられるが、実際には楽々解けてしまう。

この状況を解決したのが、J. H. Wilkinson[2] の研究である。彼は**後退誤差解析**の手法を用いて、連立1次方程式を解く過程を徹底的に解析して、この問題にケリをつけた。彼の得た重要な結論の一つは

> ピボットの部分選択法を採用した Gauss の消去法では、ほとんどすべての場合に[3]、相対残差が小さくなることが保証される。例えば:
>
> $$\frac{\|b - Ax_*\|}{\|A\|\|x\|} \le \rho\beta^{-n},$$
>
> ただし用いた浮動小数点数の体系は β 進 n 桁であるとした。

というものである。この稿では、これを認めることにして、後はなるべく自前で議論することにしよう。

要点は、

(残念ながら) 残差が小さくても、誤差も小さいとは保証されない！

係数行列の条件数が小さければ、誤差が小さくなることが保証される。

の二点である。

後退誤差解析とは　連立1次方程式 $Ax = b$ を浮動小数点演算で「解く」と、真の解の代わりに、x に「近い」 x_* が得られるが、これは元の問題を摂動した問題、例えば $(A + \Delta A)x_* = (b + \Delta b)$ の正確な解になっているとみなすことができる。こうして

- どういう摂動問題の解になっているか
- 摂動問題の解になっているのならば、真の解をどの程度近似していることになるか

[1] ちなみに、この当時のレベルでは「高位」とは 10 元程度のことを言うらしい。現在の学生世代は、連立1次方程式を人手あるいは初期のコンピューターで解くのがどれくらい大変だったかピンと来ないであろうが、それについては、一松 [15] 第2章 12 節を読むことを強く勧めたい。

[2] Rounding errors in algebraic process, 1963

[3] この気持ちの悪い言いまわしが気になる人もいるだろうが、浮動小数点数の体系は、オーバーフローやアンダーフローを始めとする破局的なエラーが起こり得るものだから、ある程度までは仕方がない。

と考察が分解されることになる。以下の節では、この後者の部分を扱うことになる。

注意: (2016/3/6) このあたりの説明を書いたのは随分前 (1993 年に「応用数理実験 I」の講義ノートを作っていた頃) のことで、そのとき参考にしたのは今から思えばかなり古いテキストで、ちょっと甘すぎた、と思う。そのうち書き直すけれど、とりあえず簡単に説明しておく。

ρ について、例えば Forsythe-Moler [5] では、「…」のように書かれていた。何とも歯切れ悪い感じだけど、講義する立場としては、調べる時間もなくて、「…とのことです」で流していた。

大抵の場合、$\rho \le \beta$ で、そうでないのは稀、とぼかしてあるけれど、Trefethen-Bau [36] (1997) では、ρ が最悪どれくらい大きくなり得るか教科書らしくすっきり説明してある ($\rho = 2^{n-1}$ だったっけ？)。von Neumann と Goldstine はそういうところまで見通していたということなのかな (Neumann-Goldstine [38] (1947), Goldstine-Neumann [8] (1951) を読んでみよう)。

この辺は、Higham-Higham [14] (1989), Wright [39] (1993), Foster [6] (1994), Foster [7] (1998) などの論文を読んでみるべきだ。

(中尾・渡部 [40] の MATLAB の例には、[6] の例がちょろっと出て来ている。)

Gauss の消去法は、大抵の場合にうまく計算できるけれど、ぽしゃることもある。ぽしゃっているかどうかは事後にチェックできる (ρ がいくつになるか分かる) みたいだけれど、現在のほとんどすべてのコードはそれをしていない。[36] で、Gauss の消去法についてぶつぶつ書いてあるのも、むべなるかな。

一度この問題について、ある程度きちんと調べておく必要があると思う (もしかして、ひょっとして、直接法といえば消去法、という教科書の書き方ががらっと変わる可能性がある)。

歴史は現在進行形だ。 ■

5.2 条件数の定義

定義 5.2.1 (正則行列の条件数) 正則行列 A と、任意の行列ノルム $\|\cdot\|$ に対して、

$$\|A\| \cdot \|A^{-1}\|$$

を A の**条件数** (**condition number**) と呼び、$\mathrm{cond}(A)$ で表わす:

$$\mathrm{cond}(A) \overset{\text{def.}}{=} \|A\| \cdot \|A^{-1}\|.$$

ノルムとして p 乗ノルム $\|\cdot\|_p$ 等を用いるときは、$\mathrm{cond}_p(A)$ で表わす。

命題 5.2.2 (条件数の基本的性質) A を n 次正則行列とするとき、

(1) $\mathrm{cond}(A) \ge 1$.

(2) $\mathrm{cond}(A) = \mathrm{cond}(A^{-1})$.

(3) $\forall \alpha \in \mathbf{C} \setminus \{0\}$ に対して $\mathrm{cond}(\alpha A) = \mathrm{cond}(A)$.

証明

(1)
$$\mathrm{cond}(A) = \|A\|\|A^{-1}\| \ge \|AA^{-1}\| = \|I\| \ge 1.$$

(2) $(A^{-1})^{-1} = A$ であるから、

$$\mathrm{cond}(A^{-1}) = \|A^{-1}\|\|(A^{-1})^{-1}\| = \|A^{-1}\|\|A\| = \|A\|\|A^{-1}\| = \mathrm{cond}(A).$$

(3) $(\alpha A)^{-1} = \alpha^{-1}A^{-1}$ であることに注意すれば

$$\mathrm{cond}(\alpha A) = \|\alpha A\|\|(\alpha A)^{-1}\| = |\alpha|\|A\||\alpha|^{-1}\|A^{-1}\| = \|A\|\|A^{-1}\| = \mathrm{cond}(A). \quad \blacksquare$$

注意 5.2.3 上の証明から、ノルムを一つ定めて、それに関する条件数を考えるとき、最小値は $\mathrm{cond}(I)$ である。任意のスカラー行列 αI の条件数は $\mathrm{cond}(I)$ に等しい。後で導入する、応用上重要な「作用素ノルム」においては、$\mathrm{cond}(I) = 1$ であるから、作用素ノルムの条件数は 1 以上であることが分かる。

命題 5.2.4 A, B ともに n 次正方行列とするとき

$$\mathrm{cond}(AB) \le \mathrm{cond}(A)\mathrm{cond}(B).$$

証明

$$\mathrm{cond}(AB) = \|AB\|\|(AB)^{-1}\| = \|AB\|\|B^{-1}A^{-1}\| \le \|A\|\|B\|\|B^{-1}\|\|A^{-1}\| = \mathrm{cond}(A)\mathrm{cond}(B). \quad \blacksquare$$

命題 5.2.5 (1) 任意の正則な対角行列 $D = \mathrm{diag}[a_1, a_2, \cdots, a_n]$ について $\mathrm{cond}_p(D) = \max|a_i|/\min|a_i|$ $(1 \le p \le \infty)$.

(2) unitary 行列 U について $\mathrm{cond}_2(U) = 1$.

(3) T を n 次 unitary 行列とするとき、任意の $A \in GL(n; \mathbf{C})$ に対して

$$\mathrm{cond}_2(A) = \mathrm{cond}_2(AT) = \mathrm{cond}_2(TA) = \mathrm{cond}_2(T^*AT).$$

証明

(1) ($1 \le p < \infty$ の場合のみ示す。$p = \infty$ の場合も同様である。)

行列ノルム $\|\cdot\|_p$ はベクトルのノルム $\|\cdot\|_p$ と両立するから、命題 4.3.9 より

$$\max_{1\le j\le n} |a_j| = r(D) \le \|D\|_p$$

が得られる。一方

$$\begin{aligned}\|Dx\|_p &= \left(\sum_{j=1}^n |a_j x_j|^p\right)^{1/p} \\ &\le \left(\sum_{j=1}^n (\max_{1\le j\le n}|a_j||x_j|)^p\right)^{1/p} = \max_{1\le j\le n}|a_j| \left(\sum_{j=1}^n |x_j|^p\right)^{1/p} = \max_{1\le j\le n}|a_j|\|x\|_p\end{aligned}$$

より

$$\|D\|_p \le \max_{1\le j\le n}|a_j|.$$

ゆえに

$$\|D\|_p = \max_{1\le j\le n} |a_j|.$$

同じことを $D^{-1} = \mathrm{diag}(1/a_1, \cdots, 1/a_n)$ に対して行って

$$\|D^{-1}\|_p = \max_{1\le j\le n} \frac{1}{|a_j|} = \frac{1}{\displaystyle\min_{1\le j\le n} |a_j|}.$$

ゆえに

$$\mathrm{cond}_p(D) = \|D\|_p \|D^{-1}\|_p = \frac{\displaystyle\max_{1\le j\le n} |a_j|}{\displaystyle\min_{1\le j\le n} |a_j|}.$$

(2) U をユニタリ行列 ($UU^* = U^*U = I$) とすると、U^{-1} もユニタリ行列であり、ユニタリ行列はベクトルの 2 乗ノルムを変えないので、$\|U\|_2 = \|U^{-1}\|_2 = 1$. ゆえに $\mathrm{cond}_2(U) = \|U\|_2\|U^{-1}\|_2 = 1$.

(3) 略 (). ■

行列 A の特異値とは、行列 AA^* の固有値の非負平方根のことであると定義したことを思いだそう。

定理 5.2.6 (スペクトル・ノルム $\|\cdot\|_2$ に関する条件数の性質) $A \in GL(n; \mathbf{C})$ とする。

(1) A のスペクトル・ノルム $\|\cdot\|_2$ に関する条件数 $\mathrm{cond}_2(A)$ は A の最大特異値と最小特異値の比である:

$$\mathrm{cond}_2(A) = \frac{A \text{ の特異値の最大値}}{A \text{ の特異値の最小値}}.$$

(2) 特に A が Hermite 行列である場合

$$\mathrm{cond}_2(A) = \frac{A \text{ の固有値の絶対値の最大値}}{A \text{ の固有値の絶対値の最小値}}.$$

証明

(1) すでに見たように

$$\|A\|_2 = A \text{ の最大特異値}, \quad \|A^{-1}\|_2 = \frac{1}{A \text{ の最小特異値}}$$

であるから、

$$\mathrm{cond}_2(A) = \frac{A \text{ の最大特異値}}{A \text{ の最小特異値}}.$$

(2) すでに見たように A が Hermite 行列であれば、

$$\|A\|_2 = r(A) = A \text{ の固有値の絶対値の最大値}.$$

A^{-1} $(= A^*)$ も Hermite 行列であるから

$$\begin{aligned} \|A^{-1}\|_2 &= r(A^{-1}) = A^{-1} \text{ の固有値の絶対値の最大値} \\ &= A \text{ の固有値の逆数の絶対値の最大値} \\ &= (A \text{ の固有値の絶対値の最小値})^{-1}. \end{aligned}$$

ゆえに

$$\mathrm{cond}_2(A) = \|A\|\|A^{-1}\| = \frac{A \text{ の固有値の絶対値の最大値}}{A \text{ の固有値の絶対値の最小値}}. \quad ■$$

一般の正則行列 A については、

$$\|A\|_2 = \sqrt{r(A^*A)}, \quad \|A^{-1}\|_2 = \sqrt{r((A^{-1})^*A^{-1})}$$

より

$$\mathrm{cond}_2(A) = \sqrt{r(A^*A)}\sqrt{r((A^{-1})^*A^{-1})}.$$

命題 5.2.7 (最大固有値と最小固有値の比と条件数の関係) 行列のノルムが、少なくとも一つのベクトルのノルムと両立しているならば、そのノルムに関する条件数 $\mathrm{cond}(\cdot)$ は、次式を満す:

$$\mathrm{cond}(A) \geq \frac{A\text{ の固有値の絶対値の最大値}}{A\text{ の固有値の絶対値の最小値}}.$$

証明 命題 4.3.9 より

$$\|A\| \geq r(A) = \max_{\lambda\in\sigma(A)} |\lambda|.$$

同様に

$$\|A^{-1}\| \geq \max_{\mu\in\sigma(A^{-1})} |\mu| = \max_{\lambda\in\sigma(A)} |\lambda^{-1}| = \frac{1}{\min\limits_{\lambda\in\sigma(A)} |\lambda|}.$$

ゆえに

$$\mathrm{cond}(A) \geq \frac{\max\limits_{\lambda\in\sigma(A)} |\lambda|}{\min\limits_{\lambda\in\sigma(A)} |\lambda|}. \quad \blacksquare$$

注意 5.2.8 この命題から、一般に (行列が Hermite であってもなくても) 固有値の絶対値の比が大きければ条件数も大きいが、逆は一般には真ではない。

例 5.2.9 (最大固有値と最小固有値の比は小さいが条件数は大きな行列) (この例は Strang [35] による。) 固有値の絶対値の最大値と最小値の比は 1 であるが、条件数は大きな行列の例.

$$A = \begin{pmatrix} 1 & 100 \\ 0 & 1 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1 & -100 \\ 0 & 1 \end{pmatrix},$$

$$A^TA = \begin{pmatrix} 1 & 100 \\ 100 & 10001 \end{pmatrix}, \quad (A^{-1})^TA^{-1} \begin{pmatrix} 1 & -100 \\ -100 & 10001 \end{pmatrix}.$$

ともに特性多項式は

$$\lambda^2 - 10002\lambda + 1 = 0$$

なので

$$\lambda = 5001 \pm \sqrt{5001^2 - 1} \fallingdotseq 10002, 0.$$

$$\mathrm{cond}_2(A) = \sqrt{r(A^TA)r((A^{-1})^TA^{-1})} \fallingdotseq 10002.$$

(この行列を係数とする連立 1 次方程式の丸め誤差の増幅の様子は次節で取り上げる。) ■

5.3 条件数の現れる近似公式

命題 5.3.1 (右辺の摂動) $A \in GL(n;\mathbf{C})$, $\Delta A \in M(n;\mathbf{C})$, $b \in \mathbf{C}^n \setminus \{0\}$, $\Delta b \in \mathbf{C}^n$ について

$$Ax = b, \quad A(x+\Delta x) = b + \Delta b$$

が成り立っているとき、

$$\frac{\|\Delta x\|}{\|x\|} \le \operatorname{cond}(A)\frac{\|\Delta b\|}{\|b\|}.$$

証明 方程式を辺々引き算すると

$$A\Delta x = \Delta b$$

となるから

$$\Delta x = A^{-1}\Delta b.$$

ゆえに

$$\|\Delta x\| \le \|A^{-1}\|\|\Delta b\|.$$

一方 $Ax = b$ より

$$\|b\| \le \|A\|\|x\|$$

であるから

$$\frac{1}{\|x\|} \le \frac{\|A\|}{\|b\|}.$$

まとめると

$$\frac{\|\Delta x\|}{\|x\|} \le \|A^{-1}\|\|\Delta b\| \cdot \frac{\|A\|}{\|b\|} = \operatorname{cond}(A)\frac{\|\Delta b\|}{\|b\|}. \quad \blacksquare$$

命題 5.3.2 (係数行列の摂動 (i)) $Ax = b$, $(A+\Delta A)(x+\Delta x) = b$ とするとき

$$\frac{\|\Delta x\|}{\|x+\Delta x\|} \le \operatorname{cond}(A)\frac{\|\Delta A\|}{\|A\|}.$$

証明

$$(A+\Delta A)(x+\Delta x) = b$$

の両辺に左から A^{-1} をかけると

$$(I + A^{-1}\Delta A)(x+\Delta x) = A^{-1}b.$$

これから

$$(x+\Delta x) + A^{-1}\Delta A(x+\Delta x) = x.$$

移項して

$$\Delta x = -A^{-1}\Delta A(x+\Delta x).$$

ゆえに

$$\|\Delta x\| \le \|A^{-1}\|\|\Delta A\|\|x+\Delta x\| = \operatorname{cond}(A)\frac{\|\Delta A\|}{\|A\|}\|x+\Delta x\|.$$

後は割り算すればよい。$\blacksquare$

上の命題は証明が簡単だが、結果の左辺がいわゆる誤差とはちょっと違っているところが今一つである。

命題 5.3.3 (係数行列の摂動 (ii)) $A \in GL(n;\mathbf{C})$, $\Delta A \in M(n;\mathbf{C})$, $\|A^{-1}\Delta A\| < 1$, $Ax = b$, $(A+\Delta A)(x+\Delta x) = b$ とするとき

$$\frac{\|\Delta x\|}{\|x\|} \le \frac{\mathrm{cond}(A)}{1-\|A^{-1}\Delta A\|}\frac{\|\Delta A\|}{\|A\|}.$$

証明 読みやすさのため、ΔA を E と書くことにする。

$$(A+E)(x+\Delta x) = b = Ax$$

から

$$(A+E)\Delta x = -Ex.$$

仮定から

$$A+E = A(I+A^{-1}E)$$

も正則である。

$$A^{-1}E = F, \quad (I+F)^{-1} = G$$

とおくと、$I = G + FG$ から

$$\|G\|(1-\|F\|) \le \|I\| = 1$$

であり、

$$\Delta x = -(A+E)^{-1}Ex = -GA^{-1}Ex$$

から

$$\frac{\|\Delta x\|}{\|x\|} \le \|G\|\|A^{-1}\|\|E\| \le \mathrm{cond}(A)\frac{\|E\|}{\|A\|}\frac{1}{1-\|A^{-1}E\|}. \blacksquare$$

系 5.3.4 (係数行列の摂動 (iii)) 上の命題の仮定に加えて $\|A^{-1}\|\|\Delta A\| < 1$ ならば

$$\frac{\|\Delta x\|}{\|x\|} \le \frac{\mathrm{cond}(A)}{1-\mathrm{cond}(A)\|\Delta A\|/\|A\|}\frac{\|\Delta A\|}{\|A\|}.$$

証明 上の命題の結果の分母について

$$\|A^{-1}\Delta A\| \le \|A^{-1}\|\|\Delta A\| < 1$$

より

$$\begin{aligned}\frac{\|\Delta x\|}{\|x\|} &\le \frac{\mathrm{cond}(A)}{1-\|A^{-1}\Delta A\|}\frac{\|\Delta A\|}{\|A\|} \le \frac{\mathrm{cond}(A)}{1-\|A^{-1}\|\|\Delta A\|}\frac{\|\Delta A\|}{\|A\|} \\ &= \frac{\mathrm{cond}(A)}{1-\mathrm{cond}(A)\|\Delta A\|/\|A\|}\frac{\|\Delta A\|}{\|A\|}. \blacksquare\end{aligned}$$

行列 (より一般には線形作用素) の「等比級数」のことを Neumann 級数と呼ぶが、これについては次の命題が基本的である。

補題 5.3.5 (Neumann 級数) $G \in M(N;K)$ が $\|G\| < 1$ を満たすならば、

$$C = \sum_{n=0}^{\infty} G^n$$

はノルム収束し (i.e., $\displaystyle\sum_{n=0}^{\infty} \|G^n\| < \infty$ — $M(N;K)$ は Banach 空間なので、級数そのものも収束する)、

$$(I-G)C = C(I-G) = I$$

が成り立つ。すなわち

$$I - G \in GL(N;K), \quad (I-G)^{-1} = C.$$

さらに

$$\|(I-G)^{-1}\| \le \frac{1}{1-\|G\|}.$$

証明 まず

$$\|G^n\| \le \|G\|^n \quad (n \in \mathbf{N})$$

であるから、$\displaystyle\sum_{n=0}^{\infty} \|G^n\|$ は収束等比級数 $\displaystyle\sum_{n=0}^{\infty} \|G\|^n$ を優級数に持つので収束する。ゆえに $\displaystyle\sum_{n=0}^{\infty} G^n$ はノルム収束する。行列空間 $M(N;K)$ は Banach 空間なので $\displaystyle\sum_{n=0}^{\infty} G^n$ は収束する。行列の積が連続写像であることに注意すれば

$$C(I-G) = C - CG = \sum_{n=0}^{\infty} G^n - \left(\sum_{n=0}^{\infty} G^n\right) G = \sum_{n=0}^{\infty} G^n - \sum_{n=0}^{\infty} (G^n \cdot G) = \sum_{n=0}^{\infty} G^n - \sum_{n=1}^{\infty} G^n = I.$$

これから $I - G \in GL(N;K)$, $(I-G)^{-1} = C$. ゆえに $(I-G)C = I$.

$$\|(I-G)^{-1}\| = \|C\| \le \sum_{n=0}^{\infty} \|G\|^n = \frac{1}{1-\|G\|}.\blacksquare$$

注意 5.3.6 $\|G\| < 1$ の代りに $r(G) < 1$ としても、上の補題の最後の $\|(I-G)^{-1}\|$ の評価以外は成立する。実際、$\|G\| < 1$ を満たす行列ノルム $\|\cdot\|$ の存在が命題 **??** から分かり、$M(N;K)$ 上のすべてのノルムが互いに同値であるから。あるいは、ノルムを全く用いない「線形代数的」な証明も可能である[4]。■

注意 5.3.7 自明な系として、「$\|G\| < 1$ ならば $D = \displaystyle\sum_{n=0}^{\infty} (-G)^n$ はノルム収束して、$I + G \in GL(N;K)$, $(I+G)^{-1} = D$, $\|(I+G)^{-1}\| \le 1/(1-\|G\|)$.」

Neumann 級数の補題は、「単位行列に十分小さな摂動を与えたものは正則行列である」と読むことができるが、単位行列の代りに任意の正則行列についての命題にしたものが次の定理である。

[4] $r(G) < 1 \Leftrightarrow \sigma(G) \subset B(0;1) \Leftrightarrow \sigma(I-G) \subset B(1;1) \Longrightarrow I-G$ の固有値 $\neq 0 \Leftrightarrow I-G \in GL(n;\mathbf{C})$.

定理 5.3.8 (Banach の摂動定理) $A \in GL(N;\mathbf{C})$, $\Delta A \in M(N;\mathbf{C})$, $\|A^{-1}\Delta A\| < 1$ とすれば、$A+\Delta A \in GL(N;\mathbf{C})$ で、

$$\|(A+\Delta A)^{-1}\| \le \frac{\|A^{-1}\|}{1-\|A^{-1}\Delta A\|}.$$

証明 $G = A^{-1}\Delta A$ とおくと、仮定より $\|G\| < 1$ であるから、Neumann 級数の補題から $I+G \in GL(N;K)$,

$$\|(I+G)^{-1}\| \le \frac{1}{1-\|G\|}$$

を得る。$A+\Delta A = A(I+A^{-1}\Delta A) = A(I+G)$ であるから、$A+\Delta A \in GL(N;K)$,

$$(A+\Delta A)^{-1} = (I+G)^{-1}A^{-1}.$$

ゆえに

$$\|(A+\Delta A)^{-1}\| \le \|(I+G)^{-1}\|\|A^{-1}\| \le \frac{\|A^{-1}\|}{1-\|G\|} = \frac{\|A^{-1}\|}{1-\|A^{-1}\Delta A\|}.\blacksquare$$

さて、連立1次方程式の摂動の話に戻って、より一般的な状況に適用できるのは次の定理である (上のいくつかの命題はこの定理の系ともみなせる)。これは Wilkinson による後退誤差解析の基礎と言えるが、大石・Rump 流の誤差解析においても重要である。

定理 5.3.9 (右辺と係数行列の摂動) $A \in GL(N;\mathbf{C})$, $\Delta A \in M(N;\mathbf{C})$, $Ax=b$, $(A+\Delta A)(x+\Delta x) = b+\Delta b$, $\|A^{-1}\Delta A\| < 1$ とすれば、

$$\frac{\|\Delta x\|}{\|x\|} \le \frac{\mathrm{cond}(A)}{1-\mathrm{cond}(A)\|A^{-1}\Delta A\|}\left(\frac{\|\Delta A\|}{\|A\|}+\frac{\|\Delta b\|}{\|b\|}\right). \tag{5.1}$$

証明 Banach の摂動定理から $A+\Delta A \in GL(N;\mathbf{C})$,

$$\|(A+\Delta A)^{-1}\| \le \frac{\|A^{-1}\|}{1-\|A^{-1}\Delta A\|}$$

が成り立つ。

$(A+\Delta A)(x+\Delta x) = b+\Delta b$ と $Ax=b$ を引き算すると

$$\Delta A x + (A+\Delta A)\Delta x = \Delta b.$$

これを Δx について解くと

$$\Delta x = (A+\Delta A)^{-1}(\Delta b - \Delta A x).$$

両辺のノルムをとって

$$\|\Delta x\| \le \|(A+\Delta A)^{-1}\|\|\Delta b - \Delta A x\| \le \frac{\|A^{-1}\|}{1-\|A^{-1}\Delta A\|}(\|\Delta b\| + \|\Delta A\|\|x\|).$$

辺々を $\|x\|$ で割ると

$$\frac{\|\Delta x\|}{\|x\|} \le \frac{\|A^{-1}\|}{1-\|A^{-1}\Delta A\|}\left(\|\Delta A\| + \frac{\|\Delta b\|}{\|x\|}\right) = \frac{\|A\|\|A^{-1}\|}{1-\|A^{-1}\Delta A\|}\left(\frac{\|\Delta A\|}{\|A\|}+\frac{\|\Delta b\|}{\|A\|\|x\|}\right).$$

後は $\|b\| \le \|A\|\|x\|$ に注意すれば、

$$\frac{\|\Delta x\|}{\|x\|} \le \frac{\|A\|\|A^{-1}\|}{1-\|A^{-1}\Delta A\|}\left(\frac{\|\Delta A\|}{\|A\|}+\frac{\|\Delta b\|}{\|b\|}\right) = (5.1) \text{ の右辺}.\blacksquare$$

注意 5.3.10 森 [20] にある証明の出だしは次のようになっている (上の証明の方が明快であると信じているが、こちらはこちらで面白い)。—- 「一般に行列 $I+G$ が正則であれば

$$(I+G)^{-1} = I - G(I+G)^{-1}$$

が成立する[5]。これから

$$\|(I+G)^{-1}\| \le 1 + \|G\| \cdot \|(I+G)^{-1}\|$$

ゆえに

$$(1-\|G\|)\|(I+G)^{-1}\| \le 1$$

となる。$\|G\| < 1$ であれば割り算して

$$\|(I+G)^{-1}\| \le \frac{1}{1-\|G\|}.\blacksquare$$

命題 5.3.11 (逆行列の近似精度) A, $A+\Delta A$ を正則行列とするとき

(1)

$$\frac{\|(A+\Delta A)^{-1} - A^{-1}\|}{\|(A+\Delta A)^{-1}\|} \le \operatorname{cond}(A)\frac{\|\Delta A\|}{\|A\|}.$$

(2)

$$\frac{\|(A+\Delta A)^{-1} - A^{-1}\|}{\|A^{-1}\|} \le \operatorname{cond}(A)\frac{\|\Delta A\|}{\|A\|}\left(1 + O(\|\Delta A\|)\right).$$

証明 (山本-北川 [45])

(1)

$$\begin{aligned}(A+\Delta A)^{-1} - A^{-1} &= (A+\Delta A)^{-1}[A-(A+\Delta A)]A^{-1}\\ &= -(A+\Delta A)^{-1}\Delta A A^{-1}\end{aligned}$$

より

$$\begin{aligned}\|(A+\Delta A)^{-1} - A^{-1}\| &\le \|(A+\Delta A)^{-1}\|\|\Delta A\|\|A^{-1}\|\\ &= \|(A+\Delta A)^{-1}\|\operatorname{cond}(A)\frac{\|\Delta A\|}{\|A\|}.\end{aligned}$$

(2) 上で見たように

$$(A+\Delta A)^{-1} - A^{-1} = -(A+\Delta A)^{-1}\Delta A A^{-1}$$

である。ここで $\Delta A \to O$ とすると、

$$\begin{aligned}(A+\Delta A)^{-1} &= [A(I+A^{-1}\Delta A)]^{-1}\\ &= (I+A^{-1}\Delta A)^{-1}A^{-1}\\ &= \left(\sum_{m=0}^{\infty}(-A^{-1}\Delta A)^m\right)A^{-1}\\ &= (I+B)A^{-1},\end{aligned}$$

[5]形式的には簡単: $\dfrac{1}{1+G} = \dfrac{1+G}{1+G} - \dfrac{G}{1+G} = 1 - \dfrac{G}{1+G}$.

ただし

$$B \stackrel{\text{def.}}{=} \sum_{m=1}^{\infty}(-A^{-1}\Delta A)^m = -A^{-1}\Delta A + (A^{-1}\Delta A)^2 - \cdots$$

とおいた。ゆえに[6]

$$(A+\Delta A)^{-1} \le (1+\|B\|)\|A^{-1}\| = (1+O(\|\Delta A\|))\|A^{-1}\|.$$

これから

$$\begin{aligned}
\|(A+\Delta A)^{-1} - A^{-1}\| &\le \|(A+\Delta A)^{-1}\|\|\Delta A\|\|A^{-1}\| \\
&\le (1+O(\|\Delta A\|))\|A^{-1}\|\|\Delta A\|\|A^{-1}\| \\
&\le (1+O(\|\Delta A\|))\mathrm{cond}(A)\frac{\|\Delta A\|}{\|A\|}\|A^{-1}\|.
\end{aligned}$$

よって与式を得る。■

例 5.3.12 (前節の Strang の例の続き: 最大/最小 $=1$ だが、条件数大きく、摂動の影響が大きい)

$$A = \begin{pmatrix} 1 & 100 \\ 0 & 1 \end{pmatrix},$$

とすると

$$A^{-1} = \begin{pmatrix} 1 & -100 \\ 0 & 1 \end{pmatrix}.$$

このとき、

$$b = \begin{pmatrix} 100 \\ 1 \end{pmatrix} \quad\Longrightarrow\quad x = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$b' = \begin{pmatrix} 100 \\ 0 \end{pmatrix} \quad\Longrightarrow\quad x' = \begin{pmatrix} 100 \\ 0 \end{pmatrix}.$$

つまり (ここから Strang [35] (正確にはその邦訳) を引用)

> b で 1% の変化が x では 100 の変化となっている。倍率は 100^2 である。$\mathrm{cond}_2(A)$ はこの倍率の上界を与えるはずであったから、$\mathrm{cond}_2(A)$ は少なくとも 10,000 でなければならぬ。これらの行列でやっかいな、直感的には A が大きくなれば A^{-1} は小さくなると考えるが、実際にはこれに反して、A の非対角要素が大きくなると A^{-1} も増大することである。

例 5.3.13 (Hermite でなければ、固有値の比が小さくとも、悪条件になりうる) (山本-北川 [45] から。) n 次正方行列

$$A = \begin{pmatrix} 1 & 2 & & & \\ & 1 & 2 & & \\ & & \ddots & \ddots & \\ & & & 1 & 2 \\ & & & & 1 \end{pmatrix}$$

の固有値はすべて 1 であるので、最大固有値と最小固有値の比は 1 であるが、

$$\mathrm{cond}_2(A) \ge 2^{n-1}.$$

[6] うーん。$\Delta A \to O$ のとき確かにそうだが。もうちょっと考えよう。

実際、

$$x \overset{\text{def.}}{=} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \Delta x \overset{\text{def.}}{=} \begin{pmatrix} 2^{n-1} \\ -2^{n-2} \\ \vdots \\ (-1)^{j-1}2^{n-j} \\ \vdots \\ (-1)^{n-2}2 \\ (-1)^{n-1} \end{pmatrix}$$

とおくと

$$Ax = x(\overset{\text{def.}}{=} b), \quad A\Delta x = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ (-1)^{n-1} \end{pmatrix} (\overset{\text{def.}}{=} \Delta b)$$

となる。命題 5.3.1 の結論の不等式

$$\frac{\|\Delta x\|_2}{\|x\|_2} \le \operatorname{cond}_2(A) \frac{\|\Delta b\|_2}{\|b\|_2}$$

に代入すると

$$\frac{\|\Delta x\|_2}{1} \le \operatorname{cond}_2(A) \frac{1}{1}.$$

ゆえに

$$\operatorname{cond}_2(A) \ge \|\Delta x\|_2 = \sqrt{(2^{n-1})^2 + \cdots + 2^2 + 1} \ge 2^{n-1}.$$

(ちょっと粗い評価だと思うけど、面白い。) ■

5.4　条件数を見積もるアルゴリズム

松尾・杉原・森 [24] を参照せよ。これは $\|A^{-1}\|$ の評価の話にもなっているので、真面目に勉強する価値があるようである。

5.5　Wilkinson の丸め誤差解析

例 5.5.1 (敏感な方程式 — Forsythe による)

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix}, \quad b = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}$$

とするとき $Ax = b$ の解は

$$x = \begin{pmatrix} 1.00 \\ -1.00 \end{pmatrix}.$$

ところで

$$\Delta b = \begin{pmatrix} 0 \\ 10^{-6} \end{pmatrix}$$

として、右辺を $b + \Delta b$ にすると、つまり $A(x + \Delta x) = b + \Delta b$ とすると、

$$x + \Delta x = \begin{pmatrix} 0.440 \\ -0.220 \end{pmatrix}.$$

$\det A = 10^{-6}$, $\mathrm{cond}(A) \fallingdotseq 10^6$ であるから、わずかの摂動が大きな誤差を生じさせるのは不思議ではない。反対から見ると、解を相当ずらしても残差は小さいということもできる。■

2 進 ℓ 桁の精度の浮動小数点数を持つコンピューターで $A \in GL(n;K)$ を LU 分解した結果を

$$L = \begin{pmatrix} 1 & & & \\ \ell_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ \ell_{n1} & \cdots & \ell_{n,n-1} & 1 \end{pmatrix}, \quad U = \begin{pmatrix} u_{11} & & u_{1n} \\ & \ddots & \vdots \\ & & u_{nn} \end{pmatrix}$$

$$LU = A + \Delta A, \quad (L + \Delta L)y = b, \quad (U + \Delta U)x = y$$

$$\begin{aligned} \|\Delta A\|_\infty &\le (2+\varepsilon)2^{-\ell}(1+n/2)(n-1)g, \\ \|\Delta L\|_\infty &\le 2^{-\ell}(n^2+n+2), \\ \|\Delta U\|_\infty &\le 2^{-\ell}(n^2+n+2)g, \end{aligned}$$

ただし ε はごく小さな正数、g は計算途中に現われる $A^{(r)}$ の成分の最大絶対値で、

$$|a_{ij}| \le a \quad \Longrightarrow \quad g \le 2^{n-1}a$$

という評価が成り立つ。これは一般には改良できないという意味でベストの不等式である。しかし、ほとんどの場合にこれは相当な過剰評価 (悲観的な評価 — 例えば $n = 100$ として考えてみよう) で、大抵の場合

$$g \le 8a$$

が成り立つという。また実際に計算する場合、g の後天的評価は容易である。

5.6 逆行列のノルム評価を利用する誤差解析 (大石, Rump 流)

次の命題は当たり前すぎるせいか、単独の命題として載っている本が少ないくらいであるが[7]、実は重要である。

命題 5.6.1 (誤差のノルムは逆行列のノルムと残差のノルムの積で押さえられる) $A \in GL(n;K)$, $b \in K^n$ に対して、 $x^* = A^{-1}b$ とおくとき、$\forall \widetilde{x} \in \mathbf{R}^n$ に対して

$$\|x^* - \widetilde{x}\| \le \|A^{-1}\|\|b - A\widetilde{x}\|.$$

証明

$$x^* - \widetilde{x} = A^{-1}A(x^* - \widetilde{x}) = A^{-1}(Ax^* - A\widetilde{x}) = A^{-1}(b - A\widetilde{x})$$

であるから

$$\|x^* - \widetilde{x}\| \le \|A^{-1}\|\|b - A\widetilde{x}\|. ■$$

次の命題は Neumann 級数の定理からすぐ導かれるものであり、古典的[8]であるが重要である。

[7]それとも逆行列のノルムの評価は簡単には出来ないと考えられていたせいか？
[8]例えば篠原 [33] に載っている。

定理 5.6.2 (近似逆行列を用いた逆行列の作用素ノルム評価) $A, R \in M(N;K)$ に対して $G \overset{\text{def.}}{=} I - RA$ とおくとき $\|G\| < 1$ が成り立つならば

$$A \in GL(N;K), \quad \|A^{-1}\| \le \frac{\|R\|}{1-\|G\|}.$$

特に $\forall b \in K^N$, $\forall \widetilde{x} \in K^N$ に対して

$$\|x^* - \widetilde{x}\| \le \frac{\|R(b-A\widetilde{x})\|}{1-\|G\|} \le \frac{\|R\|}{1-\|G\|}\cdot\|b-A\widetilde{x}\|, \quad \text{ただし} \quad x^* = A^{-1}b.$$

注意 5.6.3 (実際の数値計算では) 応用上は R として、A の近似逆行列を取るわけである。$\widetilde{x}$ としては $\widetilde{x} = Rb$ を取るのが自然であろう。また実際に行列のノルムを簡単に評価するためには、1 ノルム $\|\cdot\|_1$, ∞ ノルム $\|\cdot\|_p$ を採用することになるだろう。

証明 仮定 $\|G\| < 1$ より Neumann 級数 $C = \sum_{n=0}^{\infty} G^n$ は収束して、$C = (I-G)^{-1}$, さらに $\|C\| \le (1-\|G\|)^{-1}$. さて $C = (I-G)^{-1}$ より $C(I-G) = I$ で、これは

$$CRA = I$$

を意味するから、$A \in GL(N;K)$ かつ $A^{-1} = CR = (I-G)^{-1}R$. そして

$$\|A^{-1}\| = \|CR\| \le \|C\|\|R\| \le \frac{\|R\|}{1-\|G\|}.$$

さらに

$$x^* - \widetilde{x} = A^{-1}(Ax^* - A\widetilde{x}) = CR(b - A\widetilde{x})$$

であるから、

$$\|x^* - \widetilde{x}\| \le \|C\|\|R(b-A\widetilde{x})\| \le \frac{\|R(b-A\widetilde{x})\|}{1-\|G\|} \le \frac{\|R\|\|b-A\widetilde{x}\|}{1-\|G\|}. \blacksquare$$

しかし、最近の偏微分方程式の数値シミュレーションに現れるような大規模な連立 1 次方程式を解く場合は、計算量の観点から、係数行列の逆行列を求めずに済ませる方法を使うのが普通である。

逆行列を求めずに、そのノルムの評価をうまく行なうのは困難だと考えられて来たが、荻田&大石は次の定理を得た?[9]。

定理 5.6.4 (単調行列の逆行列の sup ノルムの評価 (荻田&大石?)) n 次単調行列 A, $\widetilde{y} \in \mathbf{R}^n$, $e = (1,1,\dots,1)^T$ に対して $s \overset{\text{def.}}{=} A\widetilde{y} - e$ とおいたとき、

$$\frac{\|\widetilde{y}\|_\infty}{1+\|s\|_\infty} \le \|A^{-1}\|_\infty.$$

特に $\|s\|_\infty < 1$ が成り立つならば

$$\|A^{-1}\|_\infty \le \frac{\|\widetilde{y}\|_\infty}{1-\|s\|_\infty}.$$

[9] 逆行列の大まかな評価を与える名取・塚本の方法を見ると、この手の命題の追求は昔からあったような気がするのだが…

証明 単調行列の定義から、$A \in GL(n;K)$ かつ $A^{-1} = (\widetilde{a}_{ij})$ とおくと $\widetilde{a}_{ij} \geq 0$ が成り立つ。$y^* = A^{-1}e$ とおくと、

$$\|A^{-1}\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^n |\widetilde{a}_{ij}| = \max_{1\leq i\leq n} \sum_{j=1}^n |\widetilde{a}_{ij} e_j| = \max_{1\leq i\leq n} \left| \sum_{j=1}^n \widetilde{a}_{ij} e_j \right| = \|A^{-1}e\|_\infty = \|y^*\|_\infty.$$

さて

$$s = A\widetilde{y} - e = A\widetilde{y} - Ay^* = A(\widetilde{y} - y^*)$$

より

$$A^{-1}s = \widetilde{y} - y^*. \tag{5.2}$$

これから

$$\widetilde{y} = y^* + A^{-1}s$$

であるから

$$\|\widetilde{y}\|_\infty \leq \|y^*\|_\infty + \|A^{-1}\|_\infty \|s\|_\infty = \|y^*\|_\infty + \|y^*\|_\infty \|s\|_\infty = \|y^*\|(1 + \|s\|_\infty).$$

ゆえに

$$\frac{\|\widetilde{y}\|_\infty}{1 + \|s\|_\infty} \leq \|y^*\|_\infty = \|A^{-1}\|_\infty.$$

また (5.2) から

$$y^* = \widetilde{y} - A^{-1}s.$$

したがって

$$\|y^*\|_\infty \leq \|\widetilde{y}\|_\infty + \|A^{-1}\|_\infty \|s\|_\infty = \|\widetilde{y}\|_\infty + \|y^*\|_\infty \|s\|_\infty.$$

移項して

$$(1 - \|s\|_\infty)\|y^*\|_\infty \leq \|\widetilde{y}\|_\infty.$$

もしも $\|s\|_\infty < 1$ である場合は

$$\|A^{-1}\|_\infty = \|y^*\|_\infty \leq \frac{\|\widetilde{y}\|_\infty}{1 - \|s\|_\infty}. \blacksquare$$

注意 5.6.5 もちろん応用上は $\widetilde{y}$ として、方程式 $Ay = e$ を数値計算で解いた解を取るわけである。丸め誤差が小さければ $s = A\widetilde{y} - e$ は 0 に近いと期待されるので、上の定理の右辺と左辺は近い値を取り、高精度の $\|A^{-1}\|_\infty$ の評価が得られるわけである。■

さて $\|A^{-1}\|_1 = \|(A^{-1})^T\|_\infty = \|(A^T)^{-1}\|_\infty$ に注意すると、逆行列の 1 ノルム $\|A^{-1}\|_1$ も評価できることが分かる。

系 5.6.6 (単調行列の逆行列の 1 ノルムの評価) n 次単調行列 A, $\widetilde{z} \in \mathbf{R}^n$, $e = (1, 1, \ldots, 1)^T$ に対して $t \overset{\text{def.}}{=} A^T\widetilde{z} - e$ とおいたとき、

$$\frac{\|\widetilde{z}\|_\infty}{1 + \|t\|_\infty} \leq \|A^{-1}\|_1.$$

特に $\|t\|_\infty < 1$ が成り立つならば

$$\|A^{-1}\|_1 \leq \frac{\|\widetilde{z}\|_\infty}{1 - \|t\|_\infty}.$$

応用上は $\widetilde{z}$ として $A^Tz = e$ を数値計算で解いた近似解を採用する。

5.6.1 区間演算

目標

さて、係数行列 A、定数項 b がそれぞれ区間行列、区間ベクトルの場合にどうするか考えよう。つまり

$$[\underline{A},\overline{A}][x]=[\underline{b},\overline{b}]$$

(ただし $\underline{A}$, $\overline{A}\in M(n;\mathbf{F})$ で、$\underline{b}$, $\overline{b}\in\mathbf{F}^n$ とする) を解く。これは

$$\{A^{-1}b; A\in[\underline{A},\overline{A}], b\in[\underline{b},\overline{b}]\}\subset[\underline{x},\overline{x}]$$

を満たす「なるべく狭い」区間ベクトル $[\underline{x},\overline{x}]\in I\mathbf{F}^n$ を求めるわけである。

中心と幅による区間で置き換える

まず

$$[\underline{A},\overline{A}]\subset[A_c-\delta A, A_c+\delta A]$$

を満たす A_c, $\delta A\in M(n;\mathbf{F})$ を求める (δA はなるべく小さく)。それには

`setround(up);`
$A_c := \dfrac{\overline{A}-\underline{A}}{2}+\underline{A};$
$\delta A := A_c-\underline{A};$

とすればよい。また

$$[\underline{b},\overline{b}]\subset[b_c-\delta b, b_c+\delta b]$$

を満たす b_c, $\delta b\in\mathbf{F}^n$ を求めたい (δb はなるべく小さく)。それには

`setround(up);`
$b_c := \dfrac{\overline{b}-\underline{b}}{2}+\underline{b};$
$\delta b := b_c-\underline{b};$

とする。

誤差の分解

さて

$$A\in[\underline{A},\overline{A}],\quad b\in[\underline{b},\overline{b}]$$

なる $A\in M(n;\mathbf{R})$, $b\in\mathbf{R}^n$ に対して、

$$\Delta A\overset{\text{def.}}{=}A-A_c,\quad \Delta b\overset{\text{def.}}{=}b-b_c$$

とおくと、明らかに

$$\Delta A\in[-\delta A,\delta A],\quad \Delta b\in[-\delta b,\delta b]$$

が成り立つ。

$$x^*\overset{\text{def.}}{=}A^{-1}b,\quad x_c\overset{\text{def.}}{=}A_c^{-1}b_c$$

とおき、$\widetilde{x} \in \mathbf{F}^n$ を $b - A\widetilde{x}$ が小さいと期待されるベクトルとする (実際には $A_c x = b_c$ を適当な近似計算で解いたものなど) とき、$\|x^* - \widetilde{x}\|$ を評価したい。

$$\|x^* - \widetilde{x}\| \le \|x^* - x_c\| + \|x_c - \widetilde{x}\|$$

として、右辺の各項を評価する。

$Ax = b$ の厳密解 x^* と $A_c x = b_c$ の厳密解 x_c の差 $\|x^* - x_c\|$ の評価

$\Delta x \overset{\text{def.}}{=} x^* - x_c$ とおくと、$x^* = x_c + \Delta x$ で、$Ax^* = b$ より

$$(A_c + \Delta A)(x_c + \Delta x) = (b_c + \Delta b).$$

ゆえに後退誤差解析でよく知られた定理から

$$\|x^* - x_c\| = \|\Delta x\| \le \frac{\|A_c^{-1}\|}{1 - \|A_c^{-1}\Delta A\|}\|\Delta b - \Delta A x\|.$$

$\|A_c^{-1}\| \le A_{\mathrm{inv}}$ を満たす $A_{\mathrm{inv}} \in \mathbf{F}$ が分かっていれば、

$$\|x^* - x_c\| \le \frac{A_{\mathrm{inv}}}{1 - A_{\mathrm{inv}}\|\delta A\|}(|\delta b\| + \|\delta A\|\|x\|).$$

$A_c x = b_c$ の厳密解 x_c と近似解 $\widetilde{x}$ の差 $\|x_c - \widetilde{x}\|$ の評価

$\|A_c^{-1}\|$ の適当な上からの評価と、不等式

$$\|x_c - \widetilde{x}\| \le \|A_c^{-1}\|\|b_c - A_c\widetilde{x}\|$$

を利用する。b_c, A_c, $\widetilde{x}$ は計算機内で正確に表現できているので、$\|b_c - A_c\widetilde{x}\|$ は簡単かつ高速に精度保証できる。

$\|A_c^{-1}\|$ の評価の仕方には

(1) (A_c の逆行列に近い R が何らかの方法で得られている場合) $R \in M(n; \mathbf{F})$ が $G \overset{\text{def.}}{=} I - RA_c$ とおいたとき $\|G\| < 1$ を満たすならば、定理 5.6.2 より

$$\|A_c^{-1}\| \le \frac{\|R\|}{1 - \|G\|}$$

となることを利用する。

(2) (A_c が単調行列である場合) Ogita-Oishi-?? の結果を用いる。方程式 $A_c y = e$ (ただし $e = (1, 1, \ldots, 1)^T$) の近似解 $\widetilde{y}$ を求めて、$s \overset{\text{def.}}{=} A_c\widetilde{y} - e$ とおいたとき、$\|s\|_\infty < 1$ が成り立つならば

$$\|A_c^{-1}\|_\infty \le \frac{\|\widetilde{y}\|}{1 - \|s\|_\infty}$$

となることを利用する。

(2016/3/6 追記: LU 分解を利用する方法とかもあるので、この節はリニューアル必要があると考えているけれど、なかなかそういうことが出来るようにならない…このテーマで修士論文 (礒野 [42]) 書いてもらったりしたのだけど)。

第6章　反復法

例えば、伊理 [16], 仁木-河野 [26], Varga [37] を見よ。

6.1　イントロ

連立 1 次方程式

$$Ax = b$$

をそれと同値な不動点型の方程式

$$x = Mx + b$$

に帰着させて、後者を反復公式 (漸化式 iteration formula, recurrence relation)

$$x^{(m+1)} = Mx^{(m)} + b$$

を用いて、解に収束する列 $\{x^{(m)}\}_{m=0}^{\infty}$ を作り出すことによって解くことを**反復法**と呼ぶ。近年の CG 法の系統の解法も反復法と呼ぶこともあるが、その場合は区別するために**定常反復法**と呼ぶ (CG 法の系統のことは**非定常反復法**と呼ぶ)。

注意 6.1.1 (用語「間接法」について) 反復法のことを間接法とも呼ぶ。Gauss の消去法のように、有限回の四則演算で厳密解が求まる方法を直接法というのに対比させるためか。

注意 6.1.2 (用語「緩和法」について) 反復法のことを「緩和法」と呼ぶこともあるらしい。何でもヤコビが反復法を使ったときは、反復ごとに残差 $r_i = (Ax - b)$ の第 i 行 の大きい順に次の反復を行う方法を使っていたからとか (分かったような分からないような説明だな)。

注意 6.1.3 (歴史) (この注は 仁木-河野 [26] による。) Gauss も 1823 年頃に反復法について触れているらしい。ヤコビは 1845 年の論文で後述するヤコビ法を発表した。ヤコビの弟子ザイデルはガウス・ザイデル法を発表した。

6.2　等比数列、等比級数の行列版

n 次正方行列 A について、

$$\lim_{m\to\infty} A^m = O$$

が成り立つかどうか？ある行列ノルム $\|\cdot\|$ について

$$\|A\| < 1$$

であれば十分であることは明らかであるが、スペクトル半径を用いれば必要十分条件が書ける。

定理 6.2.1 (行列の等比級数が O に収束するための必要十分条件) n 次正方行列 A について、

$$\lim_{m\to\infty} A^m = O \quad \Leftrightarrow \quad A \text{ のスペクトル半径} < 1.$$

証明　適当な正則行列 P を取り、H を Jordan 標準形に変換する。

$$P^{-1}AP = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix}.$$

ここで J_i は Jordan 細胞である。すなわち次の形をしている:

$$J = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \quad \text{または} \quad J_i = (\lambda).$$

$$P^{-1}A^mP = \begin{pmatrix} J_1^m & & \\ & \ddots & \\ & & J_r^m \end{pmatrix}$$

であり、

$$J^m = \begin{pmatrix} \lambda^m & m\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \cdots & \binom{m}{p-1}\lambda^{m-p+1} \\ & \lambda^m & m\lambda^{m-1} & & \vdots \\ & & \ddots & \ddots & \\ & & & \lambda^m & m\lambda^{m-1} \\ & & & & \lambda^m \end{pmatrix}.$$

これから

$$\lim_{m\to\infty} J^m = O \quad \Leftrightarrow \quad \lim_{m\to\infty} \binom{m}{j}\lambda^{m-j} = 0 \quad (0 \le j \le p-1) \quad \Leftrightarrow \quad |\lambda| < 1.$$

以上から明らか。■

こうして、行列の等比数列が分かったから、行列の等比級数

$$\sum_{m=0}^{\infty} A^m = \lim_{N\to\infty} \sum_{m=0}^{N} A^m = I + A + A^2 + \cdots$$

を調べよう[1]。これが収束すれば、一般項は O に収束する:

$$\lim_{m\to\infty} A^m = O.$$

ゆえに $r(A) < 1$. 実は逆も成立する。

系 6.2.2 $A \in M(n; \mathbf{C})$ のスペクトル半径 $r(A)$ が 1 より小さいならば、$I - A$ は正則で

$$(I - A)^{-1} = \sum_{m=0}^{\infty} A^m.$$

[1]関数解析で作用素の無限等比級数 $\sum_{n=0}^{\infty} A^n$ を Neumann 級数と呼ぶ。

証明　まず 0 は $I-A$ の固有値でない (もし 0 が $I-A$ の固有値ならば、1 は A の固有値となり、$r(A)\geq 1$ が導かれるが、これは仮定に矛盾する)。ゆえに $(I-A)^{-1}$ が存在する。恒等式

$$I-(I-A)(I+A+A^2+\cdots+A^m)=A^{m+1}$$

の両辺に $(I-A)^{-1}$ をかけると

$$(I-A)^{-1}-(I+A+A^2+\cdots+A^m)=(I-A)^{-1}A^{m+1}.$$

仮定 $r(A)<1$ より $m\to\infty$ のとき $A^{m+1}\to O$ であるから、上の式の右辺は O に収束する。ゆえに左辺も O に収束するので、

$$(I-A)^{-1}=\sum_{m=0}^{\infty}A^m. \quad \blacksquare$$

$\|(I-A)^{-1}\|$ の評価が欲しい場合もあるから、次の命題は上の系とは一応独立してあった方がいい。

命題 6.2.3 (Neumann 級数) $\|A\|<1$ を満すならば、$I-A$ は正則で

$$(I-A)^{-1}=\sum_{m=0}^{\infty}A^m.$$

このとき

$$\|(I-A)^{-1}\|\leq\frac{1}{1-\|A\|}.$$

証明　まず

$$\sum_{m=0}^{\infty}\|A\|^m$$

を収束する優級数として持つことから

$$\sum_{m=0}^{\infty}\|A^m\|$$

は収束する。$M(n;\mathbf{C})$ の完備性から

$$\sum_{m=0}^{\infty}A^m$$

は収束する。その和を S とおくと、

$$\|S\|\leq\sum_{m=0}^{\infty}\|A^m\|\leq\sum_{m=0}^{\infty}\|A\|^m=\frac{1}{1-\|A\|}.$$

以下 $S=(I-A)^{-1}$ であることを確かめる。$T\stackrel{\text{def.}}{=}I-A$ とおくと、

$$\begin{aligned}TS &= (I-A)\sum_{m=0}^{\infty}A^m=\sum_{m=0}^{\infty}A^m-\sum_{m=0}^{\infty}A^{m+1}=I,\\ ST &= \left(\sum_{m=0}^{\infty}A^m\right)(I-A)=\sum_{m=0}^{\infty}A^m-\sum_{m=0}^{\infty}A^{m+1}=I.\end{aligned}$$

ゆえに $S=T^{-1}$. (うるさく言うと、行列の積の連続性を用いている。) $\blacksquare$

6.3 反復法の収束原理

定理 6.3.1 $M \in M(n;\mathbf{C})$, $c \in \mathbf{C}^n$ とする。M のスペクトル半径が 1 より小さいならば、任意の初期ベクトル $x^{(0)}$ に対して反復法

$$x^{(k+1)} = Mx^{(k)} + c$$

による列 $\{x^{(k)}\}$ は、方程式

$$x = Mx + c$$

の解 (これは一意的) に収束する。

証明 $r(M) < 1$ ならば、$(I-M)^{-1}$ が存在するので、方程式

$$x = Mx + c$$

は一意解 $x^{(\infty)}$ を持つ。このとき

$$x^{(k)} - x^{(\infty)} = M(x^{(k-1)} - x^{(\infty)} = M^k(x^{(0)} - x^{\infty}).$$

ゆえに、$k \to \infty$ のとき、$x^{(k)} \to x^{(\infty)}$. ■

系 6.3.2 $M \in M(n;\mathbf{C})$, $c \in \mathbf{C}^n$ とする。ある行列ノルム $\|\cdot\|$ に対して $\|M\| < 1$ ならば、任意の初期ベクトル $x^{(0)}$ に対して反復法

$$x^{(k+1)} = Mx^{(k)} + c$$

による列 $\{x^{(k)}\}$ は、方程式

$$x = Mx + c$$

の解 (これは一意的) に収束する。

6.4 Jacobi (ヤコビ) 反復法

$A = (a_{ij})$ の対角要素は 0 でないとする:

$$a_{ii} \neq 0 \quad (i = 1, 2, \cdots, n).$$

Jacobi（ヤコビ） **反復法** (Jacobi iteration) とは

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^k \right) \quad (i = 1, 2, \cdots, n).$$

係数行列 $A = (a_{ij})$ を、その対角部分

$$D = \operatorname{diag}[a_{11}, \cdots, a_{nn}] = \begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix},$$

下三角部分

$$L = (\ell_{ij}), \quad \ell_{ij} = \begin{cases} a_{ij} & (j < i) \\ 0 & (j \geq i) \end{cases}$$

上三角部分

$$U = (u_{ij}), \quad u_{ij} = \begin{cases} 0 & (j \leq i) \\ a_{ij} & (j >) \end{cases}$$

の和に表わす:

$$A = L + D + U.$$

このとき Jacobi 反復は次のように表現できる:

$$x^{(k+1)} = D^{-1}(b - (L+U)x^{(k)}).$$

注意 6.4.1 (同時緩和法) (伊理 [16] より) 方程式系 $Ax = b$ の一つの方程式

$$\sum_{j=1}^{n} a_{ij}x_j = b_i$$

にだけ着目して他の方程式による束縛を “緩和 (relax)” して、次の近似値を計算するというやり方を、すべての変数について、“同時に” 行うというので、Jacobi 反復法のことを**同時緩和法** (**simultaneous relaxation**) と呼ぶこともある。

定理 6.4.2 A が狭義優対角行列のとき、Jacobi 反復行列のスペクトル半径は 1 より小さい (ゆえにヤコビ反復法は収束する)。

証明 (A が行に関して狭義の対角優位である場合)

$$\mu = \max_{1 \leq i \leq n} \frac{\displaystyle\sum_{1 \leq j \leq n, j \neq i} |a_{ij}|}{|a_{ii}|}$$

とおくと、$0 \leq \mu < 1$ であるが、このとき

$$\|I - BA\|_\infty = \max_i \left| \frac{1}{|a_{ii}|} \sum_{1 \leq j \leq n, j \neq i} |a_{ij}| \right| = \mu$$

であるから

$$\|\varepsilon^{(k)}\|_\infty \leq \mu^k \|\varepsilon^{(0)}\|_\infty$$

で

$$\lim_{k \to \infty} \varepsilon^{(k)} = 0.$$

(A が列に関して狭義対角優位の場合は、$\|\cdot\|_1$ について同様の議論が出来る。) ■

6.5 Gauss-Seidel (ガウス-ザイデル) 反復法

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right) \quad (i=1,2,\cdots,n)$$

$$\begin{aligned} x^{(k+1)} &= D^{-1}(b - Lx^{(k+1)} - Ux^{(k)}) \\ &= (D+L)^{-1}(b - Ux^{(k)}). \end{aligned}$$

定理 6.5.1 A が狭義優対角行列のとき、Gauss-Seidel 反復行列のスペクトル半径は 1 より小さい (ゆえに Gauss-Seidel 反復法は収束する)。

証明 (例えば、仁木-河野 [26] を見よ。) ■

定理 6.5.2 A が正定値 (Hermite) 行列のとき、Gauss-Seidel 反復法は収束する。

注意 6.5.3 Gauss-Seidel 反復法のことを**逐次緩和法** (**successive relaxation**) を呼ぶこともある。

6.6 SOR 法

SOR 法 (**Successive over-relaxation method**) とは

$$\begin{cases} \xi_i^{(k+1)} = \dfrac{1}{a_{ii}}(b_i - a_{i1}x_1^{(k+1)} - a_{i2}x_2^{(k+1)} - \cdots - a_{i,i-1}x_i^{(k+1)} - a_{i,i+1}x_{i+1}^{(k)} - \cdots - a_{in}x_n^{(k)}) \\ x_i^{(k+1)} = x_i^{(k)} + \omega(\xi_i^{(k+1)} - x_i^{(k)}) \end{cases}$$

行列、ベクトル記法で書くと

$$\begin{cases} \xi^{(k+1)} = D^{-1}(b + Ex^{(k+1)} + Fx^{(k)}) \\ x^{(k+1)} = x^{(k)} + \omega(\xi^{(k+1)} - x^{(k)}) \end{cases}$$

ω を**緩和因子**または**緩和係数**という。$\omega = 1$ のとき Gauss-Seidel 法に一致する。

$0 < \omega < 1$ のとき**過小緩和** (under-relaxation), $1 < \omega < 2$ のとき**過大緩和** (over-relaxation) という。

反復行列 M_{SOR} のスペクトル半径について

$$r(M_{\mathrm{SOR}}) \geq |\omega - 1|$$

が成り立つ (山本 [25] を見よ)。ゆえに SOR 法が収束するには $0 < \omega < 2$ が必要である。

定理 6.6.1 (Ostrowski) $A = (a_{ij})$ が実対称行列で $a_{ii} > 0$ $(1 \leq i \leq n)$ かつ $0 < \omega < 2$ とするとき、SOR 法が収束するための必要十分条件は A が正値であること。

証明 山本 [25] を見よ。 ■

第7章　最小自乗法

7.1　森-杉原-室田から

この節は、森-杉原-室田 [21] による。

7.1.1　イントロ

$A \in M(n, m; \mathbf{R})$, $b \in \mathbf{R}^n$, n 次正定値対称行列 G が与えられたとき

$$\phi(x) \overset{\text{def.}}{=} \frac{1}{2}(Ax - b, Ax - b)_G = \frac{1}{2}(Ax - b)^T G(Ax - b)$$

を最小にする $x \in \mathbf{R}^m$ を求める問題を**最小 2 乗問題**という。G は重み行列と呼ぶ。

例 7.1.1 統計データの最小 2 乗法など。この場合 G は分散・共分散行列の逆行列に相当する。

簡単のため、この節では、$G = I_n$ (n 次正方行列), $n \geq m$ の場合を考える。このとき

$$\phi(x) \overset{\text{def.}}{=} \frac{1}{2}(Ax - b, Ax - b)$$

であるから

$$\nabla \phi(x) = A^T Ax - A^t b.$$

そして

$$\phi \text{ の Hesse 行列} = A^T A$$

で、$A^T A$ は半正定値である。

結局、問題は

$$A^T Ax = A^T b$$

という方程式を解くことに帰着される。この方程式を**正規方程式**と呼ぶ。

$r \overset{\text{def.}}{=} \operatorname{rank} A < m$ のときには、$A^T A$ は特異になるが、それでも、正規方程式は任意の b に対して解 x を持つ。(なぜだろう？)

7.1.2　解法

正規方程式を直接解く

案外効率がいいが、$A^T A$ の条件数が A の条件数の 2 乗程度に大きくなるので、解の精度が悪くなる可能性が高い。

QR 分解を用いて解く

特異値分解を用いて解く

第8章　LU 分解

これについては桂田研ノートが詳しい。

8.1

この節は村田・名取・唐木 [23] を参考にした。

正方行列 A を下三角行列 L と上三角行列 U の積として

$$A = LU, \quad L = \quad U =$$

と表わすことを A の LU 分解と呼ぶ。任意の正方行列が LU 分解できるわけではない。

正則行列 A に関する連立 1 次方程式 $Ax = b$ は部分枢軸選択を用いた Gauss の消去法によって解くことが出来るが、この前進消去過程は

$$G_{n-1}P_{n-1}\cdots G_kP_k\cdots G_1P_1Ax = G_{n-1}P_{n-1}\cdots G_kP_k\cdots G_1P_1b$$

と行列表現できる。ここで P_k は第 k 段 (第 k 行による掃き出しの手続き) における部分枢軸選択による行交換を表わし、G_k は第 k 段における前進消去過程を表わす。これで上三角方程式

$$Ux = c, \quad U = G_{n-1}P_{n-1}\cdots G_kP_k\cdots G_1P_1A, \quad c = G_{n-1}P_{n-1}\cdots G_kP_k\cdots G_1P_1b$$

が得られた。これから

$$L \stackrel{\text{def.}}{=} (G_{n-1}P_{n-1}\cdots G_kP_k\cdots G_1P_1)^{-1}$$

とおけば、

$$A = LU$$

となる。U は上三角行列 (upper triangular matrix) であるが、L は一般には下三角行列ではない。A が対角優位の正則行列である場合のように、部分枢軸選択がまったく行われない (or 行わないで済ませる) とき、言い換えると

$$P_{n-1} = \cdots = P_1 = I = n \text{ 次単位行列}$$

のときに限り、

$$L = (G_{n-1}\cdots G_1)^{-1}$$

となって L は下三角行列になる。この場合は

$$A = LU$$

の成分表現

$$a_{ij} = \sum_{k=1}^{n} \ell_{ik}u_{kj}$$

に基づいて、L と U の要素を計算するアルゴリズムがある。この場合 $u_{ii} = 1$ $(1 \le i \le n)$ とするのをCrout（クラウト）法、$\ell_{ii} = 1$ $(1 \le i \le n)$ とするのをDolittle（ドーリットル）法と呼ぶ。

命題 8.1.1 $A=(a_{ij})$ が LU 分解可能であるための必要十分条件は

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ & \cdots & \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} \neq 0 \quad (1 \le k \le n)$$

である。

証明 山本 [25] に命題が書いてあるが、証明は？

上の命題から狭義優対角行列、正値対称行列は LU 分解可能であることが分かる。

8.2 Cholesky 分解

特に A が正値対称行列である場合は

$$A = LL^T$$

の形につねに LU 分解できる。これを **Cholesky分解**（コレスキー）と呼ぶ。

Cholesky 分解のアルゴリズム

(i)

$$\begin{aligned} \ell_{11} &= \sqrt{a_{11}}, \\ \ell_{i1} &= \frac{a_{i1}}{\ell_{11}} \qquad (2 \le i \le n) \end{aligned}$$

(ii) 各 $j = 2, 3, \cdots, n$ につき

$$\begin{aligned} \ell_{jj} &= \sqrt{a_{jj} - \sum_{k=1}^{j-1} \ell_{jk}^2}, \\ \ell_{ij} &= \frac{1}{\ell_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} \ell_{ik}\ell_{jk} \right) \qquad (j+1 \le i \le n). \end{aligned}$$

乗除算 $n^3/6 + O(n^2)$ 回、開平 n 回あれば Cholesky 分解できるわけである。

第9章　特殊な形の行列

9.1　行列の可約性、既約性

定義 9.1.1 (行列の可約性、既約性) $A \in M(n; \mathbf{C})$ とする。

(1) ［$n \geq 2$ のとき］適当な置換行列 P を用いて

$$PAP^T = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$$

(A_{11} は r 次正方行列, A_{22} は $n-r$ 次の正方行列, O は 零行列),

$(1 \leq r \leq n-1)$ と表わされるとき、A は**可約** (**reducible**) といい、このような置換行列が存在しないとき、A は**既約** (**ireducible**) であるという。

(2) ［$n = 1$ のとき］A の要素が 0 ならば可約であるといい、0 でないならば既約であるという。

注意 9.1.2 可約性の条件の簡単な言い換えとして、次のものがある。

可約性の自明な言い換え

自然数の集合 $\{1, 2, \cdots, n\}$ の空でない部分集合 J を適当に選ぶと

$$i \in J, \quad j \notin J \Longrightarrow a_{ij} = 0$$

とできる。

可約な係数行列を持つ連立１次方程式は、小さな問題に分解して解くことが出来る。
行列の既約性の判定は次の定理を用いればよい。

定理 9.1.3 (既約性の判定) $n \geq 2$, $A = (a_{ij}) \in M(n; \mathbf{C})$ とするとき、次の二条件は互いに同値である。

(i) A は既約である。

(ii) 任意の i, j $(i \neq j)$ に対して、適当な i_1, i_2, $\cdots$, $i_r \in \{1, 2, \cdots, n\}$ を選んで、

$$a_{ii_1}, a_{i_1 i_2}, \cdots, a_{i_r j} \neq 0$$

と出来る。

$A = (a_{ij}) \in M(n; \mathbf{C})$ に対して、平面上に相異なる点 P_1, $\cdots$, P_n を取り (以下、これらを節点と呼ぶ)、行列のすべての非零要素 a_{ij} に対して、節点 P_i から 節点 P_j に至る経路 $\xrightarrow{P_i P_j}$ を作る。これを A の**有限有向グラフ** (**finite directed graph**) という。

定義 9.1.4 (有向グラフの強連結性) 有向グラフの相異なる任意の二節点 P_i, P_j に対して、P_i と P_j を結ぶ経路の列

$$\overrightarrow{P_iP_{i_1}}, \overrightarrow{P_{i_1}P_{i_2}}, \cdots, \overrightarrow{P_{i_r}P_j}$$

が存在するならば、この有向グラフを**強連結** (**strongly connected**) と呼ぶ。

次は明らかであろう。

定理 9.1.5 $n \geq 2$, $A \in M(n; \mathbf{C})$ とするとき、次の二条件は同値である。

(i) A は既約である。

(ii) A の有向グラフ $G(A)$ は強連結である。

9.2 優対角行列

定義 9.2.1 $A = (a_{ij}) \in M(n; \mathbf{C})$ とする。

(1)

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad (i = 1, 2, \cdots, n)$$

を満すとき、A は (行に関して; 行方向の) **優対角行列** (**diagonally dominant matrix**) であるという。

(2)

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad (i = 1, 2, \cdots, n)$$

を満すとき、A は (行に関して; 行方向の) **狭義優対角行列** (**strictly diagonally dominant matrix**) であるという。

(3) A が既約かつ優対角で、少なくても一つの i_0 に対して

$$|a_{i_0 i_0}| > \sum_{\substack{1 \leq j \leq n \\ j \neq i_0}} |a_{i_0 j}| \tag{9.1}$$

をみたすとき、A は**既約優対角** (**irreducibly diagonally dominant**) という。

注意 9.2.2 「既約優対角 = 既約かつ優対角」ではないことに注意しよう。追加の条件 (9.1) が必要である。■

定理 9.2.3 A が狭義優対角行列のとき、ヤコビ反復法のスペクトル半径は 1 より小さい (ゆえにヤコビ法は収束する)。

証明 (例えば、仁木-河野 [26] を見よ。) ■

定理 9.2.4 (Levy-Desplanques) A が狭義優対角ならば A は正則である。

証明　背理法による。狭義優対角行列 $A=(a_{ij})$ が正則でないと仮定すると、$Ax=0$ を満たすベクトル $x\ (\neq 0)$ がある[1]。

$$x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{9.2}$$

とし、絶対値最大の成分を x_k とする。方程式 $Ax=0$ の第 k 成分を書くと

$$\sum_{j=1}^n a_{kj}x_j=0. \tag{9.3}$$

この等式の左辺の第 k 項のみ残して、残りは右辺に移項すると

$$a_{kk}x_k=-\sum_{j\neq k} a_{kj}x_j. \tag{9.4}$$

絶対値を取って、

$$|a_{kk}||x_k|=\left|-\sum_{j\neq k}a_{kj}x_j\right| \leqq \sum_{j\neq k}|a_{kj}||x_j| \leqq \left(\sum_{j\neq k}|a_{kj}|\right)|x_k|. \tag{9.5}$$

両辺を $|x_k|\ (>0)$ で割算して

$$|a_{kk}| \leqq \sum_{j\neq k}|a_{jk}|. \tag{9.6}$$

これは行列 A が狭義優対角であるという仮定に矛盾する。 ■

定理 9.2.5 (Taussky) A が既約優対角ならば A は正則である。

証明　(山本 [25] を参考にした。) A が正則でないと仮定して矛盾を導く。A が正則でなければ、$Ax=0,\ x\neq 0$ を満たす $x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ が存在する。このとき $|x_1|=|x_2|=\cdots=|x_n|$ ではない。実際、(9.1) を満たす i_0 に対して

$$|a_{i_0i_0}|\,|x_{i_0}|=|a_{i_0i_0}x_{i_0}|=\left|-\sum_{\substack{1\leq j\leq n\\ j\neq i_0}} a_{i_0j}x_j\right| \leq \left(\sum_{\substack{1\leq j\leq n\\ j\neq i_0}}|a_{i_0j}|\right)|x_{i_0}|$$

が成り立ち、$|x_{i_0}|$ で割って $|a_{i_0i_0}|\leq \sum_{\substack{1\leq j\leq n\\ j\neq i_0}}|a_{i_0j}|$ となるので矛盾が生じる。

　ゆえに

$$N:=\{1,2,\cdots,N\},\quad J:=\left\{k\in N\ \middle|\ |x_k|=\max_{1\leq i\leq n}|x_i|\right\}$$

[1] 「任意の n 次の実正方行列 A と、A により定まる線型写像 $f\colon \mathbf{R}^n\ni x\mapsto Ax\in\mathbf{R}^n$ に対して、以下の条件は互いに同値である。(i) A は正則である (A^{-1} が存在する), (ii) f は全単射である, (iii) f は単射である, (iv) f は全射である, (v) $\ker f=\{0\}$, (vi) $\operatorname{rank} f=n$, (vii) $\det A\neq 0$.」— この定理の証明の鍵は、次元定理「$n-\dim\ker f=\operatorname{rank} f$」である (これは準同型定理の系ですね)。

とおくとき、$\emptyset \neq J \neq N$.

実は、

$$(\forall i \in J)(\forall j \in N \setminus J) \quad a_{ij} = 0. \tag{9.7}$$

これが証明できれば、A の既約性に矛盾して、この命題の証明が完了する。

(9.7) を否定すると、

$$(\exists i_1 \in J)(\exists j_1 \in N \setminus J) \quad a_{i_1 j_1} \neq 0.$$

このとき、

$$|a_{i_1 i_1} x_i| = \left| - \sum_{\substack{1 \le j \le n \\ j \neq i_1}} a_{i_1 j} x_j \right| \le \sum_{\substack{1 \le j \le n \\ j \neq i_1}} |a_{i_1 j}| \, |x_j| = \sum_{\substack{j \in J \\ j \neq i}} |a_{i_1 j}| \, |x_j| + \sum_{j \in N \setminus J} |a_{i_1 j}| \, |x_j| \, .$$

$i_1 \in J$ であるから、任意の $j \in J$ に対して $|x_j| = |x_{i_1}|$ が成り立ち、

$$\sum_{\substack{j \in J \\ j \neq i_1}} |a_{i_1 j}| \, |x_j| = \sum_{\substack{j \in J \\ j \neq i}} |a_{ij}| \, |x_i| = \left(\sum_{\substack{j \in J \\ j \neq i_1}} |a_{i_1 j}| \right) | \, |x_{i_1}|$$

$i_1 \in J$ であるから、任意の $j \in N \setminus J$ に対して、$|x_j| < |x_{i_1}|$ が成り立ち、特に $j = j_1$ に対して $a_{i_1 j_1} \neq 0$ であるから

$$\sum_{j \in N \setminus J} |a_{i_1 j}| \, |x_j| < \sum_{j \in N \setminus J} |a_{i_1 j}| \, |x_{i_1}| = \left(\sum_{j \in N \setminus J} |a_{i_1 j}| \right) | \, |x_{i_1}| \, .$$

ゆえに

$$|a_{i_1 i_1}| \, |x_{i_1}| < \left(\sum_{\substack{1 \le j \le n \\ j \neq i_1}} |a_{i_1 j}| \right) |x_{i_1}| \, .$$

割り算して

$$|a_{i_1 i_1}| < \sum_{\substack{1 \le j \le n \\ j \neq i_1}} |a_{i_1 j}| \, .$$

これは A が優対角であることに矛盾する。 ■

例 9.2.6 (1 次元 Poisson 方程式に対する差分方程式の一意可解性) $A = \frac{1}{h^2}(2I - J) = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & - & \\ & & \ddots & \ddots & \\ & & & - & \\ & & & & \end{pmatrix}$

は既約優対角である。

実際、

(a) 優対角であることは明らか。

(b) $1 \le i \le n-1$ に対して、$a_{i,i+1} \neq 0$ であるから

$$1 \to 2 \to 3 \to \cdots \to n-1 \to n.$$

また $2 \le i \le n$ に対して、$a_{i,i-1} \ne 0$ であるから、

$$n \to n-1 \to n-2 \to \cdots \to 2 \to 1.$$

ゆえに既約である。

(c) さらに

$$|a_{11}| = 2 > 1 = sum_{\substack{1\le j\le n \\ j\ne 1}} |a_{1j}|.$$

ゆえに A は正則である。■

例 9.2.7 (2 次元 Poisson 方程式に対する差分方程式の一意可解性) 2次元長方形領域における Poisson 方程式の Dirichlet 境界値問題を差分近似して得られる連立1次方程式の係数行列も、既約優対角であるので正則である。

x 軸方向を N_x 等分、y 軸方向を N_y 等分したとし、$N := N_x - 1$, $n := (N_x - 1)(N_y - 1)$ とおく。格子点における値 $u(x_i, y_j)$ を column first で並べてベクトル化すると、係数行列は

$$A = \frac{1}{h_y^2} K_{N_y-1} \otimes I_{N_x-1} + I_{N_y-1} \otimes \frac{1}{h_x^2} K_{N_x-1}$$

ただし K_m は次式で定義される m 次正方行列とする:

$$K_m := 2I_m - J_m.$$

A の優対角性は明らかである。また1行目を見れば

$$|a_{11}| > \sum_{\substack{1\le j\le n \\ j\ne 1}} |a_{1j}|.$$

以下、既約性を示す。

第 k 対角ブロックは、対角線とその両隣が 0 でないので、

$$\mathcal{I}_k := \{(k-1)N+1, (k-1)N+2, \cdots, , (k-1)+N-1, kN\}$$

は強連結である $(k = 1, 2, \cdots, N_y - 1)$。

$k = 1, 2, \cdots, N_y - 1$ に対して、第 k 対角ブロックの右にあるブロックにより、$\mathcal{I}_k$ から $\mathcal{I}_{k+1}$ につながる。

$k = N_y, N_y - 1, \cdots, 2$ に対して、第 k 対角ブロックの左にあるブロックにより、$\mathcal{I}_k$ から $\mathcal{I}_{k-1}$ につながる。

以上から $\displaystyle\bigcup_{k=1}^{n} \mathcal{I}_k = \{1, 2, \cdots, n\}$ は強連結である。ゆえに A は既約である。■

9.3　非負行列

定義 9.3.1 (非負行列、Z 行列、L 行列、単調行列、M 行列) $A=(a_{ij})\in M(n;\mathbf{R})$ とする。

(1) 任意の i,j について $a_{ij}\geq 0$ が成り立つとき、A を**非負行列**といい、しばしば $A\geq O$ と表わす。

(2) 非対角線分がすべて 0 以下である、つまり $a_{ij}\leq 0$ $(i\neq j)$ が成り立つとき、A を **Z 行列**であるという。

(3) A が Z 行列かつ $a_{ii}>0$ $(i=1,2,\cdots,n)$ を満すとき、A を **L 行列**という。

(4) A が正則で、A^{-1} が非負行列であるとき、A を**単調行列** (**monotone matrix**) という。

(5) A が単調で Z 行列であるとき、A は **M 行列**であるという。

定理 9.3.2 (Perron-Frobenius) A が既約な非負行列であるとき、以下の (1) ～ (6) が成り立つ。

(1) A のスペクトル半径 $r(A)$ は正数で、A の単純固有値である。

(2) $r(A)$ に正の実固有ベクトル $x>0$ が対応する。
(ベクトルが正であるとは、各成分が >0 であることとする。)

(3) A の任意の要素を増加すると $r(A)$ が増加する。

(4) A の正の固有ベクトルは x に比例する。

(5) $|\lambda|=r(A)$ なる A の固有値 λ は単純である。

(6) A が正 (すなわち、すべての成分が >0) であるならば、$|\lambda|=r(A)$ なる A の固有値 λ は $r(A)$ 以外に存在しない。

証明　Varga [37], 伊理 [16]. ■

命題 9.3.3 行列 $A\in M(n;\mathbf{R})$ が単調であるための必要十分条件は

$$Ax\geq 0\Longrightarrow x\geq 0. \tag{9.8}$$

証明　A が単調であれば、$Ax\geq 0$ に $A^{-1}\geq 0$ をかけて $x\geq 0$. 逆に (9.8) が成り立つとすると、

$$Ax\leq 0\Longrightarrow x\leq 0$$

も成り立つので、$Ax=0\Longrightarrow x=0$ が導かれ、A が正則であることが分かる。すると条件 (9.8) は

$$y\geq 0\Longrightarrow A^{-1}y\geq 0$$

と書き換えられる。$y=\vec{e}_i$ $(i=1,2,\ldots,n)$ とすると、A^{-1} の各成分が 0 以上であることが示される。■

定理 9.3.4 n 次 Z 行列 $A = (a_{ij})$ が M 行列であるための必要十分条件は次の二条件が成り立つことである。

(1) $a_{ij} > 0$ $(1 \leq i \leq n)$ (つまり A は L 行列)

(2) $D = \mathrm{diag}(a_{11}, \cdots, a_{nn}) = \begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix}$, $B = I - D^{-1}A$ とおくとき、B のスペクトル半径は 1 より小さい: $r(B) < 1$.

証明 仁木-河野 [26] p.23 を見よ。■

定理 9.3.5 n 次非負行列 A, $\alpha > 0$ について、次の二条件は同値である。

(i) $\alpha > r(A)$.

(ii) $\alpha I - A$ は正則で $(\alpha I - A)^{-1} \geq O$.

証明 仁木-河野 [26] p.24 を見よ。■

定理 9.3.6 n 次非負行列 A に対して、次の二条件は同値である。

(i) $\alpha > r(A)$ で A は既約である。

(ii) $\alpha I - A$ が正則で $(\alpha I - A)^{-1} > O$.

証明 仁木-河野 [26] p.25 を見よ。■

定理 9.3.7 $A \in M(n; \mathbf{R})$ が Z 行列ならば、次の二条件は同値である。

(i) A は正則で $A^{-1} > O$.

(ii) A の対角要素が正の実数で、$D \stackrel{\text{def.}}{=} \mathrm{diag}(a_{11}, \cdots, a_{nn})$, $B \stackrel{\text{def.}}{=} I - D^{-1}A$ とおくとき、B は非負行列で $r(B) < 1$ を満たす。

証明 仁木-河野 [26] p.25 を見よ。■

命題 9.3.8 $A = (a_{ij}) \in M(n; \mathbf{R})$ について次の 2 条件は同値である。

(1) A は M 行列である。

(2) $a_{ii} > 0$ $(i = 1, 2, \ldots, n)$, $a_{ij} \leq 0$ $(i \neq j)$, $A \in GL(n; \mathbf{R})$, $A^{-1} \geq O$.

証明 略 (伊理 [16] p.267) を見よ。■

命題 9.3.9 n 次非負行列 A と正数 μ に対して、次の 6 条件は互いに同値である。

(i) $\forall b \in R^n$, $b \geq 0$ に対して、$\exists x \geq 0$ s.t. $(\mu I_n - A)x = b$.

(ii) $\exists x > 0$ s.t. $(\mu I_n - A)x > 0$.

(iii) $p \to \infty$ のとき $(\mu^{-1}A)^p \to O$.

(iv) $r(A) < \mu$. ($r(A)$ は A のスペクトル半径である。)

(v) $\mu I_n - A$ のすべての首座小行列式の値は正である。

(vi) $\mu I_n - A$ は $\mu I_n - A = LU$ (L は対角要素が 1 で非対角要素は ≤ 0 である下三角行列、U は対角要素が正で非対角要素は ≤ 0 である上三角行列) と LU 分解される。

証明 略 (伊理 [16] pp.265–266) を見よ。■

行列 B について、B が M 行列であるための必要十分条件は、B が上の命題の条件を満たす A, μ を用いて $B = \mu I - A$ と表わされることである。

9.4 差分法に現れる行列の正則性

9.4.1 Gauss の消去法にからめて

定理 9.4.1 次の条件 (i)∼ (iv) のいずれかを満たす行列 A は条件

$$a_{11}^{(1)} \neq 0, \quad a_{22}^{(2)} \neq 0, \quad \cdots, \quad a_{nn}^{(n)} \neq 0 \tag{9.9}$$

($a_{kk}^{(k)}$ が枢軸選択なしの Gauss の消去法のピボット) を満たす。従って A は正則である。

(i) ［行方向の狭義優対角行列］ $|a_{ii}| > \displaystyle\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|$ $(i = 1, 2, \cdots, n)$.

(ii) ［列方向の狭義優対角行列］ $|a_{jj}| > \displaystyle\sum_{\substack{i=1\\ i\neq j}}^{n} |a_{ij}|$ $(j = 1, 2, \cdots, n)$.

(iii) ［対称部分が正定値］ $\dfrac{1}{2}(A + A^T)$ が正定値 i.e. $x^T A x = \dfrac{1}{2} x^T (A + A^T) x > 0$ $(\forall x \neq 0)$.

(iv) ［M 行列］正則で、$a_{ij} \leq 0$ $(i \neq j)$, $A^{-1} \geq O$ (A^{-1} の各要素が非負)

9.4.2　既約優対角性について

定理 9.4.2 (Levy-Desplanques) n 次正方行列 $A = (a_{ij})$ が狭義優対角、すなわち

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad (i = 1, 2, \cdots, n)$$

を満たすとする。このとき、A は正則である。

定義 9.4.3 (行列の既約性) n 次正方行列 $A = (a_{ij})$ が既約であるとは、 $n = 1$ であるか、次の条件 (I) またはそれと同値な条件 (II) を満たすことであると定義する。

1. $i \neq j$ $(1 \leq i \leq n, 1 \leq j \leq n)$ ならば $a_{ij} \neq 0$ であるか、または $a_{ii_1}a_{i_1i_2}\cdots a_{i_r,j} \neq 0$ を満たす相異なる自然数 $i_1, \cdots, i_r$ $(1 \leq i_1, \cdots, i_r \leq n)$ が存在する。

2. 平面上に n 個の点 $P_1, \cdots, P_n$ を取り、すべての非零要素 a_{ij} に対して、P_i から P_j に向かう経路 $\xrightarrow{P_iP_j}$ を図のように作れば、得られるグラフ (有限有向グラフ) が強連結である。すなわち、任意の異なる i, j につき、P_i から P_j に至る経路の列

$$\xrightarrow{P_iP_{i_1}}, \xrightarrow{P_{i_1}P_{i_2}}, \cdots, \xrightarrow{P_{i_r}P_j} \text{が存在する。}$$

定義 9.4.4 既約な n 次正方行列 $A = (a_{ij})$ が条件

$$\begin{cases} |a_{ii}| \geq \displaystyle\sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad (1 \leq i \leq n) \\ \text{少なくとも 1 つの } i \text{ につき } |a_{ii}| > \displaystyle\sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \end{cases}$$

を満たすとき、**既約優対角行列**という。

定理 9.4.5 (Taussky) 既約優対角行列は正則である。

命題 9.4.6 狭義優対角な L 行列は M 行列である。

命題 9.4.7 既約優対角な L 行列 A は M 行列であり、$A^{-1} > 0$ を満たす。

9.5 三重対角行列についてのメモ — 特に固有値

9.5.1 差分法に出て来る三重対角行列

B を

$$B = \begin{pmatrix} 0 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 0 \end{pmatrix}$$

なる N 次正方行列とするとき、固有値と対応する固有ベクトルは

$$\mu_j = 2\cos\frac{j\pi}{N+1}, \quad \vec{w}_j = \left(\sin\frac{j\pi}{N+1}, \sin\frac{2j\pi}{N+1}, \cdots, \sin\frac{Nj\pi}{N+1}\right)^T \quad (j = 1, 2, \cdots, N).$$

(森-杉原-室田 [21] を見よ。)

あるいは、T を

$$T = \begin{pmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{pmatrix}$$

なる N 次正方行列とするとき、固有値と対応する固有ベクトルは

$$\lambda_j = -4\sin^2\left(\frac{j\pi}{2(N+1)}\right), \quad \vec{v}_j = \left(\sin\frac{j\pi}{N+1}, \sin\frac{2j\pi}{N+1}, \cdots, \sin\frac{Nj\pi}{N+1}\right)^T \quad (j = 1, 2, \cdots, N).$$

(例えば Smith [34] を見よ。)

注意 9.5.1 $f(x) = 1 + x$ とおくと、$T = f(B)$. また $g(x) = x - 1$ とおくと、$B = g(T)$. これから、B, T いずれか一方についての情報から他方の情報は簡単に導かれる。

$A \overset{\text{def.}}{=} -T$ の LDU 分解が Strang [35] に載っている。

$$\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & & & & \\ -\frac{1}{2} & 1 & & & \\ & -\frac{2}{3} & 1 & & \\ & & \ddots & \ddots & \\ & & & \frac{1-n}{n} & 1 \end{pmatrix} \begin{pmatrix} 2 & & & & \\ & \frac{3}{2} & & & \\ & & \frac{4}{3} & & \\ & & & \ddots & \\ & & & & \frac{n+1}{n} \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} & & & \\ & 1 & -\frac{2}{3} & & \\ & & 1 & \ddots & \\ & & & \ddots & \frac{1-n}{n} \\ & & & & 1 \end{pmatrix}.$$

特に

$$\det A = 2\left(\frac{3}{2}\right)\left(\frac{4}{3}\right)\cdots\left(\frac{n+1}{n}\right) = n + 1.$$

9.5.2 やや一般の場合 — 一様三項行列の固有値

N 次正方行列

$$\begin{pmatrix} a & b & & & & \\ c & a & b & & & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & & & c & a & b \\ & & & & c & a \end{pmatrix}$$

(ただし $bc \neq 0$) の固有値、固有ベクトルは

$$\begin{aligned} \lambda_j &= a + 2\sqrt{bc}\cos\frac{j\pi}{N+1}, \quad \vec{v}_j = \left(\sqrt{\frac{c}{b}}\sin\frac{j\pi}{N+1}, \left(\sqrt{\frac{c}{b}}\right)^2 \sin\frac{2j\pi}{N+1}, \cdots, \left(\sqrt{\frac{c}{b}}\right)^N \sin\frac{Nj\pi}{N+1}\right)^T \\ & \qquad (j = 1, 2, \cdots, N). \end{aligned}$$

(これも Smith [34] にある。)

確認 $a = 0$ の場合を証明すれば十分である。

$$x_i = \left(\sqrt{\frac{c}{b}}\right)^i \sin\frac{ij\pi}{N+1}$$

とすると、$2 \le i \le N-1$ なる i に対して、

$$\begin{aligned} cx_{i-1} + bx_{i+1} - \lambda_j x_i &= c\left(\sqrt{\frac{c}{b}}\right)^{i-1} \sin\frac{(i-1)j\pi}{N+1} + b\left(\sqrt{\frac{c}{b}}\right)^{i+1} \sin\frac{(i+1)j\pi}{N+1} \\ & \quad -2\sqrt{bc}\cos\frac{j\pi}{N+1}\left(\sqrt{\frac{c}{b}}\right)^i \sin\frac{ij\pi}{N+1} \\ &= c\left(\sqrt{\frac{c}{b}}\right)^{i-1}\left[\sin\frac{(i-1)j\pi}{N+1} + \sin\frac{(i+1)j\pi}{N+1} - 2\cos\frac{j\pi}{N+1}\sin\frac{ij\pi}{N+1}\right] \\ &= c\left(\sqrt{\frac{c}{b}}\right)^{i-1}\left[\sin(i-1)\theta + \sin(i+1)\theta - 2\cos\theta\sin i\theta\right] = 0. \\ & \qquad (\theta = j\pi/(N+1) \text{ とおいた。}) \end{aligned}$$

同様にして $bx_2 = \lambda_j x_1$, $cx_{N-1} = \lambda_j x_N$ も示せる。■

9.5.3 ある三重対角行列の逆行列

n 次正方行列

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

の逆行列は

$$A^{-1} = \begin{pmatrix} n & n-1 & & 2 & 1 \\ n-1 & n-1 & & 2 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 2 & 2 & \cdots & 2 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}$$

である。

($b_{ij} = n + 1 - \max\{i, j\}$ で定義される $B = (b_{ij})$ が A^{-1} に他ならない、ということ？)

もとの行列が三重対角のような疎行列であっても、その逆行列が疎であるとは限らないことの典型的な例。

9.5.4 三重対角行列の固有値

この小節の内容は山本 [25] による。

N 次正方行列

$$\begin{pmatrix} b_1 & c_1 & & & & \\ a_2 & b_2 & c_2 & & & \\ & a_3 & b_3 & c_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & & a_n & b_n \end{pmatrix}$$

について以下の命題が成り立つ。

命題 9.5.2 $a_i \neq 0$ $(2 \leq i \leq n)$ のとき、A が対角化可能であるための必要十分条件は、A の固有値がすべて相異なることである。

系 9.5.3 $a_i \neq 0$ $(2 \leq i \leq n)$ のとき、実対称三重対角行列 A の固有値は、すべて相異なる実数である。

系 9.5.4 $a_i \neq 0$ $(2 \leq i \leq n)$ のとき、実交代三重対角行列 A の固有値は、すべて相異なる純虚数である。

命題 9.5.5 A の固有値は $a_{i+1}c_i > 0$ $(1 \leq i \leq n-1)$ のとき相異なる実数 (従って A は対角化可能)、$a_{i+1}c_i <$ $(1 \leq i \leq n-1)$ のとき相異なる純虚数 (従って A は対角化可能) である。

第10章　行列の条件数について — 松尾・杉原・森論文

日本応用数理学会論文誌 Vol.7, No.3, 1997, pp.307–319 に載っている

松尾宇泰、杉原正顕、森正武、行列の条件数の推定方法の数値的評価

を引用する。

序文　「本ノートでは、1 乗ノルム $\|\cdot\|_1$ による行列の条件数 $\mathrm{cond}_1(A) = \|A\|_1\|A^{-1}\|_1$ の推定方法を取り扱う。世界的には、LINPACK で採用されている Cline et al. の方法 [4], および LAPACK で採用されている Hager の方法 [9] が有名であり、その有効性に関する研究も様々な形で行われ、現在、Hager の方法がもっとも優れているとされている [11, 13]。一方、我が国では、名取・塚本による方法 [22] が比較的良く知られており (残念ながら世界的にはほとんど知られていない)、さらに [22] では、行列のサイズが小さい場合であるが、この方法が LINPACK で採用されている方法よりも優れていることが指摘されている。そこで、ここでは、3 つの方法の性能を数値実験を通して比較し、実用にあたって方法を選択するための指針を与える。」

Hager の方法　「Hager の方法 [9] は、前 2 者のアルゴリズムとはまったく異なり、凸集合上での凸解析の考えから導かれたものである。

[1] $b := (1/n, 1/n, \cdots, 1/n)$, $\rho := 0$.

[2] $Ax = b$ を解く。

[3] もし $\|x\|_1 \le \rho$ ならば手順 [8] へ進む。さもなくば、$\rho := \|x\|_1$ として次に進む。

[4] $y_i := \mathrm{sgn}\,(x_i)$ $(i = 1, 2, \cdots, n)$

[5] 方程式 $A^T z = y$ を解く。

[6] $\max\limits_i |z_i|$ を達成する i を $i_{\max}$ とおく。

[7] もし $|z_{\max}| < z^T b$ ならば手順 [8] へ進む。さもなくば、

$$b_i := \begin{cases} 1 & (i = i_{\max}) \\ 0 & (i \ne i_{\max}) \end{cases}$$

とおいて手順 [2] に戻る。

[8] $\mathrm{cond}_1(A) = \|A\|_1\|A^{-1}\|_1 \simeq \|A\|_1\rho$ とする。

Hager のアルゴリズムは、M 行列の場合に、2 回の反復の後に真値を与え、停止することが知られている [12]。名取・塚本の方法も M 行列に対して条件数の真値を与えるが、それはピボット選択がないという前提の下に正しい。これに対して、Hager のアルゴリズムは、この全体が必要なく、M 行列に対しては常に真値を与える。」

結論部分　「$\|A^{-1}\|_1$ の推定アルゴリズムとして、3つの方法を取り上げて、数値的に性能を評価した。名取・塚本の方法と LINPACK の方法については、内部動作の別で、さらに2種類のアルゴリズムに分けて評価した。結果は、次のようにまとめられる。

- Hager の方法は、非常に良い推定値を与える。反復回数も高々 4 回程度で、LU 分解の手間と比較すれば、他の方法と比べても実用上耐えられる程度の手間ですむ。したがって、計算時間より精度を重視する場合は、Hager の方法が推奨できる。

- LINPACK の方法、および名取・塚本の方法は、いずれも条件数のオーダーを再現する程度の精度で、ほぼ同様であるが、名取・塚本の方法は行列が比較的小さい場合に良い精度を与え、行列が大きい場合には LINPACK の方法の方が優位である。したがって、行列が比較的小さく、精度をあまり重視しない用途には、アルゴリズム名取・塚本を、また行列が比較的大きく、精度をあまり重視しない用途には LINPACK の方法を推奨する。なお、内部動作の別による差異は認められなかった。従って、計算時間の面、およびプログラミングのしやすさの面で、簡単な方を用いることが推奨できる。」

第11章　有限次元版のFredholmの交代定理を目指して

(ただ今工事中)

11.1　チラシの裏で

有限次元の場合は良く知られている (本に載っている) みたいに書いてあったりするけれど…

- A. G. Ramm, “A Simple Proof of the Fredholm Alternative and a Characterization of the Fredholm Operators”, American Mathematical Monthly, 108 (2001) p. 855.

- X, Y を体 $\mathbb{K}$ 上の線形空間、$F\colon X \to Y$ を線形写像とする。$\forall \psi \in Y^*$ に対して

$$\langle F^*(\psi), x\rangle := \langle \psi, F(x)\rangle \quad (x \in X)$$

  とおくと、$F^*(\psi) \in X^*$ となる。ゆえに $F^*\colon Y^* \to X^*$ という写像が定義できる。これは明らかに線形である。

- $\dim X = n < \infty$, $\dim Y = m < \infty$ とする。X の基底 $e_1, \dots, e_n$ と、Y の基底 $f_1, \dots, f_m$ を取るとき、$\exists 1 A \in M(m, n; \mathbb{K})$ s.t.

$$(F(e_1) \cdots F(e_n)) = (f_1 \cdots f_m)A.$$

  この A を $e_1, \cdots, e_n$ と $f_1, \cdots, f_m$ に関する F の行列という。

- $\forall i \in \{1, \dots, n\}$ に対して、

$$\langle \varphi_i, e_j\rangle = \delta_{ij} \quad (j = 1, \cdots, n)$$

  を満たす $\varphi_i \in X^*$ が一意的に定まる。

- $\varphi_1, \cdots, \varphi_n$ は X^* の基底になる。これを $e_1, \cdots, e_n$ の双対基底と呼ぶ。

- $e_1, \cdots, e_n$ の双対基底を $\varphi_1, \cdots, \varphi_n$, $f_1, \cdots, f_m$ の双対基底を $\psi_1, \cdots, \psi_m$ とする。$\psi_1, \cdots, \psi_m$ と $\varphi_1, \cdots, \varphi_n$ に関する F^* の行列は A^T である:

$$(F^*(\psi_1) \cdots F^*(\psi_m)) = (\varphi_1 \cdots \varphi_n)A^T.$$

  証明: $\forall k \in \{1, \dots, n\}$ に対して

$$\langle F^*(\psi_i), e_k\rangle = \langle \psi_i, F(e_k)\rangle = \left\langle \psi_i, \sum_{j=1}^m a_{jk} f_j \right\rangle = \sum_{j=1}^m a_{jk}\langle \psi_i, f_j\rangle = \sum_{j=1}^m a_{jk}\delta_{ij} = a_{ik},$$

$$\left\langle \sum_{j=1}^n a_{ij}\varphi_j, e_k \right\rangle = \sum_{j=1}^n a_{ij}\left\langle \varphi_j, e_k\right\rangle = \sum_{j=1}^n a_{ij}\delta_{jk} = a_{ik}$$

であるから、$F^*(\psi_i) = \sum_{j=1}^{n} a_{ij}\varphi_j = (\varphi_1 \cdots \varphi_n)\begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix}$ である。■

- えーと、多分次のことが成り立つ。

  (1) F が全射 $\iff$ F^* が単射

  (2) F が単射 $\iff$ F^* が全射

- はてね。

$$\mathrm{Im}\, F^* = (\ker F)^{\perp}, \quad \mathrm{Im}\, F = {}^{\perp}(\ker F^*)$$

  が成り立つのかな。ここで $A \subset X$, $B \subset X^*$ に対して、

$$A^{\perp} := \{\varphi \in X^*; \forall x \in A \ \langle \varphi, x \rangle = 0\}, \quad {}^{\perp}B := \{x \in X; \forall \varphi \in B \ \langle \varphi, x \rangle = 0\}.$$

11.2 「F が全射 $\iff$ F^* が単射」の証明

$V_n = \langle e_1, \ldots, e_n \rangle, W_m = \langle f_1, \ldots, f_m \rangle$. $\langle e_1, \ldots, e_n \rangle, \langle f_1, \ldots, f_m \rangle$ の双対基底をそれぞれ $\langle \varphi_1, \ldots, \varphi_n \rangle$, $\langle \psi_1, \ldots, \psi_m \rangle$ とする。

$F \colon V_n \to V_m$ とする。$\exists A \in M(m, n; \mathbb{K})$ s.t.

$$(F(e_1), \ldots, F(e_n)) = (f_1, \ldots, f_m)\, A.$$

$F^* \colon W_m^* \to V_n^*$ であるが、

$$(F^*(\psi_1), \ldots, F^*(\psi_m)) = (\varphi_1, \ldots, \varphi_n)\, A^T.$$

F が全射であることは、$\mathrm{rank}\, A = m$ と同値であることは明らかであるが、次元定理 $\mathrm{rank}\, A = m - \dim\ker F$ より、F が単射であることは $\mathrm{rank}\, A = n$ と同値であることが分かる。ゆえに F^* が単射であるには、$\mathrm{rank}\, A^T = m$ であることが必要十分である。

$$\begin{aligned} F \text{ が全射} &\iff \mathrm{rank}\, A = m \\ &\iff m \le n \quad \text{and} \quad \exists i_1, \ldots, i_m \quad \text{s.t.} \quad \det(\boldsymbol{a}_{i_1}, \cdots, \boldsymbol{a}_{i_m}) \ne 0. \end{aligned}$$

一方で、

$$\begin{aligned} F \text{ が単射} &\iff \mathrm{rank}\, A^T = m \\ &\iff m \le n \quad \text{and} \quad \exists i_1, \ldots, i_m \quad \text{s.t.} \quad \det \begin{pmatrix} \boldsymbol{a}_{i_1}^* \\ \vdots \\ \boldsymbol{a}_{i_m}^* \end{pmatrix} \ne 0. \end{aligned}$$

これから F が全射 $\iff$ F^* が単射、が分かる。

何か随分迂回しているような気がしなくもないけれど、これで良いのかな？行列式は要らないか。単に $\mathrm{rank}\, A = \mathrm{rank}\, A^T$ だから、で良い。

線形代数において、A^T の位置づけが曖昧だ。まあ大学 1 年生に双対空間、双対写像なんて無理かな。

付録A　文献案内

A.1　一般的な本の紹介

A.1.1　佐武 vs. 齋藤

数学科向けの教科書としては、齋藤 [28], 佐武 [31] が定評がある。前者は取っつきやすく、後者はやや厳めしいという評が多いようであるが…

佐武 [31] は筆者の大学 1 年次の教科書であった。当時は正直なかなか消化できなかったが、後日線形代数を復習する段になって、不思議と読みやすく感じたことを覚えている。著者によると「自習書」を目指して書いたそうで、その辺とも関係があるのかもしれない。数値解析屋となった現在の目でみると、色々な応用について言及されていて、なかなか面白いと思う。「ブルバキ的」という世間一般の評に対する著者の反論 (佐武 [32]) は一読に値すると思う (公理的な扱いを後回しにしたこと、Jordan 標準形の導出に単因子論を使わなかったこと、Lie 群論への応用を意識したことなど、なるほどと思わせる)。

齋藤 [28] は、筆者は参考書として手に取ったのだが、佐武 [31] とはかなり違った書き方で参考になった。佐武 [32] を読むまで気が付かず、驚いたのだが、連立 1 次方程式論を行列式を用いず、基本変形を柱として解説するのが執筆当時としては珍しかったそうである。齋藤先生は執筆当時数値解析学者とも交流があったということを読むと、なるほどそうかと思わされる (そういえば、コンピューターによる線形計算の基礎を説明した名取 [46] に齋藤先生が推薦の言葉を載せている)。Jordan 標準形の導出は単因子論によっているのだが[1]、齋藤 [29] では違った手法を採用していて、齋藤先生としてはそれが大きな改良だと思っているそうである。

A.1.2　Jordan 標準形

筆者にとって、Jordan 標準形はなかなかの難物だった。杉浦 [43] で自習することになった。他にも色々出ているようで (伊理先生とか)、そのうち目を通してみたいと思っている。

A.1.3　自習書

梶原 [18] は、主に大学院の入試問題に現れた線形代数の問題を用いて、線形代数を一通り勉強しようという本である。大学院の入試を控える人には、知識の整理用に良いかもしれない。

A.2　ちょっと雰囲気を変えて

日本では線形代数の講義は、かつては「代数と幾何」というような名前をつけられることも多く、どちらかというと、代数、幾何のバックグラウンドを持つ人が講義を担当することが多いようであ

[1] この点、佐武先生の本よりもある意味では「普通」だったのかもしれないが、筆者が学生の時の先輩の評は「齋藤先生の本は Jordan 標準形の導出に単因子論を使っているのが特徴で…」だった。佐武先生の本の影響の大きさがしのばれる。

る。解析的な取り扱いをした Halmos [10] は目を通しておくべきであろう。なお、杉浦 [43] にも解析的な手法の解説がある。また解析の大家である加藤敏夫先生の加藤 [19] は、有名な Perturbation theory からの抜粋である。

教科書に向いているかどうか、筆者には判断ができないが (個人的には非常に面白く読めて気に入っているが)、しばしば Strang [35] は応用からの視点を大切にしたものとして、良く勧められている。

個人的には伊理 [16] (この本は書き直されて [17] になった, 2011/11/6 メモ) も応用家の気分がかいま見えて面白い。座右の書である。

やや脱線になるが、伊理正夫先生の訳された Chatelin [3] も、読んだ当時の筆者には衝撃的だった。

A.3 線形代数教育

一時期、日本で有名な教科書と言うと、齋藤 [28], 佐武 [31] が双璧であったのだが (これらは — 特に [31] の方は、古典であって、もはや普通の講義では教科書として採用することは不可能である、という意見もあるようだ。最近は川久保 [41] の評判が高いようだが、筆者はちゃんと読んだことがない。)、最近数学セミナーの特集で、著者に当時を振り返ってもらった文章が載っている (齋藤 [30], 佐武 [32])。さすがに著者自らの分析は鋭く参考になるところが多い。

同じ特集の有木 [1] はこれからの線形代数教育を考える上で大変面白い。この文章の中で予告されていた教科書 [2] も既に出版されている。

関連図書

[1] 有木 進 . 工学につながる線形代数を. 数学セミナー, 2000 年 6 月号, pp. 29–33, 2000.

[2] 有木進. 工学のための線形代数. 日本評論社, 2000.

[3] F. Chatelin. *Valeurs propres de matrices*. Masson, Paris, 1988. (邦訳) F. シャトラン著, 伊理正夫, 伊理由美 訳, 行列の固有値問題, シュプリンガー・フェアラーク東京 (1993).

[4] A.K. Cline, C.B. Moler, G.W. Stewart, and J.H. Wilkinson. An estimate for the condition number of a matrix. *SIAM J. Numer. Anal.*, Vol. 16, pp. 368–375, 1979.

[5] G. E. Forsythe and C. B. Moler. *Computer Solution of Linear Algebraic Systems*. Prentice-Hall, 1967.

[6] Leslie V Foster. Gaussian elimination with partial pivoting can fail in practice. *SIAM Journal on Matrix Analysis and Applications*, Vol. 15, No. 4, pp. 1354–1362, 1994.

[7] Leslie V. Foster. The growth factor and efficiency of gaussian elimination with rook pivoting. *Journal of Computational and Applied Mathematics*, Vol. 98, No. 1, pp. 177–194, 5 October 1998.

[8] Herman H. Goldstine and John Von Neumann. Numerical inverting of matrices of high order. ii. *Proceedings of the American Mathematical Society*, Vol. 2, No. 2, pp. 188–202, 1951.

[9] W. W. Hager. Condition estimates. *SIAM J. Stat. Comp.*, Vol. 5, pp. 311–316, 1984.

[10] Paul Richard Halmos. *Finite Dimensional Vector Spaces*. Princeton University Press, 1947.

[11] N. J. Higham. A survey of condition number estimation for triangular matrixes. *SIAM Review*, Vol. 29, pp. 575–596, 1987.

[12] N. J. Higham. Fortran codes for estimating the one-norm of a real or complex matrix, with applictions to condition estimation. *ACM Trans, Math. Soft.*, Vol. 15, pp. 381–396, 1988.

[13] N. J. Higham. Optimization by direct search in matrix computtoins. *SIAM J. Matrix Anal. Appl.*, Vol. 11, pp. 317–333, 1993.

[14] Nicholas J. Higham and Desmond J. Higham. Large growth factors in gaussian elimination with pivoting. *SIAM J. Matrix Analysis and Applications*, Vol. 10, No. 2, pp. 155–164, 1989.

[15] 一松信. 数学とコンピュータ. 共立出版, 1995.

[16] 伊理正夫. 一般線形代数. 岩波書店, 2003. 伊理正夫, 線形代数 I, II, 岩波講座応用数学, 岩波書店 (1993, 1994) の単行本化.

[17] 伊理正夫. 線形代数汎論. 朝倉書店, 2009. 「一般線形代数」のリニューアル.

[18] 梶原壌二. 新修線形代数学. 現代数学社, 1980.

[19] 加藤敏夫. 行列の摂動. シュプリンガー・フェアラーク東京, 1999.

[20] 森正武. 数値解析. 共立出版, 1973. 第 2 版が 2002 年に出版された。

[21] 森正武, 杉原正顯, 室田一雄. 線形計算. 岩波講座 応用数学. 岩波書店, 1994.

[22] M. Natori and A. Tsukamoto. A fast method for estimating the condition number of a matrix. *J. Info. Proc.*, Vol. 6, pp. 138–140, 1983.

[23] 村田健郎, 名取 亮 , 唐木幸比古. 大型数値シミュレーション. 岩波書店, 1990.

[24] 松尾宇泰, 杉原正顯, 森正武. 行列の条件数の推定方法の数値的評価. 日本応用数理学会論文誌, Vol. Vol. 7, No. No.3, pp. 307–319, 1997.

[25] 山本哲朗. 数値解析入門 [新訂版]. サイエンス社, 2003. 1976 年初版発行の定番本の待望の改訂版.

[26] 仁木 滉 , 河野敏行. 楽しい反復法. 共立出版, 1998.

[27] 杉原正顯, 室田一雄. 線形計算の数理. 岩波書店, 2009.

[28] 齋藤正彦. 線型代数入門. 東京大学出版会, 1966.

[29] 齋藤正彦. 線型代数演習. 東京大学出版会, 1985.

[30] 齋藤正彦. 線型代数教育の変遷. 数学セミナー, pp. 12–13, 2000 年 6 月号. これ以外に、数のコスモロジー, ちくま学芸文庫, 筑摩書房 (2007) に収録されている「自著を語る —『線型代数入門』」(これは東京大学出版会の UP に 2000 年 11 月に掲載されたもの) も参照せよ。

[31] 佐武一郎. 線型代数学. 裳華房, 1958, 1974. もともと『行列と行列式』という書名であったのを、テンソル代数の章を加筆した機会に改題した。

[32] 佐武一郎. 『行列と行列式』を書いた頃. 数学セミナー 2000 年 6 月号, pp. 10–11, 2000.

[33] 篠原能材. 数値解析の基礎. 日新出版, 初版 1978, 5 版 1997.

[34] G. D. Smith. *Numerical solution of partial differential equations third edition.* Clarendon Press Oxford, 1986. 第一版の邦訳が G. D. スミス著, 藤川洋一郎訳, コンピュータによる偏微分方程式の解法 新訂版, サイエンス社 (1996) である.

[35] G. S. Strang. *Linear algebra and its applications.* Academic Press, 1976.

[36] Lloyd Nicholas Trefethen and David Bau III. *Numerical Linear Algebra.* SIAM, 1997.

[37] R. S. Varga. *Matrix Iterative Analysis.* Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1962. second ed., Springer (2000). (初版邦訳) 渋谷 政昭他訳, 計算機による大型行列の反復解法, サイエンス社 (1972).

[38] John von Neumann and H. H. Goldstine. Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc.*, Vol. 53, No. 11, pp. 1021–1099, 1947. http://projecteuclid.org/euclid.bams/1183511222.

[39] Stephen J. Wright. A collection of problems for which gaussian elimination with partial pivoting is unstable. *SIAM J. Sci. Statist. Comput.*, Vol. 14, pp. 231–238, 1993.

[40] 中尾充宏, 渡部善隆. 実例で学ぶ 精度保証付き数値計算 — 理論と実装. SGC ライブラリ 85. サイエンス社, 2011. 新品では入手できない。僕の学生で読みたい人は相談に乗ります。

[41] 川久保勝夫. 線形代数学. 日本評論社, 1999.

[42] 礒野譲. 連立一次方程式の解の精度保証付き数値計算. http://nalab.mind.meiji.ac.jp/~mk/labo/report/pdf/isono-master.pdf, 2003.

[43] 杉浦光夫, 横沼健雄. Jordan 標準形・テンソル代数. 岩波書店, 1990. これは、杉浦 光夫, Jordan 標準形と単因子論 I, II, 岩波講座 基礎数学 (1976,1977) が元になっている。

[44] 山本哲朗. 行列解析の基礎 — Advanced 線形代数. サイエンス社, 2010.

[45] 山本哲朗, 北川高嗣(たかし). 数値解析演習. サイエンス社, 1991.

[46] 名取亮. 線形計算. 朝倉書店, 1993.