SGC ライブラリ- 79

行列解析の基礎

Advanced 線形代数

山本 哲朗　著

サイエンス社

まえがき

本書は線形代数の初歩を一通り学んだことがあるか，もしくは現在学びつつある人達のための副読本として，行列に関する既修の知識の再確認と通常の教科書・参考書に記載されていない興味ある話題の提供を目的としている．それらは，例えば

1. $\Phi_p(A) = \left(\sum_{i,j} |a_{ij}|^p\right)^{1/p}$ $\left(A = (a_{ij})\right)$ が行列ノルムとなるために正数 p がみたすべき条件 (定理 4.2, 4.3)
2. 半正定値 Hermite 行列に対する p 乗根の一意存在 (定理 5.7) とその計算法 (§5.4)
3. 正規行列と Hermite 行列の固有値に関する Hoffman-Wielandt の摂動定理 (定理 6.13, 6.14)
4. 非負行列の固有値・固有ベクトルに関するいわゆる Perron-Frobenius の理論 (定理 7.4–7.8)
5. P_0 行列を係数行列とする準線形方程式 $A\boldsymbol{x}+F(\boldsymbol{x}) = \boldsymbol{b}$ の解の一意存在に関する Sandberg-Willson, Jr. の定理 (定理 9.5)
6. 行列方程式 $AX + XB = C$ が任意の C に対して一意解をもつための必要十分条件 (定理 10.8) と解の積分表示 (定理 10.9)

等々，列記すれば枚挙にいとまがない．おそらく読者は本書のあちこちで類書にない話題 (概念，結果，証明等) に遭遇されることであろう．著者はひそかにこれを期待している次第である．

なお，Jordan 分解の導出には西岡久美子氏の方法[33] を用い，Hermite 行列の固有値に関するミニマックス定理，分離定理，摂動定理等の導出には，次元等式を用いる池辺八洲彦氏の鮮やかな技法[26] を参考にした．

本書が読者に行列の魅力の一端をいささかでも伝え得れば，著者の喜びこれに優るものはない．

最後に，本書原稿のタイプはプラスアイ代表 富岡乃美氏にお願いした．また出版に際し，サイエンス社編集部長田島伸彦氏，「数理科学」編集部平勢耕介氏とビーカム社佐藤亨氏の御世話になった．ここに記して各位に深甚なる謝意を表する．

2010 年 11 月

山本 哲朗

目　次

第 1 章　線形代数の基礎 (1) **1**

1.1　行列 1
1.2　行列の演算 2
1.3　行列式 4
1.4　終結式 10
1.5　線形空間 11
1.6　内積空間 17
1.7　ノルム空間 21
1.8　値域，零空間，ランク 25
1.9　アセントとディセント 30
1.10　ブロック行列と Schur 補元 31
1.11　既約行列 33
1.12　優対角行列 36
1.13　2 重確率行列と Birkhoff の定理 37

第 2 章　線形代数の基礎 (2) **39**

2.1　固有値と固有ベクトル 39
2.2　行列の 3 角化と対角化 41
2.3　Hamilton-Cayley の定理と最小多項式 45
2.4　3 重対角行列 49
2.5　特異値 50
2.6　Hermite 行列 51
2.7　正定値行列と半正定値行列 52
2.8　2 次形式 55

第 3 章　行列の分解 **59**

3.1　LU 分解 59
3.2　LL^t 分解 63
3.3　QR 分解 65
3.4　特異値分解 66
3.5　極分解 68
3.6　Jordan 分解 69

第 4 章　行列ノルムと行列関数 **77**
4.1 行列ノルム 77
4.2 ベクトルノルムに従属する行列ノルム 78
4.3 Frobenius ノルムの拡張 80
4.4 行列ノルムと固有値の関係 85
4.5 行列関数 87
4.6 行列の微分・積分 92
4.7 行列式の微分 93

第 5 章　AB と BA の固有値 **94**
5.1 AB と BA の固有値 94
5.2 可換な行列 96
5.3 半正定値 Hermite 行列の p 乗根 99
5.4 p 乗根を求めるアルゴリズム 102
5.5 Kantorovich の不等式と Rheinboldt の不等式 104

第 6 章　固有値の評価定理 **107**
6.1 Gerschgorin の定理とその周辺 107
6.2 Hermite 形式と基本不等式 110
6.3 ミニマックス定理 113
6.4 分離定理 114
6.5 摂動定理 116
6.6 Hoffman-Wielandt の定理 118
6.7 Schur 補元の固有値 121

第 7 章　非負行列 **123**
7.1 非負行列 123
7.2 正行列に対する Perron-Frobenius の定理 126
7.3 非負行列に対する Frobenius の定理 128
7.4 既約非負行列に対する Frobenius の定理 129
7.5 巡回指数 $k > 1$ の既約非負行列に対する Frobenius の定理 132

第 8 章　Z 行列と M 行列 **136**
8.1 Z 行列 136
8.2 M 行列 137
8.3 M 行列であるための条件 141
8.4 L 行列 145
8.5 非負行列に対する分離定理 145

第 9 章　P 行列と P_0 行列 **148**
9.1　P 行列 148
9.2　P_0 行列 151
9.3　Sandberg-Willson, Jr. の定理 156
9.4　Ω 行列 161

第 10 章　テンソル積 **163**
10.1　テンソル積 163
10.2　行列方程式 $X - AXB = C$ 166
10.3　行列方程式 $AX + XB = C$ 168
10.4　差分近似行列 171

第 11 章　Grassmann 積と複合行列 **174**
11.1　Grassmann 積 174
11.2　複合行列 178
11.3　複合行列の応用 182

第 12 章　線形反復法 **185**
12.1　Jacobi 法と緩和 Jacobi 法 185
12.2　Gauss-Seidel 法 187
12.3　SOR 法 (Gauss-Seidel 法の緩和) 187
12.4　整合順序行列と Young の定理 188
12.5　最急降下法 192
12.6　共役勾配法 (最急降下法の加速) 195
12.7　線形反復解の誤差分布 201

参考文献 **203**
主要 定理 (系) 一覧 **205**
索　　引 **207**

第1章　線形代数の基礎(1)

この章では，行列，行列式，線形空間，基底，ランク等，通常の線形代数学講義前半で学ぶ基礎事項を復習すると共に，既約強優対角行列，Birkhoff の定理等，後で必要となる事項についても述べる．

1.1　行列

本書では，$\mathbb{F}$ を実数の集合 $\mathbb{R}$ または複素数の集合 $\mathbb{C}$ のどちらかをあらわすものとする．mn 個の実数または複素数 a_{ij} $(\in \mathbb{F})$ の配列

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})$$

を **$m \times n$ 行列** ($m = n$ のときは n 次正方行列または n 次行列) といい，$A \in \mathbb{F}^{m\times n}$ と記す．(したがって $a_{ij} \in \mathbb{R}$ を強調するときは $A \in \mathbb{R}^{m\times n}, a_{ij} \in \mathbb{C}$ を強調するときは $A \in \mathbb{C}^{m\times n}$ と記し，それぞれ $m \times n$ 実行列，$m \times n$ 複素行列という．) a_{ij} を A の (i,j) 要素という．すべての要素がゼロの行列 A を**零行列** (ゼロ行列) といい $A = O$ であらわす．

また A の横方向を**行**，縦方向を**列**といい，

$$(a_{i1}, a_{i2}, \ldots, a_{in}) \quad \text{を } A \text{ の第 } i \text{ 行}, \qquad \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad \text{を } A \text{ の第 } j \text{ 列}$$

という．特に $1 \times n$ 行列を **n 次元行ベクトル**，$n \times 1$ 行列を **n 次元列ベクトル**という．

行列 $A = (a_{ij}) \in \mathbb{F}^{m\times n}$ に対し

$$A^{\mathrm{t}} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ & & \cdots & \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix} \quad (n \times m \text{ 行列})$$

$$A^{*} = \begin{pmatrix} \overline{a}_{11} & \overline{a}_{21} & \cdots & \overline{a}_{m1} \\ \overline{a}_{12} & \overline{a}_{22} & \cdots & \overline{a}_{m2} \\ & & \cdots & \\ \overline{a}_{1n} & \overline{a}_{2n} & \cdots & \overline{a}_{mn} \end{pmatrix} \quad (\overline{a}_{ij} \text{は } a_{ij} \text{の共役})$$

とおき，それぞれ A の**転置行列** (**transpose**)，**共役転置行列** (**conjugate transpose**) という．(A がベクトルのときは**転置ベクトル**，**共役転置ベクトル**という．) したがって A が実行列ならば $A^{*} = A^{\mathrm{t}}$ である．また $A^{\mathrm{t}} = A$ をみたす正方行列 A を**対称行列** (**symmetric matrix**)，$A^{*} = A$ をみたす正方行列 A を **Hermite** (**エルミート**) **行列** (**Hermitian matrix** または **hermitian matrix**) という．前者は実行列に，後者は複素行列に用いられるので，それを強調してそれぞれ実対称行列，複素 Hermite 行列とかくことも多い．さらに $a_{ji} = -a_{ij}$ $\forall i, j$ をみたす実正方行列 A を**歪対称** (**skew-symmetric**)

行列または**交代行列**，$\overline{a}_{ji} = -a_{ij}\ \forall i,j$ をみたす複素正方行列を**歪 Hermite (skew-Hermitian** または **skew-hermitian**) 行列という．

n 次正方行列 $A = (a_{ij}) \in \mathbb{F}^{n\times n}$ が $a_{ij} = 0\ (i \neq j)$ をみたすとき A は**対角行列**であるといい $A = \mathrm{diag}(a_{11}, \ldots, a_{nn})$ であらわす．

特に $a_{ii} = 1\quad \forall i$ のとき A を $\boldsymbol{n}$ **次単位行列**といい I または I_n であらわす．

なお，a_{ii} を**対角要素**，$a_{ij}\ (i \neq j)$ を**非対角要素**という．

定義から明らかに歪対称行列の対角要素はすべて 0 $(a_{ii} = 0\ \forall i)$ で，歪 Hermite 行列の対角要素はすべて純虚数または 0 である．

1.2 行列の演算

行列 $A = (a_{ij})$，$B = (b_{ij}) \in \mathbb{F}^{m\times n}$ とスカラー $\alpha \in \mathbb{F}$ に対し，加算 $A + B$ とスカラー積 αA を

$$A + B = (a_{ij} + b_{ij}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ & & \cdots & \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

$$\alpha A = (\alpha a_{ij}) = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ & & \cdots & \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{pmatrix}$$

により定義する．

$\alpha = -1$ のとき αA を $-A$ とかくことにすれば，$A - B = A + (-B)$ により減算も定義される．また前節で述べた歪対称行列，歪 Hermite 行列はそれぞれ $A^{\mathrm{t}} = -A$，$A^* = -A$ をみたす行列 A にほかならない．

行列の積 AB は A が $m \times p$ 行列，B が $p \times n$ 行列のときに限って

$$AB = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ & \cdots & \\ c_{m1} & \cdots & c_{mn} \end{pmatrix} = (c_{ij}), \quad c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

により定義される ($AB \in \mathbb{F}^{m\times n}$ である)．

このとき容易に確かめられるように $(AB)^{\mathrm{t}} = B^{\mathrm{t}} A^{\mathrm{t}}$，$(AB)^* = B^* A^*$ である．なお，$m \neq n$ ならば BA は定義されない．$m = p = n$ ならば AB と BA は定義されるが一般には $AB \neq BA$ である．

n 次正方行列 A, B が $AB = BA$ をみたすとき A と B は**可換**である，または**交換可能**であるという．

よく用いられる行列の名称には次のようなものがある．以下 A は正方行列とする．

(i) **直交行列 (orthogonal matrix)**：A は実行列で

$$AA^{\mathrm{t}} = A^{\mathrm{t}} A = I \quad (\text{単位行列})$$

をみたすとき A を**直交行列**という．ただし上記条件は実は過剰であって，

$$AA^{\mathrm{t}} = I \quad \text{または} \quad A^{\mathrm{t}} A = I$$

の一方だけでよい．(後でわかるように $AA^{\mathrm{t}} = I \Leftrightarrow A^{\mathrm{t}} A = I$ が成り立つ．)

(ii) **ユニタリ行列 (unitary matrix)**：A は複素行列で

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

をみたすとき A を**ユニタリ行列**という．ただしこの条件も実は過剰であって，

$$AA^* = I \text{ または } A^*A = I$$

の一方だけでよい．($AA^* = I \Leftrightarrow A^*A = I$ が成り立つことが後でわかる．)

(iii) **逆行列 (inverse)**：$AX = XA = I$ をみたす正方行列が存在するとき X を A の**逆行列**という．このような X は存在すればただ1つである．実際 Y を A の逆行列とすれば $AY = YA = I$ より $AY = I$ の両辺に左から X をかければ $XAY = X$．左辺は $(XA)Y = IY = Y$ であるから $Y = X$ である．よって $X = A^{-1}$ とかく．

ただし，この条件も実は過剰であって $AX = I$ または $XA = I$ の一方だけでよいことが後にわかる．(すなわち，$XA = I \Leftrightarrow AX = I$ が成り立つ)

(iv) **正則行列 (regular matrix または nonsingular matrix)**：A の逆行列が存在するとき A を**正則行列**という．この定義によれば，A が正則行列のとき逆行列も正則行列である．

実際 X を A の逆行列とすれば (iii) によって $XA = AX = I$，これは A が X の逆行列であることを意味するから X は正則となる．また A, B が同じ次数の正則行列ならば AB も正則行列で $(AB)^{-1} = B^{-1}A^{-1}$ である．

(v) **正規行列 (normal matrix)**：$AA^* = A^*A$ のとき A を**正規行列**という．

(vi) **置換行列 (または順列行列)(permutation matrix)**：実正方行列で各行各列に1がちょうど1つだけありその他はゼロの行列を**置換行列**という．この行列についてはさらに説明を加える．

n 個の文字 $1, 2, \ldots, n$ に対し 1 を $\sigma(1)$，2 を $\sigma(2), \ldots, n$ を $\sigma(n)$ に並べかえる写像 $\sigma : i \to \sigma(i),\quad i = 1, 2, \ldots, n$ を**置換 (順列)** といい

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \tag{1.1}$$

とあらわす．$i_1, i_2, \ldots, i_n$ が $1, 2, \ldots, n$ の並べかえ (順列) ならば

$$\sigma = \begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ \sigma(i_1) & \sigma(i_2) & \cdots & \sigma(i_n) \end{pmatrix}$$

とかいてもよい．また置換

$$\begin{pmatrix} \sigma(1) & \sigma(2) & \cdots & \sigma(n) \\ 1 & 2 & \cdots & n \end{pmatrix}$$

を σ の**逆置換**といい σ^{-1} であらわす．このとき σ に対応する置換行列 $P_\sigma = (p_{ij})$ が

$$p_{ij} = \begin{cases} 1 & (j = \sigma(i)) \\ 0 & (j \neq \sigma(i)) \end{cases}, \quad i = 1, 2, \ldots, n$$

により定義される．すなわち P_σ は各行各列に1がちょうど1つだけあり，その他はすべてゼロの行列である．逆にこのような行列に対しては1つの置換 σ が対応している．

明らかに $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$ (n 次元列ベクトル) に対して

$$P_\sigma \boldsymbol{x} = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})^{\mathrm{t}} \tag{1.2}$$

であり P_σ は $\boldsymbol{x}$ の成分 $x_1, \ldots, x_n$ を $x_{\sigma(1)}, \ldots, x_{\sigma(n)}$ に並べかえる．

また行ベクトル $\boldsymbol{y} = (y_1, \ldots, y_n)$ に対しては上の結果により

$$\begin{aligned} \boldsymbol{y} P_\sigma^{\mathrm{t}} &= (P_\sigma \boldsymbol{y}^{\mathrm{t}})^{\mathrm{t}} \\ &= (y_{\sigma(1)}, \ldots, y_{\sigma(n)}) \end{aligned} \tag{1.3}$$

であり，P_σ^{t} を右から行ベクトル $\boldsymbol{y}$ に作用させれば $\boldsymbol{y}$ の成分 $y_1, \ldots, y_n$ を $y_{\sigma(1)}, \ldots, y_{\sigma(n)}$ に並べ

かえる．

$P_\sigma^{\mathrm{t}} = (\widetilde{p}_{ij})$ とすれば

$$\widetilde{p}_{ij} = p_{ji} = \begin{cases} 1 & \left(i = \sigma(j) \text{ すなわち } j = \sigma^{-1}(i) \text{ のとき}\right) \\ 0 & \left(i \neq \sigma(j) \text{ すなわち } j \neq \sigma^{-1}(i) \text{ のとき}\right) \end{cases}$$

であるから P_σ^{t} は σ の逆置換 σ^{-1} に対応する置換行列である．すなわち $P_\sigma^{\mathrm{t}} = P_{\sigma^{-1}}$．さらに $P_\sigma P_{\sigma^{-1}} = P_{\sigma^{-1}} P_\sigma = I$ であるから $P_{\sigma^{-1}}$ は P_σ の逆行列であり $P_\sigma^{\mathrm{t}} = P_\sigma^{-1}$ である．したがって

$$P_\sigma P_\sigma^{\mathrm{t}} = P_\sigma^{\mathrm{t}} P_\sigma = I$$

が成り立つ．

■**注意**　上式を次のように考えてもよい．(1.2) と (1.3) より

$$\boldsymbol{y} P_\sigma^{\mathrm{t}} P_\sigma \boldsymbol{x} = (y_{\sigma(1)}, \ldots, y_{\sigma(n)})(x_{\sigma(1)}, \ldots, x_{\sigma(n)})^{\mathrm{t}} = \sum_{i=1}^{n} y_{\sigma(i)} x_{\sigma(i)} = \sum_{i=1}^{n} y_i x_i = \boldsymbol{y}\boldsymbol{x}$$

故に $A = P_\sigma^{\mathrm{t}} P_\sigma - I = (a_{ij})$ とおけば $\boldsymbol{y} A \boldsymbol{x} = \boldsymbol{0}$

ここで $\boldsymbol{e}_i$ を n 次単位行列の第 i 列とし，$\boldsymbol{x} = \boldsymbol{e}_j$，$\boldsymbol{y} = \boldsymbol{e}_i^{\mathrm{t}}$ とおけば $\boldsymbol{y} A \boldsymbol{x} = a_{ij}$ となるから $a_{ij} = 0 \quad \forall i, j \quad \therefore\ A = O$

$$\therefore\quad P_\sigma^{\mathrm{t}} P_\sigma = I$$

故に (iii) の終わりに記した注意によって $P_\sigma^{\mathrm{t}} = P_\sigma^{-1}$ が従う．$\boldsymbol{e}_i$ を**第 i 基本ベクトル**という．

1.3　行列式

置換

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

と

$$\tau = \begin{pmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{pmatrix} = \begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}$$

が与えられたとき σ と τ の積として新しい置換

$$\begin{aligned} \sigma\tau &= \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \begin{pmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} \begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix} \quad (\text{ただし} \tau(i_l) = k_l \text{とする}) \\ &= \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix} \end{aligned}$$

が定義される．この積は行列の場合と同様に一般に交換可能ではなく $\sigma\tau \neq \tau\sigma$ である．また $(\sigma\tau)^{-1} = \tau^{-1}\sigma^{-1}$ も成り立つ．特にすべての文字を動かさない置換

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

を**恒等置換**または**単位置換**といい 1 であらわすと $1\sigma = \sigma 1 = \sigma$ である．

さらに 2 つの文字 i と j を入れかえる置換

$$\begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & j & \cdots & i & \cdots & n \end{pmatrix}$$

を**互換**といい (i,j) であらわす．すると任意の置換は互換の積としてあらわされる．そのあらわし方はいろいろあるが，積の個数は表現の仕方に関係なく偶数個であるか奇数個であるかのどちらかであることが知られている．偶数個の互換の積としてあらわされる置換を**偶置換**，奇数個の互換の積としてあらわされる置換を**奇置換**という．恒等置換は偶数ゼロ個の置換の積とみなして偶置換と考えれば置換の総数 $n!$ のうち偶置換と奇置換の数はどちらも $\frac{1}{2}n!$ である．(ただし $n \geq 2$ とする．)

置換

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

が与えられたとき，それが偶置換であるか奇置換であるかを判定するのは簡単ではないが，1 つの簡便法として次の方法がある．まず i_1 を固定し $i_1 > i_j$ $(j \geq 2)$ なる j の個数を $\sharp i_1$ とする．次に i_2 を固定し $i_2 > i_j$ $(j \geq 3)$ なる j の個数を $\sharp i_2$ とする．以下これを続けて最後に $i_{n-1} > i_j$ $(j \geq n)$ なる j の個数を $\sharp i_{n-1}$ とする．(すなわち $i_{n-1} > i_n$ ならば $\sharp i_{n-1} = 1$, $i_{n-1} < i_n$ ならば $\sharp i_{n-1} = 0$ とする．) このとき $\sharp i_1 + \sharp i_2 + \cdots + \sharp i_{n-1}$ が偶数ならば偶置換，奇数ならば奇置換と判定するのである．

さて，置換 σ に対して σ の符号を

$$\mathrm{sgn}(\sigma) = \begin{cases} 1 & (\sigma : \text{偶置換}) \\ -1 & (\sigma : \text{奇置換}) \end{cases}$$

と定める．このとき n 次正方行列 $A = (a_{ij})$ の行列式を

$$\sum_{\sigma = \binom{1\ 2\ \cdots\ n}{j_1 j_2 \cdots\ j_n}} \mathrm{sgn}(\sigma) a_{1 j_1} a_{2 j_2} \cdots a_{n j_n}$$

により定義し，$\det(A)$, $\det A$, $|A|$, または

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

であらわす．ただし第 7 章で (i,j) 要素 a_{ij} を $|a_{ij}|$ でおきかえた行列をあらわすのに $|A|$ を用いるので，本書は行列式の記号としては $|A|$ を用いない．

行列 A は数の集合であるが行列式は実数または複素数をあらわす．また a_{ij} を変数とみなすときは，行列式は n^2 個の変数 a_{ij} に関する n 次多項式にほかならない．

例 1　$n = 2$ のときすべての置換は

$$\sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad (\text{恒等置換}), \qquad \sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = (1,2) \quad (\text{互換})$$

であり，σ_1 は偶置換，σ_2 は奇置換である．よって

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \mathrm{sgn}(\sigma_1) a_{11} a_{22} + \mathrm{sgn}(\sigma_2) a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21}$$

例 2　$n = 3$ のときすべての置換は

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$
$$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \qquad \sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

このうち $\sigma_1, \sigma_4, \sigma_5$ は偶置換，$\sigma_2, \sigma_3, \sigma_6$ は奇置換である．例えば σ_3 については

$$\sharp 2 = 1, \quad \sharp 1 = 0$$

であるから $\sharp 2 + \sharp 1 = 1$ (奇数) となって σ_3 は奇置換である．その他は各自検証されたい．
故に

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \operatorname{sgn}(\sigma_1) a_{11}a_{22}a_{33} + \operatorname{sgn}(\sigma_2) a_{11}a_{23}a_{32} + \operatorname{sgn}(\sigma_3) a_{12}a_{21}a_{33} + \operatorname{sgn}(\sigma_4) a_{12}a_{23}a_{31} + \operatorname{sgn}(\sigma_5) a_{13}a_{21}a_{32} + \operatorname{sgn}(\sigma_6) a_{13}a_{22}a_{31}$$
$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

定理 1.1 行列式は次の基本的性質をもつ．

(i) $\det A^{\mathrm{t}} = \det A$

(ii) A の第 i 行の各成分を c 倍して得られる行列の行列式は $c \det A$ に等しい．

(iii) $$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{i1} + b_{i1} & a_{i2} + b_{i2} & \cdots & a_{in} + b_{in} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ a_{i1} & \cdots & a_{in} \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ b_{i1} & \cdots & b_{in} \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

(iv) A の 1 つの行を何倍かして他の行に加えても行列式の値は変わらない．

(v) A の 2 つの行を入れかえれば行列式の符号が変わる．

(vi) したがって A の 2 つの行が等しければ $\det A = 0$

(vii) (i) により (ii)～(vi) の性質は列についても成り立つ．

(viii) 単位行列の行列式は 1 である．

(ix) A, B を同じ次数の正方行列とすれば $\det(AB) = \det A \det B$.

【証明】 (i)～(viii) は行列式の定義より容易に導かれる．(ix) は後出の定理 1.4 に含まれる．

証明終 ∎

n 次行列 $A = (a_{ij})$ に対し A の第 $i_1, \ldots, i_r$ 行と第 $j_1, \ldots, j_r$ 列の交差する要素を取り出してつくる行列を $A \begin{pmatrix} i_1 & \cdots & i_r \\ j_1 & \cdots & j_r \end{pmatrix}$ であらわす．すなわち

$$A \begin{pmatrix} i_1 & \cdots & i_r \\ j_1 & \cdots & j_r \end{pmatrix} = \begin{pmatrix} a_{i_1\, j_1} & \cdots & a_{i_1\, j_r} \\ & \cdots & \\ a_{i_r\, j_1} & \cdots & a_{i_r\, j_r} \end{pmatrix}$$

特に $A \begin{pmatrix} i_1 & \cdots & i_r \\ i_1 & \cdots & i_r \end{pmatrix}$ を A の ***r* 次主小行列 (principal submatrix)**，$A \begin{pmatrix} 1 & 2 & \cdots & r \\ 1 & 2 & \cdots & r \end{pmatrix}$ を ***r* 次首座小行列 (leading principal submatrix)** という．したがって

$$\det A\begin{pmatrix} i_1 & \cdots & i_r \\ i_1 & \cdots & i_r \end{pmatrix} \quad と \quad \det A\begin{pmatrix} 1 & 2 & \cdots & r \\ 1 & 2 & \cdots & r \end{pmatrix}$$

をそれぞれ **r 次主小行列式 (principal minor)**，**r 次首座小行列式 (leading principal minor)** という．また A の第 i 行と第 j 列を取り除いてえられる $n-1$ 次行列の行列式に $(-1)^{i+j}$ をかけたものを A の (i,j) **余因子 (cofactor)** または a_{ij} の余因子といい $\triangle_{ij}$ であらわす．

定理 1.2 (余因子展開) n 次行列 $A=(a_{ij})$ に対し次が成り立つ．

(i) $\det A = a_{i1}\triangle_{i1} + a_{i2}\triangle_{i2} + \cdots + a_{in}\triangle_{in}$ (第 i 行に関する展開)

(ii) $\det A = a_{1j}\triangle_{1j} + a_{2j}\triangle_{2j} + \cdots + a_{nj}\triangle_{nj}$ (第 j 列に関する展開)

【証明】 行列式の定義より導かれるが詳細は省略する． 証明終 ∎

定理 1.3 前定理の記号の下で

(i) $$\sum_{k=1}^{n} a_{ik}\triangle_{jk} = \sum_{k=1}^{n} a_{kj}\triangle_{ki} = \begin{cases} \det A & (i=j \text{ のとき}) \\ 0 & (i\neq j \text{ のとき}) \end{cases}$$

(ii) $\det A \neq 0$ のとき A は正則で

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} \triangle_{11} & \triangle_{21} & \cdots & \triangle_{n1} \\ \vdots & \vdots & & \vdots \\ \triangle_{1n} & \triangle_{2n} & \cdots & \triangle_{nn} \end{pmatrix} = (\alpha_{ij}), \quad \alpha_{ij} = \frac{\triangle_{ji}}{\det A}$$

【証明】 (i) $i=j$ のときは定理 1.2 にほかならない．$i\neq j$ のときは A の第 j 行を第 i 行でおきかえた行列 $\widetilde{A}$ を j 行につき余因子展開すれば $\widetilde{A}$ の (j,k) 余因子は $\triangle_{jk}$ であるから $\sum_{k=1}^{n} a_{ik}\triangle_{jk} = \det\widetilde{A} = 0$. 他も同様である．

(ii) は (i) から従う． 証明終 ∎

系 1.3.1 n 次行列 A に対し次の 2 条件は同値である．

(i) A が正則である． (ii) $\det A \neq 0$

【証明】 (i) ⇒ (ii) $AA^{-1}=I$ より両辺の行列式をとれば $\det A \det A^{-1} = 1$

$$\therefore \quad \det A \neq 0$$

(ii) ⇒ (i) 定理 1.3 (ii) による． 証明終 ∎

$\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}}$, $\boldsymbol{b}=(b_1,\ldots,b_n)^{\mathrm{t}}$ とする．$x_1,\ldots,x_n$ に関する n 元連立 1 次方程式 $A\boldsymbol{x}=\boldsymbol{b}$ の解は，定理 1.3 (ii) によって，$\det A\neq 0$ のとき

$$\begin{aligned} \boldsymbol{x} &= A^{-1}\boldsymbol{b} \\ &= \frac{1}{\det A}\begin{pmatrix} \triangle_{11} & \triangle_{21} & \cdots & \triangle_{n1} \\ \vdots & \vdots & & \vdots \\ \triangle_{1n} & \triangle_{2n} & \cdots & \triangle_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{\det A}\begin{pmatrix} \sum_{k=1}^{n} b_k\triangle_{k1} \\ \vdots \\ \sum_{k=1}^{n} b_k\triangle_{kn} \end{pmatrix} \end{aligned}$$

$$\therefore \quad x_i = \frac{1}{\det A}\sum_{k=1}^{n} b_k\triangle_{ki}, \quad i=1,2,\ldots,n$$

上式右辺の分子は n 次行列式

$$\begin{vmatrix} a_{11} & \cdots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \cdots & a_{1n} \\ & \cdots & & \cdots & & \cdots & \\ a_{n1} & \cdots & a_{n\,i-1} & b_n & a_{n\,i+1} & \cdots & a_{nn} \end{vmatrix}$$

を第 i 列に関して展開したものであるから

$$x_i = \frac{1}{\det A} \begin{vmatrix} a_{11} & \cdots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \cdots & a_{1n} \\ & \cdots & & \cdots & & \cdots & \\ a_{n1} & \cdots & a_{n\,i-1} & b_n & a_{n\,i+1} & \cdots & a_{nn} \end{vmatrix}, \quad i = 1, 2, \ldots, n$$

これを **Cramer (クラメルまたはクラーメル) の公式**という．昔，著者の学生時代にはクラーメルと呼ばれていたがその後なぜかクラメルまたはクラメールと呼ばれるようになり，最近では“クラーメル”が復活してきたようである．残念ながら，正しい呼び方は本人しかわからない．

以上，行列式の基礎を復習した．ここで定理 1.1 (ix) の拡張について触れておく．$A = (a_{ij}) \in \mathbb{F}^{m\times n}$, $B = (b_{ij}) \in \mathbb{F}^{n\times m}$ ならば AB は m 次正方行列となって $\det(AB)$ は定義されるが $m \neq n$ のとき $\det A$, $\det B$ は定義されない．

このような場合の定理 1.1 (ix) に相当する定理として Binet-Cauchy(ビネ・コーシー) の定理が知られている．まず $m > n$ のときは

$$AB = \begin{pmatrix} a_{11} & \cdots & a_{1n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} & \underbrace{0 \quad \cdots \quad 0}_{m-n} & & \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \Bigg\}{\scriptstyle m-n} = \widetilde{A}\widetilde{B}$$

と AB を m 次正方行列 $\widetilde{A}$, $\widetilde{B}$ の積にあらわして定理 1.1 (ix) を使えば

$$\det(AB) = \det(\widetilde{A}\widetilde{B}) = \det\widetilde{A}\det\widetilde{B} = 0$$

よって $m \leq n$ のときを考える．

定理 1.4 (Binet-Cauchy) $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times m}$ かつ $m \leq n$ ならば

$$\det(AB) = \sum_{1\leq k_1<\cdots<k_m\leq n} \det A \begin{pmatrix} 1 & 2 & \cdots & m \\ k_1 & k_2 & \cdots & k_m \end{pmatrix} \det B \begin{pmatrix} k_1 & k_2 & \cdots & k_m \\ 1 & 2 & \cdots & m \end{pmatrix} \tag{1.4}$$

右辺の和は $1, 2, \ldots, n$ から相異なる m 個を取り出す組み合わせの数 $\binom{n}{m} = \dfrac{n!}{m!(n-m)!}$ だけある．

【証明】

$$\det(AB) = \begin{vmatrix} \sum_{\alpha_1=1}^n a_{1\alpha_1}b_{\alpha_1 1} & \cdots & \sum_{\alpha_m=1}^n a_{1\alpha_m}b_{\alpha_m m} \\ \vdots & & \vdots \\ \sum_{\alpha_1=1}^n a_{m\alpha_1}b_{\alpha_1 1} & \cdots & \sum_{\alpha_m=1}^n a_{m\alpha_m}b_{\alpha_m m} \end{vmatrix}$$

$$= \sum_{\alpha_1,\ldots,\alpha_m=1}^{n} \begin{vmatrix} a_{1\alpha_1}b_{\alpha_1 1} & \cdots & a_{1\alpha_m}b_{\alpha_m m} \\ \vdots & & \vdots \\ a_{m\alpha_1}b_{\alpha_1 1} & \cdots & a_{m\alpha_m}b_{\alpha_m m} \end{vmatrix}$$

$$= \sum_{\alpha_1,\ldots,\alpha_m=1}^{n} \begin{vmatrix} a_{1\alpha_1} & \cdots & a_{1\alpha_m} \\ \vdots & & \vdots \\ a_{m\alpha_1} & \cdots & a_{m\alpha_m} \end{vmatrix} b_{\alpha_1 1}\cdots b_{\alpha_m m} \tag{1.5}$$

行列式の性質定理 1.1 (vi) によって，ある i, j $(i \neq j)$ につき $\alpha_i = \alpha_j$ ならば

$$\det A \begin{pmatrix} 1 & 2 & \cdots & m \\ \alpha_1 & \alpha_2 & \cdots & \alpha_m \end{pmatrix} = \begin{vmatrix} a_{1\alpha_1} & \cdots & a_{1\alpha_m} \\ \vdots & & \vdots \\ a_{m\alpha_1} & \cdots & a_{m\alpha_m} \end{vmatrix} = 0$$

であるから (1.5) において $\alpha_1, \ldots, \alpha_m$ は相異なるもののみ意味をもつ．そのような 1 つの組み合わせ $\{\alpha_1, \ldots, \alpha_m\}$ を $1 \leq k_1 < \cdots < k_m \leq n$ とし，これに対応する項を集めれば $(j_1, \ldots, j_m)$ を $k_1, \ldots, k_m$ の順列として (1.5) の対応する和は

$$\sum_{(j_1,\ldots,j_m)} \begin{vmatrix} a_{1k_1} & \cdots & a_{1k_m} \\ \vdots & & \vdots \\ a_{mk_1} & \cdots & a_{mk_m} \end{vmatrix} \operatorname{sgn} \begin{pmatrix} k_1 & \cdots & k_m \\ j_1 & \cdots & j_m \end{pmatrix} b_{j_1 1}b_{j_2 2}\cdots b_{j_m m}$$

$$= \begin{vmatrix} a_{1k_1} & \cdots & a_{1k_m} \\ \vdots & & \vdots \\ a_{mk_1} & \cdots & a_{mk_m} \end{vmatrix} \cdot \sum_{(j_1,\ldots,j_m)} \operatorname{sgn} \begin{pmatrix} k_1 & \cdots & k_m \\ j_1 & \cdots & j_m \end{pmatrix} b_{j_1 1}b_{j_2 2}\cdots b_{j_m m}$$

$$= \begin{vmatrix} a_{1k_1} & \cdots & a_{1k_m} \\ \vdots & & \vdots \\ a_{mk_1} & \cdots & a_{mk_m} \end{vmatrix} \cdot \begin{vmatrix} b_{k_1 1} & \cdots & b_{k_1 m} \\ \vdots & & \vdots \\ b_{k_m 1} & \cdots & b_{k_m m} \end{vmatrix}$$

よって (1.5) は

$$\sum_{1\leq k_1<\cdots<k_m\leq n} \begin{vmatrix} a_{1k_1} & \cdots & a_{1k_m} \\ & \cdots & \\ a_{mk_1} & \cdots & a_{mk_m} \end{vmatrix} \cdot \begin{vmatrix} b_{k_1 1} & \cdots & b_{k_1 m} \\ & \cdots & \\ b_{k_m 1} & \cdots & b_{k_m m} \end{vmatrix}$$

$$= \sum_{1\leq k_1<\cdots<k_m\leq n} \det A \begin{pmatrix} 1 & 2 & \cdots & m \\ k_1 & k_2 & \cdots & k_m \end{pmatrix} \det B \begin{pmatrix} k_1 & k_2 & \cdots & k_m \\ 1 & 2 & \cdots & m \end{pmatrix}$$

に等しく (1.4) が成り立つ．　　証明終 ∎

注意　$m = n$ の場合，(1.4) の右辺は

$$\det A \begin{pmatrix} 1 & 2 & \cdots & m \\ 1 & 2 & \cdots & m \end{pmatrix} \det B \begin{pmatrix} 1 & 2 & \cdots & m \\ 1 & 2 & \cdots & m \end{pmatrix} = \det A \det B$$

であり，定理 1.1 (ix) を証明したことになっている．

1.4 終結式

2 つの (実係数または複素係数) 多項式

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$$
$$g(x) = b_0x^m + b_1x^{m-1} + \cdots + b_m, \quad a_0b_0 \neq 0$$

が与えられたとき，$m+n$ 次行列式

$$R(f,g) = \begin{vmatrix} a_0 & a_1 & \cdots & a_n & & & & \\ & a_0 & a_1 & \cdots & a_n & & & \\ & & \ddots & & & & & \\ & & & a_0 & a_1 & \cdots & a_{n-1} & a_n \\ b_0 & b_1 & \cdots & b_m & & & & \\ & b_0 & b_1 & \cdots & b_m & & & \\ & & \ddots & & & & & \\ & & & b_0 & b_1 & \cdots & b_{m-1} & b_m \end{vmatrix} \begin{matrix} \\ \\ \\ \Big\} m \\ \\ \\ \\ \Big\} n \end{matrix}$$

を f, g の**終結式 (resultant)** または **Sylvester (シルベスター) の行列式**という．

定理 1.5 高々 $m-1$ 次の多項式 $\varphi(x)$ と高々 $n-1$ 次の多項式 $\psi(x)$ を適当に選んで，恒等式

$$\varphi(x)f(x) + \psi(x)g(x) = R(f,g) \tag{1.6}$$

が成り立つようにできる．

【証明】 $R(f,g)$ の第 $1, 2, \ldots, m+n-1$ 列に，順次 $x^{n+m-1}, x^{n+m-2}, \ldots, x$ をかけて最後の列に加えれば

$$R(f,g) = \begin{vmatrix} a_0 & a_1 & \cdots & a_n & & & & x^{m-1}f(x) \\ & a_0 & a_1 & \cdots & a_n & & & x^{m-2}f(x) \\ & & \ddots & & & & & \vdots \\ & & & a_0 & a_1 & \cdots & a_{n-1} & f(x) \\ b_0 & b_1 & \cdots & b_m & & & & x^{n-1}g(x) \\ & b_0 & b_1 & \cdots & b_m & & & x^{n-2}g(x) \\ & & \ddots & & & & & \vdots \\ & & & b_0 & b_1 & \cdots & b_{m-1} & g(x) \end{vmatrix}$$

よって

$$\varphi(x) = \begin{vmatrix} a_0 & a_1 & \cdots & a_n & & & & x^{m-1} \\ & a_0 & a_1 & \cdots & a_n & & & x^{m-2} \\ & & \ddots & & & & & \vdots \\ & & & a_0 & a_1 & \cdots & a_{n-1} & 1 \\ b_0 & b_1 & \cdots & b_m & & & & 0 \\ & b_0 & b_1 & \cdots & b_m & & & 0 \\ & & \ddots & & & & & \vdots \\ & & & b_0 & b_1 & \cdots & b_{m-1} & 0 \end{vmatrix}$$

$$\psi(x)=\begin{vmatrix} a_0 & a_1 & \cdots & a_n & & & & 0 \\ & a_0 & a_1 & \cdots & a_n & & & 0 \\ & & \ddots & & & & & \vdots \\ & & & a_0 & a_1 & \cdots & a_{n-1} & 0 \\ b_0 & b_1 & \cdots & b_m & & & & x^{n-1} \\ & b_0 & b_1 & \cdots & b_m & & & x^{n-2} \\ & & \ddots & & & & & \vdots \\ & & & b_0 & b_1 & \cdots & b_{m-1} & 1 \end{vmatrix}$$

とおけば，$\varphi(x)$ と $\psi(x)$ はそれぞれ高々 $m-1$ 次，$n-1$ 次の多項式で (1.6) をみたす．　証明終 ∎

系 1.5.1　$f(x)$ と $g(x)$ が共通因子をもつための必要十分条件は $R(f,g)=0$ である．

【証明】　$f(x)$ と $g(x)$ が共通因子をもつときはその共通因子の零点 (すなわち f と g の共通根) を α として (1.6) の左辺に代入すれば

$$\varphi(\alpha)f(\alpha)=\psi(\alpha)g(\alpha)=0 \quad \therefore \quad R(f,g)=0$$

逆に $R(f,g)=0$ ならば (1.6) より

$$\varphi(x)f(x)=-\psi(x)g(x)$$

よって右辺の多項式 $\psi(x)g(x)$ は $f(x)$ で割り切れねばならない．ところが $\psi(x)$ の次数は高々 $n-1$ 次で $f(x)$ の次数 n より低いから $g(x)$ は $f(x)$ と共通因子をもたねばならない．　証明終 ∎

系 1.5.2　$f(x)$ と $g(x)$ が互いに素 (すなわち共通因子をもたない) ならば

$$\varphi(x)f(x)+\psi(x)g(x)=1 \quad (\deg\varphi<m,\ \deg\psi<n) \tag{1.7}$$

をみたす多項式 $\varphi(x)$ と $\psi(x)$ が存在する．ただし $\deg\varphi$ は φ の次数をあらわす．

【証明】　f と g が互いに素ならば系 1.5.1 によって $R(f,g)\neq 0$ であるから (1.6) の φ, ψ に対し

$$\widetilde{\varphi}(x)=\frac{1}{R(f,g)}\varphi(x), \quad \widetilde{\psi}(x)=\frac{1}{R(f,g)}\psi(x)$$

とおけばこれが (1.7) の φ, ψ になる．　証明終 ∎

系 1.5.3　$f(x)$ が重根をもつための必要十分条件は $R(f,f')=0$ である．

■**注意**　系 1.5.2 は s 個の多項式に拡張される．すなわち $f_1(x),\ldots,f_s(x)$ が互いに素な多項式ならば

$$\varphi_1(x)f_1(x)+\cdots+\varphi_s(x)f_s(x)=1$$

となるような多項式 $\varphi_1(x),\ldots,\varphi_s(x)$ が存在する．証明は [4] にある．

1.5　線形空間

$\mathbb{F}=\mathbb{R}$ または $\mathbb{C}$ とする．$\mathbb{F}$ の元を成分とする n 次ベクトル

$$\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}}, \quad x_i\in\mathbb{F}$$

の集合を $\mathbb{F}^n$ であらわす．すなわち

$$\mathbb{F}^n=\left\{\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}} \;\middle|\; x_i\in\mathbb{F},\ 1\leq i\leq n\right\}=\mathbb{F}^{n\times 1}$$

(また $\mathbb{F}^{1\times n}$ もしばしば同じ記号 $\mathbb{F}^n$ であらわす．)

このとき $\boldsymbol{x} = (x_1,\dots,x_n)^{\mathrm{t}}$, $\boldsymbol{y} = (y_1,\dots,y_n)^{\mathrm{t}} \in \mathbb{F}^n$ とスカラー $\alpha \in \mathbb{F}$ に対し加算 $\boldsymbol{x}+\boldsymbol{y} = (x_1+y_1,\dots,x_n+y_n)^{\mathrm{t}}$ とスカラー積 $\alpha\boldsymbol{x} = (\alpha x_1,\dots,\alpha x_n)^{\mathrm{t}}$ が定義され次が成り立つ．

(i) $\boldsymbol{x}+\boldsymbol{y} = \boldsymbol{y}+\boldsymbol{x}$ $(\boldsymbol{x},\boldsymbol{y}\in\mathbb{F}^n)$
(ii) $(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z} = \boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$ $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\mathbb{F}^n)$
(iii) $\boldsymbol{0} = (0,\dots,0)^{\mathrm{t}}$ とおけば $\boldsymbol{x}+\boldsymbol{0} = \boldsymbol{x}$ $(\boldsymbol{x}\in\mathbb{F}^n)$ (零元の存在)
(iv) $\boldsymbol{x}+(-1)\boldsymbol{x} = \boldsymbol{0}$ $(\boldsymbol{x}\in\mathbb{F}^n)$ (逆元 $\boldsymbol{x}' = (-1)\boldsymbol{x} = -\boldsymbol{x}$ の存在)
(v) $1\cdot\boldsymbol{x} = \boldsymbol{x}$ $(\boldsymbol{x}\in\mathbb{F}^n)$
(vi) $(\alpha\beta)\boldsymbol{x} = \alpha(\beta\boldsymbol{x})$ $(\boldsymbol{x}\in\mathbb{F}^n,\ \alpha,\beta\in\mathbb{F})$
(vii) $\alpha(\boldsymbol{x}+\boldsymbol{y}) = \alpha\boldsymbol{x}+\alpha\boldsymbol{y}$ $(\boldsymbol{x}\in\mathbb{F}^n,\ \alpha\in\mathbb{F})$
(viii) $(\alpha+\beta)\boldsymbol{x} = \alpha\boldsymbol{x}+\beta\boldsymbol{x}$ $(\boldsymbol{x}\in\mathbb{F}^n,\ \alpha,\beta\in\mathbb{F})$

これを抽象化して，一般な線形空間 (ベクトル空間) が次のように定義される．

定義 1.1 集合 X が $\mathbb{F}$ 上の**線形空間** (または**ベクトル空間**) であるとは任意の $x,y\in X$ とスカラー $\alpha\in\mathbb{F}$ に対して加算 $x+y\in X$ とスカラー積 $\alpha\cdot x\in X$ が定義され次が成り立つときをいう．

(i) $x+y = y+x$ $(x,y\in X)$
(ii) $(x+y)+z = x+(y+z)$ $(x,y,z\in X)$
(iii) 任意の $x\in X$ に対して $x+0=x$ となる元 $0\in X$ が存在する．この元を**零元** (ゼロ元) という．
(iv) 任意の $x\in X$ に対して $x+x'=0$ となる $x'\in X$ が存在する．x' を x の逆元といい $-x$ であらわす．
(v) $1\cdot x = x$ $(x\in X)$
(vi) $(\alpha\beta)x = \alpha(\beta x)$ $(x\in X,\alpha,\beta\in\mathbb{F})$
(vii) $\alpha(x+y) = \alpha x+\alpha y$ $(x,y\in X,\alpha\in\mathbb{F})$
(viii) $(\alpha+\beta)x = \alpha x+\beta x$ $(x\in X,\alpha,\beta\in\mathbb{F})$

上記定義において $\mathbb{F}=\mathbb{R}$ のとき**実線形空間**，$\mathbb{F}=\mathbb{C}$ のとき**複素線形空間**という．

例 1 $\mathbb{F}^{m\times n} = \left\{A=(a_{ij}) \;\middle|\; a_{ij}\in\mathbb{F},\ 1\le i\le m,\ 1\le j\le n\right\}$ とおくとき $X=\mathbb{F}^{m\times n}$ は通常の演算によって $\mathbb{F}$ 上の線形空間である．

例 2 有限閉区間 $[a,b]$ 上の実数値連続関数の全体を $C[a,b]$ であらわす．
$f,g\in C[a,b]$, $\alpha\in\mathbb{R}$ に対し $f+g$ と αf を

$$(f+g)(t) = f(t)+g(t),\quad (\alpha f)(t) = \alpha\cdot f(t)$$

と定義すれば $X=C[a,b]$ は実線形空間となる．

このように，定義 1.1 により定義される線形空間はいろいろな空間を含む．

定義 1.2 (1 次独立と 1 次従属) $\mathbb{F}$ 上の線形空間 X の元 $x_1,\dots,x_k$ が $\mathbb{F}$ 上 **1 次独立** (**linearly independent**) であるとは

$$\alpha_1x_1+\cdots+\alpha_kx_k = 0\ (\alpha_1,\cdots,\alpha_k\in\mathbb{F}) \Rightarrow\ \alpha_1=\cdots=\alpha_k=0$$

のときをいう．また 1 次独立でないとき **1 次従属** (**linearly dependent**) であるという．すなわち $x_1,\dots,x_k$ が 1 次従属であるとはことごとくは 0 でない $\alpha_1,\dots,\alpha_k\in\mathbb{F}$ を適当に選んで $\alpha_1x_1+\cdots+\alpha_kx_k=0$ とできることである．

定義 1.3 (線形空間の次元) $\mathbb{F}$ 上の線形空間 X 内の 1 次独立な元の最大数が有限値 n のとき X は n **次元線形空間**であるといい $\dim X = n$ とかく．このような n が存在しないとき X は**無限次元空間**であるといい $\dim X = \infty$ とかく．

定義 1.4 (基底) $\mathbb{F}$ 上の n 次元線形空間 X において，次の条件をみたす $x_1, \ldots, x_n \in X$ を X の**基底 (basis)** という．

(i) $x_1, \ldots, x_n$ は $\mathbb{F}$ 上 1 次独立．

(ii) X の任意の元 x は $x_1, \ldots, x_n$ の 1 次結合として書ける．すなわち

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n$$

をみたす $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ が存在する．($x_1, \ldots, x_n$ を X の**生成元**という．)

なお，X の元 $x_1, \ldots, x_l$ の 1 次結合 $\sum_{i=1}^{l} \alpha_i x_i (\alpha_i \in \mathbb{F})$ の全体を

$$\mathrm{span}\{x_1, \ldots, x_l\}$$

であらわす．すると (ii) は $X = \mathrm{span}\{x_1, \ldots, x_n\}$ を意味する．$x_1, \ldots, x_n$ が X の基底ならば $\dim X = n$ であるが，基底の選び方は 1 通りではない．例えば $X = \mathbb{R}^n$ において，$\boldsymbol{e}_i$ を第 i 単位基本ベクトル (p.4 注意参照) とするとき，$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ は X の基底であるが，

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{pmatrix}, \ \boldsymbol{u}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{pmatrix}, \ldots, \ \boldsymbol{u}_{n-1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 1 \end{pmatrix}, \ \boldsymbol{u}_n = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

も $\mathbb{R}^n$ の異なる基底である (検証は容易)．

$\mathbb{F}$ 上の n 次元線形空間 X の 1 つの基底 $x_1, \ldots, x_n$ を固定し，X の任意の元 y を

$$y = c_1 x_1 + \cdots + c_n x_n, \ c_i \in \mathbb{F}$$

とあらわすとき，対応 $y \mapsto (c_1, \ldots, c_n)^{\mathrm{t}}$ によって y と $\eta = (c_1, \ldots, c_n)^{\mathrm{t}} \in \mathbb{F}^n$ を同一視することができるから，$\mathbb{F}$ 上の n 次元線形空間 X は $\mathbb{F}^n$ と同じとみてよい．

さて，X, Y を 2 つの $\mathbb{F}$ 上の線形空間とする ($X = Y$ でもよい)．このとき写像 $f : X \to Y$ が**線形 (linear)** であるとは，任意の $x, y \in X$ とスカラー $\alpha \in \mathbb{F}$ に対して

(i) $f(x + y) = f(x) + f(y)$

(ii) $f(\alpha x) = \alpha f(x)$

が成り立つときをいう．いま

$$\dim X = n, \ \dim Y = m$$

かつ X, Y の基底をそれぞれ $\{x_1, \ldots, x_n\}$, $\{y_1, \ldots, y_m\}$ とし

$$x = \sum_{i=1}^{n} c_i x_i = (x_1, \ldots, x_n)\boldsymbol{c}, \quad \boldsymbol{c} = (c_1, \ldots, c_n)^{\mathrm{t}} \in \mathbb{F}^n \tag{1.8}$$

$$f(x_i) = \sum_{j=1}^{m} a_{ji} y_j \quad (a_{ij} \in \mathbb{F}, \ 1 \le i \le n, \ 1 \le j \le m) \tag{1.9}$$

とおけば

$$f(x)=f\Big(\sum_{i=1}^{n}c_ix_i\Big)=\sum_{i=1}^{n}c_if(x_i)=\sum_{i=1}^{n}c_i\sum_{j=1}^{m}a_{ji}y_j=\sum_{j=1}^{m}\Big(\sum_{i=1}^{n}a_{ji}c_i\Big)y_j$$

$$=(y_1,\dots,y_m)\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{m1}&\cdots&a_{mn}\end{pmatrix}\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}=(y_1,\dots,y_m)A\boldsymbol{c} \tag{1.10}$$

ただし

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{m1}&\cdots&a_{mn}\end{pmatrix}$$

とおいた．一方 $f(x)\in Y=\mathrm{span}\{y_1,\dots,y_m\}$ であるから

$$f(x)=\sum_{j=1}^{m}d_jy_j=(y_1,\dots,y_m)\boldsymbol{d}\quad(d_j\in\mathbb{F},\ \boldsymbol{d}=(d_1,\dots,d_m)^{\mathrm{t}}) \tag{1.11}$$

と書けば，(1.10) と (1.11) より

$$(y_1,\dots,y_m)(A\boldsymbol{c}-\boldsymbol{d})=\boldsymbol{0}$$

$\boldsymbol{\alpha}=A\boldsymbol{c}-\boldsymbol{d}=(\alpha_1,\dots,\alpha_m)^{\mathrm{t}}$ とおけば上式は $\sum_{j=1}^{m}\alpha_jy_j=0$ と書けるから，$y_1,\dots,y_m$ の 1 次独立性により $\alpha_j=0\ \forall j$．よって

$$\boldsymbol{d}=A\boldsymbol{c} \tag{1.12}$$

結局，線形写像 f と X,Y の基底が与えられれば $m\times n$ 行列 $A=(a_{ij})$ が (1.9) により一意に確定し，f は (1.12) により (1.8) の係数ベクトル $\boldsymbol{c}$ を (1.11) の係数ベクトル $\boldsymbol{d}$ に移す．このようにして $\mathbb{F}$ 上の抽象的な有限次元線形空間 X の上の線形演算は $\mathbb{R}^{m\times n}$ または $\mathbb{C}^{m\times n}$ 上の行列演算に帰着されることがわかる．A を f の**行列表現**という．

定義 1.5 (部分空間) S を $\mathbb{F}$ 上の線形空間 X の部分集合とするとき，S の任意の元 x,y とスカラー $\alpha\in\mathbb{F}$ に対して

$$x+y\in S,\quad \alpha x\in S$$

が成り立つならば S はまた $\mathbb{F}$ 上の線形空間である．S を X の**部分空間 (subspace)** という．

定理 1.6 S を X の r 次元空間とする．$x_1,\dots,x_k\in S\ (1\le k<r)$ が 1 次独立ならば適当な $r-k$ 個の元 $x_{k+1},\dots,x_r\in S$ をつけ加えて $x_1,\dots,x_r$ が S の基底となるようにできる．

【証明】 S の元 x を適当に選んで $x_1,\dots,x_k,x$ が 1 次独立であるようにできる．実際，もしそのような x が存在しないとすれば任意の $x\in S$ に対し $x_1,\dots,x_k,x$ は 1 次従属で $(c_0,c_1,\dots,c_k)\neq(0,\dots,0)$ なる $c_i\in\mathbb{F}$ を適当に選べば

$$c_0x+c_1x_1+\cdots+c_kx_k=0$$

ここで $c_0=0$ ならば $x_1,\dots,x_k$ の 1 次独立性により $c_1=\cdots=c_k=0$．これは矛盾であるから $c_0\neq0$ である．このとき

$$x=-\frac{1}{c_0}(c_1x_1+\cdots+c_kx_k)=\alpha_1x_1+\cdots+\alpha_kx_k\quad\Big(\alpha_i=-\frac{c_i}{c_0},\ 1\le i\le k\Big)$$

これは $x_1,\dots,x_k$ が S の基底であることを意味し，$\dim S=r>k$ に矛盾する．よって $x_1,\dots,x_k,x$

が1次独立であるような $x \in S$ を取り出して $x_{k+1} = x$ とおく．$k+1 < r$ ならばこの議論をくり返すことにより元 $x_{k+1}, \ldots, x_r$ をつけ加えて $x_1, \ldots, x_k, x_{k+1}, \ldots, x_r$ が1次独立 (したがって S の基底) であるようにできる．　証明終 ∎

定理 1.7 $\mathbb{F}$ 上の有限次元線形空間 X の部分空間 S_1, S_2 が $S_1 \subsetneqq S_2$ ならば

$$\dim S_1 < \dim S_2$$

【証明】$\dim S_1 = r$ とし，$x_1, \ldots, x_r$ を S_1 の基底とするとき，$x \in S_2 \setminus S_1$ なる x に対して $x_1, \ldots, x_r, x$ は1次独立である．なぜならば

$$c_1 x_1 + \cdots + c_r x_r + cx = 0 \quad (c_1, \ldots, c_r,\ c \in \mathbb{F}) \tag{1.13}$$

とするとき，$c \neq 0$ ならば

$$x = -\frac{1}{c}(c_1 x_1 + \cdots + c_r x_r) \ \in \operatorname{span}\{x_1, \ldots, x_r\} = S_1$$

となって矛盾を引き起こすから $c = 0$ でなければならず，このとき (1.13) より $c_1 = \cdots = c_r = 0$．よって $x_1, \ldots, x_r, x$ は1次独立で

$$\dim S_2 \geq r + 1 > r = \dim S_1$$　証明終 ∎

S_1 と S_2 を $\mathbb{F}$ 上の線形空間 X の部分空間とするとき

$$S_1 + S_2 = \left\{ x = x_1 + x_2 \;\middle|\; x_i \in S_i,\ i = 1, 2 \right\}$$
$$S_1 \cap S_2 = \left\{ x \;\middle|\; x \in S_1 \text{かつ} x \in S_2 \right\}$$

はどちらも X の部分空間である．このとき次が成り立つ．

定理 1.8 (次元等式 (dimension identity)) 部分空間 S_1 と S_2 に対し

$$\dim(S_1 + S_2) = \dim S_1 + \dim S_2 - \dim(S_1 \cap S_2)$$

【証明】 $\dim(S_1 \cap S_2) = r$, $\dim S_1 = s$, $\dim S_2 = t$ とおく．$x_1, \ldots, x_r$ を $S_1 \cap S_2$ の基底とすれば適当な $y_{r+1}, \ldots, y_s \in S_1$ を選んで

$$x_1, \ldots, x_r, y_{r+1}, \ldots, y_s$$

が S_1 の基底であるようにできる (定理 1.6)．同様に $z_{r+1}, \ldots, z_t \in S_2$ を適当に選んで

$$x_1, \ldots, x_r, z_{r+1}, \ldots, z_t$$

が S_2 の基底であるようにできる．このとき

$$x_1, \ldots, x_r, y_{r+1}, \ldots, y_s, z_{r+1}, \ldots, z_t \tag{1.14}$$

が $S_1 + S_2$ の基底であることを示せばよい．次の (i), (ii) を示す．
(i) (1.14) が1次独立であること．

$$\alpha_1 x_1 + \cdots + \alpha_r x_r + \beta_{r+1} y_{r+1} + \cdots + \beta_s y_s + \gamma_{r+1} z_{r+1} + \cdots + \gamma_t z_t = 0 \ (\alpha_i, \beta_j, \gamma_k \in \mathbb{F})$$

とすれば

$$\alpha_1 x_1 + \cdots + \alpha_r x_r + \beta_{r+1} y_{r+1} + \cdots + \beta_s y_s = -(\gamma_{r+1} z_{r+1} + \cdots + \gamma_t z_t) \ (= u \text{とおく})$$

上式左辺は S_1 の元であり，右辺は S_2 の元であるから $u \in S_1 \cap S_2$.

よって $u = \sum_{i=1}^{r} c_i x_i$ $(c_i \in \mathbb{F})$ と書けて

$$\alpha_1 x_1 + \cdots + \alpha_r x_r + \beta_{r+1} y_{r+1} + \cdots + \beta_s y_s = \sum_{i=1}^{r} c_i x_i \tag{1.15}$$

かつ

$$-(\gamma_{r+1} z_{r+1} + \cdots + \gamma_t z_t) = \sum_{i=1}^{r} c_i x_i \tag{1.16}$$

(1.15) より　　$\beta_{r+1} = \cdots = \beta_s = 0,$

(1.16) より　　$\gamma_{r+1} = \cdots = \gamma_t = c_1 = \cdots = c_r = 0 \quad \therefore\ \alpha_1 = \alpha_2 = \cdots = \alpha_r = 0,$

よって (1.14) は 1 次独立である.

(ii) (1.14) が $S_1 + S_2$ の生成元であること.

$x \in S_1 + S_2$ とすれば $x = u_1 + u_2$, $u \in S_1$, $u_2 \in S_2$ と書ける．このとき

$$u_1 = c_1 x_1 + \cdots + c_r x_r + c_{r+1} y_{r+1} + \cdots + c_s y_s$$
$$u_2 = d_1 x_1 + \cdots + d_r x_r + d_{r+1} z_{r+1} + \cdots + d_t z_t$$

とあらわせば

$$\begin{aligned} u_1 + u_2 &= (c_1 + d_1)x_1 + \cdots + (c_r + d_r)x_r + c_{r+1} y_{r+1} + \cdots + c_s y_s \\ &\qquad + d_{r+1} z_{r+1} + \cdots + d_t z_t \\ &\in \operatorname{span}\{x_1, \ldots, x_r, y_{r+1}, \ldots, y_s, z_{r+1}, \ldots, z_t\} \subseteq S_1 + S_2 \end{aligned}$$

$$\therefore\quad S_1 + S_2 \subseteq \operatorname{span}\{x_1, \ldots, x_r, y_{r+1}, \ldots, y_s, z_{r+1}, \ldots, z_t\} \subseteq S_1 + S_2$$

$$\therefore\quad S_1 + S_2 = \operatorname{span}\{x_1, \ldots, x_r, y_{r+1}, \ldots, y_s, z_{r+1}, \ldots, z_t\}$$

証明終 ∎

定義 1.6 (直和)　$S = S_1 + S_2$ において各 $x \in S$ を $x = x_1 + x_2$, $x_i \in S_i$, $i = 1, 2$ とただ 1 通りにあらわされるとき，部分空間 S は部分空間 S_1 と S_2 の**直和**であるといい $S = S_1 \oplus S_2$，または $S = S_1 + S_2$ (直和) と書く.

定理 1.9　$S = S_1 + S_2$ において，$S = S_1 \oplus S_2 \ \Leftrightarrow S_1 \cap S_2 = \{0\}$

【証明】　$\Rightarrow$ $x \in S_1 \cap S_2$ とする．$x \neq 0$ ならば

$$x = x + 0 = 0 + x$$

は $x = x_1 + x_2$, $x_i \in S_i$, $i = 1, 2$ の形の異なる表現であるから仮定に反する．故に $x = 0$

$\Leftarrow$　$x = x_1 + x_2 = x_1' + x_2'$ $(x_1, x_1' \in S_1,\ x_2, x_2' \in S_2)$ とするとき

$$x_1 - x_1' = x_2' - x_2 \ (= v \text{ とおく})$$

左辺は S_1 の元, 右辺は S_2 の元であるから $v \in S_1 \cap S_2 = \{0\} \quad \therefore\ v = 0$ すなわち $x_1 = x_1'$, $x_2 = x_2'$.

証明終 ∎

一般に $S_1, \ldots, S_m$ を部分空間とするとき，部分空間

$$S = S_1 + \cdots + S_m = \left\{ x = x_1 + \cdots + x_m \,\middle|\, x_i \in S_i,\ 1 \leq i \leq m \right\} \tag{1.17}$$

が定義される．(1.17) における $x = x_1 + \cdots + x_m$ のあらわし方が，ただ 1 通りであるとき S は

$S_1, \ldots, S_m$ の直和であるという．直和であることを

$$S = S_1 \oplus \cdots \oplus S_m,\ S = S_1 + \cdots + S_m \text{ (直和)}$$

などであらわす．

定理 1.10 $S = S_1 + \cdots + S_m$ において

$$S = S_1 \oplus \cdots \oplus S_m \Leftrightarrow (S_1 + \cdots + S_{k-1}) \cap S_k = \{0\} \quad (2 \leq k \leq m) \tag{1.18}$$

【証明】 $\Rightarrow$ 仮にある k $(2 \leq k \leq m)$ に対して $(S_1 + \cdots + S_{k-1}) \cap S_k \neq \{0\}$ とすれば

$$(S_1 + \cdots + S_{k-1}) \cap S_k \ni x \neq 0$$

なる x がある．$x \in S_1 + \cdots + S_{k-1}$ であるから $x = x_1 + \cdots + x_{k-1}$ $(x_i \in S_i,\ 1 \leq i \leq k-1)$ とすれば

$$x = x_1 + \cdots + x_{k-1} + 0 + \cdots + 0 \quad (x_k = x_{k+1} = \cdots = x_m = 0) \tag{1.19}$$

また $x \in S_k$ であるから

$$x = 0 + \cdots + 0 + x_k + 0 + \cdots + 0 \quad (x_k = x,\ x_i = 0\ (i \neq k)) \tag{1.20}$$

(1.19) と (1.20) は $x = x_1 + \cdots + x_m \in S_1 + \cdots + S_m$ の異なる表現であるから直和であるとの仮定に反する．

$\Leftarrow$ $(S_1 + \cdots + S_{k-1}) \cap S_k = \{0\}$ $(2 \leq k \leq m)$ のとき $S = S_1 + \cdots + S_m$ の各元 x のあらわし方は1通りであることを m に関する帰納法により証明する．

$m = 2$ のときは定理 1.9 により証明済みである．故に $m-1$ のとき成り立つとし

$$x = x_1 + \cdots + x_m = x_1' + \cdots + x_m' \in S = S_1 + \cdots + S_m,\ x_i, x_i' \in S_i$$

を x の2通りの表現とし，$\widetilde{x}_i = x_i - x_i'$ とおけば $\widetilde{x}_i \in S_i$ $(1 \leq i \leq m)$ かつ

$$\widetilde{x}_i + \cdots + \widetilde{x}_{m-1} = -\widetilde{x}_m \ (= v \text{ とおく})$$

左辺は $S_1 + \cdots + S_{m-1}$ の元で，右辺は S_m の元であるから仮定によって

$$v \in (S_1 + \cdots + S_{m-1}) \cap S_m = \{0\} \quad \therefore\ v = 0$$

よって $\widetilde{x}_m = 0$ かつ

$$\widetilde{x}_1 + \cdots + \widetilde{x}_{m-1} = 0 \tag{1.21}$$

(1.21) は $S_1 + \cdots + S_{m-1}$ の元0の表現であるから帰納法の仮定により表現は一意的であり $\widetilde{x}_1 = \cdots = \widetilde{x}_{m-1} = 0$ $\therefore\ x_i = \widetilde{x}_i$ $(1 \leq i \leq m-1)$ また $\widetilde{x}_m = 0$ より $x_m = x_m'$ よって m のとき成り立ち帰納法が完了する． 証明終 ∎

1.6 内積空間

n 次元ベクトル $\boldsymbol{x} = (x_1, \ldots, x_m)^{\mathrm{t}}$ と $\boldsymbol{y} = (y_1, \ldots, y_m)^{\mathrm{t}}$ の内積は $\boldsymbol{x}, \boldsymbol{y}$ が実ベクトルのとき

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} x_i y_i$$

複素ベクトルのとき

$$(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_i\overline{y}_i \quad (\overline{y}_i \text{は } y_i \text{の共役複素数})$$

として定義される．また有限閉区間 $[a,b]$ 上の連続関数 $f:[a,b]\to\mathbb{F}$ がつくる $\mathbb{F}$ 上の線形空間 $X=C[a,b]$ (この空間は無限次元空間である) において，$f,g\in X$ の内積は

$$(f,g)=\begin{cases}\displaystyle\int_a^b f(t)g(t)dt & (\mathbb{F}=\mathbb{R}\ \text{のとき})\\ \displaystyle\int_a^b f(t)\overline{g}(t)dt & (\mathbb{F}=\mathbb{C}\ \text{のとき})\end{cases}$$

により定義される．このとき $C[a,b]$ は**内積空間**となる．

定義 1.7 (内積空間) 実線形空間 X に次の条件をみたす写像 $(\ ,\ ):X\times X\to\mathbb{R}$ が定義されているとき $(\ ,\ )$ を**内積**といい，$(X,(\ ,\ ))$ を**実内積空間**という．以下内積 $(\ ,\ )$ を省略して単に X を**実内積空間**という．また，この場合，以下に述べる複素内積に対応して内積を**実内積**という．

(i) $(x,x)\geq 0\ (x\in X)\ ;\ (x,x)=0\Leftrightarrow x=0$
(ii) $(x,y)=(y,x)\quad (x,y\in X)$
(iii) $(x_1+x_2,y)=(x_1,y)+(x_2,y)\quad (x_1,x_2,y\in X)$
(iv) $(\alpha x,y)=\alpha(x,y)\quad (\alpha\in\mathbb{R},\ x,y\in X)$

一方 X が複素線形空間 $(\mathbb{F}=\mathbb{C})$ ならば，対応する複素内積 $(\ ,\ ):X\times X\to\mathbb{C}$ は次の条件をみたすものとして定義される．

(i)′ $(x,x)\geq 0\ (x\in X);(x,x)=0\Leftrightarrow x=0$
(ii)′ $(x,y)=\overline{(y,x)}\quad (x,y\in X)$
(iii)′ $(x_1+x_2,y)=(x_1,y)+(x_2,y)\quad (x_1,x_2,y\in X)$
(iv)′ $(\alpha x,y)=\alpha(x,y)\quad (\alpha\in\mathbb{C},\ x,y\in X)$

この内積をもつ複素線形空間 X を**複素内積空間**という．

なお，複素内積空間では，(ii)′ と (iii)′ より

$$\begin{aligned}(x,y_1+y_2)&=\overline{(y_1+y_2,x)}=\overline{(y_1,x)+(y_2,x)}=\overline{(y_1,x)}+\overline{(y_2,x)}\\&=(x,y_1)+(x,y_2)\end{aligned}$$

また (ii)′ と (iv)′ より

$$(x,\alpha y)=\overline{(\alpha y,x)}=\overline{\alpha(y,x)}=\overline{\alpha}\cdot\overline{(y,x)}=\overline{\alpha}(x,y)$$

が導かれる．

内積空間 X の部分集合 S が線形空間としての部分空間をなすとき S は X の**部分空間**であるという．

定義 1.8 (直交補空間) $\mathbb{F}$ 上の内積空間 X とその部分空間 S が与えられたとき

$$S^{\perp}=\left\{x\in X\ \middle|\ (x,y)=0\ \forall y\in S\right\}$$

は 1 つの部分空間をなす．これを S の**直交補空間 (orthogonal complement)** という．明らかに，$S=X$ のとき $S^{\perp}=\{0\}$，$S=\{0\}$ のとき $S^{\perp}=X$ である．

定義 1.9 (正規直交系と正規直交底) $\mathbb{F}$ 上の内積空間 X の元 $\varphi_1,\ldots,\varphi_k$ が**正規直交系 (orthonormal system)** をなすとは

$$(\varphi_i,\varphi_j)=\delta_{ij}=\begin{cases}1 & (i=j)\\ 0 & (i\neq j)\end{cases}$$

のときをいう．δ_{ij} は **Kronecker (クロネッカー) の記号**と呼ばれる．また $\dim X=n$ のとき n 個の正規直交系 $\varphi_1,\ldots,\varphi_n$ は X の基底となるから，これを**正規直交底**または**正規直交基底 (orthonormal basis)** という．

実際，$\varphi_1,\ldots,\varphi_n$ を正規直交系とし，$\sum_{j=1}^{n}c_j\varphi_j=0$ $(c_j\in\mathbb{F})$ とすれば $0=\left(\sum_{j=1}^{n}c_j\varphi_j,\ \varphi_i\right)=\sum_j c_j(\varphi_j,\ \varphi_i)=c_i$ $(1\le i\le n)$，よって $\varphi_1,\ldots,\varphi_n$ は 1 次独立である．

正規直交底は常に存在する．実際 1 次独立な元 $u_1,\ldots,u_m$ から正規直交系 $\varphi_1,\ldots,\varphi_m$ を構成する **Gram-Schmidt(グラム・シュミット)** の方法があるからである．これは次の方法である．

[Gram-Schmidt の直交化法]

まず $u_1,\ldots,u_m$ が 1 次独立ならば $u_i\neq 0$ $(1\le i\le m)$ であることに注意する．

$$\begin{aligned}
&v_1=u_1\ (v_1\neq 0\text{ に注意}), && \varphi_1=\frac{1}{\sqrt{(v_1,v_1)}}v_1\\
&v_2=u_2-(u_2,\varphi_1)\varphi_1\ (v_2\neq 0\text{ に注意}), && \varphi_2=\frac{1}{\sqrt{(v_2,v_2)}}v_2\\
&\quad\vdots\\
&v_{i+1}=u_{i+1}-\sum_{j=1}^{i}(u_{i+1},\varphi_j)\varphi_j\ (v_{i+1}\neq 0\text{ に注意}), && \varphi_{i+1}=\frac{1}{\sqrt{(v_{i+1},v_{i+1})}}v_{i+1}\\
&\quad\vdots
\end{aligned}$$

とすれば $\varphi_1,\ldots,\varphi_{i+1}$ は正規直交系をなす．実際，帰納法により，$\varphi_1,\ldots,\varphi_i$ が正規直交系をなせば

$$\begin{aligned}(v_{i+1},\varphi_k)&=(u_{i+1},\varphi_k)-\sum_{j=1}^{i}(u_{i+1},\varphi_j)(\varphi_j,\varphi_k)\\&=(u_{i+1},\varphi_k)-(u_{i+1},\varphi_k)=0\quad(1\le k\le i)\end{aligned}\tag{1.22}$$

また $v_{i+1}=0$ とすれば

$$\begin{aligned}u_{i+1}=\sum_{j=1}^{i}(u_{i+1},\varphi_j)\varphi_j&\in\operatorname{span}\{\varphi_1,\ldots,\varphi_i\}\\&=\operatorname{span}\{v_1,\ldots,v_i\}\\&=\operatorname{span}\{u_1,\ldots,u_i\}\end{aligned}$$

となって矛盾を引き起こすから $v_{i+1}\neq 0$．よって v_{i+1} から φ_{i+1} は確かに構築できて，$1\le k\le i$ のとき

$$(\varphi_{i+1},\varphi_k)=\frac{1}{\sqrt{(v_{i+1},v_{i+1})}}(v_{i+1},\varphi_k)=0\quad(\text{(1.22) による})$$

明らかに $(\varphi_{i+1},\varphi_{i+1})=1$ であるから $\varphi_1,\ldots,\varphi_i,\varphi_{i+1}$ は正規直交系をなす．

特に $u_1,\ldots,u_n$ を X の基底としてこれに Gram-Schmidt の直交化を施せば正規直交底 $\varphi_1,\ldots,\varphi_n$ が得られる．

X の 1 次独立な元 $u_1,\ldots,u_m$ から正規直交系 $\varphi_1,\ldots,\varphi_m$ をつくる上記の操作は

$$\varphi_i = \sum_{k=1}^{i} c_{ki} u_k, \quad i = 1, 2, \ldots, m$$

なる m 次上3角行列

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ & c_{22} & & c_{2m} \\ & & \ddots & \vdots \\ & & & c_{mm} \end{pmatrix}, \ c_{ii} > 0 \ (1 \le i \le m) \tag{1.23}$$

を構成する方法である ($c_{ii} = \dfrac{1}{\sqrt{(v_i, v_i)}}$ に注意). C は正則で $C^{-1} = (\gamma_{ij})$ も上3角行列である.

$$(\varphi_1, \ldots, \varphi_m) = (u_1, \ldots, u_m)C \quad \text{より} \quad (u_1, \ldots, u_m) = (\varphi_1, \ldots, \varphi_m)C^{-1}$$

$$\therefore \quad u_i = \sum_{k=1}^{i} \gamma_{ki} \varphi_k = \sum_{k=1}^{m} \gamma_{ki} \varphi_k \ (\because \ \gamma_{ki} = 0 \ (k > i))$$

$$\therefore \quad (u_i, \ u_j) = \Big(\sum_{k=1}^{m} \gamma_{ki} \varphi_k, \ \sum_{l=1}^{m} \gamma_{lj} \varphi_l \Big) = \sum_{k=1}^{m} \gamma_{ki} \overline{\gamma}_{kj} \quad (\because \ (\varphi_k, \ \varphi_l) = \delta_{kl})$$

$$\therefore \quad \overline{(u_i, \ u_j)} = \sum_{k=1}^{m} \overline{\gamma}_{ki} \cdot \gamma_{kj}$$

右辺は $(C^{-1})^* C^{-1}$ の (i,j) 要素をあらわすから

$$\big(\overline{(u_i, u_j)}\big) = (C^{-1})^* C^{-1} = (CC^*)^{-1}$$

行列 $G = \big((u_i, u_j)\big)$ は $u_1, \ldots, u_m$ のつくる **Gram 行列**と呼ばれる.

定理 1.11 X を $\mathbb{F}$ 上の n 次元線形空間, S をその部分空間とすれば

$$X = S \oplus S^{\perp}, \quad n = \dim S + \dim S^{\perp}$$

【証明】 $\dim S = r$ とし S の基底 $u_1, \ldots, u_r$ をとる. 必要ならば Gram-Schmidt の直交化を施して $u_1, \ldots, u_r$ は正規直交底としてよい.
このとき $x \in X$ に対して $c_i = (x, u_i)$, $y = \sum_{i=1}^{r} c_i u_i$ とおけば $y \in S$ である. ここで $z = x - y$ とおけば

$$(z, u_i) = (x, u_i) - (y, u_i) = c_i - c_i = 0 \quad (1 \le i \le r)$$

$$\therefore \quad (z, u) = 0 \quad \forall u \in \operatorname{span}\{u_1, \ldots, u_r\} = S \quad \therefore \quad z \in S^{\perp}$$

よって

$$x = y + z, \quad y \in S, \ z \in S^{\perp}$$

$$\therefore \quad X \subseteq S + S^{\perp} \subseteq X \quad \therefore \quad X = S + S^{\perp}$$

さらに $v \in S \cap S^{\perp}$ とすれば $v \in S^{\perp}$ より $(v, u) = 0 \ \ \forall u \in S$. 特に $u = v$ とおけば $(v, v) = 0$. これより $v = 0$ を得る. よって $S \cap S^{\perp} = \{0\}$. したがって

$$X = S + S^{\perp} \ (\text{直和})$$

である. さらに定理 1.8 より $n = \dim S + \dim S^{\perp}$ を得る. 証明終 ∎

系 1.11.1 定理 1.11 の仮定の下で

$$(S^{\perp})^{\perp} = S$$

【証明】　定理 1.11 により $\dim S + \dim S^{\perp} = n$. また部分空間 $S^{\perp}$ に定理 1.11 を適用して $X = S^{\perp} + (S^{\perp})^{\perp}$ (直和) かつ $n = \dim S^{\perp} + \dim(S^{\perp})^{\perp}$

$$\therefore \quad \dim(S^{\perp})^{\perp} = \dim S \tag{1.24}$$

一方 $S \subseteq (S^{\perp})^{\perp}$ である. なぜならば, $x \in S$ ならば $(y,x) = 0 \quad \forall y \in S^{\perp}$ これは $S^{\perp}$ の任意の元が x と直交していることを意味するから $x \in (S^{\perp})^{\perp}$. x は S の任意の元であったから

$$S \subseteq (S^{\perp})^{\perp} \tag{1.25}$$

(1.24), (1.25) と定理 1.7 より $S = (S^{\perp})^{\perp}$ を得る.　証明終 ∎

1.7 ノルム空間

$\mathbb{F}$ 上の線形空間 X 上の写像 $||\cdot|| : X \to \mathbb{R}$ が X 上の**ノルム**であるとは $x, y \in X$ とスカラー $\alpha \in \mathbb{F}$ に対し次の条件をみたすときをいう.

(i)　$||x|| \geq 0 \quad (x \in X)$; $||x|| = 0 \Leftrightarrow x = 0$
(ii)　$||\alpha x|| = |\alpha|\, ||x||$
(iii)　$||x + y|| \leq ||x|| + ||y||$　(3 角不等式)

定義 1.10 (**ノルム空間**)　ノルム $||\cdot||$ が定義された線形空間 $(X, ||\cdot||)$ を**ノルム空間** (**normed space**) または**ノルム線形空間** (**normed linear space**) という. ノルム $||\cdot||$ を省略して単に X をノルム空間ということも多い.

ノルム空間 $(X, ||\cdot||)$ においては次が成り立つ

$$\Big|\, ||x|| - ||y|| \,\Big| \leq ||x - y|| \leq ||x|| + ||y|| \quad (x, y \in X)$$

実際, ノルムの条件 (iii) において x を $x - y$ でおきかえれば

$$||x|| = ||(x - y) + y|| \leq ||x - y|| + ||y|| \tag{1.26}$$

$$\therefore \quad ||x|| - ||y|| \leq ||x - y|| \tag{1.27}$$

ここで x と y の役割を入れかえれば

$$||y|| - ||x|| \leq ||y - x|| = ||-(x - y)|| = ||x - y|| \tag{1.28}$$

(1.27) と (1.28) より

$$\Big|\, ||x|| - ||y|| \,\Big| \leq ||x - y||$$

また, ノルムの条件 (ii) と (iii) より

$$||x - y|| = ||x + (-y)|| \leq ||x|| + ||-y|| = ||x|| + ||y||$$

いま $\dim X = n < \infty$ とし, X の基底 $u_1, \ldots, u_n$ を固定する. このとき X の任意の元 x と a は

$$x = \xi_1 u_1 + \cdots + \xi_n u_n, \quad a = \alpha_1 u_1 + \cdots + \alpha_n u_n \tag{1.29}$$

と書けて

$$x - a = (\xi_1 - \alpha_1) u_1 + \cdots + (\xi_n - \alpha_n) u_n$$

$$\therefore \quad ||x - a|| \leq |\xi_1 - \alpha_1|\, ||u_1|| + \cdots + ||\xi_n - \alpha_n||\, ||u_n||$$

$$\leq \max_i |\xi_i - \alpha_i| M \quad (M = ||u_1|| + \cdots + ||u_n||)$$

故に $\xi_i \to \alpha_i$ $(1 \leq i \leq n)$ のとき $||x - a|| \to 0$ したがって $||x||$ は x の展開係数 $\xi_1, \ldots, \xi_n$ の連続関数である．

有限次元空間では次の重要な 2 つの定理が成り立つ．

定理 1.12 (有限次元ノルムの同値性) $||\cdot||$ と $||\cdot||'$ を有限次元線形空間 X における 2 つのノルムとするとき，適当な正の定数 m と M をとれば

$$m||x||' \leq ||x|| \leq M||x||' \quad \forall x \in X \tag{1.30}$$

(M と m は x に無関係な定数であることに注意)

【証明】 $\dim X = n$ とし X の基底 $u_1, \ldots, u_n$ を固定する．$x \in X$ を (1.29) のようにあらわすとき，$||x||_\infty = \max_i |\xi_i|$ は X 上の 1 つのノルムである．また $\varphi(x) = ||x||$ は x の展開係数 $\xi_1, \ldots, \xi_n$ の連続関数であることはすでに示した．したがって $\varphi(x)$ は有界閉集合 $S = \left\{ x \in X \;\middle|\; ||x||_\infty = 1 \right\}$ の上で最大値 κ_1 と最小値 κ_2 をとる．明らかに $\kappa_1 > 0$, $\kappa_2 > 0$. このとき任意の $x(\neq 0) \in X$ に対して $\frac{x}{||x||_\infty} \in S$ であるから

$$\kappa_2 \leq \left|\left| \frac{1}{||x||_\infty} x \right|\right| \leq \kappa_1$$

$$\therefore \quad \kappa_2 ||x||_\infty \leq ||x|| \leq \kappa_1 ||x||_\infty \tag{1.31}$$

(1.31) は $x = 0$ に対しても成り立つ．同様に適当な正定数 κ_1', κ_2' $(\kappa_1' \leq \kappa_2')$ を定めて次が成り立つ．

$$\kappa_2' ||x||_\infty \leq ||x||' \leq \kappa_1' ||x||_\infty \tag{1.32}$$

(1.31) と (1.32) より

$$\frac{\kappa_2}{\kappa_1'} ||x||' \leq ||x|| \leq \frac{\kappa_1}{\kappa_2'} ||x||'$$

故に $m = \dfrac{\kappa_2}{\kappa_1'}$, $M = \dfrac{\kappa_2}{\kappa_1'}$ とおけば (1.30) を得る． 証明終 ∎

注意 定理 1.12 は無限次元空間では成り立たない (山本 [6] 参照)．

定理 1.13 (Cauchy 列の収束) 有限次元ノルム空間 $(X, ||\cdot||)$ における Cauchy (コーシー) 列 $\{x_k\}$ は $k \to \infty$ のとき X の元に収束する．すなわち $||x_k - k_l|| \to 0$ $(k, l \to \infty)$ ならば適当な $x \in X$ が存在して $||x_k - x|| \to 0$ $(k \to \infty)$ となる．これを

$$\lim_{k \to \infty} x_k = x \quad \text{または} \quad x_k \to x \ (k \to \infty)$$

と書く．

【証明】 $\dim X = n$ とし X の基底を $u_1, \ldots, u_n$ とする．$x = \sum \xi_i u_i \in X$ に対し $||x||_\infty' = \max |\xi_i|$ とおく．定理 1.12 を 2 つのノルム $||\cdot||$ と $||\cdot||_\infty'$ に適用して

$$m||x||_\infty' \leq ||x|| \leq M||x||_\infty' \quad \forall\, x \in X \tag{1.33}$$

をみたす正定数 m, M がある．このとき

$$m||x_k - x_l||_\infty' \leq ||x_k - x_l|| \leq M||x_k - x_l||_\infty'$$

$$\therefore \quad \lim_{k,l\to\infty} ||x_k - x_l|| = 0 \Leftrightarrow \lim_{k,l\to\infty} ||x_k - x_l||'_\infty = 0$$

いま x_k を

$$x_k = \sum_{i=1}^{n} \xi_i^{(k)} u_i$$

とあらわすとき

$$x_k - x_l = \sum_{i=1}^{n} (\xi_i^{(k)} - \xi_i^{(l)}) u_i$$

よって $k, l \to \infty$ のとき

$$||x_k - x_l||'_\infty = \max_i |\xi_i^{(k)} - \xi_i^{(l)}| \to 0$$

したがって，各 i につき $\{\xi_i^{(k)}\}$, $k = 1, 2, \ldots$ は実数 ($\mathbb{F} = \mathbb{R}$) または複素数 ($\mathbb{F} = \mathbb{C}$) の Cauchy 列であるから，$k \to \infty$ のときある $\xi_i \in \mathbb{F}$ に収束する．$x = \sum_{i=1}^{n} \xi_i u_i$ とおけば，$k \to \infty$ のとき

$$||x_k - x||'_\infty = \max_i |\xi_i^{(k)} - \xi_i| \to 0$$

で (1.33) より

$$||x_k - x|| \le M||x_k - x||'_\infty \to 0 \quad (k \to \infty) \qquad \therefore \quad ||x_k - x|| \to 0 \quad (k \to \infty)$$

証明終 ∎

さて内積空間 $(X, (,))$ が与えられたとき

$$||x|| = \sqrt{(x, x)}, \ x \in X \tag{1.34}$$

とおいて X 上の 1 つのノルム $||\cdot||$ が定義される．実際，ノルムの条件 (i), (ii) は内積の定義から明らかである．(iii) をみるには，$x, y \in X$ に対し

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

$$= \begin{cases} ||x||^2 + ||y||^2 + 2(x, y) & (\mathbb{F} = \mathbb{R} \text{ のとき}) \\ ||x||^2 + ||y||^2 + 2\mathrm{Re}(x, y) & (\mathbb{F} = \mathbb{C} \text{ のとき}) \end{cases}$$

であるから，$\mathbb{F} = \mathbb{R}$ のとき

$$(x, y) \le ||x|| \cdot ||y||$$

$\mathbb{F} = \mathbb{C}$ のとき

$$\mathrm{Re}(x, y) \le ||x|| \cdot ||y||$$

を示せばよい．これは次の有名な不等式から導かれる．

定理 1.14 (Cauchy-Schwarz の不等式) 内積空間 X において

$$|(x, y)| \le ||x|| \cdot ||y|| \quad (x, y \in X)$$

【証明】 $y = 0$ のときは不等式は明らかであるから $y \neq 0$ としてよい．このとき

$$\alpha = -(x, y)/||y||^2$$

とおけば

$$0 \le (x + \alpha y, x + \alpha y) = ||x||^2 + (x, \alpha y) + (\alpha y, x) + |\alpha|^2 ||y||^2$$

$$= \begin{cases} ||x||^2 + 2\alpha(x,y) + |\alpha|^2||y||^2 & (\mathbb{F}=\mathbb{R}\text{ のとき}) \\ ||x||^2 + 2\mathrm{Re}(\alpha y, x) + |\alpha|^2||y||^2 & (\mathbb{F}=\mathbb{C}\text{ のとき}) \end{cases}$$

$$= ||x||^2 - \frac{|(x,y)|^2}{||y||^2}$$

$$\therefore \quad |(x,y)|^2 \leq ||x||^2 \cdot ||y||^2$$

したがって

$$|(x,y)| \leq ||x|| \cdot ||y||$$

等号は $x+\alpha y=0$ すなわち $x=-\alpha y$ のときに限る. 証明終 ∎

定理 1.14 を内積空間 $\mathbb{F}^n$ と $x=(x_1,\ldots,x_n)^{\mathrm{t}}$, $y=(y_1,\ldots,y_n)^{\mathrm{t}} \in \mathbb{F}^n$ に適用すれば

$$|(x,y)| = \begin{cases} \left|\sum_{i=1}^n x_i y_i\right| \leq \sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n y_i^2} & (\mathbb{F}=\mathbb{R}\text{ のとき}) \\ \left|\sum_{i=1}^n x_i \overline{y}_i\right| \leq \sqrt{\sum_{i=1}^n |x_i|^2}\sqrt{\sum_{i=1}^n |y_i|^2} & (\mathbb{F}=\mathbb{C}\text{ のとき}) \end{cases}$$

また, 内積空間 $X=C[a,b]$ に適用すれば, $f,g \in X$ に対し

$$|(f,g)| = \begin{cases} \left|\int_a^b f(t)g(t)dt\right| \leq \sqrt{\int_a^b f(t)^2 dt}\sqrt{\int_a^b g(t)^2 dt} & (\mathbb{F}=\mathbb{R}\text{ のとき}) \\ \left|\int_a^b f(t)\overline{g}(t)dt\right| \leq \sqrt{\int_a^b |f(t)|^2 dt}\sqrt{\int_a^b |g(t)|^2 dt} & (\mathbb{F}=\mathbb{C}\text{ のとき}) \end{cases}$$

線形空間 X に定義されるノルムはいろいろ考えられるが例えば n 次ベクトル $\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}} \in \mathbb{F}^n$ に対して

$$||\boldsymbol{x}||_p = \left(|x_1|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}} \quad (1 \leq p \leq \infty) \tag{1.35}$$

$$||\boldsymbol{x}||_\infty = \max_i |x_i| \tag{1.36}$$

とおけばこれらは $\mathbb{F}^n$ 上のノルムとなる. (1.35) がノルムであることは **Minkowski** (ミンコフスキーまたはミンコウスキー) **の不等式**

$$\left(\sum_{i=1}^n |x_i+y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p\right)^{\frac{1}{p}} \quad (1 \leq p \leq \infty)$$

から従う (証明は補題 4.2 参照). また (1.36) がノルムであることは明らか.

$\mathbb{F}^n$($\mathbb{F}^{n\times 1}$ または $\mathbb{F}^{1\times n}$) 上のノルムは**ベクトルノルム**と呼ばれるが, (1.35) は p ベクトルノルムまたは p ノルム, (1.36) は**最大値ノルム**または**最大ノルム**と呼ばれる. ($p=2$ のときは特に **Euclid** (ユークリッド) **ノルム**と呼ばれる.)

同様に (1.35), (1.36) を関数空間 $X=C[a.b]$ に移せば, $f \in C[a,b]$ に対するノルムが

$$||f||_p = \left(\int_a^b |f(t)|^p dt\right)^{\frac{1}{p}} \quad (1 \leq p < \infty)$$

$$||f||_\infty = \max_{a\leq t\leq b} |f(t)|$$

により定義される．この場合 $X = C[a.b]$ は無限次元空間であるから，ベクトルノルムと区別して $||f||_p$, $||f||_\infty$ を**関数ノルム**という．(ただしそれぞれ f の p ノルム (または L^p ノルム)，f の最大値ノルムと呼ぶことも多い．) なお，ベクトルノルムとの混同をさけて $||f||_\infty$ を $||f||_{[a,b]}$ と記すこともある．行列ノルムについては，§4.1 において詳述する．

1.8 値域，零空間，ランク

$m \times n$ 行列 $A = (a_{ij}) \in \mathbb{F}^{m\times n}$ は $\mathbb{F}^n$ から $\mathbb{F}^m$ への写像とみなされる．このとき

$$\mathcal{R}(A) = \{A\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{F}^n\} \quad と \quad \mathcal{N}(A) = \{\boldsymbol{x} \in \mathbb{F}^n \mid A\boldsymbol{x} = \boldsymbol{0}\}$$

はそれぞれ $\mathbb{F}^m$ と $\mathbb{F}^n$ の部分空間をなす．$\mathcal{R}(A)$ を A の**値域 (range)**, $\mathcal{N}(A)$ を A の**零空間 (null space)** または**核 (kernel)** という．

定理 1.15 任意の $m \times n$ 行列 $A \in \mathbb{F}^{m\times n}$ に対して

$$\dim \mathcal{R}(A) + \dim \mathcal{N}(A) = n$$

【証明】$\dim \mathcal{N}(A) = r$ とし，$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r \in \mathbb{F}^n$ を $\mathcal{N}(A)$ の基底とする．定理 1.6 により $\boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_n \in \mathbb{F}^n$ をこれにつけ加えて

$$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r, \boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_n$$

を $\mathbb{F}^n$ の基底とすることができる．このとき $\{A\boldsymbol{u}_j\}$, $r+1 \le j \le n$ は $\mathcal{R}(A)$ の基底である．実際，次の (i), (ii) が成り立つ．

(i) $\{A\boldsymbol{u}_j\}$, $r+1 \le j \le n$ は 1 次独立である．なぜならば，

$$\sum_{j=r+1}^{n} \alpha_j A\boldsymbol{u}_j = \boldsymbol{0} \quad ならば \quad A\Big(\sum_{j=r+1}^{n} \alpha_j \boldsymbol{u}_j\Big) = \boldsymbol{0}$$

$$\therefore \quad \sum_{j=r+1}^{n} \alpha_j \boldsymbol{u}_j \in \mathcal{N}(A)$$

よって

$$\sum_{j=r+1}^{n} \alpha_j \boldsymbol{u}_j = \beta_1 \boldsymbol{u}_1 + \cdots + \beta_r \boldsymbol{u}_r \quad (\beta_1, \ldots, \beta_r \in \mathbb{F})$$

と書けて

$$\beta_1 \boldsymbol{u}_1 + \cdots + \beta_r \boldsymbol{u}_r + (-\alpha_{r+1})\boldsymbol{u}_{r+1} + \cdots + (-\alpha_n)\boldsymbol{u}_n = \boldsymbol{0}$$

$$\therefore \quad \beta_1 = \cdots = \beta_r = -\alpha_{r+1} = \cdots = -\alpha_n = 0 \quad (\because\ \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n は\ \mathbb{F}^n の基底)$$

$$\therefore \quad \alpha_j = 0 \quad (r+1 \le j \le n)$$

(ii) $\mathcal{R}(A) = \mathrm{span}\{A\boldsymbol{u}_{r+1}, \ldots, A\boldsymbol{u}_n\}$ である．実際，任意の $\boldsymbol{x} \in \mathbb{F}^n$ に対して

$$\boldsymbol{x} = c_1 \boldsymbol{u}_1 + \cdots + c_n \boldsymbol{u}_n \quad (c_1, \cdots, c_n \in \mathbb{F})$$

とあらわすとき，両辺に左から A をかければ

$$A\boldsymbol{x} = c_{r+1} A\boldsymbol{u}_{r+1} + \cdots + c_n A\boldsymbol{u}_n \quad (\because\ A\boldsymbol{u}_j = 0\ (1 \le j \le r))$$

$$\in \mathrm{span}\{A\boldsymbol{u}_{r+1}, \ldots, A\boldsymbol{u}_n\}$$

$$\therefore\quad \mathcal{R}(A)\subseteq \mathrm{span}\{A\boldsymbol{u}_{r+1},\dots,A\boldsymbol{u}_n\}\subseteq\mathcal{R}(A)\quad\therefore\quad \mathcal{R}(A)=\mathrm{span}\{A\boldsymbol{u}_{r+1},\dots,A\boldsymbol{u}_n\}$$

故に，$\{A\boldsymbol{u}_j\}$, $r+1\le j\le n$ は $\mathcal{R}(A)$ の基底であり

$$\dim\mathcal{R}(A)=n-r=n-\dim\mathcal{N}(A)\quad\therefore\quad \dim\mathcal{R}(A)+\dim\mathcal{N}(A)=n$$

証明終 ∎

定理 1.16 $A\in\mathbb{F}^{m\times n}$ とする

(i) $\mathbb{F}=\mathbb{R}$ (すなわち $m\times n$ 実行列) $\Rightarrow$ $\mathcal{N}(A)=\mathcal{R}(A^{\mathrm{t}})^{\perp}$ かつ $\mathcal{R}(A)=\mathcal{N}(A^{\mathrm{t}})^{\perp}$

(ii) $\mathbb{F}=\mathbb{C}$ (すなわち $m\times n$ 複素行列) $\Rightarrow$ $\mathcal{N}(A)=\mathcal{R}(A^{*})^{\perp}$ かつ $\mathcal{R}(A)=\mathcal{N}(A^{*})^{\perp}$

【証明】 (i) A を $m\times n$ 実行列とし A の第 $1,\dots,$ 第 m 行を

$$\boldsymbol{a}^{(1)}=(a_{11},\dots,a_{1n}),\dots,\ \boldsymbol{a}^{(m)}=(a_{m1},\dots,a_{mn})$$

とすれば $\boldsymbol{a}^{(1)^{\mathrm{t}}},\dots,\boldsymbol{a}^{(m)^{\mathrm{t}}}$ は A^{t} の第 $1,\dots,$ 第 m 列であり

$$\mathcal{R}(A^{\mathrm{t}})=\mathrm{span}\{\boldsymbol{a}^{(1)^{\mathrm{t}}},\dots,\boldsymbol{a}^{(m)^{\mathrm{t}}}\}$$

よって

$$\begin{aligned}\boldsymbol{x}=(x_1,\dots,x_n)^{\mathrm{t}}\in\mathcal{N}(A)&\Leftrightarrow \boldsymbol{a}^{(i)}\boldsymbol{x}=\boldsymbol{0}\quad(1\le i\le m)\\&\Leftrightarrow(\boldsymbol{x},\boldsymbol{a}^{(i)^{\mathrm{t}}})=\boldsymbol{0}\quad(1\le i\le m)\\&\Leftrightarrow\boldsymbol{x}\in[\mathrm{span}\{\boldsymbol{a}^{(1)^{\mathrm{t}}},\dots,\boldsymbol{a}^{(m)^{\mathrm{t}}}\}]^{\perp}\\&\Leftrightarrow\boldsymbol{x}\in\mathcal{R}(A^{\mathrm{t}})^{\perp}\\\therefore\quad\mathcal{N}(A)&=\mathcal{R}(A^{\mathrm{t}})^{\perp}\end{aligned}$$

また A を A^{t} でおきかえれば $\mathcal{N}(A^{\mathrm{t}})=\mathcal{R}(A)^{\perp}\quad\therefore\ \mathcal{N}(A^{\mathrm{t}})^{\perp}=\mathcal{R}(A)$ (系 1.11.1 による．)

(ii) A が複素行列の場合には内積は複素内積で

$$\begin{aligned}\boldsymbol{x}\in\mathcal{N}(A)&\Leftrightarrow\boldsymbol{a}^{(i)}\boldsymbol{x}=\boldsymbol{0}\quad(1\le i\le m)\\&\Leftrightarrow\boldsymbol{x}^{*}\boldsymbol{a}^{(i)^{*}}=\boldsymbol{0}\quad(1\le i\le m)\\&\Leftrightarrow(\boldsymbol{a}^{(i)^{*}},\boldsymbol{x})=\boldsymbol{0}\quad(1\le i\le m)\\&\Leftrightarrow\boldsymbol{x}\in[\mathrm{span}\{\boldsymbol{a}^{(1)^{*}},\dots,\boldsymbol{a}^{(m)^{*}}]^{\perp}=\mathcal{R}(A^{*})^{\perp}\\\therefore\quad\mathcal{N}(A)&=\mathcal{R}(A^{*})^{\perp}\end{aligned}$$

最後に A を A^{*} でおきかえれば $\mathcal{N}(A^{*})=\mathcal{R}(A)^{\perp}$

よって系 1.11.1 により $\mathcal{N}(A^{*})^{\perp}=\big(\mathcal{R}(A)^{\perp}\big)^{\perp}=\mathcal{R}(A)$

証明終 ∎

系 1.16.1 $A\in\mathbb{F}^{m\times n}$ のとき $\dim\mathcal{R}(A)=\dim\mathcal{R}(A^{\mathrm{t}})$

【証明】 (i) $\mathbb{F}=\mathbb{R}$ のとき定理 1.15 と定理 1.16 によって

$$\dim\mathcal{R}(A)=n-\dim\mathcal{N}(A)=n-\dim\mathcal{R}(A^{\mathrm{t}})^{\perp}$$

ここで定理 1.11 を $S=\mathcal{R}(A^{\mathrm{t}})$ に適用すれば，上式の右辺は $\dim\mathcal{R}(A^{\mathrm{t}})$ に等しい．

(ii) $\mathbb{F}=\mathbb{C}$ のときも

$$\begin{aligned}\dim\mathcal{R}(A)&=n-\dim\mathcal{N}(A)&&(\text{定理 1.15})\\&=n-\dim\mathcal{R}(A^{*})^{\perp}&&(\text{定理 1.16})\end{aligned}$$

$$= \dim \mathcal{R}(A^*) \qquad \text{(定理 1.11)}$$

さらに $\dim \mathcal{R}(A^*) = r$ とし

$$A^*\boldsymbol{u}_1, \ldots, A^*\boldsymbol{u}_r$$

を $\mathcal{R}(A^*)$ の基底とすれば $A^{\mathrm{t}}\overline{\boldsymbol{u}}_1, \ldots, A^{\mathrm{t}}\overline{\boldsymbol{u}}_r$ が $\mathcal{R}(A^{\mathrm{t}})$ の基底となるから

$$\dim \mathcal{R}(A^*) = \dim \mathcal{R}(A^{\mathrm{t}})$$

よってこの場合も $\dim \mathcal{R}(A) = \dim \mathcal{R}(A^{\mathrm{t}})$ $(= \dim \mathcal{R}(A^*))$ である. 証明終 ∎

定義 1.11 (行列のランク) $A \in \mathbb{F}^{m\times n}$ に対し $\dim \mathcal{R}(A)$ を**行列のランク** (または**階数**) といい $\mathrm{rank}A$ であらわす. 系 1.16.1 により $\mathrm{rank}A = \mathrm{rank}A^{\mathrm{t}}$ である.

系 1.16.2 $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times l}$ ならば

$$\mathrm{rank}(AB) \leq \min\{\mathrm{rank}A,\ \mathrm{rank}B\}$$

【証明】 $\mathcal{R}(AB) \subseteq \mathcal{R}(A)$ であるから $\mathrm{rank}(AB) \leq \mathrm{rank}A$ である. また, 系 1.16.1 によって

$$\mathrm{rank}(AB) = \mathrm{rank}\big((AB)^{\mathrm{t}}\big) = \mathrm{rank}(B^{\mathrm{t}}A^{\mathrm{t}})$$
$$\leq \mathrm{rank}B^{\mathrm{t}} = \mathrm{rank}B$$
$$\therefore\quad \mathrm{rank}(AB) \leq \min\{\mathrm{rank}A,\ \mathrm{rank}B\}$$

証明終 ∎

系 1.16.3 $A \in \mathbb{F}^{m\times n}$ とする. n 元 m 立 1 次方程式

$$A\boldsymbol{x} = \boldsymbol{0} \quad (\boldsymbol{x} \in \mathbb{F}^n) \tag{1.37}$$

の解 $\boldsymbol{x}$ の全体は $(n - \mathrm{rank}A)$ 次元の線形空間をなす. これを (1.37) の**解空間 (solution space)** という.

【証明】 (1.37) の解の全体は部分空間 $\mathcal{N}(A)$ (A の零空間) であり, 定理 1.15 によって

$$\dim \mathcal{N}(A) = n - \dim \mathcal{R}(A) = n - \mathrm{rank}A$$

である. 証明終 ∎

$m \times n$ 行列 A の第 1〜第 n 列を $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ $(\in \mathbb{F}^m)$ とすれば

$$\mathcal{R}(A) = \left\{A\boldsymbol{x} \;\middle|\; \boldsymbol{x} \in \mathbb{F}^n\right\}$$
$$= \left\{x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n \;\middle|\; \boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} \in \mathbb{F}^n\right\} \quad = \mathrm{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\} \tag{1.38}$$

であるから

$$\mathrm{rank}A = r \Leftrightarrow \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\text{のうち 1 次独立なものの最大数が } r$$

また $\mathrm{rank}A = \mathrm{rank}A^{\mathrm{t}}$ であるから

$$\mathrm{rank}A = r \Leftrightarrow A\text{ の }m\text{ 個の行ベクトルのうち 1 次独立なものの最大数が } r$$

ということもできる.

定理 1.17 $A \in \mathbb{F}^{m\times n}$ とする. 次の 2 条件は同値である.

(i) $\mathrm{rank}A = r$

(ii) A の r 次部分行列 S で $\det S \neq 0$ なるものが存在し, $r < n$ ならば S を含むいかなる $r+1$ 次部分行列 T も $\det T = 0$ である.

【証明】 (i) ⇒ (ii) (1.38) において，一般性を失うことなく $\boldsymbol{a}_1,\dots,\boldsymbol{a}_r$ が 1 次独立であるとすれば $\boldsymbol{a}_1,\dots,\boldsymbol{a}_r$ は $\mathcal{R}(A)$ の基底である．以下これを仮定し，

$$A=(a_{ij}),\ S=\begin{pmatrix} a_{11} & \cdots & a_{1r} \\ & \cdots & \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}=A\begin{pmatrix} 1 & 2 & \cdots & r \\ 1 & 2 & \cdots & r \end{pmatrix}$$

とする．$r=n$ ならば $S=A$ で

$$\dim\mathcal{N}(S)=n-\dim\mathcal{R}(S)=n-\dim\mathcal{R}(A)=n-r=0 \quad \therefore\ \ \mathcal{N}(S)=\{\boldsymbol{0}\}$$

よって S は正則で $\det S\neq 0$.

$r<n$ ならば S を含む $r+1$ 次小行列 T は $T=A\begin{pmatrix} 1 & 2 & \cdots & r & s \\ 1 & 2 & \cdots & r & t \end{pmatrix}$ $(s,t>r)$ の形である．A の第 1〜第 r 列 $\boldsymbol{a}_1,\dots,\boldsymbol{a}_r$ は $\mathcal{R}(A)$ の基底であるから $\boldsymbol{a}_t=(a_{1t},a_{2t},\dots,a_{mt})^{\mathrm{t}}$ は $\boldsymbol{a}_1,\dots,\boldsymbol{a}_r$ の 1 次結合として

$$\boldsymbol{a}_t=\sum_{j=1}^{r}c_j\boldsymbol{a}_j \quad (c_j\in\mathbb{F})$$

とあらわされる．このとき T の第 1〜第 r 列にそれぞれ $-c_1,\dots,-c_r$ をかけて $(r+1)$ 列に加えると

$$\det T=\det\begin{pmatrix} a_{11} & \cdots & a_{1r} & a_{1t}-\sum_{j=1}^{r}c_ja_{1j} \\ & \cdots & & \\ a_{r1} & \cdots & a_{rr} & a_{rt}-\sum_{j=1}^{r}c_ja_{rj} \\ a_{s1} & \cdots & a_{sr} & a_{st}-\sum_{j=1}^{r}c_ja_{sj} \end{pmatrix}=\det\left(\begin{array}{ccc|c} & S & & 0 \\ & & & \vdots \\ & & & 0 \\ \hline a_{s1} & \cdots & a_{sr} & 0 \end{array}\right)$$

$$=0 \text{ (最終列に関して展開)}$$

(ii) ⇒ (i) 再び一般性を失うことなく (必要ならば行と列の並べかえをして) $S=A\begin{pmatrix} 1 & 2 & \cdots & r \\ 1 & 2 & \cdots & r \end{pmatrix}$ としてよい．この S は A の r 次**首座小行列**と呼ばれる (§1.3 参照)．次に

$$B=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{r1} & \cdots & a_{rn} \end{pmatrix}=(\boldsymbol{b}_1,\dots,\boldsymbol{b}_n)\in\mathbb{F}^{r\times n}$$

とおくと $S=(\boldsymbol{b}_1,\dots,\boldsymbol{b}_r)$ である．仮定 $\det S\neq 0$ より S の各列は 1 次独立である．

$r=n$ ならば $S=A$ で $\det S\neq 0$ は $\det A\neq 0$ を意味し $\mathrm{rank}A=\dim\mathcal{R}(A)=n$ である．$r<n$ ならば，$t>r$ なる t を任意に選ぶとき，$\det S\neq 0$ より適当な $d_1,\dots,d_r\in\mathbb{F}$ を選んで

$$\boldsymbol{b}_t=(a_{1t},\dots,a_{rt})^{\mathrm{t}}=\sum_{j=1}^{r}d_j\boldsymbol{b}_j$$

と書ける．このとき A の第 t 列 $\boldsymbol{a}_t=(a_{1t},\dots,a_{mt})^{\mathrm{t}}$ も

$$\boldsymbol{a}_t=\sum_{j=1}^{r}d_j\boldsymbol{a}_j$$

と書ける．実際，もしそうでないとすれば，ある k につき

$$a_{kt} \neq d_1 a_{k1} + \cdots + d_r a_{kr}$$

となるが，T の $r+1$ 次小行列 $\widetilde{T} = A\begin{pmatrix} 1 & 2 & \cdots & r & k \\ 1 & 2 & \cdots & r & t \end{pmatrix}$ に対し

$$\det \widetilde{T} = \det \begin{pmatrix} & & & a_{1t} - \sum_{j=1}^{r} d_j a_{1j} \\ & S & & \vdots \\ & & & a_{rt} - \sum_{j=1}^{r} d_j a_{rj} \\ a_{k1} & \cdots & a_{kr} & a_{kt} - \sum_{j=1}^{r} d_j a_{kj} \end{pmatrix}$$

$$= \det \left(\begin{array}{ccc|c} & & & 0 \\ & S & & \vdots \\ & & & 0 \\ \hline a_{k1} & \cdots & a_{kr} & \delta \end{array}\right) \quad \left(\text{ただし } \delta = a_{kt} - \sum_{j=1}^{r} d_j a_{kj} \right)$$

$$= (\det S)\delta \neq 0$$

これは仮定に反する．

よって $\mathcal{R}(A) = \mathrm{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\} = \mathrm{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r\}$ かつ $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ は 1 次独立である（$\because$ $\sum_{j=1}^{r} c_j \boldsymbol{a}_j = 0$, $c_j \in \mathbb{F}$ ならば $\sum_{j=1}^{r} c_j \boldsymbol{b}_j = 0$ であり，$\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r$ は 1 次独立であるから $c_j = 0$ $(1 \leq j \leq r)$ となる．）

結局 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ は $\mathcal{R}(A)$ の基底であり $\dim \mathcal{R}(A) = r$ すなわち $\mathrm{rank} A = r$ を得る．　　証明終 ∎

定理 1.18　$A \in \mathbb{F}^{m \times n}$ ならば

$$\mathrm{rank} A = \begin{cases} \mathrm{rank}(A^{\mathrm{t}} A) & (\mathbb{F} = \mathbb{R} \text{ のとき}) \\ \mathrm{rank}(A^{*} A) & (\mathbb{F} = \mathbb{C} \text{ のとき}) \end{cases}$$

【証明】　$\mathrm{rank} A = r$ とすれば A の列 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の中に r 個の 1 次独立なベクトルがある．それらを一般性を失うことなく $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ とする．

(i) $\mathbb{F} = \mathbb{R}$ のとき

$$G = A^{\mathrm{t}} A = \begin{pmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & \cdots & (\boldsymbol{a}_1, \boldsymbol{a}_n) \\ & \cdots & \\ (\boldsymbol{a}_n, \boldsymbol{a}_1) & \cdots & (\boldsymbol{a}_n, \boldsymbol{a}_n) \end{pmatrix}$$

とおく．G は $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ のつくる n 次の Gram 行列である（§1.6 参照）．

$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ が 1 次独立のとき G の r 次首座小行列 $G_r = G\begin{pmatrix} 1 & 2 & \cdots & r \\ 1 & 2 & \cdots & r \end{pmatrix}$ は正則である．実際 $\boldsymbol{c} = (c_1, \ldots, c_r)^{\mathrm{t}}$ とおけば列ベクトル $G_r \boldsymbol{c}$ の第 i 成分は

$$\sum_{j=1}^{r}(\boldsymbol{a}_i,\boldsymbol{a}_j)c_j=(\boldsymbol{a}_i,\sum_{j=1}c_j\boldsymbol{a}_j)$$

よって $\boldsymbol{u}=\sum_{j=1}^{r}c_j\boldsymbol{a}_j$ とおけば

$$G_r\boldsymbol{c}=\boldsymbol{0}\Leftrightarrow(\boldsymbol{a}_i,\boldsymbol{u})=0\ (1\le i\le r)$$
$$\Leftrightarrow \boldsymbol{u}\in\mathrm{span}\{\boldsymbol{a}_1,\dots,\boldsymbol{a}_r\}^{\perp}=\mathcal{R}(A)^{\perp}$$

一方 $\boldsymbol{u}\in R(A)$ であるから $\boldsymbol{u}\in R(A)^{\perp}\cap R(A)=\{\boldsymbol{0}\}\quad\therefore\ \boldsymbol{u}=\boldsymbol{0}$

$$\therefore\quad c_j=0\quad\forall j\quad(\because\ \boldsymbol{a}_1,\dots,\boldsymbol{a}_r:1\text{ 次独立})$$

したがって $G_r\boldsymbol{c}=\boldsymbol{0}$ より $\boldsymbol{c}=\boldsymbol{0}$ が結論されて G_r は正則である.

$$\therefore\quad \mathrm{rank}G\ge r$$

一方 $\mathrm{rank}G=\mathrm{rank}(A^{\mathrm{t}}A)\le\mathrm{rank}A=r$ (系 1.16.2) であるから $\mathrm{rank}G=r$ を得る.
(ii) $\mathbb{F}=\mathbb{C}$ のとき
A^*A の (i,j) 要素は $\boldsymbol{a}_i^*\boldsymbol{a}_j=(\boldsymbol{a}_j,\boldsymbol{a}_i)$ (複素内積) であるから

$$A^*A=\begin{pmatrix}(\boldsymbol{a}_1,\boldsymbol{a}_1)&\cdots&(\boldsymbol{a}_1,\boldsymbol{a}_n)\\&\cdots&\\(\boldsymbol{a}_n,\boldsymbol{a}_1)&\cdots&(\boldsymbol{a}_n,\boldsymbol{a}_n)\end{pmatrix}^{\mathrm{t}}$$

よって系 1.16.1 により

$$\mathrm{rank}(A^*A)=\mathrm{rank}\begin{pmatrix}(\boldsymbol{a}_1,\boldsymbol{a}_1)&\cdots&(\boldsymbol{a}_n,\boldsymbol{a}_1)\\&\cdots&\\(\boldsymbol{a}_1,\boldsymbol{a}_n)&\cdots&(\boldsymbol{a}_n,\boldsymbol{a}_n)\end{pmatrix}$$

以下 (i) と殆んど同じ議論が通用して, 上式は $\mathrm{rank}A$ に等しい. 証明終 ∎

■**注意** Gram 行列の重要な性質については定理 2.14 に述べる.

1.9 アセントとディセント

$A\in\mathbb{F}^{n\times n}$ とする. $A^0=I$ (n 次単位行列) は正則であるから $\mathcal{N}(A^0)=\{\boldsymbol{0}\}$ で

$$\{\boldsymbol{0}\}=\mathcal{N}(A^0)\subseteq\mathcal{N}(A)\subseteq\cdots\subseteq\mathcal{N}(A^m)\subseteq\mathcal{N}(A^{m+1})\subseteq\cdots\subseteq\mathbb{F}^n$$

である. したがって定理 1.7 によってある k につき $\mathcal{N}(A^k)=\mathcal{N}(A^{k+1})$ となる. このとき

$$\mathcal{N}(A^m)=\mathcal{N}(A^k)\quad(m\ge k)\tag{1.39}$$

が成り立つ. 実際, $\mathcal{N}(A^k)=\mathcal{N}(A^{k+1})$ かつ $\boldsymbol{x}\in\mathcal{N}(A^{k+2})$ とすれば $A^{k+2}\boldsymbol{x}=0$

$$\therefore\quad A^{k+1}(A\boldsymbol{x})=0\quad\therefore\quad A\boldsymbol{x}\in\mathcal{N}(A^{k+1})=\mathcal{N}(A^k)\quad\therefore\quad A^{k+1}\boldsymbol{x}=A^k(A\boldsymbol{x})=0$$
$$\therefore\quad \boldsymbol{x}\in\mathcal{N}(A^{k+1})$$

これより $\mathcal{N}(A^{k+2})\subseteq\mathcal{N}(A^{k+1})$ が得られる. これと $\mathcal{N}(A^{k+1})\subseteq\mathcal{N}(A^{k+2})$ を併せて $\mathcal{N}(A^{k+2})=\mathcal{N}(A^{k+1})$. 以下, 帰納的に (1.39) が示される.

定義 1.12 (**アセント**) $\mathcal{N}(A^m)=\mathcal{N}(A^{m+1})$ をみたす最小の整数 $m\ge0$ を A の**アセント** (**ascent, 昇数**) といい, $\alpha(A)$ であらわす.

同様に

$$\mathbb{F}^n = \mathcal{R}(A^0) \supseteq \mathcal{R}(A) \supseteq \cdots \supseteq \mathcal{R}(A^m) \supseteq \mathcal{R}(A^{m+1}) \supseteq \cdots \supseteq \{\mathbf{0}\}$$

より，ある k につき $\mathcal{R}(A^{k+1}) = \mathcal{R}(A^k)$ である．このとき (1.39) と類似に次が成り立つ．

$$\mathcal{R}(A^m) = \mathcal{R}(A^k) \quad (m \geq k) \tag{1.40}$$

定義 1.13 (ディセント)　$\mathcal{R}(A^m) = \mathcal{R}(A^{m+1})$ となる最小の整数 $m \geq 0$ を A のディセント (**descent**, **降数**) といい，$\delta(A)$ であらわす．

実は任意の n 次行列 A に対して $\alpha(A) = \delta(A)$ である．なぜならば任意の整数 $m \geq 0$ に対して定理 1.15 により

$$\dim \mathcal{R}(A^m) + \dim \mathcal{N}(A^m) = n$$
$$\dim \mathcal{R}(A^{m+1}) + \dim \mathcal{N}(A^{m+1}) = n$$

辺に引いて $\dim \mathcal{R}(A^m) - \dim \mathcal{R}(A^{m+1}) = \dim \mathcal{N}(A^{m+1}) - \dim \mathcal{N}(A^m)$．これより $\alpha(A) = \delta(A)$ は明らかであろう．したがって

$$\begin{aligned} A : \text{正則} &\Leftrightarrow \alpha(A) = 0 \quad (\because\ \mathcal{N}(A) = \mathcal{N}(A^0) = \{\mathbf{0}\}) \\ &\Leftrightarrow \delta(A) = 0 \\ &\Leftrightarrow \mathcal{R}(A) = \mathbb{F}^n \quad (\because\ \mathcal{R}(A) = \mathcal{R}(A^0) = \mathbb{F}^n) \end{aligned}$$

これを写像の言葉でいえば $A : \mathbb{F}^n \to \mathbb{F}^n$ が 1 対 1 (単射) であることと A が上への写像 $\mathcal{R}(A) = \mathbb{F}^n$ (全射) であることは同値である．

注意　アセントとディセントは無限次元線形空間の線形作用素 (線形写像) T に対しても定義されるが，このとき $\alpha(T)$, $\delta(T)$ は必ずしも有限ではなく，$\alpha(T) = \delta(T)$ は一般に成り立たない．(詳細は Taylor [17] を参照されたい．)

1.10　ブロック行列と Schur 補元

行列 $A \in \mathbb{F}^{m\times n}$ を

$$A = \begin{pmatrix} A_{11} & A_{11} & \cdots & A_{1l} \\ A_{21} & A_{22} & \cdots & A_{2l} \\ & & \cdots & \\ A_{k1} & A_{k2} & \cdots & A_{kl} \end{pmatrix} \quad (A_{ij}\text{は } m_i \times n_j\text{次},\ \sum_{i=1}^{k} m_i = m,\ \sum_{j=1}^{l} n_j = n) \tag{1.41}$$

と kl 個の小行列に分解したとき，A を**ブロック行列**，各 A_{ij} を**ブロック小行列**という．特に $m = n$, $k = l$ で A_{ii} が n_i 次正方かつ $A_{ij} = O$ $(i \neq j)$ のとき

$$A = \begin{pmatrix} A_{11} & & \\ & \ddots & \\ & & A_{nn} \end{pmatrix}$$

を**ブロック対角行列**という．

さて，(1.41) において $k = l = 2$ としたブロック行列

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad (A_{11}\text{は } p\text{次},\ A_{22}\text{は } q\text{次},\ p+q=n) \tag{1.42}$$

において，A_{11} は正則とし $\widetilde{A_{11}} = A_{22} - A_{21}A_{11}^{-1}A_{12}$ とおく．$\widetilde{A_{11}} \in \mathbb{F}^{q\times q}$ を A_{11} の **Schur** (シュア) **補元** (**complement**) という．

定理 1.19 (1.42) において A_{11} を正則とするとき，次が成り立つ．

(i) $\det A = \det A_{11} \det \widetilde{A_{11}}$

(ii) A が正則 $\Leftrightarrow$ $\widetilde{A_{11}}$ が正則

(iii) A が正則のとき

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\widetilde{A_{11}}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\widetilde{A_{11}}^{-1} \\ -\widetilde{A_{11}}^{-1}A_{21}A_{11}^{-1} & \widetilde{A_{11}}^{-1} \end{pmatrix}$$

【証明】 (i) 等式

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I_p & -A_{11}^{-1}A_{12} \\ O & I_q \end{pmatrix} = \begin{pmatrix} A_{11} & O \\ A_{21} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix}$$

の両辺の行列式をとれば

$$\det A = \det A_{11} \cdot \det(A_{22} - A_{21}A_{11}^{-1}A_{12}) = \det A_{11} \det \widetilde{A_{11}}$$

(ii) は (i) より明らかである．

(iii)

$$G = \begin{pmatrix} A_{11}^{-1} & O \\ -A_{21}A_{11}^{-1} & I_q \end{pmatrix}$$

とおけば G は正則で

$$GA = \begin{pmatrix} I_p & A_{11}^{-1}A_{12} \\ O & \widetilde{A_{11}} \end{pmatrix}$$

さらに

$$H = \begin{pmatrix} I_p & -A_{11}^{-1}A_{12}\widetilde{A_{11}}^{-1} \\ O & \widetilde{A_{11}}^{-1} \end{pmatrix}$$

とおけば，$H(GA) = I_n$ となるから

$$A^{-1} = HG = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\widetilde{A_{11}}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\widetilde{A_{11}}^{-1} \\ -\widetilde{A_{11}}^{-1}A_{21}A_{11}^{-1} & \widetilde{A_{11}}^{-1} \end{pmatrix}$$

同様に定理 1.19 の双対として次も成り立つ． 証明終 ∎

定理 1.19′ (1.42) において A_{22} を正則とし $\widetilde{A_{22}} = A_{11} - A_{12}A_{22}^{-1}A_{21}$ とおく．($\widetilde{A_{22}}$ を A_{22} の **Schur 補元**という．) このとき

(i) $\det A = \det A_{22} \det \widetilde{A_{22}}$

(ii) A が正則 $\Leftrightarrow$ $\widetilde{A_{22}}$ が正則

(iii) A が正則のとき

$$A^{-1} = \begin{pmatrix} \widetilde{A_{22}}^{-1} & -\widetilde{A_{22}}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}\widetilde{A_{22}}^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}\widetilde{A_{22}}^{-1}A_{12}A_{22}^{-1} \end{pmatrix}$$

【証明】 前定理の証明と同じであるが念のため証明を与える.

(i)
$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I_p & O \\ -A_{22}^{-1}A_{21} & I_q \end{pmatrix} = \begin{pmatrix} \widetilde{A_{22}} & A_{12} \\ O & A_{22} \end{pmatrix}$$

より両辺の行列式をとればよい.

(ii) は (i) より明らかである.

(iii)
$$K = \begin{pmatrix} I_p & -A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{pmatrix}, \qquad L = \begin{pmatrix} \widetilde{A_{22}}^{-1} & O \\ -A_{22}^{-1}A_{21}\widetilde{A_{22}}^{-1} & I_q \end{pmatrix}$$

とおけば

$$KA = \begin{pmatrix} \widetilde{A_{22}} & O \\ A_{22}^{-1}A_{21} & I_q \end{pmatrix}, \qquad LKA = \begin{pmatrix} I_p & \\ & I_q \end{pmatrix} = I_n$$

故に

$$A^{-1} = LK = \begin{pmatrix} \widetilde{A_{22}}^{-1} & -\widetilde{A_{22}}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}\widetilde{A_{22}}^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}\widetilde{A_{22}}^{-1}A_{12}A_{22}^{-1} \end{pmatrix}$$

証明終 ∎

1.11 既約行列

§1.2 において, $1,2,\ldots,n$ の置換

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

に対応する置換行列 P_σ を定義した. 以下 P_σ を P であらわす.

命題 1.1 P を置換 σ に対応する n 次置換行列, $A = (a_{ij}) \in \mathbb{F}^{n\times n}$ とするとき

(i) $PP^{\mathrm{t}} = P^{\mathrm{t}}P = I$ (したがって $P^{\mathrm{t}} = P^{-1}$)

(ii) $PAP^{\mathrm{t}} = PAP^{-1} = (a_{\sigma(i)\sigma(j)})$ (すなわち (i,j) 要素が $a_{\sigma(i)\sigma(j)}$)

【証明】 (i) これはすでに §1.2 で述べた. 実際 $P = (p_{ij})$ は

$$p_{ij} = \begin{cases} 1 & (j = \sigma(i) \text{ のとき}) \\ 0 & (j \neq \sigma(i) \text{ のとき}) \end{cases}$$

であるから PP^{t} の (i,j) 要素は

$$\sum_{k=1}^{n} p_{ik}p_{jk} = p_{i\sigma(i)}p_{j\sigma(i)} = \begin{cases} 1 & (j = i) \\ 0 & (j \neq i) \end{cases}$$

これは $PP^{\mathrm{t}} = I$ であることを示す. $P^{\mathrm{t}}P$ についても同様である.

(ii) PA の (i,j) 要素は

$$\sum_{k=1}^{n} p_{ik}a_{kj} = p_{i\sigma(i)}a_{\sigma(i)j} = a_{\sigma(i)j}$$

よって $(PA)P^{\mathrm{t}}$ の (i,j) 要素は

$$\sum_{k=1}^{n} a_{\sigma(i)k}p_{jk} = a_{\sigma(i)\sigma(j)}p_{j\sigma(j)} = a_{\sigma(i)\sigma(j)} \quad \therefore \quad PAP^{\mathrm{t}} = (a_{\sigma(i)\sigma(j)})$$

証明終 ∎

定義 1.14 n 次行列 $A \in \mathbb{F}^{n\times n}$ が**可約** (**reducible**) であるとは $n=1$ のとき $A=O$, また $n \geq 2$ のとき置換行列 P を定めて

$$PAP^{\mathrm{t}} = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix} \quad (A_{11}, A_{22}\text{は正方行列}) \tag{1.43}$$

とできるときをいう. (1.43) は

$$PAP^{\mathrm{t}} = \begin{pmatrix} A_{11} & O \\ A_{21} & A_{22} \end{pmatrix}$$

でおきかえてもよい. A が可約でないとき**既約** (**irreducible**) であるという.

命題 1.2 $n \geq 2$ のとき, A が可約であるための必要十分条件は適当なインデックス集合 $J \subsetneqq N = \{1,2,\ldots,n\}$ を定めて

$$i \in J \text{ かつ } j \notin J \text{ のとき } a_{ij} = 0$$

とできることである. 特に適当な i に対して

$$a_{ij} = 0 \quad \forall j \neq i$$

となるならば A は可約である.

【証明】 (i) 必要条件 : A を可約とする. P を置換 σ に対応する置換行列とし (1.43) における A_{11}, A_{22} の次数をそれぞれ r, s $(r+s=n)$ とする.

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & r & r+1 & \cdots & n \\ j_1 & j_2 & \cdots & j_r & i_1 & \cdots & i_s \end{pmatrix}$$

$$J = \{i_1, i_2, \ldots, i_s\},\ J^c = \{j_1, j_2, \ldots, j_r\},\ J \cup J^c = \{1,2,\ldots,n\}$$

とすれば PAP^{t} の $(r+p,\ q)$ 要素は命題 1.1 (ii) によって

$$0 = (PAP^{\mathrm{t}})_{(r+p,q)} = a_{\sigma(r+p)\sigma(q)} = a_{i_p j_q} \quad (1 \leq p \leq s,\ 1 \leq q \leq r)$$

すなわち $i \in J,\ j \notin J$ のとき $a_{ij} = 0$ となる.

(ii) 十分条件 : $i \in J,\ j \notin J$ のとき $a_{ij} = 0$ とする. また

$$J = \{i_1, i_2, \ldots, i_{n-r}\},\ J^c = \{j_1, j_2, \ldots, j_r\},\ J \cup J^c = \{1,2,\ldots,n\}$$

とし, 置換

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & r & r+1 & r+2 & \cdots & n \\ j_1 & j_2 & \cdots & j_r & i_1 & i_2 & \cdots & i_r \end{pmatrix}$$

に対応する置換行列を P とすれば, (i) と同様に PAP^{t} の $(r+p,\ q)$ 要素は

$$a_{\sigma(r+p)\sigma(q)} = a_{i_p j_q} = 0 \quad \forall i_p \in J,\ j_q \in J^c$$

これは PAP^{t} が (1.43) の形であることを示す. よって A は可約である.

特にある i につき $a_{ij} = 0\ \forall j \neq i$ ならば $J = \{i\}$ は

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

をみたすから A は可約である. なお, この場合 (1.43) における $P = P_\sigma$ は

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & i-1 & i & i+1 & \cdots & n \\ 1 & 2 & \cdots & i-1 & n & i+1 & \cdots & i \end{pmatrix} = (i,n) \text{ (互換)}$$

に対応し，A_{11} は $n-1$ 次，A_{22} は 1 次の行列である． 証明終 ∎

定義 1.15 (有限有向グラフ) $P_1,\ldots,P_n$ を平面上の n 個の点として，n 次行列 $A=(a_{ij})$ の要素が $a_{ij}\neq 0$ のとき P_i から P_j へ向かう矢印 $P_i \to P_j$ により P_i と P_j を結ぶ．このようにして得られる有向グラフを**有限有向グラフ**という．

定理 1.20 $n \geq 2$ とする．n 次行列 A が既約であるための必要十分条件は有限有向グラフが強連結であること，すなわち任意の i,j $(i\neq j)$ につき P_i から P_j へ至るパスの列

$$P_i \to P_{i_1}, P_{i_1} \to P_{i_2}, \cdots, P_{i_r} \to P_j$$

が存在することである．

【証明】 (i) 必要条件：A を既約とする．仮にある i,j につき P_i から P_j に至るパスの列がないとし

$$J = \{ l \mid P_i \text{から} P_l \text{に至るパスの列が存在しない} \}$$
$$K = J^c = \{ k \mid P_i \text{から} P_k \text{に至るパスの列が存在する} \}$$

とおくと $J\neq\emptyset$ ($\because$ $j\in J$) かつ $K\neq\emptyset$ ($\because$ A は既約であるから $a_{im}\neq 0$ なる $m(\neq i)$ がある．(命題 1.2 の後半の主張の対偶をとる．))

このとき $\alpha\in K$, $\beta\notin K$ ならば $a_{\alpha\beta}=0$ である．(仮にある $k\in K$ と $l\in J=K^c$ に対して $a_{kl}\neq 0$ とすれば P_k は P_l と結べる．よって $P_i \to P_{i_1} \to P_{i_2} \to \cdots \to P_{i_r} \to P_k \to P_l$ なるパスがあり $l\in K$．これは $l\in J$ と矛盾する．) これは A が可約であることを意味する．

(ii) 十分条件：任意の i,j $(i\neq j)$ につき P_i から P_j へ至るパスの列が存在すると仮定する．このとき A が可約ならば，適当なインデックス集合 J $(J\cup J^c=\{1,2,\ldots,n\},\ J^c\neq\emptyset)$ が存在して

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

故に対偶をとって，$i\in J$ かつ $a_{ij}\neq 0$ ならば $j\in J$ となる．いま

$$J=\{i_1,i_2,\ldots,i_s\},\ J^c=\{j_1,\ldots,j_t\},\ s+t=n$$

とすれば，仮定によって $i_p\in J$, $j_q\in J^c$ に対し P_{i_p} から P_{j_q} へ至るパスの列

$$P_{i_p} \to P_{k_1} \to P_{k_2} \to \cdots \to P_{k_{r-1}} \to P_{j_q}$$

がある．このとき

$$i_p\in J, \quad a_{i_pk_1}\neq 0 \quad \text{より} \quad k_1\in J$$
$$k_1\in J, \quad a_{k_1k_2}\neq 0 \quad \text{より} \quad k_2\in J$$
$$\cdots$$

以下これを続けて

$$k_{r-1}\in J, \quad a_{k_{r-1}j_q}\neq 0 \quad \text{より} \quad j_q\in J$$

となるが，$j_q\in J^c$ であったからこれは矛盾である．よって A は既約である． 証明終 ∎

1.12 優対角行列

n 次行列 $A=(a_{ij})\in\mathbb{F}^{n\times n}$ が

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

をみたすとき**優対角行列**または**対角優位行列** (**diagonally dominant matrix**) といい

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

をみたすとき**狭義優対角行列** (**strictly diagonally dominant matrix**) という．また既約な優対角行列が少なくとも 1 つの i_0 につき

$$|a_{i_0i_0}|>\sum_{\substack{j=1\\ j\neq i_0}}^{n}|a_{i_0j}| \tag{1.44}$$

をみたすとき**既約優対角行列** (**irreducibly diagonally dominant matrix**) という．

上記既約優対角行列の定義は，Varga [19] による伝統的な定義であるが，この行列は既約な優対角行列のことではないから注意が必要である．本書では Serre [15] に従い (1.44) をみたす優対角行列を**強優対角行列** (**strongly diagonally dominant matrix**) と呼ぶことにし，以下，従来のまぎらわしい用語「既約優対角行列」の代わりに「**既約強優対角行列**」を用いることにする．

定理 1.21 狭義優対角行列および既約強優対角行列は正則である．

【証明】 (i) A を狭義優対角行列とする．仮に A が正則でないとすれば $A\boldsymbol{x}=\boldsymbol{0}$ をみたすベクトル $\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}}\neq\boldsymbol{0}$ がある．ここで $\max_i|x_i|=|x_k|$ とすれば $|x_k|>0$ で

$$a_{kk}x_k=-\sum_{\substack{j=1\\ j\neq k}}^{n}a_{kj}x_j$$

$$\therefore\quad |a_{kk}|\cdot|x_k|=\Big|-\sum_{j\neq k}a_{kj}x_j\Big|\leq\sum_{j\neq k}|a_{kj}|\cdot|x_j|\leq\Big(\sum_{j\neq k}|a_{kj}|\Big)|x_k|$$

$$\therefore\quad |a_{kk}|\leq\sum_{j\neq k}|a_{kj}|$$

これは A が狭義優対角行列であることに矛盾する．よって A は正則である．

(ii) A を既約強優対角行列とする．仮に A が正則でないとすれば再び $A\boldsymbol{x}=\boldsymbol{0}$ をみたす $\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}}\neq\boldsymbol{0}$ がある．このとき $|x_1|=\cdots=|x_n|$ となることはない．なぜならば，もしこのようなことがおきれば (1.44) をみたす i_0 につき

$$|a_{i_0i_0}|\cdot|x_{i_0}|\leq\sum_{j\neq i_0}|a_{i_0j}|\cdot|x_j|=\Big(\sum_{j\neq i_0}|a_{i_0j}|\Big)|x_{i_0}|$$

かつ $|x_{i_0}|>0$ であるから $|a_{i_0i_0}|\leq\sum_{j\neq i_0}|a_{i_0j}|$ これは矛盾である．

故に $|x_1|=\cdots=|x_n|$ となることはなく

$$J=\big\{k\ \big|\ |x_k|=\max_i|x_i|\big\}$$

は $N=\{1,2,\ldots,n\}$ の真部分集合である．このとき

$$i\in J,\ j\notin J \Rightarrow a_{ij}=0 \tag{1.45}$$

を示そう．仮に上式が成り立たないとすれば，ある $i\in J$ と $j_0\notin J$ に対して $a_{ij_0}\neq 0$ である．すると $j_0\notin J$ より $|x_{j_0}|<||x||_\infty=|x_i|$ $\therefore$ $|a_{ij_0}||x_{j_0}|<|a_{ij_0}|\cdot|x_i|$

$$\therefore\ \sum_{j\notin J}|a_{ij}||x_j|<\Big(\sum_{j\notin J}|a_{ij}|\Big)|x_i|$$

$$\therefore\ |a_{ii}||x_i|=|a_{ii}x_i|=\Big|-\sum_{j\neq i}a_{ij}x_j\Big|=\Big|\sum_{\substack{j\in J\\ j\neq i}}a_{ij}x_j+\sum_{j\notin J}a_{ij}x_j\Big|$$

$$\leq\sum_{\substack{j\in J\\ j\neq i}}|a_{ij}||x_j|+\sum_{j\notin J}|a_{ij}||x_j|<\Big(\sum_{\substack{j\in J\\ j\neq i}}|a_{ij}|+\sum_{j\notin J}|a_{ij}|\Big)|x_i|=\Big(\sum_{j\neq i}|a_{ij}|\Big)|x_i|$$

$$\therefore\ |a_{ii}|<\sum_{j\neq i}|a_{ij}|$$

これは行列 A の優対角性に反する．よって (1.45) が成り立ち A は可約である．これは仮定に反するから不合理である．よって A は正則でなければならない．　証明終 ∎

1.13　2重確率行列と Birkhoff の定理

n 次行列 $A=(a_{ij})$ が

$$\sum_{j=1}^{n}a_{ij}=1,\ \sum_{i=1}^{n}a_{ij}=1,\ a_{ij}\geq 0\ \forall i,j$$

をみたすとき A を **2重確率行列 (doubly stochastic matrix)** という．n 次の2重確率行列の全体を S_n であらわし，$\boldsymbol{e}=(1,1,\ldots,1)^{\mathrm{t}}\in\mathbb{R}^n$ とおけば，A の要素がすべて非負のとき

$$A\in S_n \Leftrightarrow\ (A\boldsymbol{e}=\boldsymbol{e}\quad\text{かつ}\quad \boldsymbol{e}^{\mathrm{t}}A=\boldsymbol{e}^{\mathrm{t}})$$

明らかに置換行列 P_σ は2重確率行列である．

定義 1.16 (凸集合と凸1次結合)　線形空間 X の部分集合 S が**凸集合**であるとは

$$x,y\in S\Rightarrow\ \theta x+(1-\theta)y\in S,\ 0\leq\theta\leq 1$$

をみたすときをいう．$\theta x+(1-\theta)y$ は x と y の**凸1次結合**と呼ばれる．また，一般に

$$x_1,\ldots,x_n\in S,\ \theta_1\geq 0,\ldots,\ \theta_n\geq 0,\ \theta_1+\cdots+\theta_n=1$$

のとき $\theta_1x_1+\cdots+\theta_nx_n$ を $x_1,\ldots,x_n$ の凸1次結合という．

さて標題の定理は次のように述べられる (初学者は証明を読み飛ばしてもよい)．

定理 1.22 (Birkhoff (バーコフ) の定理)　n 次2重確率行列の全体 S_n は $n!$ 個の置換行列の凸1次結合の全体である．すなわち $P_1,\ldots,P_m(m=n!)$ を n 次置換行列とすれば，それらは凸集合 S_n の頂点をなし，$A\in S_n$ は

$$A=\sum_{j=1}^{m}\theta_jP_j,\quad\theta_1\geq 0,\ldots,\theta_m\geq 0,\quad\sum_{j=1}^{m}\theta_j=1$$

の形に書ける．

【証明】 いろいろな証明が知られているが，以下に Hoffman-Wielandt [25] による幾何学的証明を記す．(この鮮やかな証明は定理 6.13 (Hoffman-Wielandt の定理) の証明に補題として用いられたものであるが，著者の知る限り，なぜか国内外の成書にこの証明を紹介したものはない．) $n=1$ のときは明らかであるから $n \geq 2$ とする．

$$S_n = \left\{ X=(x_{ij}) \in \mathbb{R}^{n\times n} \;\middle|\; \sum_{i=1}^{n} x_{ij} = 1,\ \sum_{j=1}^{n} x_{ij} = 1,\ x_{ij} \geq 0,\ i,j = 1,2,\ldots,n \right\}$$

の元 X を実 n^2 次元空間の点と考えれば，S_n は有界閉凸多面体であり，S_n は $2n$ 個の超平面

$$\sum_{i=1}^{n} x_{ij} = 1, \quad \sum_{j=1}^{n} x_{ij} = 1 \tag{1.46}$$

と n^2 個の半空間 $x_{ij} \geq 0$, $1 \leq i,\ j \leq n$ の交わりである．ただし (1.46) の $2n$ 個の方程式は 1 次従属で，$2n-1$ 個が 1 次独立である．例えば $\sum_{i=1}^{n} x_{in} = 1$ は他の $2n-1$ 個の式より

$$\sum_{i=1}^{n} x_{in} = \sum_{i=1}^{n}\Big(1 - \sum_{j=1}^{n-1} x_{ij}\Big) = n - \sum_{i=1}^{n}\Big(\sum_{j=1}^{n-1} x_{ij}\Big) = n - \sum_{j=1}^{n-1}\Big(\sum_{i=1}^{n} x_{ij}\Big) = n-(n-1) = 1$$

として導かれる．よって S_n の任意の頂点 X は少なくとも $n^2-(2n-1) = (n-1)^2$ 個の半空間の境界上になければならない．故に少なくとも $(n-1)^2$ 個の (i,j) について

$$x_{ij} = 0 \tag{1.47}$$

これは任意の頂点 $X=(x_{ij})$ の少なくとも 1 つの行は $n-1$ 個がゼロで残りの 1 つが 1 であることを意味する．(なぜならば，もしすべての行が高々 $n-2$ 個のゼロしか含まないならば，全体で高々 $n(n-2) = n^2-2n$ $(< (n-1)^2)$ 個のゼロしか含まないことになり (1.47) に矛盾する．)

いまそれが第 i 行で起こり $x_{ij}=1$ とすれば X の第 i 行と j 列は x_{ij} を除いてすべてゼロである．この i 行と j 列を取り除いて得られる $(n-1)$ 次行列を X_{n-1} とすれば X_{n-1} も $n-1$ 次の 2 重確率行列で，それはまた S_{n-1} の 1 つの頂点である．(仮に頂点でないとすれば

$$X_{n-1} = \theta_1 X_{n-1}^{(1)} + \theta_2 X_{n-1}^{(2)}, \quad \theta_1 \geq 0,\ \theta_2 \geq 0,\ \theta_1+\theta_2 = 1 \tag{1.48}$$

をみたす 2 重確率行列 $X_{n-1}^{(1)}, X_{n-1}^{(2)} \in S_{n-1}$ がある．ここで $n-1$ 次行列 Y_{n-1} から n 次行列 Y を，Y の第 i 行と j 列を取り除いたものが Y_{n-1} と一致し，かつ (i,j) 要素のみ 1 で i 行と j 列のその他の要素はすべて 0 なるものとして構成する．この操作を ψ_{ij} : $Y_{n-1} \to Y$ とおけば ψ_{ij} は一意に定まり ψ_{ij} は $\psi_{ij} X_{n-1} = X$ をみたす．いま (1.48) において $X^{(1)} = \psi_{ij} X_{n-1}^{(1)}$, $X^{(2)} = \psi_{ij} X_{n-1}^{(2)}$ とおけば $X^{(1)}, X^{(2)} \in S_n$ で

$$X = \theta_1 X^{(1)} + \theta_2 X^{(2)}$$

と書けるから X は S_n の頂点ではない．これは不合理である．)

以下これをくり返せば，X の各行各列は 1 がちょうど 1 つ，その他は 0 であり，これは，S_n の頂点 X は置換行列であることを意味する．逆に置換行列の凸 1 次結合が 2 重確率行列であることは明らかであるから，これで定理の証明が完了する． 証明終 ∎

注意 定理 1.22 の純代数的証明は Marcus-Minc [14], Zhang [21], Mirsky [31], Taussky-Marcus [39] 等にある．ただし，[14], [21] の証明は [39] と本質的に同じものである．

第2章　線形代数の基礎(2)

線形代数学の中心をなすのは固有値の概念である．本章では行列の固有値・固有ベクトルに関する基礎事項を述べると共に固有値と密接な関係をもつ特異値の概念を説明する．

2.1　固有値と固有ベクトル

n 次行列 $A=(a_{ij})\in\mathbb{F}^{n\times n}$ に対し、$A-\lambda I$ が正則でないような実数または複素数 λ を A の**固有値** (**eigenvalue**) という．系 1.3.1 によって λ は $\det(A-\lambda I)=0$ の根である．

$$\begin{aligned}\det(A-\lambda I)&=(-1)^n\det(\lambda I-A)\\&=(-1)^n\begin{vmatrix}\lambda-a_{11}&-a_{12}&\cdots&-a_{1n}\\-a_{21}&\lambda-a_{22}&\cdots&-a_{2n}\\\vdots&\vdots&\ddots&\vdots\\-a_{n1}&-a_{n2}&\cdots&\lambda-a_{nn}\end{vmatrix}\qquad(2.1)\\&=(-1)^n\{\lambda^n-(a_{11}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^n\det A\}\qquad(2.2)\end{aligned}$$

(2.2) は (2.1) の行列式を展開して得られる．

$$f(\lambda)=\lambda^n-(a_{11}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^n\det A$$

を A の**特性多項式** (**characteristic polynomial**) といい

$$\det(A-\lambda I)=0\quad\text{または}\quad f(\lambda)=\det(\lambda I-A)=0$$

を**特性方程式** (**characteristic equation**) という．

V を任意の n 次正則行列とすれば

$$\det(V^{-1}AV-\lambda I)=\det(V^{-1}(A-\lambda I)V)=\det(A-\lambda I)$$

であるから $V^{-1}AV$ と A の固有値は等しい．また n 次代数方程式は重複度も込めて n 個の根をもつから，n 次行列 A は n 個の固有値 $\lambda_1,\ldots,\lambda_n$ をもつ．特性方程式 $f(\lambda)=0$ に対する根と係数の関係によって

$$\begin{aligned}\lambda_1+\lambda_2+\cdots+\lambda_n&=a_{11}+a_{22}+\cdots+a_{nn}\qquad(2.3)\\\lambda_1\lambda_2\cdots\lambda_n&=\det A\end{aligned}$$

(2.3) の右辺を A の**トレース** (**trace**) といい $\mathrm{tr}(A)$ と書く．同じ次数の行列 $A=(a_{ij})$ と $B=(b_{ij})$ に対し

$$\mathrm{tr}(A+B)=\mathrm{tr}(A)+\mathrm{tr}(B)$$

$$\mathrm{tr}(AB)=\mathrm{tr}(BA)$$

が成り立つ．前者は明らかであるが後者は

$$\mathrm{tr}(AB) = \sum_{i=1}^{n}\Big(\sum_{k=1}^{n} a_{ik}b_{ki}\Big) = \sum_{k=1}^{n}\Big(\sum_{i=1}^{n} a_{ik}b_{ki}\Big) = \sum_{k=1}^{n}\Big(\sum_{i=1}^{n} b_{ki}a_{ik}\Big)$$
$$= \sum_{k=1}^{n}(BA)_{kk} = \mathrm{tr}(BA)$$

よりわかる．また $V^{-1}AV$ と A は同一の固有値をもつから次式も成り立つ．

$$\mathrm{tr}(V^{-1}AV) = \mathrm{tr}(A)$$

さて λ が A の固有値ならば $A-\lambda I$ は正則でないから $\mathrm{rank}(A-\lambda I) < n$ である．系 1.16.3 によって $(A-\lambda I)\boldsymbol{x} = \boldsymbol{0}$ の解 $\boldsymbol{x}$ の全体 (解空間 $\mathcal{N}(A-\lambda I)$) は

$$n - \mathrm{rank}(A-\lambda I)$$

次元の線形空間をつくる．$\mathcal{N}(A-\lambda I)$ の元 $\boldsymbol{x} \neq \boldsymbol{0}$ を λ に対応する**固有ベクトル (eigenvector)** という．すなわち A の固有値 λ と固有ベクトル $\boldsymbol{x}$ は

$$A\boldsymbol{x} = \lambda\boldsymbol{x},\ \boldsymbol{x} \neq \boldsymbol{0}$$

をみたすものとして定義される．また $\mathcal{N}(A-\lambda I)$ を λ に対する**固有空間 (eigenspace)** という．

なお，$A_\lambda = A - \lambda I$, $\alpha = \alpha(A_\lambda)$ (A_λのアセント §1.9 参照) とすれば

$$\{\boldsymbol{0}\} \subsetneqq \mathcal{N}(A_\lambda) \subsetneqq \mathcal{N}(A_\lambda^2) \subsetneqq \cdots \subsetneqq \mathcal{N}(A_\lambda^{\alpha}) = \mathcal{N}(A_\lambda^{\alpha+1}) = \cdots$$

である．ただし $\mathcal{N} \subsetneqq \mathcal{M}$ は集合 $\mathcal{N}$ が集合 $\mathcal{M}$ の 真部分集合であることをあらわす．$\mathcal{N}(A_\lambda^m)$ $(m \leq \alpha)$ を m 位 (またはランク m) の**広義固有空間**または**一般固有空間 (generalized eigenspace)** という．

固有値の重複度として次の 2 種がある．

(i) 代数的重複度
特性方程式 $f(\lambda) = 0$ の根として重複度 m のとき，m を λ の**代数的重複度 (algebraic multiplicity)** という．

(ii) 幾何学的重複度
$\dim\mathcal{N}(A-\lambda I) = m'$ のとき m' を λ の**幾何学的重複度 (geometric multiplicity)** という．

例

$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ のとき特性多項式は $\det(\lambda I - A) = (\lambda-2)^3$ であるから固有値 $\lambda = 2$ の代数的重複度は 3 である．

一方 $\dim\mathcal{N}(A-2I) = 2$ (各自検証されたい) であるから $\lambda = 2$ の幾何学的重複度は 2 である．

さて，A の固有値は一般に複素数であるが，実対称行列 $(A^{\mathrm{t}} = A)$，Hermite 行列 $(A^* = A)$ の固有値は実数である．また直交行列 $(A^{\mathrm{t}} = A^{-1})$，ユニタリ行列 $(A^* = A^{-1})$ の固有値は単位円周上にある．これらを次に示す．

定理 2.1 実対称行列および複素 Hermite 行列の固有値は実数である．

【証明】 A を複素 Hermite 行列とし

$$A\boldsymbol{x} = \lambda\boldsymbol{x},\ \boldsymbol{x} = (x_1,\ldots,x_n)^{\mathrm{t}} \neq \boldsymbol{0} \tag{2.4}$$

とする．両辺に左から共役転置ベクトル $\boldsymbol{x}^* = (\bar{x}_1, \ldots, \bar{x}_n)$ をかけて

$$\boldsymbol{x}^* A \boldsymbol{x} = \lambda \boldsymbol{x}^* \boldsymbol{x} \tag{2.5}$$

一方 (2.4) の両辺の共役転置をとれば

$$\boldsymbol{x}^* A^* = \bar{\lambda} \boldsymbol{x}^*$$

右から $\boldsymbol{x}$ をかけて

$$\boldsymbol{x}^* A^* \boldsymbol{x} = \bar{\lambda} \boldsymbol{x}^* \boldsymbol{x} \tag{2.6}$$

上式の左辺は $\boldsymbol{x}^* A \boldsymbol{x}$ に等しいから (2.5), (2.6) より $\lambda \boldsymbol{x}^* \boldsymbol{x} = \bar{\lambda} \boldsymbol{x}^* \boldsymbol{x}$．$\boldsymbol{x}^* \boldsymbol{x} = \sum_{i=1}^{n} |x_i|^2 > 0$ であるから $\lambda = \bar{\lambda}$ よって λ は実数である．

A が実対称の場合の証明は上の証明に含まれる (A^* を A^{t} でおきかえればよい)．　証明終 ∎

定理 2.2 直交行列およびユニタリ行列の固有値は絶対値 1 をもつ．

【証明】 $A\boldsymbol{x} = \lambda \boldsymbol{x}$, $\boldsymbol{x} \neq \boldsymbol{0}$ において A がユニタリ行列 ($A^* A = I$) ならば，両辺に左から A^* をかけて

$$A^* A \boldsymbol{x} = \lambda A^* \boldsymbol{x} \quad \therefore \quad \boldsymbol{x} = \lambda A^* \boldsymbol{x}$$

最後の式の両辺に左から $\boldsymbol{x}^*$ をかけて

$$\boldsymbol{x}^* \boldsymbol{x} = \lambda \boldsymbol{x}^* A^* \boldsymbol{x} \tag{2.7}$$

一方 $A\boldsymbol{x} = \lambda \boldsymbol{x}$ の共役転置をとれば $\boldsymbol{x}^* A^* = \bar{\lambda} \boldsymbol{x}^*$

$$\therefore \quad \boldsymbol{x}^* A^* \boldsymbol{x} = \bar{\lambda} \boldsymbol{x}^* \boldsymbol{x} \quad \therefore \quad \lambda \boldsymbol{x}^* A^* \boldsymbol{x} = \lambda \bar{\lambda} \boldsymbol{x}^* \boldsymbol{x} \tag{2.8}$$

(2.7) と (2.8) より　$\lambda \bar{\lambda} \boldsymbol{x}^* \boldsymbol{x} = \boldsymbol{x}^* \boldsymbol{x}$　　$\therefore \quad |\lambda|^2 = 1$ ($\because$ $\boldsymbol{x}^* \boldsymbol{x} > 0$)

この証明は A が直交行列 ($A^{\mathrm{t}} A = I$) の場合にも適用される．　証明終 ∎

▒**注意**　相異なる固有値に対応する固有ベクトルは 1 次独立であることが知られているが，この事実を含めてさらに精密な結果を定理 2.7 に与える．

2.2 行列の 3 角化と対角化

定理 2.3 (Schur (シュア) の定理)　任意の n 次行列 A は適当なユニタリ行列 U をえらんで

$$U^* A U = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \quad (\text{上 3 角行列}) \tag{2.9}$$

とできる．ここに $\lambda_1, \ldots, \lambda_n$ は A の固有値である．特に A が実行列でその固有値がすべて実数ならば U は直交行列でおきかえられる．

【証明】 A の次数 n に関する帰納法により示す．

$n = 1$ のときは明らかであるから $n > 1$ かつ $n - 1$ のとき成り立つとする．

λ_1 を A の 1 つの固有値とし対応する長さ 1 の固有ベクトル $\boldsymbol{x}_1$ をとる．このとき定理 1.6 により，$n-1$ 個の n 次元ベクトル $\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n$ をえらんで

$$\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n$$

が $\mathbb{F}^n$ の基底であるようにできる．これらに必要ならば Gram-Schmidt の直交化を施して正規直交系 $\boldsymbol{v}_1=\boldsymbol{x}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_n$ をつくる．すると，これらを第 1〜第 n 列とする行列 $V=(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)$ はユニタリ行列であり

$$AV=V\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & *\end{pmatrix}=V\left(\begin{array}{c|c}\lambda_1 & \boldsymbol{c} \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right)\quad\left(\begin{array}{l}B\text{ は }n-1\text{ 次行列}\\ \boldsymbol{c}\text{ は }n-1\text{ 次元行ベクトル}\end{array}\right)$$

$V^*=V^{-1}$ を左から乗じて

$$V^*AV=\left(\begin{array}{c|c}\lambda_1 & \boldsymbol{c} \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right)$$

ここで $n-1$ 次行列 B に帰納法の仮定を用いれば適当な $n-1$ 次ユニタリ行列 V_B をえらんで

$$V_B^*BV_B=\begin{pmatrix}\beta_1 & & & * \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_{n-1}\end{pmatrix}\quad(\text{上 3 角行列})$$

とできる．ただし $\beta_1,\ldots,\beta_{n-1}$ は B の固有値である．

$$\widetilde{V}=\begin{pmatrix}1 & \\ & V_B\end{pmatrix},\ U=V\widetilde{V}$$

とおけば

$$\begin{aligned}U^*AU=\widetilde{V}^*V^*AV\widetilde{V}&=\begin{pmatrix}1 & \\ & V_B^*\end{pmatrix}\begin{pmatrix}\lambda_1 & \boldsymbol{c} \\ & B\end{pmatrix}\begin{pmatrix}1 & \\ & V_B\end{pmatrix}\\ &=\begin{pmatrix}\lambda_1 & \boldsymbol{c} \\ & V_B^*B\end{pmatrix}\begin{pmatrix}1 & \\ & V_B\end{pmatrix}=\left(\begin{array}{c|c}\lambda_1 & \boldsymbol{c}V_B \\ \hline 0 & \\ \vdots & V_B^*BV_B \\ 0 & \end{array}\right)\\ &=\left(\begin{array}{c|ccc}\lambda_1 & \multicolumn{3}{c}{\boldsymbol{c}V_B} \\ \hline 0 & \beta_1 & & \\ \vdots & & \ddots & * \\ 0 & & & \beta_{n-1}\end{array}\right)\quad(\text{上 3 角行列})\end{aligned}$$

となって帰納法が完了する．

さらに

$$\det(U^*AU-\lambda I_n)=\begin{vmatrix}\lambda_1-\lambda & & & * \\ & \beta_1-\lambda & & \\ & & \ddots & \\ & & & \beta_{n-1}-\lambda\end{vmatrix}=(\lambda_1-\lambda)(\beta_1-\lambda)\cdots(\beta_{n-1}-\lambda)$$

であるから $\beta_1,\ldots,\beta_{n-1}$ は A の固有値で $\lambda_1,\lambda_2=\beta_1,\ldots,\lambda_n=\beta_{n-1}$ は A の固有値をつくす．特に上の証明において A が実行列で A の固有値がすべて実数ならば $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$ は実ベクトルにとることができて V は直交行列であり B も実行列となる．B の固有値 $\beta_1,\ldots,\beta_{n-1}$ も実数であるから V_B も直交行列にとることができて U も直交行列となる．　証明終 ∎

系 2.3.1　A が (複素) Hermite 行列 $(A^*=A)$ ならばユニタリ行列 U をえらんで

$$U^*AU=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}$$

とできる．これをユニタリ行列による**対角化**という．特に A が実対称行列 $(A^{\mathrm{t}}=A)$ ならば U は直交行列にとることができる．

【証明】　A が (複素) Hermite 行列ならば

$$(U^*AU)^*=U^*A^*U=U^*AU$$

となって U^*AU も Hermite 行列である．従って (2.9) の右辺は対角行列でなければならない．

証明終 ∎

系 2.3.2　A が正規行列 $(AA^*=A^*A)$ であるための必要十分条件は適当なユニタリ行列 U により U^*AU を対角行列にできることである．

【証明】　(i) 必要条件：A を正規行列とする．定理 2.3 により適当なユニタリ行列 U をえらんで

$$U^*AU=S=\begin{pmatrix}s_{11} & s_{12} & \cdots & s_{1n}\\ & \ddots & & \vdots\\ & & \ddots & \vdots\\ 0 & & & s_{nn}\end{pmatrix}\quad (i>j\text{ ならば }s_{ij}=0)$$

と上 3 角化するとき $A=USU^*,\ A^*=US^*U^*$

故に

$$A:\text{正規}\Rightarrow AA^*=A^*A\Rightarrow SS^*=S^*S\Rightarrow\sum_{k=1}^{n}s_{ik}\cdot\overline{s}_{jk}=\sum_{k=1}^{n}\overline{s}_{ki}\cdot s_{kj}\quad\forall i,\ j$$

$i>j$ のとき $s_{ij}=0$ であるから，上式は

$$\sum_{\substack{i\le k\\ j\le k}}s_{ik}\overline{s}_{jk}=\sum_{\substack{k\le i\\ k\le j}}\overline{s}_{ki}s_{kj}$$

特に $i=j$ とおけば

$$\sum_{i\le k}|s_{ik}|^2=\sum_{k\le i}|s_{ki}|^2\quad\therefore\quad\sum_{i<k}|s_{ik}|^2=\sum_{k<i}|s_{ki}|^2$$

したがって

$$i=1 \text{ として } \sum_{k=2}^{n}|s_{1k}|^2=0 \qquad \therefore\ s_{12}=\cdots=s_{1n}=0$$

$$i=2 \text{ として } \sum_{k=3}^{n}|s_{2k}|^2=\sum_{k<2}|s_{k2}|^2=|s_{12}|^2=0 \qquad \therefore\ s_{23}=\cdots=s_{2n}=0$$

以下同様にして $s_{ij}=0\ (1\le i<j\le n)$ となり S は対角行列となる．

(ii) 十分条件：適当なユニタリ行列 U により

$$U^*AU=\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \quad (\text{対角行列})$$

とできたとすれば

$$\begin{aligned} U^*AA^*U &= (U^*AU)(U^*A^*U) \\ &= \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}\begin{pmatrix} \overline{d}_1 & & \\ & \ddots & \\ & & \overline{d}_n \end{pmatrix}=\begin{pmatrix} |d_1|^2 & & \\ & \ddots & \\ & & |d_n|^2 \end{pmatrix} \\ &= (U^*AU)^*(U^*AU)=U^*A^*AU \end{aligned}$$

両辺に左から U を，右から U^* を掛けて $AA^*=A^*A$ を得る．よって A は正規である．　証明終 ∎

■**注意**　Hermite行列は正規行列であるから系 2.3.1 は系 2.3.2 の特別な場合とみなせる．なお，Hermite行列以外の正規行列としては，ユニタリ行列，直交行列，歪 Hermite 行列等がある．

定理 2.4　A,B を n 次 (複素)Hermite 行列とする．B の固有値がすべて正ならば適当な正則行列 S をえらんで

$$S^*AS=D(\text{対角行列}) \quad \text{かつ} \quad S^*BS=I_n(\text{単位行列})$$

とできる．

【証明】　系 2.3.1 により適当なユニタリ行列 U をえらんで

$$U^*BU=\begin{pmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_n \end{pmatrix}$$

と対角化できる．ただし $\beta_1,\ldots,\beta_n$ は B の固有値である．仮定によって $\beta_i>0\ \ \forall i$ であるから $\rho_i=1/\sqrt{\beta_i}$ とおき $V=\mathrm{diag}(\rho_1,\ldots,\rho_n)$ とおけば

$$V^*(U^*BU)V=I_n$$

$H=V^*U^*AUV$ は Hermite 行列であるから適当なユニタリ行列 W をえらんで $W^*HW=\mathrm{diag}(d_1,\ldots,d_n)\ (=D\text{ とおく})$ と対角化できる．ここに $d_1,\ldots,d_n$ は H の固有値である．

このとき $S=UVW$ とおけば

$$S^*AS=W^*V^*U^*AUVW=W^*HW=D$$

かつ

$$S^*BS = W^*V^*U^*BUVW = W^*(V^*U^*BUV)W = W^*I_nW = W^*W = I_n$$

証明終 ∎

2.3 Hamilton-Cayley の定理と最小多項式

定理 2.5 (Hamilton-Cayley (ハミルトン・ケイリー) の定理) n 次行列 A の特性多項式を $f_A(\lambda) = \det(\lambda I - A)$ とおけば $f_A(A) = O$

【証明】 A の固有値を $\lambda_1, \ldots, \lambda_n$ とすれば $f_A(\lambda) = (\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$ と書ける.
定理 2.3 によって適当なユニタリ行列 U をえらんで

$$U^*AU = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

とできる. このとき

$$U^*f_A(A)U = U^*\prod_{i=1}^{n}(A - \lambda_i I)U = \prod_{i=1}^{n}(U^*AU - \lambda_i I)$$

ここで

$$A_1 \equiv U^*AU - \lambda_1 I = \begin{pmatrix} 0 & & & * \\ & \lambda_2 - \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_n - \lambda_1 \end{pmatrix}$$

$$A_2 \equiv A_1(U^*AU - \lambda_2 I)$$

$$= \begin{pmatrix} 0 & & & * \\ & \lambda_2 - \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_n - \lambda_1 \end{pmatrix} \begin{pmatrix} \lambda_1 - \lambda_2 & & & * \\ & 0 & & \\ & & \ddots & \\ & & & \lambda_n - \lambda_2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & & & * \\ & 0 & & \\ & & \prod_{i=1}^{2}(\lambda_3 - \lambda_i) & \\ & & & \ddots \\ & & & \prod_{i=1}^{2}(\lambda_n - \lambda_i) \end{pmatrix}$$

$$A_3 \equiv A_2(U^*AU - \lambda_3 I)$$

$$= \begin{pmatrix} 0 & & & & * \\ & 0 & & & \\ & & \prod_{i=1}^{2}(\lambda_3-\lambda_i) & & \\ & & & \ddots & \\ & & & & \prod_{i=1}^{2}(\lambda_n-\lambda_i) \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_3 & & & & * \\ & \lambda_2-\lambda_3 & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & \lambda_n-\lambda_3 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & & & & * \\ & 0 & & & \\ & & 0 & & \\ & & & \prod_{i=1}^{3}(\lambda_4-\lambda_i) & \\ & & & & \ddots & \\ & & & & & \prod_{i=1}^{3}(\lambda_n-\lambda_i) \end{pmatrix}$$

以下これを続けて

$$U^* f_A(A) U = A_{n-1}(U^*AU - \lambda_n I)$$

$$= \begin{pmatrix} 0 & & & * \\ & \ddots & & \\ & & 0 & \\ & & & \prod_{i=1}^{n-1}(\lambda_n-\lambda_i) \end{pmatrix} \begin{pmatrix} \lambda_1-\lambda_n & & & * \\ & \lambda_2-\lambda_n & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} = O$$

よって $f_A(A) = O$. 証明終 ∎

注意 1 定理 2.5 は，Cayley-Hamilton の定理とも呼ばれる．Macduffee [13] によれば，定理 2.5 は Hamilton (1853) が 4 元数に対して証明し，Cayley (1858) は $n = 2, 3$ に対して示した由であるので，本書は [13] に従って Hamilton に敬意を表し Hamilton-Cayley の定理と呼ぶ．

定理 2.5 によって，n 次行列 A が与えられれば高々 n 次の多項式 $p(x)$ で $p(A) = 0$ をみたすものが存在する．最高次の係数を 1 と正規化した多項式の中で

$$\varphi(x) = x^m + a_1 x^{m-1} + \cdots + a_m \tag{2.10}$$

を $\varphi(A) = 0$ をみたす最小次数の多項式とすれば $\varphi(x)$ は A により一意的に定まる．

実際 $\widetilde{\varphi}(x) = x^m + \widetilde{a}_1 x^{m-1} + \cdots + \widetilde{a}_m$ を $\widetilde{\varphi}(A) = 0$ かつ $\widetilde{\varphi}(x) \neq \varphi(x)$ なるものとすれば $\widetilde{a}_j = a_j$ $(1 \leq j \leq k-1)$, $\widetilde{a}_k \neq a_k$ をみたす k $(1 \leq k \leq m)$ がある．このとき

$$\varphi(x) = \frac{1}{\widetilde{a}_k - a_k}\big(\widetilde{\varphi}(x) - \varphi(x)\big)$$

は高々 $m-1$ 次の多項式で最高次の係数は 1 かつ $\varphi(A) = 0$ をみたし，仮定に反する．この多項式 $\varphi(x)$ を A の **最小多項式 (minimal polynomial)** という．

A の特性多項式 $f_A(x) = \det(xI - A)$ は最小多項式 $\varphi(x)$ により割り切れる．なぜならば $f_A(x)$ を $\varphi(x)$ で割ったときの商を $q(x)$，剰余を $r(x)$ とすれば

$$f_A(x) = q(x)\varphi(x) + r(x) \quad (r(x) \text{ の次数は} \varphi(x) \text{ の次数より小})$$

$$\therefore \quad r(A) = f_A(A) - q(A)\varphi(A) = O \quad \therefore \quad r(x) = 0$$

さて，A の相異なる固有値を $\lambda_1,\dots,\lambda_s$ とすれば定理 2.5 の証明より $\varphi(x)$ は

$$\varphi(x)=(x-\lambda_1)^{m_1}\cdots(x-\lambda_s)^{m_s}\quad(m_1,\dots,m_s\text{は自然数で}\sum_{i=1}^{s}m_i=m)$$

の形である．このとき $A_i=A-\lambda_iI$ とおくと次式が成り立つ．

$$\mathbb{F}^n=\mathcal{N}(A_1^{m_1})\oplus\cdots\oplus\mathcal{N}(A_s^{m_s}) \tag{2.11}$$

これは次の定理の第 1 式から従う (定理 2.6 の証明の最後に示す)．

定理 2.6 ([44]) 任意の自然数 $n_1,\dots,n_s$ に対して次が成り立つ．

$$\mathcal{N}(A_1^{n_1}\cdots A_s^{n_s})=\mathcal{N}(A_1^{n_1})\oplus\cdots\oplus\mathcal{N}(A_s^{n_s})$$
$$\mathcal{R}(A_1^{n_1}\cdots A_s^{n_s})=\mathcal{R}(A_1^{n_1})\cap\cdots\cap\mathcal{R}(A_s^{n_s})$$

【証明】 簡単のため $\widetilde{A}_1=A_1^{n_1}$, $\widetilde{A}_2=A_2^{n_2}\cdots A_s^{n_s}$, $\mathcal{N}_i=\mathcal{N}(\widetilde{A}_i)$, $i=1,\ 2$ とおく．

まず $\widetilde{A}_1\widetilde{A}_2=\widetilde{A}_2\widetilde{A}_1$ に注意して $\mathcal{N}(\widetilde{A}_1\widetilde{A}_2)=\mathcal{N}_1+\mathcal{N}_2$(直和) を示す．

(i) $\mathcal{N}_1+\mathcal{N}_2\subseteq\mathcal{N}(\widetilde{A}_1\widetilde{A}_2)$ なること：$\boldsymbol{x}\in\mathcal{N}_1+\mathcal{N}_2$ とすれば $\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2$, $\boldsymbol{x}_i\in\mathcal{N}_i$, $i=1,\ 2$ と書けるから

$$\begin{aligned}\widetilde{A}_1\widetilde{A}_2\boldsymbol{x}&=\widetilde{A}_1\widetilde{A}_2\boldsymbol{x}_1+\widetilde{A}_1\widetilde{A}_2\boldsymbol{x}_2\\&=\widetilde{A}_2(\widetilde{A}_1\boldsymbol{x}_1)+\widetilde{A}_1(\widetilde{A}_2\boldsymbol{x}_2)=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}\\\therefore\quad\boldsymbol{x}\in\mathcal{N}(\widetilde{A}_1\widetilde{A}_2)\quad&\therefore\quad\mathcal{N}_1+\mathcal{N}_2\subseteq\mathcal{N}(\widetilde{A}_1\widetilde{A}_2)\end{aligned}$$

(ii) $\mathcal{N}(\widetilde{A}_1\widetilde{A}_2)\subseteq\mathcal{N}_1+\mathcal{N}_2$ なること：

$$f_1(z)=(z-\lambda_1)^{n_1},\ f_2(z)=\prod_{i=2}^{s}(z-\lambda_i)^{n_i}$$

とおけば $f_1(z)$ と $f_2(z)$ は互いに素な多項式であるから系 1.5.2 によって

$$h_1(z)f_1(z)+h_2(z)f_2(z)=1$$

をみたす多項式 $h_1(z)$, $h_2(z)$ が存在する．このとき

$$h_1(A)f_1(A)+h_2(A)f_2(A)=I$$

故に任意の $\boldsymbol{x}\in\mathbb{F}^n$ に対して

$$\boldsymbol{x}=I\boldsymbol{x}=h_1(A)f_1(A)\boldsymbol{x}+h_2(A)f_2(A)\boldsymbol{x}=h_1(A)\widetilde{A}_1\boldsymbol{x}+h_2(A)\widetilde{A}_2\boldsymbol{x} \tag{2.12}$$

ここで $\boldsymbol{x}\in\mathcal{N}(\widetilde{A}_1\widetilde{A}_2)$ ならば

$$h_1(A)\widetilde{A}_1\boldsymbol{x}\in\mathcal{N}_2\quad(\because\ \widetilde{A}_2(h_1(A)\widetilde{A}_1\boldsymbol{x})=h_1(A)(\widetilde{A}_1\widetilde{A}_2\boldsymbol{x})=\boldsymbol{0})$$

かつ

$$h_2(A)\widetilde{A}_2\boldsymbol{x}\in\mathcal{N}_1\quad(\because\ \widetilde{A}_1(h_2(A)\widetilde{A}_2\boldsymbol{x})=h_2(A)(\widetilde{A}_1\widetilde{A}_2\boldsymbol{x})=\boldsymbol{0})$$

故に (2.12) より $\boldsymbol{x}\in\mathcal{N}_1+\mathcal{N}_2\quad\therefore\quad\mathcal{N}(\widetilde{A}_1\widetilde{A}_2)\subseteq\mathcal{N}_1+\mathcal{N}_2$

(iii) $\mathcal{N}_1 \cap \mathcal{N}_2 = \{\boldsymbol{0}\}$ なること：

$\boldsymbol{x} \in \mathcal{N}_1 \cap \mathcal{N}_2$とすれば　　$\widetilde{A}_1\boldsymbol{x} = \boldsymbol{0}$　かつ　$\widetilde{A}_2\boldsymbol{x} = \boldsymbol{0}$

よって (2.12) によって　　$\boldsymbol{x} = \boldsymbol{0}$　を得る．　∴　$\mathcal{N}_1 \cap \mathcal{N}_2 = \{\boldsymbol{0}\}$

(i), (ii), (iii) より　$\mathcal{N}(\widetilde{A}_1\widetilde{A}_2) = \mathcal{N}_1 \oplus \mathcal{N}_2$．したがってまた

$$\mathcal{N}_2 = \mathcal{N}(A_2^{n_2}) \oplus \mathcal{N}(A_3^{n_3} \cdots A_s^{n_s})$$

以下これをくり返せば

$$\mathcal{N}(A_1^{n_1} \cdots A_s^{n_s}) = \mathcal{N}(A_1^{n_1}) \oplus \mathcal{N}(A_2^{n_2}) \oplus \cdots \oplus \mathcal{N}(A_s^{n_s})$$

次に

$$\mathcal{R}(\widetilde{A}_1\widetilde{A}_2) = \mathcal{R}(\widetilde{A}_1) \cap \mathcal{R}(\widetilde{A}_2)$$

を示そう．$\mathcal{R}(\widetilde{A}_1\widetilde{A}_2) \subseteq \mathcal{R}(\widetilde{A}_1) \cap \mathcal{R}(\widetilde{A}_2)$ は明らかであるから $\supseteq$ を示せばよい．

$\boldsymbol{x} \in \mathcal{R}(\widetilde{A}_1) \cap \mathcal{R}(\widetilde{A}_2)$ とすれば適当な $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in \mathbb{F}^n$ が存在して

$$\boldsymbol{x} = \widetilde{A}_1\boldsymbol{x}_1 = \widetilde{A}_2\boldsymbol{x}_2$$

と書ける．すると (2.12) より

$$\begin{aligned}\boldsymbol{x} &= h_1(A)\widetilde{A}_1\boldsymbol{x} + h_2(A)\widetilde{A}_2\boldsymbol{x} = h_1(A)\widetilde{A}_1(\widetilde{A}_2\boldsymbol{x}_2) + h_2(A)\widetilde{A}_2(\widetilde{A}_1\boldsymbol{x}_1) \\ &= \widetilde{A}_1\widetilde{A}_2\big(h_1(A)\boldsymbol{x}_2 + h_2(A)\boldsymbol{x}_1\big) \quad \in \mathcal{R}(\widetilde{A}_1\widetilde{A}_2)\end{aligned}$$

$$\therefore\ \mathcal{R}(\widetilde{A}_1) \cap \mathcal{R}(\widetilde{A}_2) \subseteq \mathcal{R}(\widetilde{A}_1\widetilde{A}_2) \subseteq \mathcal{R}(\widetilde{A}_1) \cap \mathcal{R}(\widetilde{A}_2) \quad \therefore\ \mathcal{R}(\widetilde{A}_1\widetilde{A}_2) = \mathcal{R}(\widetilde{A}_1) \cap \mathcal{R}(\widetilde{A}_2)$$

以下これをくり返せば

$$\begin{aligned}\mathcal{R}(A_1^{n_1}A_2^{n_2} \cdots A_s^{n_s}) &= \mathcal{R}(A_1^{n_1}) \cap \mathcal{R}(A_2^{n_2} \cdots A_s^{n_s}) \\ &= \mathcal{R}(A_1^{n_1}) \cap \{\mathcal{R}(A_2^{n_2}) \cap \mathcal{R}(A_3^{n_3} \cdots A_s^{n_s})\} \\ &= \cdots \\ &= \mathcal{R}(A_1^{n_1}) \cap \mathcal{R}(A_2^{n_2}) \cap \cdots \cap \mathcal{R}(A_s^{n_s})\end{aligned}$$

証明終 ∎

【(2.11) の証明】　A の最小多項式 $\varphi(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_s)^{m_s}$ に対しては $\varphi(A) = (A-\lambda_1 I)^{m_1} \cdots (A-\lambda_s I)^{m_s} = O$ であるから定理 2.6 の第 1 式より (2.11) が得られる．　証明終 ∎

注意 2　A が複素行列ならば (2.11) は

$$\mathbb{C}^n = \mathcal{N}(A_1^{m_1}) \oplus \cdots \oplus \mathcal{N}(A_s^{m_s})$$

であるが，A が実行列で固有値がすべて実数ならば

$$\mathbb{R}^n = \mathcal{N}(A_1^{m_1}) \oplus \cdots \oplus \mathcal{N}(A_s^{m_s})$$

となる．

注意 3　m_i は A_i のアセントに等しい．

ここで §2.1 の終わりに注意として述べておいた「相異なる固有値に対応する固有ベクトルの 1 次独立性」を定理として述べる．

定理 2.7 A の相異なる固有値を $\lambda_1,\ldots,\lambda_s$, $A_i = A-\lambda_i I$ のアセントを m_i, $\dim \mathcal{N}(A_i)=\sigma_i$ とする.

(i) 全空間 $\mathbb{F}^n$ はランク m_i の広義（一般）固有空間 $\mathcal{N}(A_i^{m_i})$ の直和である.

(ii) λ_i に対応する 1 次独立な固有ベクトルを $\boldsymbol{x}_1^{(i)},\ldots,\boldsymbol{x}_{\sigma_i}^{(i)}$ とするとき

$$\boldsymbol{x}_1^{(1)},\ldots,\boldsymbol{x}_{\sigma_1}^{(1)},\boldsymbol{x}_1^{(2)},\ldots,\boldsymbol{x}_{\sigma_2}^{(2)},\ldots,\boldsymbol{x}_1^{(s)},\ldots,\boldsymbol{x}_{\sigma_s}^{(s)}$$

は 1 次独立である.

【証明】 $\mathcal{N}(A_i)\subseteq\mathcal{N}(A_i^{m_i})$ であるから，定理 1.6 によって $\mathcal{N}(A_i)$ の基底 $\boldsymbol{x}_1^{(i)},\ldots,\boldsymbol{x}_{\sigma_i}^{(i)}$ を $\mathcal{N}(A_i^{m_i})$ の基底の一部となるようにすることができる．したがって (2.11) より定理 2.7 が得られる． 証明終 ∎

2.4 3 重対角行列

次の形の行列を **3 重対角行列**または **3 対角行列** (**tridiagonal matrix**) という.

$$A=\begin{pmatrix} b_1 & c_1 & & \\ a_1 & b_2 & c_2 & \\ & \ddots & \ddots & \ddots \\ & & a_{n-1} & b_n \end{pmatrix} \tag{2.13}$$

定理 2.8 (山本 [6]) 3 重対角行列 (2.13) において $a_i\neq 0\ \ \forall i$ とするとき，A が対角化可能である (すなわち適当な正則行列 V により $V^{-1}AV$ が対角行列とできる) 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.

【証明】 A を対角化可能と仮定する．k 個の勝手な実数または複素数 $\alpha_1,\ldots,\alpha_k$ をとり，I を n 次単位行列として

$$\widetilde{A}=(A-\alpha_1 I)\cdots(A-\alpha_k I)$$

をつくれば，簡単な計算により，$k<n$ ならば $\widetilde{A}$ の $(k+1,1)$ 要素は $a_1a_2\cdots a_k\neq 0$. したがって $\widetilde{A}\neq O$ である．よって A の最小多項式の次数は $n-1$ 次以下ではありえずちょうど n 次でなければならない．したがって A の固有値は相異なる.

逆に A の固有値 $\lambda_1,\ldots,\lambda_n$ が相異なれば対応する n 個の固有 (列) ベクトルを並べてつくった行列 V は正則 (定理 2.7) で，$AV=V\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ をみたす．したがって $V^{-1}AV$ は対角行列で A は V により対角化される． 証明終 ∎

系 2.8.1 (2.13) において A は実対称行列 ($a_i=c_i\ \forall i$) とする．$a_i\neq 0\ \ \forall i$ ならば A の固有値はすべて相異なる実数である.

【証明】 実対称行列は直交行列により対角化可能でありその固有値は実数であるから． 証明終 ∎

系 2.8.2 $a_ic_i>0\ \ \forall i$ ならば (2.13) の行列 A の固有値はすべて相異なる実数である.

【証明】

$$d_i = \begin{cases} 1 & (i=1) \\ \displaystyle\prod_{k=1}^{i-1} \sqrt{\frac{a_k}{c_k}} & (i>1) \end{cases}, \quad D = \mathrm{diag}(d_1, \ldots, d_n)$$

とおけば $D^{-1}AD$ は実対称 3 重対角行列となる (各自検証されたい). よって系 2.8.1 を適用すればよい.

証明終 ∎

注意 (2.13) において $a_i = a$, $b_i = b$, $c_i = c$ かつ $ac > 0$ ならば, A の固有値 λ_j と固有ベクトル $\boldsymbol{x}^{(j)}$ は

$$\lambda_j = b + 2\sqrt{ac}\cos\frac{j\pi}{n+1},$$

$$\boldsymbol{x}^{(j)} = \left[\sin\frac{j\pi}{n+1},\ \sqrt{\frac{a}{c}}\sin\frac{2j\pi}{n+1},\ \ldots,\ \sqrt{\frac{a}{c}}^{n-1}\sin\frac{nj\pi}{n+1}\right]^{\mathrm{t}}, \quad 1 \le j \le n$$

で与えられる. (証明は山本 [6] 参照).

2.5 特異値

A を n 次行列 ($A \in \mathbb{F}^{n\times n}$) とするとき, (複素) Hermite 行列 A^*A (A が実行列ならば実対称行列 $A^{\mathrm{t}}A$) の固有値はすべて非負である. 実際 A^*A の固有値 μ と固有ベクトル $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} \neq \boldsymbol{0}$ をとれば $A^*A\boldsymbol{x} = \mu\boldsymbol{x}$ である. 両辺に左から $\boldsymbol{x}^*$ をかけて $\boldsymbol{x}^*A^*A\boldsymbol{x} = \mu\boldsymbol{x}^*\boldsymbol{x}$ $\therefore$ $(A\boldsymbol{x}, A\boldsymbol{x}) = \mu(\boldsymbol{x}, \boldsymbol{x})$ (ただし $(,)$ は複素内積である). $(\boldsymbol{x}, \boldsymbol{x}) = \|\boldsymbol{x}\|_2^2$ $(= \sum_{i=1}^n |x_i|^2) > 0$ であるから

$$\mu = \frac{\|A\boldsymbol{x}\|_2^2}{\|\boldsymbol{x}\|_2^2} \ge 0$$

$\sqrt{\mu}$ を A の**特異値 (singular value)** という. 通常 A^*A の固有値を $\alpha_1^2, \ldots, \alpha_n^2$ ($\alpha_i \ge 0$ $\forall i$) とあらわすので, A の特異値は $\alpha_1, \ldots, \alpha_n$ であらわされる.

すると (2.3) によって

$$\mathrm{tr}(A^*A) = \alpha_1^2 + \cdots + \alpha_n^2 = \sum_{i,j=1}^n |a_{ij}|^2 = \mathrm{tr}(AA^*)$$

である (A^*A の (i,i) 要素は $\sum_{j=1}^n |a_{ij}|^2$ であることに注意).

定理 2.3 から次の美しい定理が得られる.

定理 2.9 (Schur) $A = (a_{ij}) \in \mathbb{F}^{n\times n}$ の固有値を $\lambda_1, \ldots, \lambda_n$, 特異値を $\alpha_1, \ldots, \alpha_n$ とすれば

$$\sum_{i=1}^n |\lambda_i|^2 \le \sum_{i=1}^n {\alpha_i}^2 = \sum_{i,j=1}^n |a_{ij}|^2$$

【証明】 定理 2.3 により適当なユニタリ行列 U をえらんで

$$U^*AU = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix} (= S = (s_{ij}) \text{ とおく})$$

とできる.

$$S \cdot S^* = (U^*AU)(U^*AU)^* = U^*AA^*U$$

$$\therefore\ \mathrm{tr}(SS^*) = \mathrm{tr}(U^*AA^*U) = \mathrm{tr}(AA^*) = \sum_{i,j=1}^{n} |a_{ij}|^2$$

左辺は

$$\sum_{i,j=1}^{n} |s_{ij}|^2 = \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i<j} |s_{ij}|^2$$

に等しいから

$$\sum_{i=1}^{n} |\lambda_i|^2 \leq \sum_{i,j=1}^{n} |a_{ij}|^2$$

等号が成り立つのは $\sum_{i<j} |s_{ij}|^2 = 0$ のとき，すなわち S が対角行列のときに限る．系 2.3.2 によって，これは A が正規行列であることにほかならない． 証明終 ∎

注意 定理 2.9 の一般化として次が知られている．

(i) 任意の非負実数 $p \geq 0$ に対して

$$\sum_{i=1}^{n} |\lambda_i|^p \leq \sum_{i=1}^{n} \alpha_i{}^p \quad \text{(Weyl [40])}$$

(ii) $1 \leq p \leq 2$ ならば

$$\sum_{i=1}^{n} \alpha_i{}^p \leq \sum_{i,j=1}^{n} |a_{ij}|^p \quad \text{(Petri-Ikramov [35])}\ (p = 2 \text{ のときは等号である．})$$

2.6 Hermite 行列

A を n 次の (複素) Hermite 行列とすれば任意の列ベクトル $\boldsymbol{x} \in \mathbb{C}^n$ に対し

$$(A\boldsymbol{x}, \boldsymbol{x}) = (\boldsymbol{x}, A^*\boldsymbol{x}) = (\boldsymbol{x}, A\boldsymbol{x}) = \overline{(A\boldsymbol{x}, \boldsymbol{x})}$$

よって $(A\boldsymbol{x}, \boldsymbol{x})$ は実数である．この逆が成り立つことを次に示す．

定理 2.10 n 次複素行列 A に対し次の条件は同値である．

(i) A は Hermite 行列である．
(ii) 任意の $\boldsymbol{x} \in \mathbb{C}^n$ に対し $(A\boldsymbol{x}, \boldsymbol{x})$ は実数である．

【証明】 (i)⇒(ii) は上に示したから (ii)⇒(i) を示せばよい．
任意の $\boldsymbol{x} \in \mathbb{C}$ に対し $(A\boldsymbol{x}, \boldsymbol{x})$ が実数ならば複素内積の定義によって

$$(A\boldsymbol{x}, \boldsymbol{x}) = \overline{(A\boldsymbol{x}, \boldsymbol{x})} = \overline{(\boldsymbol{x}, A^*\boldsymbol{x})} = (A^*\boldsymbol{x}, \boldsymbol{x})$$

$$\therefore \qquad \big((A - A^*)\boldsymbol{x}, \boldsymbol{x}\big) = \boldsymbol{0} \quad \forall \boldsymbol{x} \in \mathbb{C}^n$$

よって $B = A - A^*$ とおくとき

$$(B\boldsymbol{x}, \boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \mathbb{C}^n \Rightarrow B = O \tag{2.14}$$

を示せばよい．以下 $B = (b_{ij})$ とおき $\boldsymbol{e}_i$ を n 次単位行列の第 i 列 (第 i 基本単位ベクトル) とすれば，

$\boldsymbol{x}=\boldsymbol{e}_i$ のとき $(B\boldsymbol{x},\boldsymbol{x})=(B\boldsymbol{e}_i,\boldsymbol{e}_i)=b_{ii}$ であるから

$$b_{ii}=0 \quad \forall i \tag{2.15}$$

次に $\boldsymbol{x}=\boldsymbol{e}_i+\boldsymbol{e}_j$ $(i\neq j)$ とおけば

$$\begin{aligned}(B\boldsymbol{x},\boldsymbol{x}) &= \big(B(\boldsymbol{e}_i+\boldsymbol{e}_j),\boldsymbol{e}_i+\boldsymbol{e}_j\big)=(B\boldsymbol{e}_i,\boldsymbol{e}_i)+(B\boldsymbol{e}_i,\boldsymbol{e}_j)+(B\boldsymbol{e}_j,\boldsymbol{e}_i)+(B\boldsymbol{e}_j,\boldsymbol{e}_j)\\ &= b_{ii}+b_{ij}+b_{ji}+b_{jj}=b_{ij}+b_{ji} \quad ((2.15)\text{ による})\\ \therefore\ & b_{ij}+b_{ji}=0 \quad (i\neq j)\end{aligned} \tag{2.16}$$

さらに $\boldsymbol{x}=\boldsymbol{e}_i\sqrt{-1}+\boldsymbol{e}_j$ $(i\neq j)$ とすれば

$$\begin{aligned}(B\boldsymbol{x},\boldsymbol{x}) &= \big(B(\boldsymbol{e}_i\sqrt{-1}+\boldsymbol{e}_j),\boldsymbol{e}_i\sqrt{-1}+\boldsymbol{e}_j\big)\\ &= (B\boldsymbol{e}_i\sqrt{-1},\boldsymbol{e}_i\sqrt{-1})+(B\boldsymbol{e}_i\sqrt{-1},\boldsymbol{e}_j)+(B\boldsymbol{e}_j,\boldsymbol{e}_i\sqrt{-1})+(B\boldsymbol{e}_j,\boldsymbol{e}_j)\\ &= b_{ii}+b_{ij}\sqrt{-1}-b_{ji}\sqrt{-1}+b_{jj}\\ &= (b_{ij}-b_{ji})\sqrt{-1} \quad \forall i,j\ (i\neq j) \quad ((2.15)\text{ による})\\ \therefore\ & b_{ij}-b_{ji}=0 \quad \forall i,j\ (i\neq j)\end{aligned} \tag{2.17}$$

(2.16) + (2.17) により $b_{ij}=0 \quad \forall i,j\ (i\neq j)$

$$\therefore \quad B=O \quad \text{すなわち} \quad A=A^* \qquad \text{証明終} \ \blacksquare$$

▒**注意**　A を実行列とするとき，定理 2.10 において条件 (i) と (ii) をそれぞれ「A は対称行列である」「任意の $\boldsymbol{x}\in\mathbb{R}^n$ に対し $(A\boldsymbol{x},\boldsymbol{x})$ は実数である」でおきかえることはできない．実際

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

は $(A\boldsymbol{x},\boldsymbol{x})=0 \quad \forall \boldsymbol{x}=(x_1,x_2)^{\mathrm{t}}\in\mathbb{R}^2$ をみたすが A は実対称でない．

2.7　正定値行列と半正定値行列

A を n 次複素行列とするとき，$\boldsymbol{x}$ を複素列ベクトルとして

$$(A\boldsymbol{x},\boldsymbol{x})=\boldsymbol{x}^*A\boldsymbol{x}>0 \quad \forall \boldsymbol{x}(\neq\boldsymbol{0})\in\mathbb{C}^n \tag{2.18}$$

が成り立つとき A は**正定値行列 (positive definite matrix)** と呼ばれる．また

$$(A\boldsymbol{x},\boldsymbol{x})=\boldsymbol{x}^*A\boldsymbol{x}\geq 0 \quad \forall \boldsymbol{x}\in\mathbb{C}^n \tag{2.19}$$

のとき**半正定値行列 (positive semi-definite matrix)** という．

(2.18) から "$A\boldsymbol{x}=\boldsymbol{0}\Rightarrow\boldsymbol{x}=\boldsymbol{0}$" が従うから，複素正定値行列は正則である．また定理 2.10 により複素正定値行列および複素半正定値行列は必然的に Hermite 行列である．しかし (複素) 正定値 Hermite 行列，(複素) 半正定値 Hermite 行列等と念を入れた表現がなされることも多い．

A が n 次実行列の場合には

$$(A\boldsymbol{x},\boldsymbol{x})=\boldsymbol{x}^{\mathrm{t}}A\boldsymbol{x}>0 \quad \forall \boldsymbol{x}(\neq\boldsymbol{0})\in\mathbb{R}^n \tag{2.20}$$

のとき A を正定値 (実) 行列といい，

$$(A\boldsymbol{x},\boldsymbol{x})=\boldsymbol{x}^{\mathrm{t}}A\boldsymbol{x}\geq 0 \quad \forall \boldsymbol{x}\in\mathbb{R}^n \tag{2.21}$$

のとき半正定値(実)行列という．明らかに正定値実行列は正則である．しかし(2.20)または(2.21)から $A^{\mathrm{t}} = A$ は従わない(§2.6 注意参照)．

次が成り立つ．

定理 2.11 A を n 次実行列とするとき

(i) A が正定値行列 $\Leftrightarrow$ 実対称行列 $A + A^{\mathrm{t}}$ が正定値

(ii) A が半正定値行列 $\Leftrightarrow$ 実対称行列 $A + A^{\mathrm{t}}$ が半正定値

【証明】 (i) A を n 次正定値実行列とするとき

$$
\begin{aligned}
&(A\boldsymbol{x}, \boldsymbol{x}) > 0 \quad \forall \boldsymbol{x} \neq \boldsymbol{0} \\
\therefore\ &(A^{\mathrm{t}}\boldsymbol{x}, \boldsymbol{x}) = (\boldsymbol{x}, A\boldsymbol{x}) = (A\boldsymbol{x}, \boldsymbol{x}) \\
\therefore\ &\big((A + A^{\mathrm{t}})\boldsymbol{x}, \boldsymbol{x}\big) = 2(A\boldsymbol{x}, \boldsymbol{x}) > 0 \quad \forall \boldsymbol{x}(\neq 0) \in \mathbb{R}^n
\end{aligned}
$$

よって $A + A^{\mathrm{t}}$ は正定値である．

逆に $A + A^{\mathrm{t}}$ が正定値行列ならば

$$
\begin{aligned}
&2(A\boldsymbol{x}, \boldsymbol{x}) = \big((A + A^{\mathrm{t}})\boldsymbol{x}, \boldsymbol{x}\big) > 0 \quad \forall \boldsymbol{x}(\neq 0) \in \mathbb{R}^n \\
\therefore\ &(A\boldsymbol{x}, \boldsymbol{x}) > 0 \quad \forall \boldsymbol{x}(\neq 0) \in \mathbb{R}^n
\end{aligned}
$$

よって A は正定値実行列である．

(ii) 同様である．　　証明終 ∎

さらに

定理 2.12 A を Hermite 行列とする．次の条件は同値である．

(i) A は正定値(半正定値)である．

(ii) A の固有値は全て正(非負)である．

【証明】 A を Hermite 行列とすれば適当なユニタリ行列により

$$
U^* A U = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \Lambda \quad (\lambda_1, \ldots, \lambda_n \text{は } A \text{ の固有値})
$$

と対角化できる。(A が実対称ならば U は直交行列でおきかえられるのであった．)

ここで $\boldsymbol{x} \in \mathbb{C}^n$, $\boldsymbol{y} = U^*\boldsymbol{x} = (y_1, \ldots, y_n)^{\mathrm{t}}$ とおけば

$$
(A\boldsymbol{x},\ \boldsymbol{x}) = (U\Lambda U^*\boldsymbol{x},\ \boldsymbol{x}) = (\Lambda\boldsymbol{y},\ \boldsymbol{y}) = \sum_{i=1}^{n} \lambda_i |y_i|^2
$$

$\boldsymbol{x} \neq \boldsymbol{0} \Leftrightarrow \boldsymbol{y} \neq \boldsymbol{0}$ であるから上式より

$$
\begin{aligned}
A \text{ が正定値} \quad &\Leftrightarrow\ (A\boldsymbol{x}, \boldsymbol{x}) > 0 && \forall \boldsymbol{x} \neq \boldsymbol{0} \\
&\Leftrightarrow\ \sum_{i=1}^{n} \lambda_i |y_i|^2 > 0 && \forall \boldsymbol{y} \neq \boldsymbol{0} \\
&\Leftrightarrow\ \lambda_i > 0 && \forall i
\end{aligned}
$$

同様に

$$\begin{aligned} A\text{が半正定値} &\Leftrightarrow (A\boldsymbol{x},\boldsymbol{x}) \geq 0 \quad \forall \boldsymbol{x} \\ &\Leftrightarrow \sum_{i=1}^{n} \lambda_i |y_i|^2 \geq 0 \quad \forall \boldsymbol{y} \\ &\Leftrightarrow \lambda_i \geq 0 \quad \forall i \end{aligned}$$

証明終 ∎

系 2.12.1 A が正定値 Hermite 行列ならば $\det A > 0$

【証明】 $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$ (§2.1 参照) であるから. 証明終 ∎

系 2.12.2 任意の n 次行列 $A \in \mathbb{C}^{n\times n}$ に対し A^*A は半正定値行列であり，特に A が正則ならば正定値行列である.

【証明】 §2.5 において A^*A の固有値は全て非負であることをみた. よって Hermite 行列 A^*A は半正定値である. 特に A が正則ならば定理 1.18 によって

$$\mathrm{rank}(A^*A) = \mathrm{rank}A = n$$

よって A^*A はゼロ固有値をもたない. (あるいは，A が正則ならば A^* も正則. したがって A^*A も正則でゼロ固有値をもたないといってもよい.) 証明終 ∎

定義 2.1 ($A \succ B$ と $A \succcurlyeq B$)
A, B を同じ次数の Hermite 行列で $A - B$ が正定値行列のとき $A \succ B$ とかく. また $A - B$ が半正定値行列のとき $A \succcurlyeq B$ とかく. したがって A が正定値ならば $A \succ O$, また半正定値ならば $A \succcurlyeq O$ とかく ($B = O$ の場合である).

定理 2.13 A, B を n 次 Hermite 行列とするとき次が成り立つ.

(i) $A \succcurlyeq O$ かつ $B \succcurlyeq O$ ならば $A + B \succcurlyeq O$
(ii) $A \succ O$ かつ $B \succcurlyeq O$ ならば $A + B \succ O$
(iii) $AB = BA$, $A \succ O$ かつ $B \succcurlyeq O$ ならば $AB \succcurlyeq O$

【証明】 (i) と (ii) は

$$\big((A+B)\boldsymbol{x},\ \boldsymbol{x}\big) = (A\boldsymbol{x},\ \boldsymbol{x}) + (B\boldsymbol{x},\ \boldsymbol{x})$$

と仮定 $A \succcurlyeq O$, $B \succcurlyeq O$ または $A \succ O$, $B \succcurlyeq O$ より従う.
(iii) $AB = BA$ であるから $(AB)^* = B^*A^* = BA = AB$ 故に AB は Hermite 行列である. また $A \succ O$ より $A^{-1} \succ O$ ($\because$ 定理 2.12 を Hermite 行列 A^{-1} に適用する).

よって $AB\boldsymbol{x} = \lambda \boldsymbol{x}$, $\boldsymbol{x} \neq \boldsymbol{0}$ ならば $B\boldsymbol{x} = \lambda A^{-1}\boldsymbol{x}$ $\therefore$ $(B\boldsymbol{x},\boldsymbol{x}) = \lambda(A^{-1}\boldsymbol{x},\boldsymbol{x})$ かつ $(A^{-1}\boldsymbol{x},\boldsymbol{x}) \succ O$ したがって $B \succcurlyeq O$ のとき

$$\lambda = \frac{(B\boldsymbol{x},\boldsymbol{x})}{(A^{-1}\boldsymbol{x},\boldsymbol{x})} \geq 0$$

であるから，定理 2.12 によって $AB \succcurlyeq O$ を得る. 証明終 ∎

注意 (iii) において $A \succ O$ は $A \succcurlyeq O$ でおきかえることができる. (§5.2 系 5.5.1 参照)

ここで §1.6 で述べた Gram 行列の性質を示しておく.

定理 2.14 (Gram 行列の性質) $(X,(,))$ を $\mathbb{F}$ 上の内積空間とし, $u_1,\dots,u_r \in X$ のつくる Gram 行列を G とする. すなわち

$$G=\begin{pmatrix} (u_1,u_1) & \cdots & (u_1,u_r) \\ \vdots & & \vdots \\ (u_r,u_1) & \cdots & (u_r,u_r) \end{pmatrix}$$

このとき

(i) $G \succcurlyeq O$

(ii) $u_1,\dots,u_r$ が 1 次独立のとき, かつそのときに限って $G \succ O$

(iii) したがって G が正則となるのは $u_1,\dots,u_r$ が 1 次独立のとき, かつそのときに限る.

【証明】 G は $\mathbb{F}=\mathbb{R}$ のとき実対称で, $\mathbb{F}=\mathbb{C}$ のとき複素 Hermite 行列である.
さて, $\boldsymbol{c}=(c_1,\dots,c_r)^{\mathrm{t}} \in \mathbb{F}^r$ に対して

$$\begin{aligned}(G\boldsymbol{c},\boldsymbol{c}) &= \begin{cases} \displaystyle\sum_{i=1}^{r} c_i\Big(\sum_{j=1}^{r}(u_i,u_j)c_j\Big) & (\mathbb{F}=\mathbb{R}\text{ のとき}) \\ \displaystyle\sum_{i=1}^{r} \bar{c}_i\Big(\sum_{j=1}^{r}(u_i,u_j)c_j\Big) & (\mathbb{F}=\mathbb{C}\text{ のとき}) \end{cases} \\ &= \begin{cases} \displaystyle\Big(\sum_{i=1}^{r} c_iu_i,\ \sum_{j=1}^{r} c_ju_j\Big) & (\mathbb{F}=\mathbb{R}\text{ のとき}) \\ \displaystyle\Big(\sum_{i=1}^{r} \bar{c}_iu_i,\ \sum_{j=1}^{r} \bar{c}_ju_j\Big) & (\mathbb{F}=\mathbb{C}\text{ のとき}) \end{cases} \\ &= (u,u) \end{aligned} \tag{2.22}$$

$$\text{ただし}\quad u=\begin{cases} c_1u_1+\cdots+c_ru_r & (\mathbb{F}=\mathbb{R}\text{ のとき}) \\ \bar{c}_1u_1+\cdots+\bar{c}_ru_r & (\mathbb{F}=\mathbb{C}\text{ のとき}) \end{cases}$$

よって

(i) $(G\boldsymbol{c},\boldsymbol{c}) \geq 0 \quad \forall \boldsymbol{c} \in \mathbb{F}^r$ したがって $G \succcurlyeq O$

(ii) 特に $u_1,\dots u_r$ が 1 次独立で $\boldsymbol{c} \neq \boldsymbol{0}$ ならば $u=\sum_{i=1}^{r} c_iu_i \neq 0$. よって (2.22) より $(G\boldsymbol{c},\boldsymbol{c})>0 \quad \forall \boldsymbol{c} \neq \boldsymbol{0} \quad \therefore\ G \succ O$. 逆に $G \succ O$ ならば, 任意の $\boldsymbol{c}=(c_1,\dots,c_r) \neq \boldsymbol{0}$ に対して (2.22) より $u \neq O$

これは $u_1,\dots,u_r$ が 1 次独立であることを意味している.

(iii) は (i) と (ii) より明らかである. 証明終 ∎

2.8 2 次形式

$A=(a_{ij})$ を n 次実対称行列とし, n 個の変数 $x_1,\dots,x_n$ に関する実係数の斉次 2 次式

$$\boldsymbol{x}^{\mathrm{t}}A\boldsymbol{x}=(A\boldsymbol{x},\boldsymbol{x})=\sum_{i,j=1}^{n} a_{ij}x_ix_j \tag{2.23}$$

を **2 次形式 (quadratic form)** という.

定理 2.15 (Sylvester (シルベスター) の慣性則第 1 型) 2 次形式 (2.23) に適当な正則変換 $\boldsymbol{x} = P\boldsymbol{y}$ (P：実正則行列，$\boldsymbol{y} = (y_1, \ldots, y_n)^{\mathrm{t}}$) を施せば

$$(A\boldsymbol{x}, \boldsymbol{x}) = \sum_{i=1}^{s} {y_i}^2 - \sum_{i=1}^{t} {y_{s+i}}^2 \tag{2.24}$$

となる．ここに s と t は A により一意的に定まる．$(s,\ t)$ を A の**符号数 (signature)** という．また $(s,\ t,\ n-s-t)$ を A の**慣性数 (inertia)** といい $\mathrm{In}(A)$ であらわす．この記号はすべての固有値が実数である一般な行列 A に対しても適用される．

【証明】 A は実対称であるから適当な直交行列 T により

$$T^{\mathrm{t}}AT = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad (\lambda_1, \cdots, \lambda_n \text{は } A \text{ の固有値})$$

とできる．一般性を失うことなく

$$\lambda_1 > 0, \cdots, \lambda_s > 0, \lambda_{s+1} < 0, \cdots, \lambda_{s+t} < 0, \lambda_{s+t+1} = \cdots = \lambda_n = 0$$

としてよい．このとき

$$P = T \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & & & & & & & \\ & \ddots & & & & & & \\ & & \frac{1}{\sqrt{\lambda_s}} & & & & & \\ & & & \frac{1}{\sqrt{-\lambda_{s+1}}} & & & & \\ & & & & \ddots & & & \\ & & & & & \frac{1}{\sqrt{-\lambda_{s+t}}} & & \\ & & & & & & 1 & \\ & & & & & & & \ddots \\ & & & & & & & & 1 \end{pmatrix}$$

とおけば P は正則で

$$P^{\mathrm{t}}AP = \begin{pmatrix} I_s & & & \\ & -I_t & & \\ & & 0 & \\ & & & \ddots \\ & & & & 0 \end{pmatrix} \quad (I_s, I_t \text{はそれぞれ } s \text{ 次, } t \text{ 次の単位行列})$$

$$(A\boldsymbol{x},\ \boldsymbol{x}) = (AP\boldsymbol{y},\ P\boldsymbol{y}) = (P^{\mathrm{t}}AP\boldsymbol{y},\ \boldsymbol{y}) = \sum_{i=1}^{s} y_i^2 - \sum_{i=1}^{t} y_{s+i}^2$$

s と t は A により一意的に定まる．実際 P と異なる正則変換 $\boldsymbol{x} = Q\boldsymbol{z}$ (Q：実正則行列) によって

$$(A\boldsymbol{x}, \boldsymbol{x}) = \sum_{i=1}^{\sigma} z_i^2 - \sum_{i=1}^{\tau} z_{\sigma+i}^2$$

すなわち

$$Q^{\mathrm{t}}AQ=\begin{pmatrix} I_\sigma & & \\ & -I_\tau & \\ & & O \end{pmatrix}$$

とすれば

$$s+t=\mathrm{rank}A=\sigma+\tau \tag{2.25}$$

仮に $s>\sigma$ として $P^{-1}=(\hat{p}_{ij})$, $Q^{-1}=(\hat{q}_{ij})$, $m=\sigma+(n-s)$ とおくとき n 元 m 立 1 次方程式

$$\begin{cases} \displaystyle\sum_{j=1}^{n}\hat{p}_{ij}x_j=0 \quad (s+1\leq i\leq n) \\ \displaystyle\sum_{j=1}^{n}\hat{q}_{ij}x_j=0 \quad (1\leq i\leq \sigma) \end{cases}$$

は非自明解 $\boldsymbol{x}^{(0)}$ をもつ.

$$\left(\begin{array}{l} \because \text{(係数行列のランク)}\leq m \text{ で解空間の次元は少なくとも} \\ n-m=n-(\sigma+(n-s))=s-\sigma>0 \text{ (系 1.16.3)} \end{array}\right)$$

いま

$$\hat{\boldsymbol{y}}=P^{-1}\boldsymbol{x}^{(0)},\quad \hat{\boldsymbol{z}}=Q^{-1}\boldsymbol{x}^{(0)}$$

とおけば $\hat{\boldsymbol{y}}\neq\boldsymbol{0}$, $\hat{\boldsymbol{z}}\neq\boldsymbol{0}$ かつ

$$\hat{\boldsymbol{y}}=\begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_s \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ \hat{\boldsymbol{z}}=\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \hat{z}_{\sigma+1} \\ \vdots \\ \hat{z}_n \end{pmatrix}$$

したがって

$$(A\boldsymbol{x}^{(0)},\ \boldsymbol{x}^{(0)})=(AP\hat{\boldsymbol{y}},\ P\hat{\boldsymbol{y}})=(P^{\mathrm{t}}AP\hat{\boldsymbol{y}},\ \hat{\boldsymbol{y}})=\sum_{i=1}^{s}\hat{y}_i^2$$

かつ

$$(A\boldsymbol{x}^{(0)},\ \boldsymbol{x}^{(0)})=(AQ\hat{\boldsymbol{z}},\ Q\hat{\boldsymbol{z}})=(Q^{\mathrm{t}}AQ\hat{\boldsymbol{z}},\ \hat{\boldsymbol{z}})=-\sum_{i=1}^{\tau}\hat{z}_{\sigma+i}^2$$

$$\therefore\ \sum_{i=1}^{s}\hat{y}_i^2=-\sum_{i=1}^{\tau}\hat{z}_{\sigma+i}^2\quad\therefore\ \sum_{i=1}^{s}\hat{y}_i^2+\sum_{i=1}^{\tau}\hat{z}_{\sigma+i}^2=0$$

よって $\hat{y}_1=\cdots=\hat{y}_s=0$ となり $\hat{\boldsymbol{y}}=\boldsymbol{0}$ これは $\hat{\boldsymbol{y}}\neq\boldsymbol{0}$ に矛盾する. $\quad\therefore\ s\leq\sigma$

同様に $s<\sigma$ としても矛盾を生じる. (今度は連立 1 次方程式

$$\begin{cases} \displaystyle\sum_{j=1}^{n}\hat{q}_{ij}x_j=0 \quad (\sigma+1\leq i\leq n) \\ \displaystyle\sum_{j=1}^{n}\hat{p}_{ij}x_j=0 \quad (1\leq i\leq s) \end{cases}$$

を考えればよい．このとき方程式の個数は $m=(n-\sigma)+s=n-(\sigma-s)<n$．よって解空間の次元は少なくとも $n-m=\sigma-s>0$ であり上記連立1次方程式は非自明解 $\boldsymbol{x}^{(0)}$ をもつ．以下同様．)

結局 $s=\sigma$ でなければならない．したがって (2.25) によって $t=\tau$ を得る． 証明終 ∎

系 2.15.1 (Sylvester の慣性則第 2 型) 任意の実正則行列 V に対し $\mathrm{In}(V^{\mathrm{t}}AV)=\mathrm{In}(A)$

【証明】 $\mathrm{In}(A)=(s,\ t,\ n-s-t)$, $\mathrm{In}(V^{\mathrm{t}}AV)=(\sigma,\ \tau,\ n-\sigma-\tau)$ とおく．

定理 2.15 によって，適当な正則変換 $\boldsymbol{u}=Q\boldsymbol{z}$ により

$$(V^{\mathrm{t}}AV\boldsymbol{u},\ \boldsymbol{u})=\sum_{i=1}^{\sigma}z_i^2-\sum_{i=1}^{\tau}z_{\sigma+i}^2$$

$$\therefore\ (V^{\mathrm{t}}AVQ\boldsymbol{z},\ Q\boldsymbol{z})=\sum_{i=1}^{\sigma}z_i^2-\sum_{i=1}^{\tau}z_{\sigma+i}^2$$

$$\boldsymbol{x}=VQ\boldsymbol{z}\quad\text{とおけば}\quad(A\boldsymbol{x},\ \boldsymbol{x})=\sum_{i=1}^{\sigma}z_i^2-\sum_{i=1}^{\tau}z_{\sigma+i}^2$$

よって再び定理 2.15 により $\sigma=s,\ \tau=t\quad\therefore\ \mathrm{In}(V^{\mathrm{t}}AV)=\mathrm{In}(A)$ 証明終 ∎

注意 1 A が (複素) Hermite 行列の場合には複素内積を用い (2.23) の代わりに

$$(A\boldsymbol{x},\ \boldsymbol{x})=\boldsymbol{x}^*A\boldsymbol{x}=\sum_{i,j=1}^{n}a_{ij}\bar{x}_ix_j$$

を考え，これを **Hermite 形式**または **Hermite 2 次形式**という．

この場合にも定理 2.15 は (2.24) を

$$(A\boldsymbol{x},\ \boldsymbol{x})=\sum_{i=1}^{s}|y_i|^2-\sum_{i=1}^{t}|y_{s+i}|^2$$

として成り立つ．また任意の正則行列 W に対し $\mathrm{In}(W^*AW)=\mathrm{In}(A)$ である．

注意 2 定理 2.15 の証明より

$$\mathrm{In}(A)=(s,\ t,\ n-s-t)\Leftrightarrow A\text{ の正の固有値が }s\text{ 個で負の固有値が }t\text{ 個}$$

を意味する．

第3章　行列の分解

この章では，LU 分解，QR 分解，特異値分解，極分解，Jordan 分解等与えられた行列を考察しやすい行列の積としてあらわすいくつかの結果を述べる．これらは連立 1 次方程式や行列の固有値問題を取り扱う場合の基礎となる．

3.1　LU 分解

$$L=\begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix},\quad U=\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix} \tag{3.1}$$

の形の行列 L と U をそれぞれ**下 3 角行列 (lower triangular matrix)**，**上 3 角行列 (upper triangular matrix)** という．

与えられた正則行列 $A\in\mathbb{F}^{n\times n}$ を下 3 角行列 L と上 3 角行列 U の積としてあらわすことを A の **LU 分解 (LU decomposition)** という．このとき A の正則性から $l_{ii}\neq 0,\ u_{ii}\neq 0\quad\forall i$ である．特に $l_{ii}=1\quad\forall i$ のとき L を**単位下 3 角行列**，$u_{ii}=1\quad\forall i$ のとき U を**単位上 3 角行列**という．

定理 3.1　n 次正則行列 $A=(a_{ij})$ が LU 分解可能であるための必要十分条件は

$$\det A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}=\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}\neq 0\quad(1\leq k\leq n) \tag{3.2}$$

である．特に L を単位下 3 角行列に限定すれば A の LU 分解は (もし可能ならば) 一意に定まる．

【証明】　(i) 必要条件：n 次正則行列 A が (3.1) の形の行列 L,U により

$$A=LU \tag{3.3}$$

と書けたとすれば，A の正則性より $\det A=\det L\det U\neq 0$．ここで

$$\det L=l_{11}\cdots l_{nn},\quad \det U=u_{11}\cdots u_{nn}$$

であるから $l_{ii}\neq 0,\ u_{ii}\neq 0\quad\forall i$.

さらに (3.3) の両辺の k 次首座小行列を比較すれば

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}=\begin{pmatrix} l_{11} & & \\ \vdots & \ddots & \\ l_{k1} & \cdots & l_{kk} \end{pmatrix}\begin{pmatrix} u_{11} & \cdots & u_{1k} \\ & \ddots & \vdots \\ & & u_{kk} \end{pmatrix}\quad(1\leq k\leq n)$$

両辺の行列式をとれば

$$\det A \begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix} = (l_{11} \cdots l_{kk})(u_{11} \cdots u_{kk}) \neq 0$$

である．

(ii) 十分条件：(3.2) を仮定し A の LU 分解が可能であることを示そう．証明は n に関する帰納法による．

$n=1$ のときは明らかであるから $n>1$ かつ $n-1$ のとき成り立つとすれば帰納法の仮定によって

$$\det A \begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix} \neq 0 \quad (1 \leq k \leq n-1) \tag{3.4}$$

ならば下 3 角行列と上 3 角行列

$$L_{n-1} = \begin{pmatrix} l_{11} & & \\ \vdots & \ddots & \\ l_{n-1\,1} & \cdots & l_{n-1\,n-1} \end{pmatrix}, \quad U_{n-1} = \begin{pmatrix} u_{11} & \cdots & u_{1\,n-1} \\ & \ddots & \vdots \\ & & u_{n-1\,n-1} \end{pmatrix}$$

を適当にえらんで

$$A \begin{pmatrix} 1 & 2 & \cdots & n-1 \\ 1 & 2 & \cdots & n-1 \end{pmatrix} = L_{n-1} U_{n-1}$$

と書ける．仮定により左辺は正則であるから L_{n-1} と U_{n-1} は正則で $l_{ii} \neq 0$, $u_{ii} \neq 0$ $(1 \leq i \leq n-1)$ である．

このとき適当な定数 $x_1, \ldots, x_{n-1}, x_n$ と $y_1, \ldots, y_{n-1}, y_n$ を定めて

$$A = \left(\begin{array}{c|c} L_{n-1} & \mathbf{0} \\ \hline x_1 \cdots x_{n-1} & x_n \end{array}\right) \left(\begin{array}{c|c} & y_1 \\ U_{n-1} & \vdots \\ & y_{n-1} \\ \hline 0 \cdots 0 & y_n \end{array}\right) \tag{3.5}$$

とできることを示せばよい．まず $1 \leq j \leq n-1$ のとき (3.5) の両辺の (n,j) 要素を比較して

$$a_{nj} = x_1 u_{1j} + x_2 u_{2j} + \cdots + x_{n-1} u_{n-1\,j}$$

ここで　$j=1$ として　$a_{n1} = x_1 u_{11}$　$\therefore\ x_1 = \dfrac{a_{n1}}{u_{11}}$

$j=2$ として　$a_{n2} = x_1 u_{12} + x_2 u_{22}$　$\therefore\ x_2 = \dfrac{1}{u_{22}}(a_{n2} - x_1 u_{12})$

以下これを $j=n-1$ まで続けて

$$a_{n\,n-1} = x_1 u_{1\,n-1} + \cdots + x_{n-2} u_{n-2\,n-1} + x_{n-1} u_{n-1\,n-1}$$

$$\therefore\quad x_{n-1} = \frac{1}{u_{n-1\,n-1}}(a_{n\,n-1} - x_1 u_{1\,n-1} - \cdots - x_{n-2} u_{n-2\,n-1})$$

次に $1 \leq i \leq n-1$ として (3.5) の (i,n) 要素を比較すれば

$$a_{in} = l_{i1} y_1 + l_{i2} y_2 + \cdots + l_{ii} y_i$$

$i=1$ として　$a_{1n} = l_{11} y_1$　$\therefore\ y_1 = \dfrac{a_{1n}}{l_{11}}$

$i=2$ として　$a_{2n} = l_{21} y_1 + l_{22} y_2$　$\therefore\ y_2 = \dfrac{1}{l_{22}}(a_{2n} - l_{21} y_1)$

以下これを $i=n-1$ までくり返して

$$a_{n-1\,n} = l_{n-1}y_1 + \cdots + l_{n-1\,n-1}y_{n-1}$$

$$\therefore \quad y_{n-1} = \frac{1}{l_{n-1\,n-1}}(a_{n-1\,n} - l_{n-1\,1}y_1 - \cdots - l_{n-1\,n-2}y_{n-1})$$

最後に (3.5) の両辺の行列式をとれば

$$\det A = (l_{11}\cdots l_{n-1\,n-1}x_n)(u_{11}\cdots u_{n-1\,n-1}y_n) \quad \therefore \quad x_n y_n = \frac{\det A}{\prod_{i=1}^{n-1}(l_{ii}u_{ii})} \neq 0$$

x_n, y_n のえらび方には自由度があるから x_n を任意に定めればそれに応じて上式により y_n が確定する．これで帰納法が完了し A の LU 分解が完成した．

特に L を単位下 3 角行列に限ればこのような L, U は 1 通りしかない．実際 $A = LU = \widetilde{L}\widetilde{U}$ を 2 通りの LU 分解とし L と $\widetilde{L}$ を単位下 3 角とすれば

$$\widetilde{L}^{-1}L = \widetilde{U}U^{-1}$$

左辺は単位下 3 角で右辺は上 3 角であるから両辺は対角成分が 1 の対角行列，すなわち単位行列でなければならない．

$$\therefore \ \widetilde{L}^{-1}L = \widetilde{U}U^{-1} = I_n \quad \therefore \quad L = \widetilde{L},\ U = \widetilde{U}$$

故に A の LU 分解はこのとき一意的である． 証明終 ∎

定理 3.1 によって，(3.2) をみたさない正則行列 A は LU 分解できない．例えば

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

は LU 分解できない (各自検証されたい)．しかしながら次に示すようにどのような正則行列も A の行を適当に並べかえれば LU 分解可能である．

定理 3.2 A を任意の正則行列とする．適当な置換行列 P をえらべば PA は LU 分解可能である．

【証明 (Rose1975)】 A を n 次正則行列とし，次数 n に関する帰納法による．$n = 1$ のときは明らかであるから $n > 1$ とし $n-1$ のとき成り立つとする．A は正則であるから $A = (a_{ij})$ の第 1 列には少なくとも 1 つの非零要素がある．それを $\alpha = a_{i1}$ とし第 i 行と第 1 行を入れかえる置換行列を P_1 とすると

$$P_1 A = \begin{pmatrix} \alpha & \boldsymbol{v} \\ \boldsymbol{u} & B \end{pmatrix},\ \boldsymbol{u} = \begin{pmatrix} a_{21} \\ \vdots \\ a_{i-1\,1} \\ a_{11} \\ a_{i+1\,1} \\ \vdots \\ a_n \end{pmatrix},\ \boldsymbol{v} = (a_{i2}, \ldots, a_{in})$$

と書ける．これを

$$P_1 A = \begin{pmatrix} 1 & \boldsymbol{0} \\ \dfrac{\boldsymbol{u}}{\alpha} & I_{n-1} \end{pmatrix} \begin{pmatrix} \alpha & \boldsymbol{v} \\ \boldsymbol{0} & B^{(2)} \end{pmatrix},\ B^{(2)} = B - \frac{1}{\alpha}\boldsymbol{u}\boldsymbol{v}$$

と書けば，左辺は正則であるから $\begin{pmatrix} \alpha & \boldsymbol{v} \\ \boldsymbol{0} & B^{(2)} \end{pmatrix}$ は正則．したがって $(n-1)$ 次行列 $B^{(2)}$ も正則でなければならない．故に帰納法の仮定により適当な置換行列 $P^{(2)}$ をえらんで

$$P^{(2)}B^{(2)} = L^{(2)}U^{(2)} \quad (L^{(2)} : (n-1) \text{ 次下 3 角},\ U^{(2)} : (n-1) \text{ 次上 3 角})$$

と分解できる．ここで

$$P_2 = \begin{pmatrix} 1 & \\ & P^{(2)} \end{pmatrix}$$

とおけば

$$P_2P_1A = \begin{pmatrix} 1 & 0 \\ P^{(2)}\dfrac{\boldsymbol{u}}{\alpha} & P^{(2)} \end{pmatrix}\begin{pmatrix} \alpha & \boldsymbol{v} \\ 0 & B^{(2)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ P^{(2)}\dfrac{\boldsymbol{u}}{\alpha} & I_{n-1} \end{pmatrix}\begin{pmatrix} \alpha & \boldsymbol{v} \\ 0 & P^{(2)}B^{(2)} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ \dfrac{1}{\alpha}P^{(2)}\boldsymbol{u} & I_{n-1} \end{pmatrix}\begin{pmatrix} \alpha & \boldsymbol{v} \\ 0 & L^{(2)}U^{(2)} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ \dfrac{1}{\alpha}P^{(2)}\boldsymbol{u} & L^{(2)} \end{pmatrix}}_{\text{下 3 角}}\underbrace{\begin{pmatrix} \alpha & \boldsymbol{v} \\ 0 & U^{(2)} \end{pmatrix}}_{\text{上 3 角}}$$

よって $P = P_1P_2$ が求める置換行列で PA の LU 分解がえられた．　　証明終 ∎

注意 1 (LDV 分解)　定理 3.1 の分解において

$$d_i = u_{ii},\ D = \mathrm{diag}(d_1, \ldots, d_n),\ V = D^{-1}U$$

とおけば

$$A = LDV \ (L : \text{下 3 角},\ D : \text{対角},\ V : \text{単位上 3 角})$$

と書ける．これを正則行列 A の **LDV 分解 (LDV decomposition)** という．

L を単位下 3 角行列に限定すればこのような分解は一意的である．なお V の代わりに U を用い上記分解を A の **LDU 分解**ということもある．その場合には L を単位下 3 角行列とするのが普通である．したがって LDU 分解では L, U はそれぞれ単位下 3 角，単位上 3 角とする．

なお，定理 3.1 を補う次の結果を記しておく．

定理 3.3　n 次行列 A に対し次の 2 条件は同値である．

(i)　$\det A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix} > 0 \quad (1 \le k \le n)$

(ii)　対角要素が全て正の下 3 角行列 L と対角要素がすべて正の上 3 角行列 U が存在して $A = LU$ と書ける．

【証明】 (i) $\Rightarrow$ (ii)　$A = (a_{ij})$ とおき n に関する帰納法による．$n = 1$ ならば明らかであるから，$n > 1$ かつ $n-1$ のとき成り立つとし

$$A = \left(\begin{array}{c|c} A_{n-1} & \boldsymbol{a} \\ \hline \boldsymbol{b} & a_{nn} \end{array}\right),\quad A_{n-1} = A\begin{pmatrix} 1 & 2 & \cdots & n-1 \\ 1 & 2 & \cdots & n-1 \end{pmatrix}$$

とおく．A_{n-1} に帰納法の仮定を用いて

$A_{n-1} = L_{n-1}U_{n-1}$ (L_{n-1}, U_{n-1}はそれぞれ下3角，上3角行列で対角要素はすべて正)

と書ける．このとき $\widetilde{A_{n-1}} = a_{nn} - \boldsymbol{b}A_{n-1}^{-1}\boldsymbol{a}$ とおいて

$$
\begin{aligned}
A &= \begin{pmatrix} I_{n-1} & 0 \\ \boldsymbol{b}A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \boldsymbol{a} \\ 0 & \widetilde{A_{n-1}} \end{pmatrix} \\
&= \begin{pmatrix} I_{n-1} & 0 \\ \boldsymbol{b}U_{n-1}^{-1}L_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} L_{n-1}U_{n-1} & \boldsymbol{a} \\ 0 & \widetilde{A_{n-1}} \end{pmatrix} \\
&= \begin{pmatrix} L_{n-1} & 0 \\ \boldsymbol{b}U_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} U_{n-1} & L_{n-1}^{-1}\boldsymbol{a} \\ 0 & \widetilde{A_{n-1}} \end{pmatrix}
\end{aligned}
$$

とかく．($\widetilde{A_{n-1}}$ はスカラーで A_{n-1} の Schur 補元 (§1.10) であることに注意．)

上式の両辺の行列式をとれば

$$\det A = (\det L_{n-1})(\det U_{n-1})\det(\widetilde{A_{n-1}}) = \det(A_{n-1})\widetilde{A_{n-1}}$$

仮定によって $\det A > 0$ かつ $\det A_{n-1} > 0$ であるから

$$\widetilde{A_{n-1}} = \frac{\det A}{\det A_{n-1}} > 0$$

よって

$$A = LU,\ L = \begin{pmatrix} L_{n-1} & 0 \\ \boldsymbol{b}U_{n-1}^{-1} & 1 \end{pmatrix},\ U = \begin{pmatrix} U_{n-1} & \boldsymbol{a} \\ 0 & \widetilde{A}_{n-1} \end{pmatrix}$$

は求める A の LU 分解である．

(ii) $\Rightarrow$ (i) $A = LU$ (L, U の対角要素がすべて正) と LU 分解できるとすれば

$$A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix} = L\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}U\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix} \quad (1 \le k \le n)$$

であるから両辺の行列式をとればよい．　　証明終 ∎

注意 2 (LL^t 分解)　A が実正定値対称行列ならば定理 3.3 の条件 (i) はみたされて (ii) が成り立つが，このときはもっと強く $A = LL^t$ なる分解 (LL^t 分解) ができる．これを次節に述べよう．

3.2 LL^t 分解

A を実正定値対称行列 ($A^t = A,\ \boldsymbol{x}^tA\boldsymbol{x} > 0 \quad \forall \boldsymbol{x}(\neq \boldsymbol{0}) \in \mathbb{R}^n$) とすれば $\boldsymbol{x} = (x_1, \ldots, \boldsymbol{x}_k, 0, \ldots, 0)^t$, $\widetilde{\boldsymbol{x}} = (x_1, \ldots, x_k)^t$ とおくとき

$$\widetilde{\boldsymbol{x}}^tA_k\widetilde{\boldsymbol{x}} = \boldsymbol{x}^tA\boldsymbol{x} > 0 \quad \forall \widetilde{\boldsymbol{x}}(\neq \boldsymbol{0}) \in \mathbb{R}^k \tag{3.6}$$

ただし $A_k = A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}$ (A の k 次首座小行列)

したがって A_k は正定値対称行列であり定理 3.3 の条件 (i) がみたされる．故に A は対角成分がすべて正の下3角行列 L と対角成分がすべて正の上3角行列 U の積に分解されるがさらに精密な次の定理が成り立つ．

定理 3.4 A が実正定値対称行列ならば

(i) 各 k $(1 \leq k \leq n)$ につき A_k は正定値対称行列である.

(ii) 対角要素がすべて正の下3角行列 L をえらんで $A = LL^{\mathrm{t}}$ と書ける. この分解は L が単位下3角でなくても一意的である. これを A の **Cholesky (コレスキー) 分解**という.

【証明】 (i) (3.6) より明らかである.

(ii) $A = L \cdot L^{\mathrm{t}}$, $L = \begin{pmatrix} l_{11} & & \\ \vdots & \ddots & \\ l_{n1} & \cdots & l_{nn} \end{pmatrix}$, $l_{ii} > 0 \quad \forall i$ とおいて l_{ij} を順次定めればよい.

$$A = \begin{pmatrix} A_{n-1} & \boldsymbol{a} \\ \boldsymbol{a}^{\mathrm{t}} & a_{nn} \end{pmatrix}, \quad L = \begin{pmatrix} L_{n-1} & \\ \boldsymbol{c} & \alpha \end{pmatrix}, \quad \boldsymbol{c} = (l_{n1}, \ldots, l_{n\,n-1}), \quad \alpha = l_{nn}$$

とおき, n に関する帰納法を用いる.

$$A = LL^{\mathrm{t}} \text{より} \quad A_{n-1} = L_{n-1}L_{n-1}^{\mathrm{t}}, \quad \boldsymbol{a} = L_{n-1}\boldsymbol{c}^{\mathrm{t}}, \quad a_{nn} = \boldsymbol{c}\boldsymbol{c}^{\mathrm{t}} + \alpha^2$$

A_{n-1} は $n-1$ 次正定値対称であるから帰納法の仮定によって対角要素がすべて正の下3角行列 L_{n-1} は一意に定まる. したがって $\boldsymbol{c}$ も $(L_{n-1}^{-1}\boldsymbol{a})^{\mathrm{t}}$ として一意に定まる. このとき

$$\begin{aligned} \alpha^2 &= a_{nn} - \boldsymbol{c} \cdot \boldsymbol{c}^{\mathrm{t}} = a_{nn} - (L_{n-1}^{-1}\boldsymbol{a})^{\mathrm{t}} L_{n-1}^{-1}\boldsymbol{a} \\ &= a_{nn} - \boldsymbol{a}^{\mathrm{t}} A_{n-1}^{-1}\boldsymbol{a} = \widetilde{A_{n-1}} \quad (A_{n-1}\text{の Schur 補元}) \\ &= \det A / \det A_{n-1} \quad (\text{定理 1.19 (i) による.}) \\ &> 0 \end{aligned}$$

よって $\alpha = \sqrt{\widetilde{A_{n-1}}}$ も一意に定まり, 対角要素が正の下3角行列 L がただ1つ確定する. 証明終 ∎

注意 1 前節注意1で述べた A の LDV 分解を用いて

$$A = LDV \quad (L: \text{下3角}, \ D: \text{対角}, \ V: \text{単位上3角})$$

とあらわすとき

$$A = A^{\mathrm{t}} = V^{\mathrm{t}}D^{\mathrm{t}}L^{\mathrm{t}} = V^{\mathrm{t}}DL^{\mathrm{t}}$$

V^{t} は単位下3角で DL^{t} は上3角であるから分解の一意性 (定理 3.1) により

$$V^{\mathrm{t}} = L, \quad DL^{\mathrm{t}} = DV$$

よって

$$A = LDL^{\mathrm{t}} \tag{3.7}$$

A は正定値対称行列であったから $D = \mathrm{diag}(d_{11}, \ldots, d_n)$ の対角成分はすべて正 (これは Sylvester の慣性則によって $\mathrm{In}(D) = \mathrm{In}(LDL^{\mathrm{t}}) = \mathrm{In}(A) = (n, 0, 0)$ からもわかる) であるから

$$\sqrt{D} = \mathrm{diag}(\sqrt{d_1}, \ldots, \sqrt{d_n}), \quad \widetilde{L} = L\sqrt{D}$$

とおけば $A = \widetilde{L}\widetilde{L}^{\mathrm{t}}$ を得る. $\widetilde{L}$ をあらためて L と書けば (ii) の分解が得られる. ただしこの分解では $\widetilde{L}$ (したがって L) は必ずしも単位下3角ではない.

注意 2　同じ議論によって正定値 (複素) Hermite 行列ならば $A = LL^*$ または $A = R^*R$ (R：上 3 角) と分解できる (後者は $R = L^*$ とおけばよい).

3.3 QR 分解

長さ 1 の n 次元実 (列) ベクトル $\boldsymbol{a}$ が与えられたとき，n 次行列

$$P = I - 2\boldsymbol{a}\cdot\boldsymbol{a}^{\mathrm{t}} \tag{3.8}$$

は $P^{\mathrm{t}} = P$ かつ $P^2 = I$ をみたす．後者は直接計算により

$$P^2 = (I - 2\boldsymbol{a}\cdot\boldsymbol{a}^{\mathrm{t}})(I - 2\boldsymbol{a}\cdot\boldsymbol{a}^{\mathrm{t}}) = I - 4\boldsymbol{a}\cdot\boldsymbol{a}^{\mathrm{t}} + 4\boldsymbol{a}\cdot(\boldsymbol{a}^{\mathrm{t}}\boldsymbol{a})\boldsymbol{a}^{\mathrm{t}} = I \quad (\because\ \boldsymbol{a}^{\mathrm{t}}\boldsymbol{a} = 1)$$

として確かめられる．したがって P は実対称な直交行列である．これを**基本直交行列**という．$\boldsymbol{x}, \boldsymbol{y}$ を同じ長さの相異なるベクトル ($\boldsymbol{x} \neq \boldsymbol{y}$ かつ $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2$) とするとき (3.8) において

$$\boldsymbol{a} = \frac{\boldsymbol{x} - \boldsymbol{y}}{\|\boldsymbol{x} - \boldsymbol{y}\|_2}$$

とすれば $P\boldsymbol{x} = \boldsymbol{y}$ となる．実際

$$P\boldsymbol{x} = \boldsymbol{x} - 2\frac{(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{t}}}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}\boldsymbol{x} = \boldsymbol{x} - 2\frac{(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{t}}\boldsymbol{x}}{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}(\boldsymbol{x}-\boldsymbol{y})$$

かつ

$$\begin{aligned}\|\boldsymbol{x}-\boldsymbol{y}\|_2^2 &= (\boldsymbol{x}-\boldsymbol{y})^{\mathrm{t}}(\boldsymbol{x}-\boldsymbol{y}) = \boldsymbol{x}^{\mathrm{t}}\boldsymbol{x} - 2\boldsymbol{x}^{\mathrm{t}}\boldsymbol{y} + \boldsymbol{y}^{\mathrm{t}}\boldsymbol{y}\\ &= 2(\boldsymbol{x}^{\mathrm{t}}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{t}}\boldsymbol{y}) = 2\boldsymbol{x}^{\mathrm{t}}(\boldsymbol{x}-\boldsymbol{y}) = 2(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{t}}\cdot\boldsymbol{x}\end{aligned}$$

に注意すれば $P\boldsymbol{x} = \boldsymbol{x} - (\boldsymbol{x}-\boldsymbol{y}) = \boldsymbol{y}$ を得る．

Householder (ハウスホルダー) はこの形の行列を用いて実対称行列 A を実対称 3 重対角行列に相似変換する有名な方法を提案したので P は **Householder 行列**とも呼ばれる．この行列を用いて次の定理を証明しよう．

定理 3.5　A を n 次実行列とすれば

$$A = QR$$

をみたす直交行列 Q と上 3 角行列 $R = (r_{ij})$　$(r_{ij} = 0 (i > j),\ r_{ii} \geq 0 \quad \forall i)$ が存在する．これを A の **QR 分解**という．A が複素行列ならば Q はユニタリ行列でおきかえられる．

【証明】　$n = 1$ のときは明らかであるから $n > 1$ とする．まず $A = (a_{ij})$ の第 1 列

$$\boldsymbol{x} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}$$

に対し $P_1\boldsymbol{x} = \boldsymbol{y},\ \boldsymbol{y} = (r_{11}, 0, \ldots, 0)^{\mathrm{t}},\ r_{11} = \|\boldsymbol{x}\|_2$ をみたす基本直交行列 P_1 を定める．(もし最初から $\boldsymbol{x}$ が $\boldsymbol{y}$ の形であれば $P_1 = I_n$ とおき何もしない．) このとき

$$P_1 A = \begin{pmatrix} r_{11} & \\ 0 & \\ \vdots & * \\ 0 & \end{pmatrix} \quad (= (a_{ij}^{(1)}) \text{ とおく})$$

次に P_1A の $n-1$ 次主小行列 $(P_1A)\begin{pmatrix} 2 & 3 & \cdots & n \\ 2 & 3 & \cdots & n \end{pmatrix}$ の第1列を考え

$$\boldsymbol{x}^{(1)} = \begin{pmatrix} a_{22}^{(1)} \\ \vdots \\ a_{n2}^{(1)} \end{pmatrix},\ \boldsymbol{y}^{(1)} = \begin{pmatrix} r_{22} \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ r_{22} = \|\boldsymbol{x}^{(1)}\|_2$$

に対し $P^{(1)}\boldsymbol{x}^{(1)} = \boldsymbol{y}^{(1)}$ をみたす $n-1$ 次基本直交行列 $P^{(1)}$ を定めて

$$P_2 = \begin{pmatrix} 1 & \\ & P^{(1)} \end{pmatrix}$$

とおく．($\boldsymbol{x}^{(1)}$ が最初から $\boldsymbol{y}^{(1)}$ の形であれば $P_2 = I_n$ とき何もしない．)

このような操作をくり返せば直交行列 $P_1, P_2, \ldots, P_{n-1}$ により

$$P_{n-1}\cdots P_2P_1A = \begin{pmatrix} r_{11} & & & * \\ & r_{22} & & \\ & & \ddots & \\ & & & r_{nn} \end{pmatrix},\ r_{ii} \geq 0 \quad (1 \leq i \leq n) \tag{3.9}$$

となる．右辺の上3角行列を R とし $Q = P_1^{-1}P_2^{-1}\cdots P_{n-1}^{-1}$ とおけば求める QR 分解 $A = QR$ が得られる．なお，A が正則であれば (3.9) において $r_{ii} \neq 0 \quad \forall i$ であり，$r_{ii} \geq 0 \quad \forall i$ と併せて $r_{ii} > 0 \quad \forall i$ となる．　証明終 ∎

3.4 特異値分解

定理 3.6 (Autonne (オートン)) A を n 次実行列とすれば適当な n 次直交行列 T_1 と T_2 をえらんで

$$T_1AT_2 = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \quad (\alpha_1, \ldots, \alpha_n \text{は } A \text{ の特異値}) \tag{3.10}$$

とできる．A が複素行列ならば T_1, T_2 はユニタリ行列 U_1, U_2 でおきかえられる．

【証明】 $A^{\mathrm{t}}A$ は半正定値対称行列であるから適当な直交行列 T_2 により

$$T_2^{\mathrm{t}}A^{\mathrm{t}}AT_2 = (AT_2)^{\mathrm{t}}(AT_2) = \begin{pmatrix} \alpha_1^2 & & \\ & \ddots & \\ & & \alpha_n^2 \end{pmatrix},\ \alpha_1 \geq \cdots \geq \alpha_n \geq 0 \tag{3.11}$$

とできる (系 2.3.1)．ここで $\mathrm{rank}A = r$ とすれば $\mathrm{rank}(A^{\mathrm{t}}A) = \mathrm{rank}A = r$ (定理 1.18) であるから $\alpha_1 \geq \cdots \geq \alpha_r > 0,\ \alpha_{r+1} = \cdots = \alpha_n = 0$.

AT_2 の第1〜第 n 列を $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ とすれば (3.11) より

$$(\boldsymbol{u}_i, \boldsymbol{u}_j) = \begin{cases} \alpha_i^2 & (i=j) \\ 0 & (i \neq j) \end{cases}$$

よって $\dfrac{1}{\alpha_1}\boldsymbol{u}_1, \cdots, \dfrac{1}{\alpha_r}\boldsymbol{u}_r$ は正規直交系をなし，かつ $\boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_n = 0$

いま $\dfrac{1}{\alpha_1}\boldsymbol{u}_1^{\mathrm{t}}, \cdots, \dfrac{1}{\alpha_r}\boldsymbol{u}_r^{\mathrm{t}}$ を第1～第 r 行とする n 次直交行列を T_1 とする．(この行列の存在は定理 1.6 と Gram-Schmidt の直交化法による．) このとき

$$T_1 A T_2 = \begin{pmatrix} \alpha_1 & & & & & \\ & \ddots & & & & \\ & & \alpha_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix}$$

となる． 証明終 ∎

上記証明は $m \times n$ $(m \neq n)$ 行列に対してもそのまま通用する．すなわち，A を $m \times n$ 行列，$\mathrm{rank} A = r$ かつ $\alpha_1^2 \geq \cdots \geq \alpha_n^2$ $(\alpha_i \geq 0 \quad \forall i)$ を n 次行列 $A^{\mathrm{t}}A$ の固有値とすれば適当な m 次直交行列 T_1 と n 次直交行列 T_2 をえらんで

$$T_1 A T_2 = \left(\begin{array}{ccc|c} \alpha_1 & & & \\ & \ddots & & \\ & & \alpha_r & \\ \hline & & & \end{array}\right) \begin{matrix} 0 \\ \\ \\ 0 \end{matrix} \quad (m \times n \text{ 行列})$$

$$\begin{matrix} 0 & 0 \end{matrix}$$

$(= \Sigma \text{とおく})$

とできる．このとき

$$U = T_1^{\mathrm{t}} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m), \quad \boldsymbol{u}_i \in \mathbb{R}^m$$
$$V = T_2 = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n), \quad \boldsymbol{v}_i \in \mathbb{R}^n$$

とおけば U, V はそれぞれ m 次，n 次の直交行列で

$$A = T_1^{\mathrm{t}} \Sigma T_2^{\mathrm{t}} = U \Sigma V^{\mathrm{t}} \tag{3.12}$$
$$= \alpha_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{t}} + \cdots + \alpha_r \boldsymbol{u}_r \boldsymbol{v}_r^{\mathrm{t}} \tag{3.13}$$

これを A の**特異値分解 (singular value decomposition)** という．

▒**注意**　A が複素行列ならば上の議論において T_1, T_2 をそれぞれ m 次，n 次のユニタリ行列でおきかえ $\boldsymbol{v}_i^{\mathrm{t}}$ を $\boldsymbol{v}_i^*$ でおきかえればよい．

n 元 m 立 1 次方程式

$$A\boldsymbol{x} = \boldsymbol{b} \quad (A \in \mathbb{R}^{m \times n}) \tag{3.14}$$

に対し A の特異値分解 (3.12), (3.13) が知られれば

$$\Sigma^+ = \left(\begin{array}{ccc|c} \alpha_1^{-1} & & & \\ & \ddots & & 0 \\ & & \alpha_r^{-1} & \\ \hline & 0 & & 0 \end{array}\right) \quad (n \times m \text{ 行列})$$

$$A^+ = V\Sigma^+ U^{\mathrm{t}} = \frac{1}{\alpha_1}\boldsymbol{v}_1\boldsymbol{u}_1^{\mathrm{t}} + \cdots + \frac{1}{\alpha_r}\boldsymbol{v}_r\boldsymbol{u}_r^{\mathrm{t}} \quad (n \times m \text{ 行列})$$

とおくとき $\boldsymbol{x}^+ = A^+\boldsymbol{b}$ は

$$\|A\boldsymbol{x}^+ - \boldsymbol{b}\|_2 \leq \|A\boldsymbol{x} - \boldsymbol{b}\|_2 \quad \forall \boldsymbol{x} \in \mathbb{R}^n$$

をみたすことが知られている．$\boldsymbol{x}^+$ は (3.14) の**最小 2 乗解**と呼ばれる．

$m = n$ で A が正則ならば $A^+ = A^{-1}$ で $\boldsymbol{x}^+$ は (3.14) の厳密解にほかならない．この意味で A^+ を A の**一般逆行列 (generalized inverse)** という．

3.5 極分解

任意の複素数は極表示 $re^{i\theta}$ $(r \geq 0,\ 0 \leq \theta < 2\pi)$ をもつが，次の定理はこの事実の行列への一般化である．

定理 3.7 任意の n 次複素行列 A は適当な半正定値 Hermite 行列 H とユニタリ行列 U をえらんで

$$A = HU$$

と書ける．A が正則ならばこのような分解は一意的である．これを A の**極分解 (polar decomposition)** という．

【証明】 定理 3.6 とその注意によって

$$U_1AU_2 = \left(\begin{array}{ccc} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{array}\right) \ (= D \text{ とおく}),\ \alpha_i \geq 0 \quad \forall i$$

となる n 次ユニタリ行列 U_1 と U_2 がある．このとき

$$A = U_1^* D U_2^* = (U_1^* D U_1)(U_1^* U_2^*) = HU \ (H = U_1^* D U_1,\ U = U_1^* U_2^*)$$

§2.7 定義 2.1 の記号を用いて $D \succcurlyeq 0$ より $H \succcurlyeq 0$ であり，U はユニタリ行列である．

次に A が正則のとき分解の一意性を示そう．仮に

$$A = HU = \widetilde{H}\widetilde{U} \tag{3.15}$$

を 2 つの極分解とすれば A の正則性より H と $\widetilde{H}$ は正則で

$$H \succ 0 \quad \text{かつ} \quad \widetilde{H} \succ 0$$

さらに (3.15) より

$$\widetilde{H}^{-1}H = \widetilde{U}U^{-1} \tag{3.16}$$

$\widetilde{H}^{-1}H$ の固有値はすべて正数である．実際 $\widetilde{H}^{-1}H\boldsymbol{x} = \lambda\boldsymbol{x},\ \boldsymbol{x} \neq \boldsymbol{0}$ とすれば

$$H\boldsymbol{x} = \lambda \widetilde{H}\boldsymbol{x} \quad \therefore\ \boldsymbol{x}^* H\boldsymbol{x} = \lambda \boldsymbol{x}^* \widetilde{H}\boldsymbol{x} \quad \therefore\ \lambda = \frac{\boldsymbol{x}^* H\boldsymbol{x}}{\boldsymbol{x}^* \widetilde{H}\boldsymbol{x}} > 0 \quad (\because\ H \succ 0,\ \widetilde{H} \succ 0)$$

一方 (3.16) の右辺の行列 $\widetilde{U}U^{-1}$ はユニタリ行列であるから，定理 2.1 によりその固有値は絶対値 1 である．故に (3.16) の行列の固有値はすべて 1 でなければならない．しかし，固有値がすべて 1 のユニタリ行列は単位行列の他にはない．(なぜならば，そのような行列を B とすればユニタリ行列は正規行列であるから，系 2.3.2 によって適当なユニタリ行列 V をえらんで $V^*BV = I_n$ とできる．したがって $B = VV^* = I_n$．)

$$\therefore\ \widetilde{H}^{-1}H = \widetilde{U}U^{-1} = I_n \quad \therefore\ \widetilde{H} = H,\ \widetilde{U} = U$$

すなわち A が正則のとき極分解は一意的である．　　証明終 ∎

■**注意**　行列 A の共役転置行列 A^* を極分解し

$$A^* = HV\ (H\ :\text{半正定値},\ V\ :\text{ユニタリ})$$

と書けば

$$A = (HV)^* = V^*H^* = UH\ (U = V^*\text{はユニタリ行列})$$

これも A の極分解と呼ばれる．

3.6 Jordan 分解

この節では，n 次行列 $A \in \mathbb{F}^{n\times n}$ が与えられたとき，適当な正則行列 V をえらんで $V^{-1}AV$ を **Jordan (ジョルダン) の標準形**と呼ばれる簡単な形の行列 J に変換できることを示す．このとき $A = VJV^{-1}$ を A の **Jordan 分解**という．

さて A の特性多項式を

$$f(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{n_i}, \quad \lambda_i \neq \lambda_j\ (i \neq j), \quad \sum_{i=1}^{s} n_i = n$$

とすれば Hamilton-Cayley の定理によって $f(A) = O$ である．

故に $A_i = A - \lambda_i I_n$ とおくとき定理 2.6 によって

$$\mathbb{F}^n = \mathcal{N}(O) = \sum_{i=1}^{s} \mathcal{N}(A_i^{n_i}) \quad (\text{直和}) \tag{3.17}$$

ただし A が実行列で固有値 $\lambda_1, \ldots, \lambda_s$ がすべて実数ならば $\mathbb{F}^n = \mathbb{R}^n$ で，それ以外の場合には $\mathbb{F}^n = \mathbb{C}^n$ である．

いま $\dim \mathcal{N}(A_i^{n_i}) = \sigma_i$ とし，$\mathcal{N}(A_i^{n_i})$ の基底を

$$\boldsymbol{v}_1^{(i)}, \ldots, \boldsymbol{v}_{\sigma_i}^{(i)}$$

それらを並べてつくる $n \times \sigma_i$ 行列を $B^{(i)} = (\boldsymbol{v}_1^{(i)}, \ldots, \boldsymbol{v}_{\sigma_i}^{(i)})$ とすれば (3.17) によって $\sum_{i=1}^{s} \sigma_i = n$ かつ線形部分空間 $\mathcal{N}(A_i^{n_i})$ は A-不変 (A-invariant) すなわち $A\mathcal{N}(A_i^{n_i}) \subseteq \mathcal{N}(A_i^{n_i})$ をみたす．

$$\left(\because\ \boldsymbol{x} \in \mathcal{N}(A_i^{n_i}) \Rightarrow A_i^{n_i}(A\boldsymbol{x}) = (A - \lambda_i)^{n_i} A\boldsymbol{x} = A[(A - \lambda_i)^{n_i}\boldsymbol{x}] = A \cdot \boldsymbol{0} = \boldsymbol{0}\right)$$

よって行列 A を $\mathbb{F}^n$ の基底

$$\boldsymbol{v}_1^{(1)}, \ldots, \boldsymbol{v}_{\sigma_1}^{(1)}, \ldots, \boldsymbol{v}_1^{(s)}, \ldots, \boldsymbol{v}_{\sigma_s}^{(s)}$$

に関して表現すれば

$$A(B^{(1)},\dots,B^{(s)}) = (B^{(1)},\dots,B^{(s)}) \begin{pmatrix} C^{(1)} & & \\ & \ddots & \\ & & C^{(s)} \end{pmatrix} \tag{3.18}$$

の形になる．ただし各 $C^{(i)} = (c_{jk}^{(i)})$ は σ_i 次行列でその (j,k) 要素 $c_{jk}^{(i)}$ は

$$A\boldsymbol{v}_j^{(i)} = \sum_{k=1}^{\sigma_i} c_{kj}^{(i)} \boldsymbol{v}_k^{(i)} \quad \in \mathcal{N}(A_i^{n_i})$$

をみたす．
ここで A を線形空間 $\mathcal{N}(A_i^{n_i}) = \mathrm{span}\{\boldsymbol{v}_1^{(i)},\dots,\boldsymbol{v}_{\sigma_i}^{(i)}\}$ の上に制限すれば $(A-\lambda_i I)^{n_i}\boldsymbol{x} = \boldsymbol{0} \quad \forall \boldsymbol{x} \in \mathcal{N}(A_i^{n_i})$ より

$$(C^{(i)} - \lambda_i I_{\sigma_i})^{n_i} = O$$

故に $N_i = C^{(i)} - \lambda_i I_{\sigma_i}$ は巾零行列 (nilpotent matrix) である．(一般に n 次行列 N が巾零行列であるとはある自然数 k に対して $N^k = O$ となるときをいう．このとき N の固有値は零しかない．なぜならば $Ny = \lambda y,\ \boldsymbol{y} \neq \boldsymbol{0}$ とすれば $N^k y = \lambda^k y$ であるから $N^k = O$ ならば $\lambda^k y = \boldsymbol{0}$ $\therefore\ \lambda^k = 0$ $\therefore\ \lambda = 0$)
したがって $C^{(i)}$ の固有値は λ_i(σ_i 重根) で

$$C = \begin{pmatrix} C^{(1)} & & \\ & \ddots & \\ & & C^{(s)} \end{pmatrix}$$

の特性多項式は $\prod_{i=1}^{s}(\lambda-\lambda_i)^{\sigma_i}$ である．

一方 A の特性多項式は $f(\lambda) = \prod_{i=1}^{s}(\lambda-\lambda_i)^{n_i}$, $\lambda_i \neq \lambda_j\ (i \neq j)$ と仮定しているから $\sigma_i = n_i\ (1 \leq i \leq s)$ である．($\because\ B = (B^{(1)},\dots,B^{(s)})$ とおけば B は n 次正則行列で $AB = BC$ より $B^{-1}AB = C$．よって A と C の特性多項式は一致しなければならない．)

結局 A の (Jordan) 標準形を求める問題は巾零行列 N の標準形を求める問題に帰着される．

以下，西岡 [33] の方法により N の標準形を求める．いま δ を $N^m = O$ となる最小の非負整数とする．δ は §1.9 に述べた N のディセント $\delta(N)$ (それは同時にアセント $\alpha(N)$ に等しい) で

$$\{\boldsymbol{0}\} = R(N^\delta) \subsetneqq R(N^{\delta-1}) \subsetneqq \cdots \subsetneqq R(N^0) = \mathbb{F}^n$$

が成り立つ．さらに $N \neq O$ とする．このとき $\delta \geq 2$ である．

さて，まず $\dim \mathcal{R}(N^{\delta-1}) = s_1$ として $\mathcal{R}(N^{\delta-1})$ の基底

$$N^{\delta-1}\boldsymbol{u}_1^{(1)},\dots,N^{\delta-1}\boldsymbol{u}_{s_1}^{(1)}$$

をとる．このとき $\mathcal{R}(N^{\delta-2})$ の元

$$N^{\delta-1}\boldsymbol{u}_1^{(1)},\quad \dots,\quad N^{\delta-1}\boldsymbol{u}_{s_1}^{(1)},\quad N^{\delta-2}\boldsymbol{u}_1^{(1)},\quad \dots,\quad N^{\delta-2}\boldsymbol{u}_{s_1}^{(1)} \tag{3.19}$$

は 1 次独立である．

$\because$

$$\sum_{i=1}^{s_1} c_i N^{\delta-1}\boldsymbol{u}_i^{(1)} + \sum_{j=1}^{s_1} d_j N^{\delta-2}\boldsymbol{u}_j^{(1)} = \boldsymbol{0} \quad (c_i,\ d_j : \text{定数}) \tag{3.20}$$

とおき両辺に左から N をかければ $N^{\delta}\boldsymbol{u}_i^{(1)} = \boldsymbol{0}$ $(1 \le i \le s_1)$ より

$$\sum_{j=1}^{s_1} d_j N^{\delta-1}\boldsymbol{u}_j^{(1)} = \boldsymbol{0}$$

$\{N^{\delta-1}\boldsymbol{u}_j^{(1)}\}$, $j = 1, 2, \ldots, s_1$ は $\mathcal{R}(N^{\delta-1})$ の基底であるから 1 次独立で，$d_j = 0$ $(1 \le j \le s_1)$
よって (3.20) より $c_i = 0$ $(1 \le i \le s_1)$.

次に (3.19) を含む $\mathcal{R}(N^{\delta-2})$ の基底を

$$N^{\delta-1}\boldsymbol{u}_1^{(1)}, \ldots,\ N^{\delta-1}\boldsymbol{u}_{s_1}^{(1)},\ N^{\delta-2}\boldsymbol{u}_1^{(1)}, \ldots,\ N^{\delta-2}\boldsymbol{u}_{s_1}^{(1)},\ N^{\delta-2}\boldsymbol{u}_1^{(2)}, \ldots,\ N^{\delta-2}\boldsymbol{u}_{s_2}^{(2)} \tag{3.21}$$

とする (定理 1.6 参照 ($s_2 = 0$ の場合も起こり得る)). このとき

$$N^{\delta-1}\boldsymbol{u}_1^{(2)} = \cdots = N^{\delta-1}\boldsymbol{u}_{s_2}^{(2)} = \boldsymbol{0}$$

となるように選べる．実際 $N^{\delta-1}\boldsymbol{u}_j^{(2)} \in \mathcal{R}(N^{\delta-1})$ $(1 \le j \le s_2)$ であるから

$$N^{\delta-1}\boldsymbol{u}_j^{(2)} = \sum_{i=1}^{s_1} \alpha_{ij} N^{\delta-1}\boldsymbol{u}_i^{(1)} \ (\alpha_{ij}\text{は定数})$$

と書けて

$$\widetilde{\boldsymbol{u}}_j^{(2)} \equiv \boldsymbol{u}_j^{(2)} - \sum_{i=1}^{s_1} \alpha_{ij}\boldsymbol{u}_i^{(1)} \in \mathcal{N}(N^{\delta-1}) \quad (\because\ N^{\delta-1}\widetilde{\boldsymbol{u}}_j^{(2)} = \boldsymbol{0}) \tag{3.22}$$

故に各 $N^{\delta-2}\boldsymbol{u}_j^{(2)}$ を $N^{\delta-2}\widetilde{\boldsymbol{u}}_j^{(2)}$ でおきかえればよい．このようにしても 1 次独立性は保たれる．

$\because$

$$\sum_{i=1}^{s_1} c_i N^{\delta-1}\boldsymbol{u}_i^{(1)} + \sum_{i=1}^{s_1} d_i N^{\delta-2}\boldsymbol{u}_i^{(1)} + \sum_{j=1}^{s_2} e_j N^{\delta-2}\widetilde{\boldsymbol{u}}_j^{(2)} = \boldsymbol{0} \quad (c_i,\ d_i,\ e_j : \text{定数}) \tag{3.23}$$

とし両辺に左から N をかければ

$$\sum_{i=1}^{s_1} d_i N^{\delta-1}\boldsymbol{u}_i^{(1)} = \boldsymbol{0} \quad \therefore\ d_i = 0\ (1 \le i \le s_1)$$

次に (3.22) を (3.23) に代入して

$$\sum_{i=1}^{s_1} c_i N^{\delta-1}\boldsymbol{u}_i^{(1)} + \sum_{j=1}^{s_2} e_j (N^{\delta-2}\boldsymbol{u}_j^{(2)} - \sum_{k=1}^{s_1} \alpha_{kj} N^{\delta-2}\boldsymbol{u}_k^{(1)}) = \boldsymbol{0}$$

$$\therefore \sum_{j=1}^{s_2} e_j N^{\delta-2}\boldsymbol{u}_j^{(2)} + ((3.19)\text{のベクトルの 1 次結合}) = \boldsymbol{0}$$

$$\therefore\ e_j = 0\ (1 \le j \le s_2) \quad (\because\ \text{上式の左辺は (3.21) のベクトルの 1 次結合})$$

したがって (3.23) より $c_i = 0$ $(1 \le i \le s_1)$ を得る．

このとき $\mathcal{R}(N^{\delta-3})$ の元

$$N^{\delta-1}\boldsymbol{u}_1^{(1)}, \ldots,\ N^{\delta-1}\boldsymbol{u}_{s_1}^{(1)},\ N^{\delta-2}\boldsymbol{u}_1^{(1)}, \ldots,\ N^{\delta-2}\boldsymbol{u}_{s_1}^{(1)},\ N^{\delta-3}\boldsymbol{u}_1^{(1)}, \ldots,\ N^{\delta-3}\boldsymbol{u}_{s_1}^{(1)},$$
$$N^{\delta-2}\boldsymbol{u}_1^{(2)}, \ldots,\ N^{\delta-2}\boldsymbol{u}_{s_2}^{(2)},\ N^{\delta-3}\boldsymbol{u}_1^{(2)}, \ldots,\ N^{\delta-3}\boldsymbol{u}_{s_2}^{(2)}$$

は1次独立である．これをみるには，(これらの1次結合式) $=\mathbf{0}$ とおき両辺に左から N をかけて $N^{\delta-1}\boldsymbol{u}_j^{(2)}=\mathbf{0}$ $(1\le j\le s_2)$ 等を用い，1次結合式の係数がすべて零であることを導けばよい．これは容易であるから各自検証されたい．

次にこれらに $\mathcal{R}(N^{\delta-3})$ の元 $N^{\delta-3}\boldsymbol{u}_1^{(3)},\dots,N^{\delta-3}\boldsymbol{u}_{s_3}^{(3)}$ を追加して $\mathcal{R}(N^{\delta-3})$ の基底をつくる．すなわち $\mathcal{R}(N^{\delta-3})$ の基底として

$$N^{\delta-1}\boldsymbol{u}_i^{(1)}(1\le i\le s_1),\ N^{\delta-2}\boldsymbol{u}_i^{(1)}(1\le i\le s_1),\ N^{\delta-3}\boldsymbol{u}_i^{(1)}(1\le i\le s_1)$$
$$N^{\delta-2}\boldsymbol{u}_i^{(2)}(1\le i\le s_2),\ N^{\delta-3}\boldsymbol{u}_i^{(2)}(1\le i\le s_2)$$
$$N^{\delta-3}\boldsymbol{u}_i^{(3)}(1\le i\le s_3)$$

をとる．このとき上と同様に $N^{\delta-2}\boldsymbol{u}_i^{(3)}=\mathbf{0}$ $(1\le i\le s_3)$ としてよい．実際 $N^{\delta-2}\boldsymbol{u}_i^{(3)}\in\mathcal{R}(N^{\delta-2})$ であるから $\mathcal{R}(N^{\delta-2})$ の基底 (3.21) の1次統合として

$$N^{\delta-2}\boldsymbol{u}_i^{(3)}=\sum_{k=1}^{s_1}\alpha_{ki}N^{\delta-1}\boldsymbol{u}_k^{(1)}+\sum_{k=1}^{s_1}\beta_{ki}N^{\delta-2}\boldsymbol{u}_k^{(1)}+\sum_{k=1}^{s_2}\gamma_{ki}N^{\delta-2}\boldsymbol{u}_k^{(2)}$$
$$(\alpha_{ki},\ \beta_{ki},\ \gamma_{ki}:\text{定数})$$

と書けて

$$N^{\delta-2}(\boldsymbol{u}_i^{(3)}-\sum_{k=1}^{s_1}\alpha_{ki}N\boldsymbol{u}_k^{(1)}-\sum_{k=1}^{s_1}\beta_{ki}\boldsymbol{u}_k^{(1)}-\sum_{k=1}^{s_2}\gamma_{ki}\boldsymbol{u}_k^{(2)})=\mathbf{0}$$
$$\therefore\quad \widetilde{\boldsymbol{u}}_i^{(3)}\equiv\boldsymbol{u}_i^{(3)}-\sum_{k=1}^{s_1}\alpha_{ki}N\boldsymbol{u}_k^{(1)}-\sum_{k=1}^{s_1}\beta_{ki}\boldsymbol{u}_k^{(1)}-\sum_{k=1}^{s_2}\gamma_{k}\boldsymbol{u}_k^{(2)}\quad\in\mathcal{N}(N^{\delta-2})$$

よって各 $N^{\delta-2}\boldsymbol{u}_i^{(3)}$ $(1\le i\le s_3)$ を $N^{\delta-2}\widetilde{\boldsymbol{u}}_i^{(3)}$ $(1\le i\le s_3)$ でおきかえればよい．このようにしても1次独立性は保たれるからである (前と同様な議論をくり返せばよい)．

以下この操作を続けて $\mathbb{F}^n=\mathcal{R}(N^0)$ の基底として，

$$N^{\delta-1}\boldsymbol{u}_i^{(1)}(1\le i\le s_1),\ N^{\delta-2}\boldsymbol{u}_i^{(1)}(1\le i\le s_1),\ \cdots,\boldsymbol{u}_i^{(1)}(1\le i\le s_1)$$
$$N^{\delta-2}\boldsymbol{u}_i^{(2)}(1\le i\le s_2),\ N^{\delta-3}\boldsymbol{u}_i^{(2)}(1\le i\le s_2),\ \cdots,\boldsymbol{u}_i^{(2)}(1\le i\le s_2)$$
$$\cdots$$
$$\boldsymbol{u}_i^{(\delta)}(1\le i\le s_\delta)$$

が得られる．ただし $\boldsymbol{u}_i^{(\delta)}$ $(1\le i\le s_\delta)$ は上と同様に $N\boldsymbol{u}_i^{(\delta)}=\mathbf{0}$ $(1\le i\le s_\delta)$ となるようにとる．

ここで，

$$S_i^{(1)}=(N^{\delta-1}\boldsymbol{u}_i^{(1)},N^{\delta-2}\boldsymbol{u}_i^{(1)},\dots,N\boldsymbol{u}_i^{(1)},\boldsymbol{u}_i^{(1)})\quad(n\times\delta\text{ 行列})\ (1\le i\le s_1)$$

とおけば

$$NS_i^{(1)}=S_i^{(1)}\begin{pmatrix}0&1&&&\\&0&1&&\\&&\ddots&\ddots&\\&&&\ddots&1\\&&&&0\end{pmatrix}\quad(\delta\times\delta\text{ 行列})\ (1\le i\le s_1)$$

等に着目して基底を並べかえ

$$S_i^{(1)} = (N^{\delta-1}\boldsymbol{u}_i^{(1)},\ N^{\delta-2}\boldsymbol{u}_i^{(1)},\ldots,\ N\boldsymbol{u}_i^{(1)},\ \boldsymbol{u}_i^{(1)}) \quad (1 \le i \le s_1)$$
$$S_i^{(2)} = (N^{\delta-2}\boldsymbol{u}_i^{(2)},\ N^{\delta-3}\boldsymbol{u}_i^{(2)},\ldots,\ N\boldsymbol{u}_i^{(2)},\ \boldsymbol{u}_i^{(2)}) \quad (1 \le i \le s_2)$$
$$\vdots$$
$$S_i^{(\delta)} = (N^{\delta-\delta}\boldsymbol{u}_i^{(\delta)}) = (\boldsymbol{u}_i^{\delta}) \quad (1 \le i \le s_\delta)$$
$$S^{(1)} = (S_1^{(1)},\ldots,S_{s_1}^{(1)}),\ S^{(2)} = (S_1^{(2)},\ldots,S_{s_2}^{(2)}),\ldots,\ S^{(\delta)} = (S_1^{(\delta)},\ldots,S_{s_\delta}^{(\delta)})$$
$$S = (S^{(1)}, S^{(2)},\ldots,S^{(\delta)})$$

とおけば

$$NS = S\begin{pmatrix} J^{(1)} & & & \\ & J^{(2)} & & \\ & & \ddots & \\ & & & J^{(\delta)} \end{pmatrix}$$

ただし

$$J^{(1)} = \begin{pmatrix} J_1^{(1)} & & \\ & \ddots & \\ & & J_{s_1}^{(1)} \end{pmatrix},\ J_1^{(1)} = \cdots = J_{s_1}^{(1)} = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \quad (\delta \times \delta)$$

$$J^{(2)} = \begin{pmatrix} J_1^{(2)} & & \\ & \ddots & \\ & & J_{s_2}^{(2)} \end{pmatrix},\ J_1^{(2)} = \cdots = J_{s_2}^{(2)} = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \quad ((\delta-1) \times (\delta-1))$$

$$\vdots$$

$$J^{(\delta)} = \begin{pmatrix} J_1^{(\delta)} & & \\ & \ddots & \\ & & J_{s_\delta}^{(\delta)} \end{pmatrix},\ J_1^{(\delta)} = \cdots = J_{s_\delta}^{(\delta)} = (0) \quad (1 \times 1)$$

結局

$$S^{-1}NS = \begin{pmatrix} J^{(1)} & & & \\ & J^{(2)} & & \\ & & \ddots & \\ & & & J^{(\delta)} \end{pmatrix} \ (= K \text{ とおく})$$

を得る．これが巾零行列 $N \neq O$ の標準形である．($N = O$ のときは $K = O$ が対応する．)

故に (3.18) に戻って N と $\delta = \delta(N)$ に対する上の議論を各 $N_i = C^{(i)} - \lambda_i I_{\sigma_i}$ と $\delta_i = \delta(N_i)$ (N_i のディセント) に適用すれば $C^{(i)} = N_i + \lambda_i I_{\sigma_i}$ (N_i は σ_i 次の巾零行列で $N_i^{\sigma_i} = O$ かつ $\sigma_i = n_i$) であることに注意して

$$AB = B\begin{pmatrix} C^{(1)} & & \\ & \ddots & \\ & & C^{(s)} \end{pmatrix} = B\begin{pmatrix} \lambda_1 I_{n_1} + N_1 & & \\ & \ddots & \\ & & \lambda_s I_{n_s} + N_s \end{pmatrix}$$

$$\therefore \quad B^{-1}AB = \begin{pmatrix} \lambda_1 I_{n_1} + N_1 & & \\ & \ddots & \\ & & \lambda_s I_{n_s} + N_s \end{pmatrix}$$

故に $V = BS$ とおけば次の定理が得られる．

定理 3.8 行列 A は適当な正則行列 V により

$$V^{-1}AV = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix} = J \tag{3.24}$$

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

の形に変換される．$A = V^{-1}JV$ を A の **Jordan 分解**，J を **Jordan の標準形**，各 J_i を **Jordan ブロック**という．

例

$$A = \begin{pmatrix} -2 & 5 & 2 & -3 \\ -3 & 5 & 3 & -3 \\ -2 & 2 & 4 & -2 \\ -1 & 1 & 1 & 1 \end{pmatrix} \quad (n = 4)$$

特性多項式は $f(\lambda) = \det(\lambda I_4 - A) = (\lambda - 2)^4$ 故に固有値は $\lambda = 2$ (4 重根)．A は実行列で固有値は実数てあるから (3.17) において $\mathbb{F}^n = \mathbb{R}^n$ $(n = 4)$ である．$A_1 = A - \lambda_1 I_4$ の代わりに $N = A - 2I_4$ とおくと

$$N = \begin{pmatrix} -4 & 5 & 2 & -3 \\ -3 & 3 & 3 & -3 \\ -2 & 2 & 2 & -2 \\ -1 & 1 & 1 & -1 \end{pmatrix}, \quad N^2 = \begin{pmatrix} 0 & -4 & 8 & -4 \\ 0 & -3 & 6 & -3 \\ 0 & -2 & 4 & -2 \\ 0 & -1 & 2 & -1 \end{pmatrix}, \quad N^3 = O$$

よって $\delta(= \delta_1) = \delta(N) = 3$.

$$\mathcal{R}(N^2) = \mathrm{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ -3 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 8 \\ 6 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} -4 \\ -3 \\ -2 \\ -1 \end{pmatrix} \right\}$$

$$= \mathrm{span}\left\{\begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}\right\} = \mathrm{span}\{N^2\boldsymbol{u}_1^{(1)}\} \quad \left(\text{ただし } \boldsymbol{u}_1^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}\right)$$

であるから $\mathcal{R}(N^2)$ の基底として $N^2\boldsymbol{u}_1^{(1)}$ をとる．$s_1 = \dim\mathcal{R}(N^2) = 1$ かつ $N\boldsymbol{u}_1^{(1)} = \begin{pmatrix} 3 \\ 3 \\ 2 \\ 1 \end{pmatrix}$

よって (3.19) より

$$N^2\boldsymbol{u}_1^{(1)} = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix},\ N\boldsymbol{u}_1^{(1)} = \begin{pmatrix} 3 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

は $\mathcal{R}(N)$ の基底の一部をなす．

$$\mathcal{R}(N) = \mathrm{span}\left\{\boldsymbol{v}_1 = \begin{pmatrix} -4 \\ -3 \\ -2 \\ -1 \end{pmatrix},\ \boldsymbol{v}_2 = \begin{pmatrix} 5 \\ 3 \\ 2 \\ 1 \end{pmatrix},\ \boldsymbol{v}_3 = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 1 \end{pmatrix},\ \boldsymbol{v}_4 = \begin{pmatrix} -3 \\ -3 \\ -2 \\ -1 \end{pmatrix}\right\}$$
$$= \mathrm{span}\{\boldsymbol{v}_1, \boldsymbol{v}_4\} \quad (\because\ \boldsymbol{v}_2 = -2\boldsymbol{v}_1 + \boldsymbol{v}_4,\ \boldsymbol{v}_3 = \boldsymbol{v}_1 - 2\boldsymbol{v}_4)$$

$\boldsymbol{v}_1,\ \boldsymbol{v}_4$ は 1 次独立であるから $\dim\mathcal{R}(N) = 2$．よって $N^2\boldsymbol{u}_1^{(1)} = -\boldsymbol{v}_1$ と $N\boldsymbol{u}_1^{(1)} = -\boldsymbol{v}_4$ は $\mathcal{R}(N)$ の基底である．$\therefore\ s_2 = 0$

次に $\mathcal{R}(N^{\delta-3})$ の 1 次独立なベクトル

$$N^2\boldsymbol{u}_1^{(1)} = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix},\quad N\boldsymbol{u}_1^{(1)} = \begin{pmatrix} 3 \\ 3 \\ 2 \\ 1 \end{pmatrix},\quad \boldsymbol{u}_1^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

に新しいベクトル $\boldsymbol{v} = \boldsymbol{u}_1^{(3)}$ $(s_3 = 1)$ を追加して $\mathcal{R}(N^{\delta-3}) = \mathcal{R}(N^0) = \mathbb{R}^4$ の基底をつくる．
$\boldsymbol{v} = \boldsymbol{e}_2 = (0,1,0,0)^{\mathrm{t}}$ ととれば $N^2\boldsymbol{u}_1^{(1)}, N\boldsymbol{u}_1^{(1)}, \boldsymbol{u}_1^{(1)}, \boldsymbol{v}$ は 1 次独立である．
$\left(\because\ \det(N^2\boldsymbol{u}_1^{(1)},\ N\boldsymbol{u}_1^{(1)},\ \boldsymbol{u}_1^{(1)}, \boldsymbol{v}) \neq 0\right)$

$\boldsymbol{v}$ を $\widetilde{\boldsymbol{v}} = \widetilde{\boldsymbol{u}}_1^{(3)}$ にとりかえて $N\widetilde{\boldsymbol{v}} = \boldsymbol{0}$ となるようにするために

$$N\boldsymbol{v} = \begin{pmatrix} 5 \\ 3 \\ 2 \\ 1 \end{pmatrix} \quad \in \mathcal{R}(N)$$

を $\mathcal{R}(N)$ の基底 $N^2\boldsymbol{u}_1^{(1)}, N\boldsymbol{u}_1^{(1)}$ の 1 次統合であらわすと

$$N\boldsymbol{v} = 2N^2\boldsymbol{u}_1^{(1)} - N\boldsymbol{u}_1^{(1)} \quad \therefore\quad N(\boldsymbol{v} - 2N\boldsymbol{u}_1^{(1)} + \boldsymbol{u}_1^{(1)}) = \boldsymbol{0}$$

よって

$$\widetilde{\boldsymbol{v}} = \boldsymbol{v} - 2N\boldsymbol{u}_1^{(1)} + \boldsymbol{u}_1^{(1)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} - 2\begin{pmatrix} 3 \\ 3 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} = -\begin{pmatrix} 6 \\ 5 \\ 4 \\ 3 \end{pmatrix}$$

を $\boldsymbol{v}$ と取り換え，$\mathbb{R}^4$ の基底として

$$N^2\boldsymbol{u}_1^{(1)}, \quad N\boldsymbol{u}_1^{(1)}, \quad \boldsymbol{u}_1^{(1)}, \quad \widetilde{\boldsymbol{v}}$$

を採用する．

$$V = (N^2\boldsymbol{u}_1^{(1)},\ N\boldsymbol{u}_1^{(1)},\ \boldsymbol{u}_1^{(1)}, \widetilde{\boldsymbol{v}}) = \begin{pmatrix} 4 & 3 & 0 & -6 \\ 3 & 3 & 0 & -5 \\ 2 & 2 & 0 & -4 \\ 1 & 1 & -1 & -3 \end{pmatrix}$$

とおくと

$$NV = V\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = VK, \quad K = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.25}$$

を得る．実際検算してみると

$$NV = \begin{pmatrix} -4 & 5 & 2 & -3 \\ -3 & 3 & 3 & -2 \\ -2 & 2 & 2 & -2 \\ -1 & 1 & 1 & -1 \end{pmatrix}\begin{pmatrix} 4 & 3 & 0 & -6 \\ 3 & 3 & 0 & -5 \\ 2 & 2 & 0 & -4 \\ 1 & 1 & -1 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 4 & 3 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} = VK$$

である．

故に (3.25) より

$$V^{-1}AV = V^{-1}(N + 2I_4)V = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} J_1 & \\ & J_2 \end{pmatrix} = J$$

$$\text{ただし } J_1 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},\ J_2 = (2)$$

これが A に対する Jordan の標準形である．

第4章　行列ノルムと行列関数

この章ではベクトルノルムに従属する行列ノルムの性質を詳述すると共に

$$\Phi_p(A) = \Big(\sum_{i,j} |a_{ij}|^p\Big)^{\frac{1}{p}}, \quad A = (a_{ij})$$

は $1 \leq p \leq 2$ のとき p ベクトルノルムと両立する行列ノルムであるが，$p > 2$ または $0 < p < 1$ のとき行列ノルムでないことを示す．また Jordan 分解を用いて解析関数 $f(x)$ から行列関数 $f(A)$ を定義し微分方程式への応用例を与える．

4.1　行列ノルム

n 次行列 $A = (a_{ij}) \in \mathbb{F}^{n\times n}$ のつくる $\mathbb{F}$ 上の線形空間 X において次の条件をみたす X 上の写像 $\|\cdot\| : X \to \mathbb{R}$ を**行列ノルム (matrix norm)** という．

(i)　$\|A\| \geq 0 \quad (A \in X)$; $\|A\| = O \Leftrightarrow A = O$
(ii)　$\|\alpha A\| = |\alpha|\,\|A\| \quad (\alpha \in \mathbb{F})$
(iii)　$\|A + B\| \leq \|A\| + \|B\| \quad (A, B \in X)$
(iv)　$\|AB\| \leq \|A\| \cdot \|B\| \quad (A, B \in X)$

よく用いられる行列ノルムとして次の (a)～(d) がある．

(a)　$\|A\|_F = \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2}$　　(**Frobenius** (フロベニウス) ノルム)

(b)　$\|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$　　(**最大値ノルム**)

(c)　$\|A\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$　　(**1 ノルム**)

(d)　$\|A\|_2 = \sqrt{A^*A \text{ の最大固有値}}$　　(**スペクトルノルム**)

ただし (d) において $\mathbb{F} = \mathbb{R}$ のときは A^*A は $A^{\mathrm{t}}A$ でおきかえられる．

【(a) が行列ノルムであることの証明】　条件 (i), (ii) は明らかである．(iii) を示すために $B = (b_{ij})$ とすれば

$$\begin{aligned}\|A + B\|_F^2 &= \sum_{i,j=1}^{n} |a_{ij} + b_{ij}|^2 \\ &\leq \sum_{i,j=1}^{n} \big(|a_{ij}| + |b_{ij}|\big)\big(|a_{ij} + b_{ij}|\big)\end{aligned}$$

$$\leq \sum_{i,j=1}^{n} |a_{ij}| \cdot |a_{ij}+b_{ij}| + \sum_{i,j=1}^{n} |b_{ij}| \cdot |a_{ij}+b_{ij}| \tag{4.1}$$

$$\leq \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} \sqrt{\sum_{i,j=1}^{n} |a_{ij}+b_{ij}|^2} + \sqrt{\sum_{i,j=1}^{n} |b_{ij}|^2} \sqrt{\sum_{i,j=1}^{n} |a_{ij}+b_{ij}|^2} \tag{4.2}$$

$$= \|A\|_F \cdot \|A+B\|_F + \|B\|_F \cdot \|A+B\|_F$$

$$= \big(\|A\|_F + \|B\|_F\big)\|A+B\|_F$$

ただし (4.1) から (4.2) へ移る際に Cauchy-Schwarz の不等式を用いた.

$$\therefore \quad \|A+B\|_F^2 \leq (\|A\|_F + \|B\|_F)\|A+B\|_F$$

よって $A+B \neq O$ ならば $\|A+B\|_F > 0$ で上式の両辺を割れば

$$\|A+B\|_F \leq \|A\|_F + \|B\|_F$$

この不等式は $A+B=O$ のときも成り立つから任意の $A, B \in X$ に対して (iii) が示された.
最後に (iv) を示すために $C=AB=(c_{ij})$ とおくと Cauchy-Schwarz の不等式により

$$|c_{ij}|^2 = \Big|\sum_{k=1}^{n} a_{ik}b_{kj}\Big|^2 \leq \Big(\sum_{k=1}^{n} |a_{ik}||b_{kj}|\Big)^2 \leq \Big(\sum_{k=1}^{n} |a_{ik}|^2\Big)\Big(\sum_{k=1}^{n} |b_{kj}|^2\Big)$$

$$\therefore \quad \sum_{j=1}^{n} |c_{ij}|^2 \leq \Big(\sum_{k=1}^{n} |a_{ik}|^2\Big) \sum_{j=1}^{n}\sum_{k=1}^{n} |b_{kj}|^2 = \Big(\sum_{k=1}^{n} |a_{ik}|^2\Big)\|B\|_F^2$$

故に

$$\|AB\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |c_{ij}|^2 \leq \|A\|_F^2\|B\|_F^2 \quad \therefore \quad \|AB\|_F \leq \|A\|_F \cdot \|B\|_F$$

よって (iv) が成り立ち $\|\cdot\|_F$ は行列ノルムである.

(b), (c), (d) が行列ノルムであることの証明は次節で与える. 証明終 ∎

4.2 ベクトルノルムに従属する行列ノルム

n 次元ノルム空間 $(\mathbb{F}^n, \|\cdot\|)$ が与えられたとき, n 次行列 $A \in \mathbb{F}^{n\times n}$ に対し

$$\|A\| = \sup_{\substack{\boldsymbol{x}\in\mathbb{F}^n \\ \boldsymbol{x}\neq\boldsymbol{0}}} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\| \tag{4.3}$$

とおけば $\|A\|$ は行列ノルムである (証明容易). $\|A\|$ をベクトルノルム $\|\boldsymbol{x}\|(\boldsymbol{x}\in\mathbb{F}^n)$ に**従属する行列ノルム**またはベクトルノルムから誘導される**作用素ノルム (operator norm)** という. 関数解析学では写像のことを**作用素 (operator)** と呼ぶのでこの名が用いられる.

定理 4.1 $\|\cdot\|_\infty$, $\|\cdot\|_1$, $\|\cdot\|_2$ を $\mathbb{F}^n$ 上のベクトルノルムとする (§1.7 参照). このとき, これらに従属する行列ノルム $\|A\|_\infty$, $\|A\|_1$, $\|A\|_2$ は次で与えられる.

(i) $\|A\|_\infty = \max_{\|\boldsymbol{x}\|_\infty=1} \|A\boldsymbol{x}\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$

(ii) $\|A\|_1 = \max_{\|\boldsymbol{x}\|_1=1} \|A\boldsymbol{x}\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$

(iii) $\|A\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|A\boldsymbol{x}\|_2 = \alpha_1$ (A の最大特異値)

【証明】 以下 $\mathbb{F}=\mathbb{C}$ とする．($\mathbb{F}=\mathbb{R}$ のときも同様である．)

(i) $\|\boldsymbol{x}\|_\infty=1$ とするとき

$$\|A\boldsymbol{x}\|_\infty=\max_i\Big|\sum_{j=1}^n a_{ij}x_j\Big|\le\max_i\sum_j|a_{ij}|\cdot|x_j|\le\max_i\sum_j|a_{ij}| \tag{4.4}$$

ここで $\max\limits_i\sum\limits_{j=1}^n|a_{ij}|=\sum\limits_{j=1}^n|a_{kj}|$ として，$\boldsymbol{y}=(y_1,\dots,y_n)^{\mathrm{t}}\in\mathbb{C}^n$ を

$$y_j=\begin{cases}\dfrac{\bar a_{kj}}{|a_{kj}|} & (a_{kj}\neq0\text{ のとき})\\ 0 & (a_{kj}=0\text{ のとき})\end{cases}\quad(1\le j\le n)$$

と定めれば，$A\neq O$ のとき $\|\boldsymbol{y}\|_\infty=1$ で

$$A\boldsymbol{y}=\Big(\sum_{j=1}^n a_{1j}y_j,\dots,\sum_{j=1}^n a_{nj}y_j\Big)^{\mathrm{t}}$$

$$\Big|\sum_{j=1}^n a_{ij}y_j\Big|\le\sum_{j=1}^n|a_{ij}|\cdot|y_j|\le\sum_{j=1}^n|a_{ij}|$$

特に $i=k$ のとき

$$\Big|\sum_{j=1}^n a_{kj}y_j\Big|=\sum_{j=1}^n|a_{kj}|=\max_i\sum_{j=1}^n|a_{ij}|$$

故に $\boldsymbol{x}=\boldsymbol{y}$ のとき (4.4) の不等号は等号となる．よって

$$\|A\|_\infty=\max_{\|\boldsymbol{x}\|_\infty=1}\|A\boldsymbol{x}\|_\infty=\sum_{j=1}^n|a_{kj}|$$

(ii) $\|\boldsymbol{y}\|_1=1$ かつ $\|A\|_1=\max_{\|\boldsymbol{x}\|_1=1}\|A\boldsymbol{x}\|_1=\|A\boldsymbol{y}\|_1$ なる $\boldsymbol{y}=(y_1,\dots,y_n)^{\mathrm{t}}$ をとれば

$$\|A\|_1=\sum_{i=1}^n\Big|\sum_{j=1}^n a_{ij}y_j\Big|\le\sum_{i=1}^n\sum_{j=1}^n|a_{ij}|\cdot|y_j|=\sum_{j=1}^n|y_j|\Big(\sum_{i=1}^n|a_{ij}|\Big)$$
$$\le\Big(\sum_{j=1}^n|y_j|\Big)\max_j\sum_{i=1}^n|a_{ij}|=\max_j\sum_{i=1}^n|a_{ij}|$$

$\max_j\sum_{i=1}^n|a_{ij}|=\sum_{i=1}^n|a_{ik}|$ とし $\boldsymbol{e}_k$ を第 k 基本単位ベクトル (n 次単位行列の第 k 列) とすれば $\|A\boldsymbol{e}_k\|_1=\sum_{i=1}^n|a_{ik}|$ であるから

$$\|A\|_1=\|A\boldsymbol{e}_k\|_1=\max_j\sum_{i=1}^n|a_{ij}|$$

(iii) $\|\boldsymbol{x}\|_2=1$ なる n 次元列ベクトル $\boldsymbol{x}$ に対して

$$\|A\boldsymbol{x}\|_2^2=(A\boldsymbol{x},\ A\boldsymbol{x})=(A^*A\boldsymbol{x},\ \boldsymbol{x})$$

系 2.12.2 によって A^*A は半正定値 Hermite 行列であり，適当なユニタリ行列 U によって

$$U^*A^*AU=\begin{pmatrix}\alpha_1^2 & & \\ & \ddots & \\ & & \alpha_n^2\end{pmatrix},\quad \alpha_1^2\ge\cdots\ge\alpha_n^2,\ \alpha_i\ge0\quad\forall i$$

と対角化できる (系 2.3.1).

このとき $y = U^*x = (y_1, \cdots, y_n)^{\mathrm{t}}$ とおけば $(\boldsymbol{y},\ \boldsymbol{y}) = (U^*\boldsymbol{x},\ U^*\boldsymbol{x}) = (UU^*\boldsymbol{x},\ \boldsymbol{x}) = (\boldsymbol{x},\ \boldsymbol{x}) = 1$ であるから

$$(A^*A\boldsymbol{x},\ \boldsymbol{x}) = (A^*AU\boldsymbol{y},\ U\boldsymbol{y}) = (U^*A^*AU\boldsymbol{y},\ \boldsymbol{y})$$

$$= \sum_{i=1}^{n} \alpha_i^2 |y_i|^2 \leq \alpha_1^2 \sum_{i=1}^{n} |y_i|^2 = \alpha_1^2 \tag{4.5}$$

$$\therefore\ \|A\boldsymbol{x}\|_2^2 = (A^*A\boldsymbol{x},\ \boldsymbol{x}) \leq \alpha_1^2$$

$$\therefore\ \|A\boldsymbol{x}\|_2 \leq \alpha_1 = \sqrt{A^*A \text{ の最大固有値}} \ =\ (\text{A の最大特異値})$$

特に $\boldsymbol{x}$ として A^*A の固有値 α_1^2 に対応する長さ 1 の固有ベクトル $\boldsymbol{v}$ をとれば (4.5) で等号が成り立つ ($\because\ (A^*A\boldsymbol{v},\ \boldsymbol{v}) = (\alpha_1^2\boldsymbol{v},\ \boldsymbol{v}) = \alpha_1^2\|\boldsymbol{v}\|_2^2 = \alpha_1^2$). 故に

$$\|A\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|A\boldsymbol{x}\|_2 = \alpha_1$$

証明終 ∎

■**注意 (ベクトルノルムと両立する行列ノルム)** ベクトルノルム $\|\boldsymbol{x}\|$ に従属する行列ノルム $\|A\|$ は (4.3) より

$$\|A\boldsymbol{x}\| \leq \|A\| \cdot \|\boldsymbol{x}\| \quad \forall \boldsymbol{x} \tag{4.6}$$

をみたす. (4.6) をみたす行列ノルム $\|A\|$ はベクトルノルム $\|\boldsymbol{x}\|$ と両立するという. Frobenius ノルム $\|A\|_F$ は Euclid(ユークリッド) ベクトルノルム $\|\boldsymbol{x}\|_2$ と両立する. 実際

$$\|A\boldsymbol{x}\|_2^2 = \sum_{i=1}^{n} \Big| \sum_{j=1}^{n} a_{ij}x_j \Big|^2 \leq \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} |a_{ij}||x_j| \Big)^2 \leq \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} |a_{ij}|^2 \Big) \Big(\sum_{j=1}^{n} |x_j|^2 \Big)$$

$$= \Big(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \Big) \|\boldsymbol{x}\|_2^2 = \|A\|_F^2 \|\boldsymbol{x}\|_2^2$$

$$\therefore\ \|A\boldsymbol{x}\|_2 \leq \|A\|_F \|\boldsymbol{x}\|_2$$

また §2.5 でみたように

$$\|A\|_F^2 = \mathrm{tr}(A^*A) = \alpha_1^2 + \cdots \alpha_n^2 \geq \alpha_1^2 = \|A\|_2^2$$

$$\therefore\ \ \|A\|_2 \leq \|A\|_F \quad \therefore\ \ \|A\boldsymbol{x}\|_2 \leq \|A\|_2\|\boldsymbol{x}\|_2 \leq \|A\|_F\|\boldsymbol{x}\|_2$$

しかし, 一般に $\|A\|_2 < \|A\|_F$ であるから, $\|A\|_F$ はベクトルノルム $\|\boldsymbol{x}\|_2$ に従属する行列ノルムではない.

4.3 Frobenius ノルムの拡張

n 次行列 $A = (a_{ij})$ に対し

$$\Phi_p(A) = \Big(\sum_{i,j=1}^{n} |a_{ij}|^p \Big)^{\frac{1}{p}}$$

とおくとき, 次が成り立つ.

定理 4.2 $1 \le p \le 2$ のとき $\Phi_p(A)$ は p ベクトルノルム $\|\boldsymbol{x}\|_p$ と両立する行列ノルムである．すなわち $\Phi_p(A)$ は §4.1 のノルムの条件 (i)～(iv) をみたし，かつ

$$\|A\boldsymbol{x}\|_p \le \Phi_p(A)\|\boldsymbol{x}\|_p \quad \forall \boldsymbol{x}$$

をみたす．

これを示すために 3 つの補題を証明する．

補題 4.1 (Hölder(ヘルダー) の不等式) $a_i \ge 0,\ b_i \ge 0,\ p > 0,\ q > 0,\ 1/p + 1/q = 1$ のとき

$$\sum_{i=1}^{n} a_i b_i \le \Big(\sum_{i=1}^{n} a_i^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^{n} b_i^q\Big)^{\frac{1}{q}} \tag{4.7}$$

等号は $a_1^p/b_1^q = \cdots = a_n^p/b_n^q$ のときに限って成り立つ．

【証明】

$$\alpha_i = \frac{a_i}{\Big(\sum_{j=1}^{n} a_j^p\Big)^{\frac{1}{p}}}, \quad \beta_i = \frac{b_i}{\Big(\sum_{j=1}^{n} b_j^q\Big)^{\frac{1}{q}}}$$

に一般化された相加平均と相乗平均の不等式

$$\alpha_i\beta_i \le \frac{1}{p}\alpha_i^p + \frac{1}{q}\beta_i^q \quad (\text{等号は} \alpha_i^p = \beta_i^q \text{のとき}) \tag{4.8}$$

を使えば

$$\sum_{i=1}^{n} \alpha_i\beta_i \le \frac{1}{p}\sum_{i=1}^{n} \alpha_i^p + \frac{1}{q}\sum_{i=1}^{n} \beta_i^q = \frac{1}{p} + \frac{1}{q} = 1 \quad (\text{等号は} \alpha_i^p = \beta_i^q \ \ \forall i \text{ のとき})$$

これより (4.7) が従う．なお，等号が成り立つための条件

$$\alpha_i^p = \beta_i^q, \quad 1 \le i \le n$$

は

$$\frac{a_1^p}{b_1^q} = \cdots = \frac{a_n^p}{b_n^q} \quad \left(= \frac{\sum_{j=1}^{n} a_j^p}{\sum_{j=1}^{n} b_j^q} \right)$$

と同値である．　　証明終 ∎

注意 1 $0 < p < 1$ のとき $a_i > 0,\ b_i > 0$ に対し (4.7) と (4.8) の不等号は逆向きになる (後述する定理 4.3 の証明参照)．

補題 4.2 (Minkowski (ミンコフスキーまたはミンコウスキー) の不等式) $a_i \ge 0,\ b_i \ge 0,\ 1 \le p < \infty$ のとき

$$\Big(\sum_{i=1}^{n} (a_i + b_i)^p\Big)^{\frac{1}{p}} \le \Big(\sum_{i=1}^{n} a_i^p\Big)^{\frac{1}{p}} + \Big(\sum_{i=1}^{n} b_i^p\Big)^{\frac{1}{p}}$$

等号が成り立つのは，$p > 1$ のときは $a_1/b_1 = \cdots = a_n/b_n$ のときに限る．($p = 1$ のときは常に等号．)

【証明】 $p=1$ のときは明らかであるから $p>1$ とする.

$$x=\Big(\sum_{i=1}^{n}a_i^p\Big)^{\frac{1}{p}},\ y=\Big(\sum_{i=1}^{n}b_i^p\Big)^{\frac{1}{p}},\ z_p=\sum_{i=1}^{n}(a_i+b_i)^p$$

とおけば Hölder の不等式 (補題 4.1) によって

$$\begin{aligned}z_p&=\sum_{i=1}^{n}a_i(a_i+b_i)^{p-1}+\sum_{i=1}^{n}b_i(a_i+b_i)^{p-1}\\&\le xz_{(p-1)q}^{\frac{1}{q}}+yz_{(p-1)q}^{\frac{1}{q}}=(x+y)z_p^{\frac{1}{q}}\quad(\because\ (p-1)q=p)\end{aligned}\tag{4.9}$$

$$\therefore\quad z_p^{1-\frac{1}{q}}\le x+y$$

$1-1/q=1/p$ であるから $z_p^{1/p}\le x+y$. これは求める不等式である.

$p>1$ のとき (4.9) において等号が成り立つための条件は

$$\frac{a_1^p}{(a_1+b_1)^{(p-1)q}}=\cdots=\frac{a_n^p}{(a_n+b_n)^{(p-1)q}}\tag{4.10}$$

および

$$\frac{b_1^p}{(a_1+b_1)^{(p-1)q}}=\cdots=\frac{b_n^p}{(a_n+b_n)^{(p-1)q}}\tag{4.11}$$

であるが $(p-1)q=p$ であるから (4.10) と (4.11) は

$$\frac{a_1}{b_1}=\cdots=\frac{a_n}{b_n}$$

と同値である. 証明終 ∎

■**注意 2** $0<p<1$ のときは $a_i>0,\ b_i>0$ に対し Minkowski の不等式は逆向きになる (定理 4.3 の証明参照).

補題 4.3 (Jensen (イエンゼン) の不等式) $a_i>0,\ n\ge 2$ とするとき

$$f(t)=(a_1^t+\cdots+a_n^t)^{\frac{1}{t}}\quad(t>0)$$

は t の狭義単調減少関数である.

【証明】 $\log f(t)=\frac{1}{t}\log\Big(\sum_{i=1}^{n}a_i^t\Big)$, $t>0$ の両辺を t で微分すると

$$\frac{f'(t)}{f(t)}=-\frac{1}{t^2}\log\Big(\sum_{i=1}^{n}a_i^t\Big)+\frac{1}{t}\,\frac{\displaystyle\sum_{i=1}^{n}a_i^t\log a_i}{\displaystyle\sum_{i=1}^{n}a_i^t}=-\frac{1}{t^2}\left[\log\Big(\sum_{i=1}^{n}a_i^t\Big)-\sum_{i=1}^{n}\theta_i\log a_i^t\right]$$

$$\left(\text{ただし}\quad\theta_i=\frac{a_i^t}{\displaystyle\sum_{j=1}^{n}a_j^t}\right)$$

$0<\theta_i<1$, $\sum_{i=1}^{n}\theta_i=1$ かつ $\varphi(t)=\log t$ は上に凸な狭義単調増加関数であるから

$$\varphi\Big(\sum_{i=1}^{n}a_i^t\Big)>\varphi\Big(\sum_{i=1}^{n}\theta_ia_i^t\Big)\ge\sum_{i=1}^{n}\theta_i\varphi(a_i^t)=\sum_{i=1}^{n}\theta_i\log(a_i^t)$$

$$\therefore \quad \frac{f'(t)}{f(t)} < 0 \quad (t > 0)$$

$f(t) > 0$ であるから $f'(t) < 0 \quad (t > 0)$ を得る. 証明終 ∎

【定理 4.2 の証明】 $p = 1$ のとき定理は明らかであるから, $1 < p \leq 2$ とする. $\Phi_p(A)$ が行列ノルムの性質

(i) $\Phi_p(A) \geq 0;\ \Phi_p(A) = 0 \Leftrightarrow A = O$

(ii) $\Phi_p(\alpha A) = |\alpha|\Phi_p(A) \quad \alpha \in \mathbb{F}$

をみたすことは明らかである. また 3 角不等式

(iii) $\Phi_p(A + B) \leq \Phi_p(A) + \Phi_p(B)$

は補題 4.2 より従う. 以下

(iv) $\Phi_p(AB) \leq \Phi_p(A)\Phi_p(B)$

を示そう.

$B = (b_{ij}),\ C = AB = (c_{ij})$ とおけば補題 4.1 によって

$$|c_{ij}| = \left|\sum_{k=1}^{n} a_{ik}b_{kj}\right| \leq \sum_{k=1}^{n} |a_{ik}| \cdot |b_{kj}| \leq \left(\sum_{k=1}^{n} |a_{ik}|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |b_{kj}|^q\right)^{\frac{1}{q}} \quad \left(\frac{1}{p} + \frac{1}{q} = 1\right) \tag{4.12}$$

$1 < p \leq 2$ より $q = p/(p-1) \geq p$ であるから, 補題 4.3 によって

$$\left(\sum_{k=1}^{n} |b_{kj}|^q\right)^{\frac{1}{q}} \leq \left(\sum_{k=1}^{n} |b_{kj}|^p\right)^{\frac{1}{p}} \tag{4.13}$$

(4.12) と (4.13) より

$$|c_{ij}| \leq \left(\sum_{k=1}^{n} |a_{ik}|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |b_{kj}|^p\right)^{\frac{1}{p}}$$

$$\therefore \sum_{i,j=1}^{n} |c_{ij}|^p \leq \left(\sum_{i=1}^{n}\sum_{k=1}^{n} |a_{ik}|^p\right) \left(\sum_{j=1}^{n}\sum_{k=1}^{n} |b_{kj}|^p\right)$$

$$\therefore \Phi_p(AB) = \Phi_p(C) \leq \Phi_p(A)\Phi_p(B)$$

以上によって $\Phi_p(A)$ は行列ノルムであるが, さらに $\boldsymbol{x} \in \mathbb{F}^n$ に対して

$$\begin{aligned}
\|A\boldsymbol{x}\|_p^p &= \sum_{i=1}^{n} \left|\sum_{j=1}^{n} a_{ij}x_j\right|^p \leq \sum_{i=1}^{n} \left(\sum_{j=1}^{n} |a_{ij}| \cdot |x_j|\right)^p \\
&\leq \sum_{i=1}^{n} \left\{ \left(\sum_{j=1}^{n} |a_{ij}|^p\right) \left(\sum_{j=1}^{n} |x_j|^q\right)^{\frac{p}{q}} \right\} \quad \text{(補題 4.1 による)} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^p \cdot \|\boldsymbol{x}\|_q^p
\end{aligned}$$

$$\therefore \|A\boldsymbol{x}\|_p \leq \left(\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^p\right)^{\frac{1}{p}} \|\boldsymbol{x}\|_q = \Phi_p(A)\|\boldsymbol{x}\|_q \tag{4.14}$$

$q \geq p$ であるから再び補題 4.3 により $\|\boldsymbol{x}\|_q \leq \|\boldsymbol{x}\|_p$ である. よって (4.14) より $\|A\boldsymbol{x}\|_p \leq \Phi_p(A)\|\boldsymbol{x}\|_p$ すなわち $\Phi_p(A)$ は $\|\boldsymbol{x}\|_p$ と両立する行列ノルムである. 証明終 ∎

定理 4.3 $p>2$ のときおよび $0<p<1$ のとき $\Phi_p(A)$ は行列ノルムでない.

【証明】 (i) $p>2$ のとき

$$A=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \cdots & & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \quad (n\text{次})$$

とすれば $A^2=nA$ かつ $\Phi_p(A)=n^{\frac{2}{p}}$ であるから, $A=B$ のとき

$$\Phi_p(AB)=\Phi_p(A^2)=\Phi_p(nA)=n\Phi_p(A)=n\cdot n^{\frac{2}{p}}>\left(n^{\frac{2}{p}}\right)^2 \quad \left(\because\ 1>\frac{2}{p}\right)$$
$$=\Phi_p(A)^2$$

よって $\Phi_p(AB)\leq\Phi_p(A)\Phi_p(B)$ は成り立たず行列ノルムの条件 (iv) をみたさない.
(ii) $0<p<1$ のときは, $\alpha>1$ に対して成り立つ不等式

$$t^\alpha-1\geq\alpha(t-1) \quad (t>0) \quad (\text{等号は } t=1 \text{ のとき})$$

において, $\alpha=1/p$, $q=p/(p-1)<0$, $t=a/b$ $(a>0,\ b>0)$ とおけば $1/p+1/q=1$ で

$$\left(\frac{a}{b}\right)^{\frac{1}{p}}-1\geq\frac{1}{p}\left(\frac{a}{b}-1\right)$$

両辺に $b>0$ を乗じて

$$b\left(\frac{a}{b}\right)^{\frac{1}{p}}-b\geq\frac{b}{p}\left(\frac{a}{b}-1\right)=\frac{a}{p}-\frac{b}{p}$$
$$\therefore\ a^{\frac{1}{p}}b^{1-\frac{1}{p}}\geq\frac{a}{p}+\left(1-\frac{1}{p}\right)b=\frac{a}{p}+\frac{b}{q} \quad \therefore\ a^{\frac{1}{p}}b^{\frac{1}{q}}\geq\frac{a}{p}+\frac{b}{q}$$

$a^{\frac{1}{p}}$, $b^{\frac{1}{q}}$ を a, b でおきかえれば

$$ab\geq\frac{1}{p}a^p+\frac{1}{q}b^q \quad (\text{等号は } a^p=b^q \text{のとき})$$

これは (4.8) と逆向きの不等式である. よって補題 4.1 と 4.2 の証明から $a_i>0$, $b_i>0$, $0<p<1$ のとき, **Hölder の逆向き不等式**

$$\sum_{i=1}^n a_ib_i\geq\left(\sum_{i=1}^n a_i^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^n b_i^q\right)^{\frac{1}{q}} \quad \left(\text{等号は}\frac{a_1^p}{b_1^q}=\cdots=\frac{a_n^p}{b_n^q}\text{のとき}\right)$$

と **Minkowski の逆向き不等式**

$$\left(\sum_{i=1}^n(a_i+b_j)^p\right)^{\frac{1}{p}}\geq\left(\sum_{i=1}^n a_i^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^n b_i^p\right)^{\frac{1}{p}} \quad \left(\text{等号は}\frac{a_1}{b_1}=\cdots=\frac{a_n}{b_n}\text{のとき}\right)$$

が得られる. したがって $a_{ij}>0$, $b_{ij}>0$ なる行列 A,B に対しては, 等号条件 (a_{ij}/b_{ij} が一定) が成り立たなければ

$$\Phi_p(A+B)>\Phi_p(A)+\Phi_p(B)$$

となり, $\Phi_p(A)$ は行列ノルムの条件 (iv) をみたさない. 証明終 ∎

4.4 行列ノルムと固有値の関係

定義 4.1 (スペクトル半径) n 次行列 $A = (a_{ij})$ の固有値を $\lambda_1, \ldots, \lambda_n$ とするとき $\rho(A) = \max_i |\lambda_i|$ を A の**スペクトル半径 (spectral radius)** という.

定理 4.4 ベクトルノルム $\|\boldsymbol{x}\|$ と両立する任意の行列ノルム $\|A\|$ に対して

$$\rho(A) \leq \|A\|$$

【証明】 $A\boldsymbol{x} = \lambda\boldsymbol{x}$, $\boldsymbol{x} \neq \boldsymbol{0}$ とすれば, 両辺のベクトルノルムをとって

$$\|A\boldsymbol{x}\| = |\lambda| \cdot \|\boldsymbol{x}\|$$

$\|A\boldsymbol{x}\| \leq \|A\| \cdot \|\boldsymbol{x}\|$ であるから $|\lambda| \cdot \|\boldsymbol{x}\| \leq \|A\| \cdot \|\boldsymbol{x}\|$
$\boldsymbol{x} \neq \boldsymbol{0}$ より $\|\boldsymbol{x}\| > 0$ であるから上式の両辺を $\|\boldsymbol{x}\|$ で割って $|\lambda| \leq \|A\|$

$$\therefore \quad \rho(A) \leq \|A\|$$

証明終 ∎

系 4.4.1 次の評価が成り立つ

(i) $\rho(A) \leq \|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$

(ii) $\rho(A) \leq \|A\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$

(iii) $\rho(A) \leq \|A\|_2 = \alpha_1 = \sqrt{\rho(A^*A)}$

定理 4.5 A を n 次行列とする. 任意に与えられた正数 ε に対して

$$\|A\| \leq \rho(A) + \varepsilon$$

をみたす行列ノルムが存在する.

【証明】 適当な正則行列 V により A を Jordan の標準形に変換し

$$V^{-1}AV = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix}, \quad J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \quad \text{または } J_i = (\lambda_i) \text{ (1 次)}$$

とあらわす. この両辺に右から

$$D = \begin{pmatrix} 1 & & & \\ & \varepsilon & & \\ & & \ddots & \\ & & & \varepsilon^{n-1} \end{pmatrix}$$

をかけ, 左から D^{-1} をかければ

$$D^{-1}(V^{-1}AV)D=\widehat{J}=\begin{pmatrix}\widehat{J}_1 & & \\ & \ddots & \\ & & \widehat{J}_m\end{pmatrix},\ \widehat{J}_i=\begin{pmatrix}\lambda_i & \varepsilon & & \\ & \ddots & \ddots & \\ & & \ddots & \varepsilon \\ & & & \lambda_i\end{pmatrix}\quad \text{または}\widehat{J}_i=\lambda_i$$

ここで $W=VD$ とおき $\|\boldsymbol{x}\|=\|W^{-1}\boldsymbol{x}\|_\infty$ とすれば $\|\boldsymbol{x}\|$ は1つのベクトルノルムであり，これに従属する行列ノルムは次をみたす．

$$\begin{aligned}\|A\|&=\max_{\|\boldsymbol{x}\|=1}\|A\boldsymbol{x}\|=\max_{\|W^{-1}\boldsymbol{x}\|_\infty=1}\|W^{-1}A\boldsymbol{x}\|_\infty=\max_{\|\boldsymbol{y}\|_\infty=1}\|W^{-1}AW\boldsymbol{y}\|_\infty\\&=\max_{\|\boldsymbol{y}\|_\infty=1}\|\widehat{J}\boldsymbol{y}\|_\infty=\|\widehat{J}\|_\infty\quad(\text{定理 4.1(i)})\\&=\max_{1\le i\le m}\|\widehat{J}_i\|_\infty\le\rho(A)+\varepsilon\end{aligned}$$

証明終 ∎

系 4.5.1 $\rho(A)<1$ ならば $\|A\|<1$ をみたす行列ノルムが存在する．

【証明】 $\rho(A)<1$ ならば $\rho(A)+\varepsilon<1$ をみたす $\varepsilon>0$ がある．このとき定理 4.5 の行列ノルムは $\|A\|<1$ をみたす．

証明終 ∎

系 4.5.2 (Oldenburger (オルデンブルガー)) $\lim_{p\to\infty}A^p=O\ \Leftrightarrow\ \rho(A)<1$

【証明】 $\lim_{p\to\infty}A^p=O$ ならば，$|\lambda|=\rho(A)$ をみたす固有値 λ と対応する固有ベクトル $\boldsymbol{x}\neq\boldsymbol{0}$ に対し $A\boldsymbol{x}=\lambda\boldsymbol{x}$，$A^p\boldsymbol{x}=\lambda^p\boldsymbol{x}$

$$\therefore\ \lambda^p=\frac{\boldsymbol{x}^*A^p\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{x}}\to0\quad(p\to\infty)\quad\therefore\ |\lambda|<1\quad\text{すなわち}\quad\rho(A)<1$$

逆に $\rho(A)<1$ とすれば系 4.5.1 より $\|A\|<1$ なる行列ノルムが存在する．このとき

$$\|A^p\|\le\|A\|^p\to0\quad(p\to\infty)\quad\text{よって}\quad\lim_{p\to\infty}A^p=O$$

証明終 ∎

系 4.5.3 ある行列ノルムで $\|A\|<1$ ならば $I-A$ は正則で

$$(I-A)^{-1}=I+A+A^2+\cdots\tag{4.15}$$

さらに $\|I\|=1$ ならば

$$\frac{1}{1+\|A\|}\le\|(I-A)^{-1}\|\le\frac{1}{1-\|A\|}\tag{4.16}$$

【証明】 $S_k=I+A+\cdots+A^k$ とおけば $k>l$ のとき

$$\begin{aligned}S_k-S_l&=A^{l+1}+\cdots+A^k\\\therefore\quad\|S_k-S_l\|&\le\|A^{l+1}\|+\cdots+\|A^k\|\le\|A\|^{l+1}+\cdots+\|A\|^k\\&\le\frac{\|A\|^{l+1}}{1-\|A\|}\to0\quad(k>l\to\infty)\end{aligned}$$

よって $\{S_k\}$ は Cauchy 列をなし，ある行列 B に収束する．$B=I+A+A^2+\cdots$ である．

このとき等式 $(I-A)S_k=I-A^{k+1}$ において $k\to\infty$ とすれば $(I-A)B=I$．よって $I-A$ は正則で

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

これで (4.15) が示された．次に $\|I\|=1$ を仮定すれば，

$$\|S_k\| \le 1 + \|A\| + \cdots + \|A\|^k$$

であることを注意する．

さて $(I-A)S_k = I - A^{k+1}$ より

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

$$\therefore \quad (I-A)^{-1} = (I-A)^{-1}A^{k+1} + S_k$$

$$\therefore \quad \|(I-A)^{-1}\| \le \|(I-A)^{-1}A^{k+1}\| + \|S_k\|$$

$$\le \|(I-A)^{-1}\| \cdot \|A\|^{k+1} + 1 + \|A\| + \cdots + \|A\|^k$$

$$\therefore \quad (1-\|A\|^{k+1})\|(I-A)^{-1}\| \le 1 + \|A\| + \cdots + \|A\|^k$$

$$\therefore \quad \|(I-A)^{-1}\| \le \frac{1}{1-\|A\|^{k+1}}(1 + \|A\| + \cdots + \|A\|^k) = \frac{1}{1-\|A\|} \tag{4.17}$$

また

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

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

$$\therefore \quad \frac{1}{1+\|A\|} \le \|(I-A)^{-1}\| \tag{4.18}$$

よって (4.17) と (4.18) より (4.16) が得られる． 証明終 ∎

4.5 行列関数

§3.6 で導いたように n 次行列 A は適当な正則行列 V により

$$V^{-1}AV = J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix} \tag{4.19}$$

と Jordan の標準形に変換される．ただし各 Jordan ブロック J_i は

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \quad \text{または } J_i = \lambda_i \ (1 \text{ 次})$$

の形で λ_i は A の固有値であるが $i \ne j$ でも $\lambda_i = \lambda_j$ は起こり得る (§3.6 の例参照)．J_i の次数を n_i 次とする．

Jordan ブロックの 1 つを

$$K = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \quad (l \text{ 次})$$

とすれば

$$K=\lambda I_l+N,\quad N=\begin{pmatrix}0&1&&\\&\ddots&\ddots&\\&&\ddots&1\\&&&0\end{pmatrix}\quad (l\text{ 次})$$

と書けて

$$N^2=\begin{pmatrix}0&0&1&&\\&\ddots&\ddots&\ddots&\\&&\ddots&\ddots&1\\&&&\ddots&0\\&&&&0\end{pmatrix},\quad N^3=\begin{pmatrix}0&0&0&1&&\\&\ddots&\ddots&\ddots&\ddots&\\&&\ddots&\ddots&\ddots&1\\&&&\ddots&\ddots&0\\&&&&\ddots&0\\&&&&&0\end{pmatrix},$$

$$\cdots,\ N^{l-1}=\begin{pmatrix}0&\cdots&0&1\\\vdots&\ddots&&0\\\vdots&&\ddots&\vdots\\0&\cdots&\cdots&0\end{pmatrix}$$

$$N^k=O\ (k\geq l)\quad (\text{したがって}\delta(N)=l)$$

$$\therefore\quad K^p=(\lambda I_l+N)^p=\lambda^p I_l+\binom{p}{1}\lambda^{p-1}N+\cdots+\binom{p}{p-1}\lambda N^{p-1}+N^p$$

$$\left(\text{ただし}\quad \binom{p}{j}=\frac{p!}{j!(p-j)!}=\frac{p(p-1)\cdots(p-j+1)}{j!}\right)$$

よって $p\geq l-1$ のとき

$$K^p=\begin{pmatrix}\lambda^p&\binom{p}{1}\lambda^{p-1}&\binom{p}{2}\lambda^{p-2}&\cdots&\binom{p}{l-1}\lambda^{p-l+1}\\&\lambda^p&\binom{p}{1}\lambda^{p-1}&\ddots&\vdots\\&&\ddots&\ddots&\binom{p}{2}\lambda^{p-2}\\&&&\ddots&\binom{p}{1}\lambda^{p-1}\\&&&&\lambda^p\end{pmatrix}\tag{4.20}$$

$p<l-1$ のとき

$$K^p=\begin{pmatrix}\lambda^p&\binom{p}{1}\lambda^{p-1}&\cdots&\binom{p}{p-1}\lambda&1&0&\cdots&0\\&\lambda^p&\binom{p}{1}\lambda^{p-1}&&\ddots&\ddots&\ddots&\vdots\\&&\ddots&\ddots&&\ddots&\ddots&0\\&&&\ddots&\ddots&&\ddots&1\\&&&&\ddots&\ddots&&\binom{p}{p-1}\lambda\\&&&&&\ddots&\ddots&\vdots\\&&&&&&\lambda^p&\binom{p}{1}\lambda^{p-1}\\&&&&&&&\lambda^p\end{pmatrix}\tag{4.21}$$

$\varphi(x)=x^p$ とおくとき

$$\frac{1}{k!}\varphi^{(k)}(x)=\binom{p}{k}x^{p-k}\quad(k\le p)\quad\text{かつ}\quad\varphi^{(k)}(x)=0\quad(k\ge p+1)$$

であるから (4.20) と (4.21) を 1 つにまとめて

$$K^p=\begin{pmatrix}\varphi(\lambda) & \frac{\varphi'(\lambda)}{1!} & \frac{\varphi''(\lambda)}{2!} & \cdots & \frac{\varphi^{(l-1)}(\lambda)}{(l-1)!}\\ & \varphi(\lambda) & \frac{\varphi'(\lambda)}{1!} & \ddots & \vdots\\ & & \ddots & \ddots & \frac{\varphi''(\lambda)}{2!}\\ & & & \ddots & \frac{\varphi'(\lambda)}{1!}\\ & & & & \varphi(\lambda)\end{pmatrix}$$

とあらわすことができる.

さて複素解析学の教えるところによれば複素平面上の開円板 $|z|<r$ で正則な関数 $f(z)$ はそこで任意回微分可能で

$$f(z)=\sum_{k=0}^{\infty}\frac{f^{(k)}(0)}{k!}z^k$$

と Taylor (テイラー) 級数に展開される.

このとき仮定 $\rho(A)<r$ の下で行列関数 $f(A)$ を

$$f(A)=\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)A^k$$

により定義する. (4.19) により $A=VJV^{-1}$, $A^k=VJ^kV^{-1}$ $(k\ge0)$ であるから

$$\begin{aligned}f(A)&=\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)(VJV^{-1})^k=V\Big(\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)J^k\Big)V^{-1}\\&=V\begin{pmatrix}\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)J_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)J_m^k\end{pmatrix}V^{-1}\\&=V\begin{pmatrix}F_1 & & \\ & \ddots & \\ & & F_m\end{pmatrix}V^{-1},\quad F_i=\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)J_i^k\ (1\le i\le m)\end{aligned}$$

ブロック対角要素 F_i は n_i 次正方行列で, $\varphi_k(x)=x^k$ とおけば上に示したことより

$$\begin{aligned}F_1&=\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)J_1^k\\&=\sum_{k=0}^{\infty}\frac{1}{k!}f^{(k)}(0)\begin{pmatrix}\varphi_k(\lambda_1) & \varphi_k'(\lambda_1) & \frac{1}{2!}\varphi_k''(\lambda_1) & \cdots & \frac{1}{(n_1-1)!}\varphi_k^{n_1-1}(\lambda_1)\\ & \varphi_k(\lambda_1) & \ddots & \ddots & \vdots\\ & & \ddots & \ddots & \vdots\\ & & & \ddots & \varphi_k'(\lambda_1)\\ & & & & \varphi_k(\lambda_1)\end{pmatrix}\end{aligned}$$

F_1 の $(1,j)$ 要素 F_{1j} は

$$\begin{aligned} F_{1j} &= \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(0) \frac{1}{(j-1)!} \varphi_k^{(j-1)}(\lambda_1) \\ &= \frac{1}{(j-1)!} \sum_{k=j-1}^{\infty} \frac{1}{k!} f^{(k)}(0) k(k-1)\cdots(k-j+2) \lambda_1^{k-j+1} \\ &= \frac{1}{(j-1)!} \sum_{k=j-1}^{\infty} \frac{f^{(k)}(0)}{(k-j+1)!} \lambda_1^{k-j+1} = \frac{1}{(j-1)!} f^{(j-1)}(\lambda_1) \end{aligned}$$

したがって

$$\begin{aligned} F_1 &= \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(0) J_1^k \\ &= \begin{pmatrix} f(\lambda_1) & f'(\lambda_1) & \frac{1}{2!} f''(\lambda_1) & \cdots \\ & f(\lambda_1) & f'(\lambda_1) & \ddots \\ & & \ddots & \ddots \\ & & & \ddots \end{pmatrix} \quad (n_1\text{次}) \end{aligned}$$

一般に

$$\begin{aligned} F_i &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} J_i^k \\ &= \begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{f''(\lambda_i)}{2!} & \cdots & \frac{1}{(n_i-1)!} f^{(n_i-1)}(\lambda_i) \\ & f(\lambda_i) & f'(\lambda_i) & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & f'(\lambda_i) \\ & & & & f(\lambda_i) \end{pmatrix} \quad (n_i\text{次}) \end{aligned}$$

例 1　$f(z) = e^z$ は $e^z = \sum_{k=0}^{\infty} \frac{1}{k!} z^k$ $(|z| < \infty)$ と展開されるから

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = V \begin{pmatrix} F_1 & & \\ & \ddots & \\ & & F_m \end{pmatrix} V^{-1}$$

$$F_i = \begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!} f''(\lambda_i) & \cdots & \frac{1}{(n_i-1)!} f^{(n_i-1)}(\lambda_i) \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & f'(\lambda_i) \\ & & & & f(\lambda_i) \end{pmatrix}$$

$$= e^{\lambda_i} \begin{pmatrix} 1 & 1 & \frac{1}{2!} & \cdots & \frac{1}{(n_i-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2!} \\ & & & \ddots & 1 \\ & & & & 1 \end{pmatrix}$$

したがって

$$e^{tA} = \sum_{k=0}^{\infty} \frac{1}{k!}(tA)^k = V \begin{pmatrix} F_1(t) & & \\ & \ddots & \\ & & F_m(t) \end{pmatrix} V^{-1}$$

ただし

$$\begin{aligned} F_i(t) &= \sum_{k=0}^{\infty} \frac{1}{k!}(tJ_i)^k \\ &= \begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{k!}(t\lambda_i)^k & \sum_{k=1}^{\infty} \frac{t}{k!}\binom{k}{1}(t\lambda_i)^{k-1} & \sum_{k=2}^{\infty} \frac{t^2}{k!}\binom{k}{2}(t\lambda_i)^{k-2} & \cdots \\ \ddots & \ddots & \ddots & \ddots \\ & \ddots & \ddots & \ddots \end{pmatrix} \\ &= e^{\lambda_i t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{n_i-1}}{(n_i-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2!} \\ & & & \ddots & t \\ & & & & 1 \end{pmatrix} \end{aligned}$$

例 2 $u(t) = e^{tA}\boldsymbol{c} = (u_1(t), \ldots, u_n(t))^{\mathrm{t}}$ は n 元連立 1 階線形常微分方程式の初期値問題

$$\begin{aligned} \frac{du_1}{dt} &= \sum_{j=1}^{n} a_{1j}u_j \\ &\vdots \\ \frac{du_n}{dt} &= \sum_{j=1}^{n} a_{nj}u_j, \quad t > 0 \\ u_1(0) &= c_1, \ldots, u_n(0) = c_n \end{aligned}$$

すなわち

$$\frac{d\boldsymbol{u}}{dt} = A\boldsymbol{u}, \quad t > 0 \tag{4.22}$$

$$\boldsymbol{u}(0) = \boldsymbol{c} = (c_1, \ldots, u_n)^{\mathrm{t}} \tag{4.23}$$

の解である．実際

$$u(t) = e^{tA}\boldsymbol{c} = \left(I + tA + \frac{1}{2!}(tA)^2 + \cdots\right)\boldsymbol{c} = \boldsymbol{c} + tA\boldsymbol{c} + \frac{1}{2!}t^2A^2\boldsymbol{c} + \cdots$$

を項別微分すれば

$$\begin{aligned}\frac{d\boldsymbol{u}}{dt} &= A\boldsymbol{c} + tA^2\boldsymbol{c} + \cdots = A(I + tA + \frac{1}{2!}(tA)^2 + \cdots)\boldsymbol{c}\\ &= Ae^{tA}\boldsymbol{c} = A\boldsymbol{u}\end{aligned}$$

かつ $\boldsymbol{u}(0) = e^{O}\boldsymbol{c} = \boldsymbol{c}$ である．したがって A の Jordan 分解を正確に求めなくても上記連立 1 階線形常微分方程式に対する解の評価として

$$\begin{aligned}\|\boldsymbol{u}(t)\|_\infty = \|e^{tA}\boldsymbol{c}\|_\infty &\le \|e^{tA}\|_\infty \cdot \|\boldsymbol{c}\|_\infty \le \left(1 + t\|A\|_\infty + \frac{t^2}{2!}\|A\|_\infty^2 + \cdots\right)\|\boldsymbol{c}\|_\infty\\ &\le e^{t\|A\|_\infty}\|\boldsymbol{c}\|_\infty \quad (t > 0)\end{aligned}$$

を得る．また初期値を $\boldsymbol{c}$ から $\widetilde{\boldsymbol{c}}$ へ変化させたときの解 $\boldsymbol{u}(t)$ の変動は

$$\|\widetilde{\boldsymbol{u}}(t) - \boldsymbol{u}(t)\|_\infty \le e^{t\|A\|_\infty}\|\widetilde{\boldsymbol{c}} - \boldsymbol{c}\|_\infty \quad (t > 0)$$

と見積ることができる．($\because$ $\boldsymbol{u}(t) = e^{tA}\boldsymbol{c}$, $\widetilde{\boldsymbol{u}}(t) = e^{tA}\widetilde{\boldsymbol{c}}$ より $\widetilde{\boldsymbol{u}}(t) - \boldsymbol{u}(t) = e^{tA}(\widetilde{\boldsymbol{c}} - \boldsymbol{c})$) このような評価を**解の定性的評価**という．

なお，(4.22), (4.23) の解は $\boldsymbol{u} = e^{tA}\boldsymbol{c}$ 以外にない．

これは常微分方程式論においてよく知られた事実であるが，念のため，次節で行列の微分・積分を定義したあとで簡単な証明を与える．

4.6 行列の微分・積分

$A(t) = (a_{ij}(t)) \in \mathbb{F}^{m\times n}$ の各要素 $a_{ij}(t)$ が t につき微分可能ならば行列 $(a'_{ij}(t))$ が定義される．この行列を $dA(t)/dt$ または $A'(t)$ とかく．

すなわち

$$\frac{dA(t)}{dt} = A'(t) = \begin{pmatrix} a'_{11}(t) & \cdots & a'_{1n}(t)\\ & \cdots & \\ a'_{m1}(t) & \cdots & a'_{mn}(t)\end{pmatrix} \in \mathbb{F}^{m\times n}$$

特に $\boldsymbol{x}(t) = \big(x_1(t), \ldots, x_n(t)\big)^{\mathrm{t}} \in \mathbb{F}^{n\times 1}$ の各成分 $x_i(t)$ が t につき微分可能ならば

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{x}'(t) = \big(x'_1(t), \ldots, x'_n(t)\big)^{\mathrm{t}}$$

である．(実は前節例 2 においてこの記号を断りなく使った．)

同様に各 i, j につき積分 $\int_a^b a_{ij}(t)dt$ が存在するとき

$$\int_a^b A(t)dt = \begin{pmatrix} \int_a^b a_{11}(t)dt & \cdots & \int_a^b a_{1n}(t)dt\\ & \cdots & \\ \int_a^b a_{m1}(t)dt & \cdots & \int_a^b a_{mn}(t)dt\end{pmatrix}$$

が定義される．特にベクトル $\boldsymbol{x}(t) = \big(x_1(t), \ldots, x_n(t)\big)^{\mathrm{t}}$ の各成分 $x_i(t)$ が積分可能のとき

$$\int_a^b \boldsymbol{x}(t)dt = \Big(\int_a^b x_1(t)dt, \ldots, \int_a^b x_n(t)dt\Big)^{\mathrm{t}}$$

である．

これらは約束事であるが，第 10 章でみるように，このような表記により数学の記述が簡単になり議論の見通しがよくなるのである．

さて前節の終わりに述べた初期値問題 (4.22), (4.23) の解の一意性を示そう．上に述べた行列の積分の定義に基づいて (4.22) の両辺を 0 から t まで積分すれば

$$\begin{aligned}\boldsymbol{u}(t) &= \boldsymbol{c} + A\int_0^t \boldsymbol{u}(t_1)dt_1 = \boldsymbol{c} + A\int_0^t \Big\{\boldsymbol{c} + A\int_0^{t_1} \boldsymbol{u}(t_2)dt_2\Big\}dt_1 \\ &= \boldsymbol{c} + tA\boldsymbol{c} + A^2\int_0^t\int_0^{t_1} \Big(\boldsymbol{c} + A\int_0^{t_2} \boldsymbol{u}(t_3)dt_3\Big)dt_2dt_1 \\ &= \boldsymbol{c} + tA\boldsymbol{c} + \frac{t^2}{2}A^2\boldsymbol{c} + A^3\int_0^t\int_0^{t_1}\int_0^{t_2} \boldsymbol{u}(t_3)dt_3dt_2dt_1\end{aligned}$$

以下これをくり返せば

$$\boldsymbol{u}(t) = \boldsymbol{c} + tA\boldsymbol{c} + \frac{t^2}{2}A^2\boldsymbol{c} + \cdots = \sum_{k=0}^{\infty}\frac{(tA)^k}{k!}\boldsymbol{c} = e^{tA}\boldsymbol{c}$$

となる．この反復は **Picard (ピカール) 反復**または **Picard の逐次代入法**と呼ばれる反復の特別な場合である．

4.7 行列式の微分

$A(t) = (a_{ij}(t)) \in \mathbb{F}^{n\times n}$ かつ $a_{ij}(t)$ が t につき微分可能ならば

$$\frac{d}{dt}\big(\det A(t)\big) = \begin{vmatrix} a'_{11}(t) & \cdots & a'_{1n}(t) \\ a_{21}(t) & \cdots & a_{2n}(t) \\ a_{31}(t) & \cdots & a_{3n}(t) \\ & \cdots & \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{vmatrix} + \begin{vmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ a'_{21}(t) & \cdots & a'_{2n}(t) \\ a_{31}(t) & \cdots & a_{3n}(t) \\ & \cdots & \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{vmatrix}$$

$$+\cdots+ \begin{vmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ & \cdots & \\ a_{n-1\ 1}(t) & \cdots & a_{n-1\ n}(t) \\ a'_{n1}(t) & \cdots & a'_{nn}(t) \end{vmatrix}$$

これは行列式の定義式 (§1.3) を積の微分を用いて項列微分すれば得られる．

第5章　AB と BA の固有値

A, B を同じ次数の正方行列とすれば $AB \neq BA$ でも AB と BA の固有値は代数的重複度も込めて一致する．また非零固有値の幾何学的重複度も相等しい．さらに $AB = BA$ ならば A と B は共通のユニタリ行列 U により同時に上3角化される．本章では AB と BA の固有値に関するその他多くの興味ある性質を述べると共にその応用として半正定値 Hermite 行列に対する p 乗根の一意存在とその計算法を述べる．

5.1　AB と BA の固有値

A, B を n 次行列とすれば一般に $AB \neq BA$ であるが，両者の固有値は完全に等しく次が成り立つ．

定理 5.1　A, B を n 次行列とするとき AB と BA の固有値は代数的重複度も込めて等しい．

【証明】　いろいろな証明が知られているが以下の巧妙な証明は Zhang [21] に見い出される．

λ を変数とする等式

$$\begin{pmatrix} I_n & -A \\ O & \lambda I_n \end{pmatrix} \begin{pmatrix} \lambda I_n & A \\ B & I_n \end{pmatrix} = \begin{pmatrix} \lambda I_n - AB & O \\ \lambda B & \lambda I_n \end{pmatrix}$$

の両辺の行列式をとれば

$$\lambda^n \det \begin{pmatrix} \lambda I_n & A \\ B & I_n \end{pmatrix} = \det(\lambda I_n - AB) \cdot \lambda^n$$

これは λ に関する恒等式であるから

$$\det \begin{pmatrix} \lambda I_n & A \\ B & I_n \end{pmatrix} = \det(\lambda I_n - AB) \tag{5.1}$$

同様に等式

$$\begin{pmatrix} I_n & O \\ -B & \lambda I_n \end{pmatrix} \begin{pmatrix} \lambda I_n & A \\ B & I_n \end{pmatrix} = \begin{pmatrix} \lambda I_n & A \\ O & \lambda I_n - BA \end{pmatrix}$$

より

$$\det \begin{pmatrix} \lambda I_n & A \\ B & I_n \end{pmatrix} = \det(\lambda I_n - BA) \tag{5.2}$$

(5.1) と (5.2) より

$$\det(\lambda I_n - AB) = \det(\lambda I_n - BA)$$

よって AB と BA の特性多項式は等しく，固有値は (代数的) 重複度も込めて等しい．　証明終 ∎

系 5.1.1 $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times m}$ かつ一般性を失うことなく $m \geq n$ とする．このとき m 次行列 AB の固有値は n 次行列 BA の固有値に 0 を $m-n$ 個つけ加えたものに等しい．

【証明】 $m = n$ の場合は定理 5.1 で済んでいるから $m > n$ とする．このとき

$$\widetilde{A} = (A \overbrace{O}^{m-n}),\ \widetilde{B} = \begin{pmatrix} B \\ O \end{pmatrix}\Big\}_{m-n}$$

とおけば $\widetilde{A}$ と $\widetilde{B}$ は $m \times m$ 行列で $\widetilde{A}\widetilde{B} = AB$ かつ

$$\widetilde{B}\widetilde{A} = \left(\begin{array}{c|c} BA & \overbrace{O}^{m-n} \\ \hline O & O \end{array}\right)\begin{matrix}\}n \\ \}m-n\end{matrix}$$

$\widetilde{A}$ と $\widetilde{B}$ に定理 5.1 を適用すれば AB の固有値は BA の固有値に 0 を $m-n$ 個つけ加えたものに等しい．

証明終 ∎

定理 5.2 $A, B \in \mathbb{F}^{n\times n}$ とする．

(i) AB と BA の固有値は代数的重複度も込めて等しい．
(ii) AB と BA の非零固有値の幾何学的重複度は等しい．
(iii) 特に AB と BA が共に Hermite 行列ならばゼロ固有値の幾何学的重複度も等しい．

【証明】 (i) 定理 5.1 で済んでいる．
(ii) $\lambda \neq 0$ を AB の固有値とする．$\dim \mathcal{N}(\lambda I - AB) = r$ とし，$\mathcal{N}(\lambda I - AB)$ の基底 $\varphi_1, \ldots, \varphi_r$ をとれば $AB\varphi_i = \lambda\varphi_i$, $1 \leq i \leq r$．このとき $B\varphi_i \neq 0 \quad \forall i$ かつ $B\varphi_1, \ldots, B\varphi_r$ は 1 次独立である．実際

$$\sum_{i=1}^{r} c_i B\varphi_i = 0 \quad (c_1, \ldots, c_r \in \mathbb{F})$$

とすれば

$$\sum_{i=1}^{r} c_i\varphi_i = \sum_{i=1}^{r} c_i\Big(\frac{1}{\lambda}AB\varphi_i\Big) = \frac{1}{\lambda}A\Big(\sum_{i=1}^{r} c_i B\varphi_i\Big) = 0$$

$\varphi_1, \ldots, \varphi_r$ は 1 次独立であるから $c_i = 0$ $(1 \leq i \leq r)$．よって $B\varphi_1, \ldots, B\varphi_r$ は 1 次独立である．

さらに

$$BA(B\varphi_i) = B(AB\varphi_i) = B(\lambda\varphi_i) = \lambda B\varphi_i$$

よって λ は BA の固有値で，$B\varphi_i, \ldots, B\varphi_r$ は λ に対応する BA の固有ベクトルである．

$$\therefore \quad \dim \mathcal{N}(\lambda I - BA) \geq r = \dim \mathcal{N}(\lambda I - AB)$$

AB と BA の役割を入れかえれば

$$\dim \mathcal{N}(\lambda I - AB) \geq r = \dim \mathcal{N}(\lambda I - BA)$$
$$\therefore \quad \dim \mathcal{N}(\lambda I - AB) = \dim \mathcal{N}(\lambda I - BA)$$

よって (ii) が成り立つ．

(iii) 一般に Hermite 行列の任意の固有値 λ の代数的重複度と幾何学的重複度は一致するから，AB と BA が Hermite 行列ならば，(i) によって両者のゼロ固有値の幾何学的重複度は共通の代数的重複度と一致する．　証明終 ∎

注意　定理 5.2(iii) は AB または BA の 1 つが Hermite 行列でなければ一般に成り立たない．例えば $n=2$,

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

のとき $AB=\begin{pmatrix}0&0\\0&0\end{pmatrix}$, $BA=\begin{pmatrix}0&0\\2&0\end{pmatrix}$ で $\dim\mathcal{N}(AB)=2\neq\dim\mathcal{N}(BA)=1$.

この場合 AB は Hermite 行列 (実対称) であるが BA は Hermite 行列でない．

定理 5.3　A,B を (複素)n 次 Hermite 行列とする．$B\succ O$ (§2.7 定義 2.1 参照) ならば AB の固有値は実数で $\mathrm{In}(AB)=\mathrm{In}(A)$ である．ただし $\mathrm{In}(A)$ は A の慣性数 (定理 2.15) をあらわす．

【証明】　$B\succ O$ (正定値) であるから §3.2 注意によって $B=LL^*=R^*R$ (L は下 3 角，R は上 3 角) と分解できる．このとき定理 5.1 によって $AB=AR^*R=(AR^*)R$ の固有値は RAR^* の固有値に等しい．RAR^* は Hermite 行列であるから Sylvester の慣性則により $\mathrm{In}(RAR^*)=\mathrm{In}(A)$．よって $\mathrm{In}(AB)=\mathrm{In}(A)$．すなわち AB と A の正，負，零の固有値の個数はそれぞれ等しい．　証明終 ∎

系 5.3.1　定理 5.3 の仮定の下で $B^{-1}A$ の固有値は実数で $\mathrm{In}(B^{-1}A)=\mathrm{In}(A)$.

【証明】　$B^{-1}A$ の固有値は AB^{-1} の固有値に等しくかつ $B^{-1}\succ O$ であるから定理 5.3 から従う．　証明終 ∎

系 5.3.2　$A\succ O$, $B\succ O$ ならば AB の固有値は実数ですべて正である．

系 5.3.3　$A\succcurlyeq O$ かつ $B\succ O$ ならば AB の固有値は実数ですべて非負である．

5.2 可換な行列

定理 5.4　A,B は n 次行列で可換 ($AB=BA$) ならば A の任意の固有値 λ に対し，B の固有値 μ と λ,μ に対応する共通な固有ベクトル $\boldsymbol{x}\neq\boldsymbol{0}$ が存在する．すなわち $A\boldsymbol{x}=\lambda\boldsymbol{x}$, $B\boldsymbol{x}=\mu\boldsymbol{x}$, $\boldsymbol{x}\neq\boldsymbol{0}$ とできる．

【証明】　λ を A の任意の固有値とし

$$W_\lambda=\mathcal{N}(A-\lambda I_n)\quad(\lambda\text{に対応する固有空間})$$
$$\dim W_\lambda=r$$

とすれば $W_\lambda=\mathrm{span}\{\boldsymbol{x}_1,\dots,\boldsymbol{x}_r\}$ をみたす基底 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_r$ が存在する．

$X=(\boldsymbol{x}_1,\dots,\boldsymbol{x}_r)$ とすれば X は $n\times r$ 行列である．このとき次の 2 つの場合が考えられる．

(i) ある i につき $B\boldsymbol{x}_i=\boldsymbol{0}$ となる場合．この場合にはゼロは B の固有値で $\boldsymbol{x}_i$ は対応する B の固有ベクトルである．このとき $\mu=0$, $\boldsymbol{x}=\boldsymbol{x}_i$ は求めるものである．

(ii) $B\boldsymbol{x}_i \neq \boldsymbol{0}$ $\forall i$ ならば

$$A(B\boldsymbol{x}_i) = B(A\boldsymbol{x}_i) = B(\lambda \boldsymbol{x}_i) = \lambda B\boldsymbol{x}_i \quad \therefore\ B\boldsymbol{x}_i \in W_\lambda \quad \forall i$$

よって，各 i につき $B\boldsymbol{x}_i$ は $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r$ の 1 次結合としてあらわされる：

$$B\boldsymbol{x}_i = \sum_{j=1}^{r} c_{ji}\boldsymbol{x}_j = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r) \begin{pmatrix} c_{1i} \\ \vdots \\ c_{ri} \end{pmatrix} \tag{5.3}$$

ここで

$$\boldsymbol{y}_i = (c_{1i}, \ldots, c_{ri})^{\mathrm{t}},\ Y = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_r) \quad (r \text{ 次行列})$$

とおけば (5.3) より $BX = XY$ である．

いま μ を r 次正方行列 Y の 1 つの固有値とし，$\boldsymbol{z} = (d_1, \ldots, d_r)^{\mathrm{t}} \neq 0$ を対応する固有ベクトルとすれば

$$B(X\boldsymbol{z}) = XY\boldsymbol{z} = X(\mu \boldsymbol{z}) = \mu(X\boldsymbol{z}) \tag{5.4}$$

X の列 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r$ は 1 次独立であるから，$\boldsymbol{z} \neq \boldsymbol{0}$ より $X\boldsymbol{z} = \sum_{j=1}^{r} d_j \boldsymbol{x}_j \neq \boldsymbol{0}$

よって (5.4) より μ は B の固有値で $X\boldsymbol{z}$ は μ に対応する固有ベクトルである．しかも

$$X\boldsymbol{z} = \sum_{j=1}^{r} d_j \boldsymbol{x}_j \in \mathrm{span}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r\} = W_\lambda$$

よって $\boldsymbol{x} = X\boldsymbol{z}$ は A の固有値 λ に対応する固有ベクトルでもある． 証明終 ∎

定理 5.5 $AB = BA$ ならば適当なユニタリ行列 U が存在して U^*AU と U^*BU を同時に上 3 角行列にできる．

【証明】 行列の次数 n に関する帰納法による．$n = 1$ のときは明らかであるから $n > 1$ とし $n-1$ のとき定理が成り立つと仮定する．前定理によって A の任意の固有値 λ_1 に対し，B の固有値 μ_1 と

$$A\boldsymbol{x} = \lambda_1 \boldsymbol{x}, \quad B\boldsymbol{x} = \mu_1 \boldsymbol{x}$$

をみたす共通の固有ベクトル $\boldsymbol{x} \neq \boldsymbol{0}$ が存在する．$\boldsymbol{x}$ を第 1 列とするユニタリ行列を V とすれば

$$AV = V \begin{pmatrix} \lambda_1 & \boldsymbol{u} \\ \boldsymbol{0} & A_1 \end{pmatrix}, \quad BV = V \begin{pmatrix} \mu_1 & \boldsymbol{v} \\ \boldsymbol{0} & B_1 \end{pmatrix}$$

$(A_1, B_1$ は $n-1$ 次行列で $\boldsymbol{u}, \boldsymbol{v}$ は $n-1$ 次元行ベクトル$)$

$$\therefore\ (V^*AV)(V^*BV) = \begin{pmatrix} \lambda_1 & \boldsymbol{u} \\ \boldsymbol{0} & A_1 \end{pmatrix} \begin{pmatrix} \mu_1 & \boldsymbol{v} \\ \boldsymbol{0} & B_1 \end{pmatrix} = \begin{pmatrix} \lambda_1\mu_1 & \lambda_1\boldsymbol{v} + \boldsymbol{u}B_1 \\ \boldsymbol{0} & A_1B_1 \end{pmatrix}$$

一方

$$\begin{aligned}(V^*AV)(V^*BV) &= V^*(AB)V = V^*(BA)V = (V^*BV)(V^*AV) \\ &= \begin{pmatrix} \mu_1 & \boldsymbol{v} \\ \boldsymbol{0} & B_1 \end{pmatrix} \begin{pmatrix} \lambda_1 & \boldsymbol{u} \\ \boldsymbol{0} & A_1 \end{pmatrix} = \begin{pmatrix} \lambda_1\mu_1 & \mu_1\boldsymbol{u} + \boldsymbol{v}A_1 \\ \boldsymbol{0} & B_1A_1 \end{pmatrix}\end{aligned}$$

よって $A_1B_1 = B_1A_1$ であり，$n-1$ 次行列 A_1, B_1 に帰納法の仮定を用いて

$$U_1^*A_1U_1=\begin{pmatrix}\lambda_2 & & *\\ & \ddots & \\ & & \lambda_n\end{pmatrix},\quad U_1^*B_1U_1=\begin{pmatrix}\mu_2 & & *\\ & \ddots & \\ & & \mu_n\end{pmatrix}$$

をみたす $n-1$ 次ユニタリ行列 U_1 が存在する．このとき

$$U=V\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & U_1\end{pmatrix}$$

とおけば，U は n 次ユニタリ行列で

$$U^*AU=\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & U_1^*\end{pmatrix}V^*AV\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & U_1\end{pmatrix}=\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & U_1^*\end{pmatrix}\begin{pmatrix}\lambda_1 & \boldsymbol{u}\\ \mathbf{0} & A_1\end{pmatrix}\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & U_1\end{pmatrix}$$

$$=\begin{pmatrix}\lambda_1 & \boldsymbol{u}U_1\\ \mathbf{0} & U_1^*A_1U_1\end{pmatrix}=\begin{pmatrix}\lambda_1 & & & *\\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$$

同様に

$$U^*BU=\begin{pmatrix}\mu_1 & & & *\\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_n\end{pmatrix}$$

も得られ帰納法が完了する．　証明終 ∎

系 5.5.1　$AB=BA$ ならば A の固有値 λ_i と B の固有値 μ_i を適当な順序で並べかえて AB の固有値を $\lambda_i\mu_i$, $(1\leq i\leq n)$ とかくことができる．したがって A,B が可換な Hermite 行列で $A\succcurlyeq O$, $B\succcurlyeq O$ ならば $AB\succcurlyeq O$ である (これは定理 2.13 (iii) の改良である)．

系 5.5.2　A,B が可換な Hermite 行列ならば A,B は適当なユニタリ行列により同時に対角化できる．すなわち適当なユニタリ行列 U をえらんで

$$U^*AU=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix},\quad U^*BU=\begin{pmatrix}\mu_1 & & \\ & \ddots & \\ & & \mu_n\end{pmatrix}$$

とできる．A,B が実行列ならば U は直交行列でおきかえられる．

定理 5.6 (定理 5.5 の一般化)　$A_1,A_2\ldots,A_s$ を互いに可換な n 次行列とすれば適当なユニタリ行列 U をえらんで U^*A_iU $(1\leq i\leq s)$ を同時に上 3 角行列にすることができる．

【証明】　s と n に関する 2 重帰納法による．$s=1$ のときは明らかである (Schur の定理 (定理 2.3) にほかならない)．よって $s>1$ とし，$s-1$ 以下のとき定理は成り立つと仮定する．また $n=1$ のときは明らかであるから $n>1$ とし，$n-1$ のとき定理は成り立つとする．A_1 の 1 つの固有値を λ_1 とし，λ_1 に対応する固有空間を W_1 とする (すなわち $W_1=W_{\lambda_1}=\mathcal{N}(A_1-\lambda_1I_n)$)．

$W_1=X$ (全空間) ならば $A_1=\lambda_1I_n$ となり A_1 は除外してよいから $s-1$ の場合に帰する．よってこの場合は定理は成り立つ．

$W_1 \subsetneqq X$ ならば，定理 1.7 により $\dim W_1 < n$ で

$$A_i\boldsymbol{x} \in W_1 \quad (1 \le i \le s) \quad \forall \boldsymbol{x} \in W_1 \tag{5.5}$$
$$\left(\because \ A_1(A_i\boldsymbol{x}) = A_i(A_1\boldsymbol{x}) = A_i(\lambda_1\boldsymbol{x}) = \lambda_1(A_i\boldsymbol{x}) \right)$$

$\dim W_1 = n_1$ として X の基底 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ を

$$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n_1} \in W_1, \quad \boldsymbol{u}_{n_1+1}, \ldots, \boldsymbol{u}_n \in W_1^{\perp}$$

とする．(必要ならば Schmidt の直交化法を用いて) これらは正規直交系としてよい (各 W_1 と $W_1^{\perp}$ で直交化を行えばよい)．

このとき $V = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ として

$$A_iV = V \begin{pmatrix} A_{i1} & * \\ & A_{i2} \end{pmatrix} \quad (A_{i1} : n_1\text{次正方行列}),\ 1 \le i \le s$$

と書ける．($\because$ (5.5) によって $A_iW_1 \subseteq W_1$)

$$\therefore \quad V^*A_iV = \begin{pmatrix} A_{i1} & * \\ & A_{i2} \end{pmatrix}$$

$A_iA_j = A_jA_i$ であるから上式より $A_{i1}A_{j1} = A_{j1}A_{i1}$ および $A_{i2}A_{j2} = A_{j2}A_{i2}$ である．$1 \le n_1 < n$ であるから帰納法の仮定によって適当な n_1 次ユニタリ行列 V_1 をえらんで $V_1^*A_{i1}V_1$ $(1 \le i \le s)$ を同時に上 3 角行列にできる．また $1 \le n_2 = n - n_1 < n$ より n_2 次ユニタリ行列 V_2 をえらんで $V_2^*A_{i2}V_2$ $(1 \le i \le s)$ を同時に上 3 角行列にできる．

$$U = \begin{pmatrix} V_1 & \\ & V_2 \end{pmatrix}$$

とおけば U はユニタリ行列で

$$U^*A_iU = \begin{pmatrix} V_1^*A_{i1}V_1 & * \\ & V_2^*A_{i2}V_2 \end{pmatrix} \quad (1 \le i \le s)$$

は上 3 角行列である．よって帰納法によりすべての自然数 s, n につき定理が成り立つ．　証明終 ∎

系 5.6.1　n 次行列 $A_1, \ldots, A_s$ が互いに可換な Hermite 行列ならば適当な n 次ユニタリ行列 U により U^*A_iU $(1 \le i \le s)$ を同時に実対角行列にできる．(特に $A_1, \ldots, A_s$ が実対称行列ならば，ユニタリ行列 U は直交行列でおきかえられる．)

【証明】　定理 5.6 の証明において上 3 角行列を対角行列でおきかえ，対角成分が実数であることに注意すればよい．　証明終 ∎

系 5.6.2　n 次行列 $A_1, \ldots, A_s$ が互いに可換な正規行列であるための必要十分条件は適当な n 次ユニタリ行列 U をえらんで U^*A_iU $(1 \le i \le s)$ を対角行列にできることである．

5.3　半正定値 Hermite 行列の p 乗根

非負実数 a に対して p 乗根 $\sqrt[p]{a} \ge 0$ が一意に定義されるが，この事実は n 次の半正定値 Hermite 行列 A に対して一般化される．いまユニタリ行列 U をえらんで A を

$$U^*AU = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \Lambda$$

と対角化する．ただし λ_i は A の固有値で $\lambda_i \geq 0 \quad \forall i$ である．このとき p を $p \geq 2$ なる自然数として

$$B = U \begin{pmatrix} \lambda_1^{\frac{1}{p}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{p}} \end{pmatrix} U^* \tag{5.6}$$

とおけば B は半正定値 Hermite 行列で

$$B^p = U\Lambda U^* = A \tag{5.7}$$

をみたす．逆に (5.7) $(B^p = A)$ をみたす半正定値 Hermite 行列 B は (5.6) 以外に存在しない．これを示すために C を $C^p = A$ をみたす半正定値 Hermite 行列とすれば C と $C^p(= A)$ は交換可能であるから系 5.5.2 によって適当なユニタリ行列 V により

$$V^*CV = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} \quad \text{かつ} \quad V^*AV = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} = D$$

と同時に対角化される．$d_1, \ldots, d_n$ は A の固有値であるから $\lambda_1, \ldots, \lambda_n$ を並びかえたものであり，適当な置換行列 $P = P_\sigma$ (σは $(1, 2, \ldots, n)$ の置換) により $D = P\Lambda P^{\mathrm{t}}$ と書ける．このとき

$$V^*AV = P\Lambda P^{\mathrm{t}}$$

よって $W = VP$ とおけば

$$W^*AW = \Lambda \tag{5.8}$$

さらに

$$W^*CW = P^* \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} P = \begin{pmatrix} \widetilde{\gamma}_1 & & \\ & \ddots & \\ & & \widetilde{\gamma}_n \end{pmatrix}, \quad \widetilde{\gamma}_i = \lambda_{\sigma(i)} \tag{5.9}$$

より

$$(W^*CW)^p = W^*C^pW = W^*AW = \Lambda \tag{5.10}$$

(5.9) と (5.10) より $\quad \widetilde{\gamma}_i^p = \lambda_i \quad \forall i \quad \therefore\ \widetilde{\gamma}_i = \lambda_i^{\frac{1}{p}}$

よって (5.9) より

$$W^*CW = \begin{pmatrix} \lambda_1^{\frac{1}{p}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{p}} \end{pmatrix}$$

$$\therefore \quad C = W \begin{pmatrix} \lambda_1^{\frac{1}{p}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{p}} \end{pmatrix} W^* \tag{5.11}$$

一方 (5.7) と (5.8) より

$$W\Lambda = AW = B^pW = (U\Lambda U^*)W \quad \therefore\ U^*W\Lambda = \Lambda U^*W$$

ここで $T = U^*W = (t_{ij})$ とおけば上式の (i, j) 要素を比較して

$$t_{ij}\lambda_j = \lambda_i t_{ij} \tag{5.12}$$

このとき任意の正数 α に対して

$$t_{ij}\lambda_j^{\alpha} = \lambda_i^{\alpha} t_{ij} \tag{5.13}$$

(実際，$\lambda_j = \lambda_i$ ならば当然成り立つが，$\lambda_j \neq \lambda_i$ ならば (5.12) より $t_{ij} = 0$ となって (5.13) が成り立つのである．) 故に (5.13) で

$$\alpha = \frac{1}{p} \quad \text{とおけば} \quad t_{ij}\lambda_i^{\frac{1}{p}} = \lambda_i^{\frac{1}{p}} t_{ij}, \quad \forall i,j$$

したがって

$$U^*W \begin{pmatrix} \lambda_1^{\frac{1}{p}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{p}} \end{pmatrix} = \begin{pmatrix} \lambda_1^{\frac{1}{p}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{p}} \end{pmatrix} U^*W$$

$$\therefore\ W \begin{pmatrix} \lambda_1^{\frac{1}{p}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{p}} \end{pmatrix} W^* = U \begin{pmatrix} \lambda_1^{\frac{1}{p}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{p}} \end{pmatrix} U^*$$

(5.11) と (5.6) より上式は $C = B$ を意味する．

以上をまとめて次の定理がえられた．

定理 5.7 A を n 次半正定値 Hermite 行列とするとき，与えられた自然数 $p \geq 2$ に対し $B^p = A$ をみたす半正定値 Hermite 行列 B がただ 1 つ存在する．これを **A の p 乗根**といい $A^{\frac{1}{p}}$ または $\sqrt[p]{A}$ であらわす．その具体的な形は $A = U\Lambda U^*$ $\left(\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),\ U : \text{ユニタリ}\right)$ と分解するとき

$$A^{\frac{1}{p}} = U \begin{pmatrix} \lambda_1^{\frac{1}{p}} & & \\ & \ddots & \\ & & \lambda_n^{\frac{1}{p}} \end{pmatrix} U^* \tag{5.14}$$

である．

注意 $p = 2$ のとき A の p 乗根 $\sqrt[p]{A}$ を $\sqrt{A}$ と書き，A の**平方根**という．これは数の場合と同様である．

系 5.7.1 n 次行列 A の極分解を $A = HU$ $(H \succcurlyeq O$ かつ U はユニタリ$)$ とするとき，A が正規 $(AA^* = A^*A)$ であるための必要十分条件は $HU = UH$ が成り立つことである．

【証明】 必要条件：A を正規とすると

$$\begin{aligned} AA^* = A^*A &\Leftrightarrow (HU)(U^*H) = (U^*H)(HU) \\ &\Leftrightarrow H^2 = U^*H^2U \\ &\Leftrightarrow UH^2 = H^2U \end{aligned} \tag{5.15}$$

$H \succcurlyeq O$ であるから適当なユニタリ行列 V をえらんで

$$V^*HV = \begin{pmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_n \end{pmatrix} \ (= B \text{ とおく}), \quad \beta_i \geq 0,\ \forall i$$

とできる．ここに $\beta_1, \ldots, \beta_n$ は H の固有値である．$H = VBV^*$ を (5.15) に代入して

$$U(VB^2V^*) = VB^2V^*U \quad \therefore \quad (V^*UV)B^2 = B^2(V^*UV)$$

$W = V^*UV = (w_{ij})$ とおけば $\quad w_{ij}\beta_j^2 = \beta_i^2 w_{ij} \quad \forall i, j$

このとき (5.12) から (5.13) を導いたのと同様にして

$$w_{ij}\beta_j = \beta_i w_{ij} \quad \forall i, j$$

が結論される．

$$\therefore \quad WB = BW \quad \therefore \quad W(V^*HV) = (V^*HV)W \quad \therefore \quad UH = HU$$

十分条件：逆に $UH = HU$ のとき $AA^* = A^*A$ の成立は明らかである．　　証明終 ∎

5.4 p 乗根を求めるアルゴリズム

前節で，半正定値 Hermite 行列 A に対する p 乗根 $A^{\frac{1}{p}}$ の一意存在を示した．しかしながら，実際に $A^{\frac{1}{p}}$ を求めたい場合，(5.14) では A の固有値 $\lambda_1, \ldots, \lambda_n$ と対応する正規直交固有ベクトル $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ を必要とする $\left(\because\ U = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)\right)$．これらを用いない簡単な反復解法を以下に示す．

【$A^{\frac{1}{p}}$ の計算法】次の 2 つのステップによる．

ステップ 1. Frobenius ノルム $\|A\|_F = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}$ を計算し，$\widetilde{A} = A/\|A\|_F$ とおく．

ステップ 2. 行列 A_k $(k = 1, 2, \ldots)$ を次式によりつくる．

$$A_0 = O, \quad A_{k+1} = A_k + \frac{1}{p}(\widetilde{A} - A_k^p) \quad (k \geq 0) \tag{5.16}$$

このとき $A_0 \preccurlyeq A_1 \preccurlyeq \cdots \preccurlyeq A_k$ (定義 2.1 参照) で $k \to \infty$ のとき A_k は $\widetilde{A}^{\frac{1}{p}}$ に収束する (証明は後述する)．したがって

$$\|A\|_F^{\frac{1}{p}} \lim_{k\to\infty} A_k = \|A\|_F^{\frac{1}{p}} \widetilde{A}^{\frac{1}{p}} = A^{\frac{1}{p}}$$

となり，反復 (5.16) が十分収束したとき $\|A\|_F^{\frac{1}{p}} A_k$ が $A^{\frac{1}{p}}$ の近似行列となる．

各ステップの意味は次のようである．

【ステップの解説】

ステップ 1. (**行列の規格化**)： $\|A\|_2$ を A のスペクトルノルムとすれば

$$(A\boldsymbol{x},\ \boldsymbol{x}) \leq \|A\|_2(\boldsymbol{x},\ \boldsymbol{x}) \leq \|A\|_F(\boldsymbol{x},\ \boldsymbol{x}) \quad \forall \boldsymbol{x}$$

$$\therefore \quad (\widetilde{A}\boldsymbol{x},\ \boldsymbol{x}) = \left(\frac{A}{\|A\|_F}\boldsymbol{x},\ \boldsymbol{x}\right) \leq (\boldsymbol{x},\ \boldsymbol{x})$$

よって $\rho(\widetilde{A}) \leq 1$ で $\widetilde{A}$ の固有値 $\widetilde{\lambda}, \ldots, \widetilde{\lambda}_n$ は $|\widetilde{\lambda}_i| \leq 1 \quad \forall i$ をみたす．したがって $\widetilde{A} \preccurlyeq I$ である．

ステップ 2. (A_k **の収束**)： まず k に関する帰納法によって次を示す．

$$O = A_0 \preccurlyeq A_1 \preccurlyeq \cdots \preccurlyeq A_k \preccurlyeq I = I_n \tag{5.17}$$

(5.16) によって $A_1 = \frac{1}{p}\widetilde{A} \succcurlyeq O = A_0$ であり，ステップ 1 によって $\widetilde{A} \preccurlyeq I$ であるから $A_1 \preccurlyeq I$ は明ら

かである．よって (5.17) は $k=1$ のとき成り立つ．次に k のとき成り立つと仮定して $k+1$ のとき成り立つことを示そう．各 A_i $(i \le k)$ は $\widetilde{A}$ の多項式であるから $\widetilde{A}, A_0, A_1, \ldots, A_k$ は互いに可換 (交換可能) な Hermite 行列であり，適当なユニタリ行列 U により

$$U^*\widetilde{A}U = \begin{pmatrix} \widetilde{\lambda}_1 & & \\ & \ddots & \\ & & \widetilde{\lambda}_n \end{pmatrix}, \quad U^*A_iU = \begin{pmatrix} \widetilde{\lambda}_1^{(i)} & & \\ & \ddots & \\ & & \widetilde{\lambda}_n^{(i)} \end{pmatrix} \quad (0 \le i \le k) \tag{5.18}$$

と対角化される．ここで $O \preccurlyeq A_i \preccurlyeq I$ $(i \le k)$ を仮定すれば $0 \le \lambda_j^{(i)} \le 1$, $1 \le j \le n$ であり

$$\begin{aligned} I - A_{k+1} &= I - A_k - \frac{1}{p}(\widetilde{A} - A_k^p) \\ &= \frac{1}{p}\{A_k^p - pA_k + (p-1)I\} + \frac{1}{p}(I - \widetilde{A}) \\ &\succcurlyeq \frac{1}{p}\{A_k^p - pA_k + (p-1)I\} \quad (\because\ I - \widetilde{A} \succcurlyeq O) \end{aligned} \tag{5.19}$$

Hermite 行列 $H_k = A_k^p - pA_k + (p-1)I$ の固有値 $\mu_1^{(k)}, \ldots, \mu_n^{(k)}$ は

$$\mu_j^{(k)} = \lambda_j^{(k)p} - p\lambda_j^{(k)} + (p-1) \quad (1 \le j \le n)$$

であるが，$0 \le t \le 1$ のとき $f(t) = t^p - pt + p - 1 \ge 0$ $(\because\ f(0) = p-1 > 0,\ f'(t) = pt^{p-1} - p \le 0,\ f(1) = 0)$ であるから $\mu_j^{(k)} \ge 0$ $\forall j$ $\therefore\ H_k \succcurlyeq O$. よって (5.19) より

$$I - A_{k+1} = \frac{1}{p}H_k \succcurlyeq O \quad \therefore\ A_{k+1} \preccurlyeq I \tag{5.20}$$

また

$$\begin{aligned} A_{k+1} - A_k &= A_k - A_{k-1} - \frac{1}{p}(A_k^p - A_{k-1}^p) \\ &= A_k - A_{k-1} - \frac{1}{p}(A_k - A_{k-1})(A_k^{p-1} + A_k^{p-2}A_{k-1} + \cdots + A_kA_{k-1}^{p-1} + A_{k-1}^{p-1}) \\ &= B_k \cdot (A_k - A_{k-1}) \end{aligned} \tag{5.21}$$

ただし

$$B_k = I - \frac{1}{p}(A_k^{p-1} + A_k^{p-2}A_{k-1} + \cdots + A_kA_{k-1}^{p-1} + A_{k-1}^{p-1})$$

とおいた．仮定によって $A_k \preccurlyeq I$, $A_{k-1} \preccurlyeq I$ かつ A_k^{p-1-j} と A_k^j は可換であるから系 5.5.1 によって $A_k^{p-1-j}A_{k-1}^j \preccurlyeq I$ $(0 \le j \le p-1)$.

よって

$$B_k = \frac{1}{p}\{(I - A_k^{p-1}) + (I - A_k^{p-2}A_{k-1}) + \cdots + (I - A_{k-1}^{p-1})\} \succcurlyeq O$$

また $B_k \succcurlyeq O$, $A_k - A_{k-1} \succcurlyeq O$ は可換であるから $B_k(A_k - A_{k-1})$ は Hermite 行列で系 5.5.1 により $B_k(A_k - A_{k-1}) \succcurlyeq O$. したがって (5.21) より $A_{k+1} - A_k \succcurlyeq O$ となり (5.20) と併せて $A_k \preccurlyeq A_{k+1} \preccurlyeq I$ を得る．よって帰納法により (5.17) は任意の $k \ge 0$ に対して成り立つ．

次に $\{A_k\}$ の収束を示すために $C_k = I - A_k$ とおくと $k \le l$ のとき $A_k \preccurlyeq A_l$ であるから $C_k \succcurlyeq C_l$ すなわち $C_k - C_l \succcurlyeq O$. また $A_kA_l = A_lA_k$ より $C_k - C_l$, C_k, C_l は互いに可換な半正定値 Hermite 行列であり $(C_k - C_l)C_k \succcurlyeq O$ かつ $C_l(C_k - C_l) \succcurlyeq O$

$$\therefore\ C_k^2 \succcurlyeq C_kC_l = C_lC_k \succcurlyeq C_l^2 \quad (k \le l)$$

故に任意の n 次元列ベクトル $\boldsymbol{x}$ に対して

$$(C_k^2\boldsymbol{x},\ \boldsymbol{x}) \geq (C_kC_l\boldsymbol{x},\ \boldsymbol{x}) \geq (C_l^2\boldsymbol{x},\ \boldsymbol{x}) \geq 0 \tag{5.22}$$

$d_k = (C_k^2\boldsymbol{x},\ \boldsymbol{x})$ は単調減少非負実数であるから $k \to \infty$ のとき収束する．収束先を d とすれば (5.22) より

$$\lim_{l>k\to\infty} (C_kC_l\boldsymbol{x},\ \boldsymbol{x}) = d \geq 0$$

よって $l > k \to \infty$ のとき

$$\begin{aligned}\|C_l\boldsymbol{x} - C_k\boldsymbol{x}\|_2^2 &= \big((C_l - C_k)\boldsymbol{x},\ (C_l - C_k)\boldsymbol{x}\big) = (C_l\boldsymbol{x},\ C_l\boldsymbol{x}) - 2(C_l\boldsymbol{x},\ C_k\boldsymbol{x}) + (C_k\boldsymbol{x},\ C_k\boldsymbol{x}) \\ &= d_l - 2(C_l\boldsymbol{x},\ C_k\boldsymbol{x}) + d_k \\ &\to d - 2d + d = 0\end{aligned}$$

故に各 $\boldsymbol{x}$ につき $\{C_k\boldsymbol{x}\}$ は Cauchy 列をなし，$k \to \infty$ のとき収束する．

$\boldsymbol{x}$ として特に基本単位ベクトル $\boldsymbol{e}_j$ (n 次単位行列の第 j 列) をとれば C_k の第 j 列 $C_k\boldsymbol{e}_j$ は $k \to \infty$ のとき収束する．これは行列 C_k の各列が $k \to \infty$ のとき収束することを意味する．収束先を C とすれば $C_k = I - A_k$ より

$$\lim_{k\to\infty} A_k = I - \lim_{k\to\infty} C_k = I - C\ (= D \text{ とおく})$$

すると (5.16) で $k \to \infty$ として $D = D + \frac{1}{p}(\widetilde{A} - D^p) \quad \therefore\ \widetilde{A} = D^p$

結局 $D = \widetilde{A}^{\frac{1}{p}}$ が示されて

$$\lim_{k\to\infty} A_k = \widetilde{A}^{\frac{1}{p}}$$

$A = \|A\|_F\widetilde{A}$ であったから

$$A^{\frac{1}{p}} = \|A\|_F^{\frac{1}{p}}\widetilde{A}^{\frac{1}{p}} = \|A\|_F^{\frac{1}{p}} \lim_{k\to\infty} A_k$$

すなわち $\|A\|_F^{\frac{1}{p}}A_k$ は $k \to \infty$ のとき $A^{\frac{1}{p}}$ に収束する．

■**注意**　(5.16) は素朴な反復であって，その収束は早いとはいえない．さらに実用に足る高速反復解法は Newton 法 (山本 [6], [7]) の適用などを含めていろいろ考えられるであろう．読者の研究課題として残しておく．

5.5 Kantorovich の不等式と Rheinboldt の不等式

1948 年 Kantorovich (カントロヴィッチ) は次の不等式を証明した．この不等式はある種の反復解法の収束解析に用いられる (1 例として §12.5 定理 12.6 の証明参照)．

定理 5.8 (Kantorovich の不等式)　A は正定値 Hermite 行列で $O \prec mI \preccurlyeq A \preccurlyeq MI$ とするとき

$$(\boldsymbol{x},\ \boldsymbol{x})^2 \leq (A\boldsymbol{x},\ \boldsymbol{x})(A^{-1}\boldsymbol{x},\ \boldsymbol{x}) \leq \frac{(M+m)^2}{4Mm}(\boldsymbol{x},\ \boldsymbol{x})^2 \quad \forall\boldsymbol{x} \tag{5.23}$$

特に A の固有値を $\lambda_1 \geq \cdots \geq \lambda_n > 0$ とするとき $M = \lambda_1,\ m = \lambda_n$ ととることができる．

【証明 (Mond (1965) による)】　左側の不等式は

$$(\boldsymbol{x},\ \boldsymbol{x})^2 = (\sqrt{A}^{-1}\cdot\sqrt{A}\boldsymbol{x},\ \boldsymbol{x})^2 = (\sqrt{A}\boldsymbol{x},\ \sqrt{A}^{-1}\boldsymbol{x})^2 \le \|\sqrt{A}\boldsymbol{x}\|^2\cdot\|\sqrt{A}^{-1}\boldsymbol{x}\|^2$$
$$= (\sqrt{A}\boldsymbol{x},\ \sqrt{A}\boldsymbol{x})\cdot(\sqrt{A}^{-1}\boldsymbol{x},\ \sqrt{A}^{-1}\boldsymbol{x}) = (A\boldsymbol{x},\ \boldsymbol{x})(A^{-1}\boldsymbol{x},\ \boldsymbol{x})$$

より従う．次に右側の不等式を示そう．
$(A-mI)(A-MI)A^{-1}$ は半負定値 Hermite 行列 ($\because\ m\le t\le M$ のとき $(t-m)(t-M)\le 0$) であるから $\|\boldsymbol{x}\|_2=1$ なる任意の $\boldsymbol{x}\in\mathbb{C}^n$ に対し

$$\big((A-mI)(A-MI)A^{-1}\boldsymbol{x},\ \boldsymbol{x}\big)\le\boldsymbol{0}\quad\therefore\ (A\boldsymbol{x},\ \boldsymbol{x})+Mm(A^{-1}\boldsymbol{x},\ \boldsymbol{x})\le M+m$$

$t=Mm(A^{-1}\boldsymbol{x},\ \boldsymbol{x})$ とおき，上式両辺に $t\ (\ge 0)$ をかければ

$$t(A\boldsymbol{x},\ \boldsymbol{x})+t^2\le(M+m)t$$
$$\therefore\ t(A\boldsymbol{x},\ \boldsymbol{x})\le(M+m)t-t^2\le\frac{1}{4}(M+m)^2$$
$$\therefore\ (A\boldsymbol{x},\ \boldsymbol{x})(A^{-1}\boldsymbol{x},\ \boldsymbol{x})\le\frac{(M+m)^2}{4Mm}\quad\forall\boldsymbol{x},\ (\boldsymbol{x},\ \boldsymbol{x})=1$$

一般の $\boldsymbol{x}$ に対しては $\boldsymbol{x}\neq\boldsymbol{0}$ のとき $\boldsymbol{x}/\sqrt{(\boldsymbol{x},\ \boldsymbol{x})}$ を上式に代入して

$$\frac{1}{(\boldsymbol{x},\ \boldsymbol{x})^2}(A\boldsymbol{x},\ \boldsymbol{x})(A^{-1}\boldsymbol{x},\ \boldsymbol{x})\le\frac{(M+m)^2}{4Mm}$$
$$\therefore\ \ (A\boldsymbol{x},\ \boldsymbol{x})(A^{-1}\boldsymbol{x},\ \boldsymbol{x})\le\frac{(M+m)^2}{4Mm}(\boldsymbol{x},\ \boldsymbol{x})^2$$

この不等式は $\boldsymbol{x}=\boldsymbol{0}$ のときも成り立つ．　証明終 ∎

その後 1965 年 Rheinboldt (レインボルト) は次の不等式を発表した．

定理 5.9 (Rheinboldt の不等式)　A,B は可換な正定値 Hermite 行列で $O\prec m_1I\preccurlyeq A\preccurlyeq M_1I,\ O\prec m_2I\preccurlyeq B\preccurlyeq M_2I$ ならば

$$(A\boldsymbol{x},\ A\boldsymbol{x})(B\boldsymbol{x},\ B\boldsymbol{x})\le\frac{(M_1M_2+m_1m_2)^2}{4M_1M_2m_1m_2}(A\boldsymbol{x},\ B\boldsymbol{x})^2\quad\forall\boldsymbol{x}\qquad(5.24)$$

【証明】　$C=AB^{-1}$ とおけば $C^*=B^{-1^*}A^*=B^{-1}A=C$ ($\because\ AB=BA$)．よって C は Hermite 行列で

$$O\prec\frac{m_1}{M_2}I\preccurlyeq C\preccurlyeq\frac{M_1}{m_2}I$$

よって定理 5.8 より

$$(C\boldsymbol{x},\ \boldsymbol{x})(C^{-1}\boldsymbol{x},\ \boldsymbol{x})\le\frac{\left(\dfrac{M_1}{m_2}+\dfrac{m_1}{M_2}\right)^2}{4\left(\dfrac{m_1}{M_2}\right)\left(\dfrac{M_1}{m_2}\right)}(\boldsymbol{x},\ \boldsymbol{x})^2$$
$$=\frac{(M_1M_2+m_1m_2)^2}{4M_1M_2m_1m_2}(\boldsymbol{x},\ \boldsymbol{x})^2\quad\forall\boldsymbol{x}\qquad(5.25)$$

ここで $\boldsymbol{x}=\sqrt{AB}\boldsymbol{y}$ とおけば $AB=BA$ より

$$\sqrt{AB}=\sqrt{BA}=\sqrt{B}\sqrt{A}=\sqrt{A}\sqrt{B}\quad\text{かつ}\quad AB^{-1}=B^{-1}A$$

よって

$$(C\boldsymbol{x},\ \boldsymbol{x}) = (AB^{-1}\sqrt{AB}\boldsymbol{y},\ \sqrt{AB}\boldsymbol{y}) = (A\boldsymbol{y},\ A\boldsymbol{y})$$
$$(C^{-1}\boldsymbol{x},\ \boldsymbol{x}) = (BA^{-1}\sqrt{AB}\boldsymbol{y},\ \sqrt{AB}\boldsymbol{y}) = (B\boldsymbol{y},\ B\boldsymbol{y})$$

かつ

$$(\boldsymbol{x},\ \boldsymbol{x}) = (\sqrt{AB}\boldsymbol{y},\ \sqrt{AB}\boldsymbol{y}) = (A\boldsymbol{y},\ B\boldsymbol{y})$$

よって (5.25) より

$$(A\boldsymbol{y},\ A\boldsymbol{y})(B\boldsymbol{y},\ B\boldsymbol{y}) \leq \frac{(M_1M_2 + m_1m_2)^2}{4M_1M_2m_1m_2}(A\boldsymbol{y},\ B\boldsymbol{y})^2 \quad \forall \boldsymbol{y}$$

これは (5.24) にほかならない　　証明終 ∎

定理 5.10　Kantorovich の右側の不等式と Rheinboldt の不等式は同値である.

【証明】　定理 5.9 の証明を読み返してみれば, (5.23) の右側の不等式を用いて (5.24) が導き出されたのであった. 逆に (5.24) から (5.23) の右側の不等式を導くために $O \prec mI \preccurlyeq C \preccurlyeq MI$ なる Hermite 行列 C をとり $A = \sqrt{C},\ B = \sqrt{C^{-1}}$ とおくと

$$O \prec \sqrt{m}I \preccurlyeq A \preccurlyeq \sqrt{M}I, \qquad O \prec \frac{1}{\sqrt{M}}I \preccurlyeq B \preccurlyeq \frac{1}{\sqrt{m}}I$$

よって定理 5.9 より

$$(\sqrt{C}\boldsymbol{x},\ \sqrt{C}\boldsymbol{x})(\sqrt{C^{-1}}\boldsymbol{x},\ \sqrt{C^{-1}}\boldsymbol{x}) \leq \frac{\left(\frac{\sqrt{M}}{\sqrt{m}} + \frac{\sqrt{m}}{\sqrt{M}}\right)^2}{4\left(\frac{\sqrt{M}}{\sqrt{m}}\right)\left(\frac{\sqrt{m}}{\sqrt{M}}\right)}(\sqrt{C}\boldsymbol{x},\ \sqrt{C^{-1}}\boldsymbol{x})^2$$
$$\therefore\ (C\boldsymbol{x},\ \boldsymbol{x})(C^{-1}\boldsymbol{x},\ \boldsymbol{x}) \leq \frac{(M+m)^2}{4Mm}(\boldsymbol{x},\ \boldsymbol{x})^2$$

これは (5.23) の右側の不等式である.　　証明終 ∎

第6章　固有値の評価定理

n 次行列 A の固有値は n 次代数方程式 (特性方程式) $\det(\lambda I - A) = 0$ の根であるが，方程式を直接解かずに，根の存在範囲を A の要素から見積ることは応用上からもまたそれ自身としても興味ある問題であり，古来多くの研究がある．この章ではまずその代表的な結果を述べる．次に Hermite 行列の固有値に関するミニマックス定理，分離定理，摂動定理と正規行列に対する Hoffman-Wielandt の定理等を述べる．

6.1　Gerschgorin の定理とその周辺

定理 6.1 (Gerschgorin (ゲルシュゴリン) の定理)　n 次行列 $A = (a_{ij}) \in \mathbb{C}^n$ に対し

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \quad R_i = \left\{ z \in \mathbb{C} \;\middle|\; |z - a_{ii}| \leq r_i \right\}$$

とおけば，A の任意の固有値は閉円板 R_i の合併 $R = \bigcup_{i=1}^{n} R_i$ に属す．集合 R を連結な領域に分けるとき，その 1 つの連結成分 Γ が m 個の R_i からなれば，Γ 内には重複度も込めてちょうど m 個の固有値がある．

【証明】　λ を A の固有値，$\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} \neq \mathbf{0}$ を λ に対応する固有ベクトルとする．$\|\boldsymbol{x}\|_\infty = \max_i |x_i| = |x_k|$ とすれば $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ の第 k 成分は

$$(a_{kk} - \lambda)x_k + \sum_{\substack{j=1 \\ j \neq k}}^{n} a_{kj} \cdot x_j = 0$$

$$\therefore \quad |a_{kk} - \lambda||x_k| = \left| -\sum_{\substack{j=1 \\ j \neq k}}^{n} a_{kj} \cdot x_j \right| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}| \cdot |x_j| \leq r_k |x_k|$$

$|x_k| > 0$ であるから上式の両辺を $|x_k|$ で割って $|a_{kk} - \lambda| \leq r_k$．したがって　$\lambda \in R_k \subseteq R$

次に後半を示そう．そのために $D = \mathrm{diag}(a_{11}, \ldots, a_{nn})$, $E = A - D$ とおき行列 $A(t) = D + tE$, $0 \leq t \leq 1$ を考える．$A(t)$ に対する Gerschgorin 円板は

$$R_i(t) = \left\{ z \in \mathbb{C} \;\middle|\; |z - a_{ii}| \leq t r_i \right\}$$

であり，パラメータ t が 0 から 1 に連続的に変化するとき $A(t)$ は D から A に連続的に変化し，その固有値 $\lambda_i(t)$ は各 i につき $\lambda_i(0) = a_{ii}$ から $\lambda_i(1) = \lambda_i$ に至る連続曲線 C_i を描く．一般性を失うことなく R の連結成分の 1 つを $\Gamma = R_1 \cup \cdots \cup R_m$ とすれば $R_i(t) \subseteq R_i$ であるから各 t につき $\Gamma(t) = R_1(t) \cup \cdots \cup R_m(t)$ は $R_{m+1} \cup \cdots \cup R_n$ と交わることはない．したがって m 個の連続曲線 $C_1, \ldots, C_m$ は $\Gamma(1) = \Gamma$ を越えて外部に出ることはなく，$C_{m+1}, \ldots, C_n$ も Γ 内に入ることはない．故に A の固有値は重複度も込めてちょうど m 個が Γ 内にある．　証明終 ∎

系 6.1.1　行列 A が狭義優対角または既約強優対角で対角要素が正 (非負) ならば固有値の実部は正 (非負) である．

【証明】 対角要素がすべて正の狭義優対角行列の Gerschgorin 円板 R_i は右半平面 $(\mathrm{Re}z > 0)$ に属するから固有値の実部は正である．また既約強優対角行列 A の Gerschgorin 円板は，対角要素が正のとき，高々原点を除いて右半平面に属するから A の固有値 λ は $\lambda \neq 0$ ならば $\mathrm{Re}\lambda > 0$ である．また定理 1.21 により A は正則であるから $\lambda = 0$ は起こりえない．対角要素がすべて非負の場合も同様である．

証明終 ∎

系 6.1.2 n 次行列 $A = (a_{ij})$ が狭義優対角で $D = \mathrm{diag}(a_{11}, \ldots, a_{nn})$ とおくとき

$$\rho(I - D^{-1}A) < 1$$

【証明】 $\widetilde{A} = I - D^{-1}A = (\widetilde{a}_{ij})$ とおけば $\widetilde{a}_{ii} = 0$ かつ

$$\sum_{j=1}^{n} |\widetilde{a}_{ij}| = \sum_{j \neq i} |\widetilde{a}_{ij}| = \frac{1}{|a_{ii}|} \sum_{j \neq i} |a_{ij}| < 1 \quad \forall i$$

$$\therefore \quad \rho(\widetilde{A}) \leq ||\widetilde{A}||_\infty < 1$$

証明終 ∎

系 6.1.3 Gerschgorin の定理において r_i を $s_i = \sum_{j=1}^{n} |a_{ji}|$ でおきかえることができる．

【証明】 A^{t} に Gerschgorin の定理を適用すればよい．

証明終 ∎

定理 6.2 (Taussky (タウスキー)) Gerschgorin の定理において A は既約であるとする．もし A の 1 つの固有値 λ が Gerschgorin 円板の合併集合 $R = \bigcup_{i=1}^{n} R_i$ の境界上にあるならば，すべての円 $|z - a_{ii}| = r_i$ $(1 \leq i \leq n)$ は λ を通る．

【証明】 $A\boldsymbol{x} = \lambda\boldsymbol{x}$, $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} \neq \boldsymbol{0}$ とし，一般性を失うことなく

$$|x_1| = 1 \geq |x_2| \geq \cdots \geq |x_n|$$

とすれば

$$|\lambda - a_{11}| \leq \sum_{j \neq 1} |a_{1j}||x_j| \leq \sum_{j=2}^{n} |a_{1j}| = r_1$$

しかし λ は R の境界上にあるから $|\lambda - a_{11}| = r_1$ をみたす．故に上式は等号で

$$|\lambda - a_{11}| = \sum_{j=2}^{n} |a_{1j}||x_j| = \sum_{j=2}^{n} |a_{1j}| = r_1$$

よって $a_{1j} \neq 0$ なる j (≥ 2) に対して $|x_j| = 1$ でなければならない．

A は既約であるから $a_{1p} \neq 0$ なる p $(\neq 1)$ は必ず存在し，このとき $|x_p| = 1$ である．よって λ は

$$|\lambda - a_{pp}| \leq \sum_{\substack{j=1 \\ j \neq p}}^{n} |a_{pj}||x_j| \leq \sum_{j \neq p} |a_{pj}| = r_p$$

をみたす．λ は R の境界上の点であるから $|\lambda - a_{pp}| = r_p$．したがって上式は等号で

$$|\lambda - a_{pp}| = \sum_{j \neq p}^{n} |a_{pj}||x_j| = \sum_{j \neq p}^{n} |a_{pj}| = r_p$$

A の既約性により $a_{pq} \neq 0$ $(q \neq p)$ なる要素は存在し，このとき $|x_q| = 1$ である．等々．

A は既約であるから任意の j に対して

$$a_{1p_1} \neq 0, a_{p_1p_2} \neq 0, \ldots, a_{p_kj} \neq 0$$

なる $p_1, p_2, \ldots, p_k$ が存在する (定理 1.20). よって上記議論をくり返して $|x_j| = 1$ かつ $|\lambda - a_{jj}| = r_j$ を得る. j は任意であったから $|x_1| = \cdots = |x_n| = 1$ かつ λ は $|\lambda - a_{ii}| = r_i \ \forall_i$ をみたす. 証明終 ∎

注意 この定理から既約強優対角行列の正則性 (定理 1.21) に対する別証明ができる. 実際 仮に A が正則でないとすれば $\lambda = 0$ は A の固有値で $R = \bigcup_{i=1}^n R_i$ の境界上にある. したがって $\lambda = 0$ は $|\lambda - a_{ii}| = r_i$, $i = 1, 2, \cdots, n$ をみたす. しかし強優対角であるから少なくとも 1 つの i につき $|a_{ii}| > r_i$ であり, これは矛盾である.

定理 6.3 (Ostrowski (オストロフスキーまたはオストロウスキー))

$$r_i = \sum_{j \neq i} |a_{ij}|, \quad s_i = \sum_{j \neq i} |a_{ji}|, \quad 0 \leq \theta \leq 1$$

$$\widetilde{R}_i = \left\{ z \in \mathbb{C} \;\middle|\; |z - a_{ii}| \leq r_i^{\theta} s_i^{1-\theta} \right\}, \quad 1 \leq i \leq n$$

とおけば A のすべての固有値は閉円板 $\widetilde{R}_i$ の合併 $\widetilde{R} = \bigcup_{i=1}^n \widetilde{R}_i$ に属する. $\widetilde{R}$ の連結成分の 1 つ $\widetilde{\Gamma}$ が m 個の $\widetilde{R}_i$ からなれば $\widetilde{\Gamma}$ 内には重複度も込めてちょうど m 個の固有値がある.

【証明】 ([14] 参照) $\theta = 1$ と $\theta = 0$ の場合は A および A^{t} に対する Gerschgorin の定理にほかならないから $0 < \theta < 1$ としてよい. またある i につき $r_i = 0$ ならば a_{ii} は A の固有値でその他の固有値は A の第 i 行 i 列を除いてえられる $n-1$ 次行列の固有値であるから $r_i \neq 0 \quad \forall i$ としてよい.

このとき $A\boldsymbol{x} = \lambda \boldsymbol{x}$, $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} \neq \boldsymbol{0}$ とすれば $(\lambda - a_{ii})x_i = \sum_{j \neq i} a_{ij} x_j$ より

$$|\lambda - a_{ii}| \cdot |x_i| \leq \sum_{j \neq i} |a_{ij}||x_j|$$

仮に $\lambda \notin \widetilde{R} = \bigcup_{i=1}^n \widetilde{R}_i$ とすれば $\lambda \notin \widetilde{R}_i \ \forall i$, すなわち $|\lambda - a_{ii}| > r_i^{\theta} s_i^{1-\theta} \quad \forall i$

$$\therefore \quad r_i^{\theta} s_i^{1-\theta} |x_i| \leq |\lambda - a_{ii}||x_i| \quad \forall i \quad (\ |x_i| = 0 \text{ の場合も考慮している.}) \tag{6.1}$$

$$\therefore \quad r_i^{\theta} s_i^{1-\theta} |x_i| \leq \sum_{j \neq i} |a_{ij}||x_j| = \sum_{j \neq i} |a_{ij}|^{\theta} |a_{ij}|^{1-\theta} |x_j| \tag{6.2}$$

$$\leq \left\{ \sum_{j \neq i} \left(|a_{ij}|^{\theta} \right)^{\frac{1}{\theta}} \right\}^{\theta} \left\{ \sum_{j \neq i} \left(|a_{ij}|^{1-\theta} |x_j| \right)^{\frac{1}{1-\theta}} \right\}^{1-\theta} \tag{6.3}$$

$$= \left(\sum_{j \neq i} |a_{ij}| \right)^{\theta} \left(\sum_{j \neq i} |a_{ij}||x_j|^{\frac{1}{1-\theta}} \right)^{1-\theta} = r_i^{\theta} \left(\sum_{j \neq i} |a_{ij}||x_j|^{\frac{1}{1-\theta}} \right)^{1-\theta}$$

ただし (6.2) から (6.3) へ移るとき Hölder の不等式 (4.7) を $p = 1/\theta$, $q = 1/(1-\theta)$ として用いた.

$$\therefore \quad s_i^{1-\theta} |x_i| \leq \left(\sum_{j \neq i} |a_{ij}||x_j|^{\frac{1}{1-\theta}} \right)^{1-\theta} \quad (\because \ r_i^{\theta} > 0)$$

$$\therefore \quad s_i |x_i|^{\frac{1}{1-\theta}} \leq \sum_{j \neq i} |a_{ij}||x_j|^{\frac{1}{1-\theta}} \tag{6.4}$$

ただし $|x_i| \neq 0$ ならば (6.1) の不等号 $\leq$ は狭義不等号 $<$ でおきかえられ, そのとき (6.4) の不等号 $\leq$ も狭義不等号 $<$ でおきかえられる.

したがって少なくとも 1 つの i につき (6.4) は狭義の不等式である.

$$\therefore \quad \sum_{i=1}^{n} s_i|x_i|^{\frac{1}{1-\theta}} < \sum_{i=1}^{n}\Big(\sum_{j\neq i}|a_{ij}||x_j|^{\frac{1}{1-\theta}}\Big)$$

$$= \sum_{i=1}^{n}\Big(\sum_{\substack{i=1\\ i\neq j}}^{n}|a_{ij}||x_j|^{\frac{1}{1-\theta}}\Big) = \sum_{i=1}^{n}|x_j|^{\frac{1}{1-\theta}}\Big(\sum_{\substack{i=1\\ i\neq j}}^{n}|a_{ij}|\Big) = \sum_{j=1}^{n}|x_j|^{\frac{1}{1-\theta}}s_j$$

これは矛盾であるから $\lambda \in \widetilde{R} = \bigcup_{i=1}^{n}\widetilde{R}_i$. 結局すくなくとも 1 つの i につき

$$|\lambda - a_{ii}| \leq r_i^{\theta}s_i^{1-\theta}$$

が成り立つ. (この結論に至る Ostrowski の証明はまるで手品をみているようである.)

後半は Gerschgorin の定理の証明と同じ議論をくり返せばよい. 証明終 ∎

系 6.3.1 $0 \leq \theta \leq 1$ かつ $|a_{ii}| > \theta r_i + (1-\theta)s_i$ $\forall i$ ならば A は正則である.

【証明】 一般化された相加平均と相乗平均の関係 ((4.8) 参照) によって $\theta r_i + (1-\theta)s_i \geq r_i^{\theta}s_i^{1-\theta}$ $\forall i$ であるから系 6.3.1 の仮定の下では閉円板 $\widetilde{R}_i$ の合併 $\widetilde{R}$ は原点を含まない. すなわち A は零固有値をもたない. 証明終 ∎

6.2 Hermite 形式と基本不等式

A を n 次の (複素)Hermite 行列とすれば適当なユニタリ行列 U をえらんで

$$U^*AU = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \Lambda \tag{6.5}$$

と対角化できる. この場合 U の各列 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ は $\lambda_1, \ldots, \lambda_n$ に対応する正規直交固有ベクトルである.

$\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$, $\boldsymbol{y} = (y_1, \ldots, y_n)^{\mathrm{t}}$, $\boldsymbol{x} = U\boldsymbol{y}$ とおいて

$$\boldsymbol{x}^*A\boldsymbol{x} = (A\boldsymbol{x},\ \boldsymbol{x}) = (U\Lambda U^*\boldsymbol{x},\ \boldsymbol{x})$$

$$= (U\Lambda U^*U\boldsymbol{y},\ U\boldsymbol{y}) = (\Lambda\boldsymbol{y},\ \boldsymbol{y}) = \sum_{i=1}^{n}\lambda_i|y_i|^2$$

よって $(\boldsymbol{x},\ \boldsymbol{x}) = (\boldsymbol{y},\ \boldsymbol{y})$ に注意して

$$(\max_i \lambda_i)\sum_{i=1}^{n}|x_i|^2 = (\max_i \lambda_i)\sum_{i=1}^{n}|y_i|^2 \geq \sum_{i=1}^{n}\lambda_i|y_i|^2 = (A\boldsymbol{x},\ \boldsymbol{x}) \geq (\min_i \lambda_i)\sum_{i=1}^{n}|y_i|^2$$

$$= (\min_i \lambda_i)\sum_{i=1}^{n}|x_i|^2 \tag{6.6}$$

定理 6.4 A を n 次 Hermite 行列, $\lambda_1, \ldots, \lambda_n$ をその固有値とすれば任意の $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ に対して次が成り立つ.

$$|(A\boldsymbol{x},\ \boldsymbol{y})|^2 \leq (\max_i |\lambda_i|)^2(\boldsymbol{x},\ \boldsymbol{x})(\boldsymbol{y},\ \boldsymbol{y}) \tag{6.7}$$

かつ

$$(\min_i \lambda_i)(\boldsymbol{x},\ \boldsymbol{x}) \leq (A\boldsymbol{x},\ \boldsymbol{x}) \leq (\max_i \lambda_i)(\boldsymbol{x},\ \boldsymbol{x}) \tag{6.8}$$

【証明】 (6.8) は (6.6) ですでに示されているから，(6.7) を示せばよい．$\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$, $\boldsymbol{y} = (y_1, \ldots, y_n)^{\mathrm{t}} \in \mathbb{C}^n$ に対して

$$(A\boldsymbol{x},\ \boldsymbol{y}) = (U\Lambda U^*\boldsymbol{x},\ \boldsymbol{y}) = (\Lambda U^*\boldsymbol{x},\ U^*\boldsymbol{y})$$

$$\boldsymbol{z} = U^*\boldsymbol{x} = (z_1, \ldots, z_n)^{\mathrm{t}},\ \boldsymbol{w} = U^*\boldsymbol{y} = (w_1, \ldots, w_n)^{\mathrm{t}}$$

とおけば $(\boldsymbol{z}, \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{x})$, $(\boldsymbol{w}, \boldsymbol{w}) = (\boldsymbol{y}, \boldsymbol{y})$ に注意して

$$(\Lambda U^*\boldsymbol{x},\ U^*\boldsymbol{y}) = (\Lambda\boldsymbol{z},\ \boldsymbol{w}) = \sum_{i=1}^{n} \lambda_i z_i w_i$$

$$\therefore\quad |(A\boldsymbol{x},\ \boldsymbol{y})|^2 = \left|\sum_{i=1}^{n} \lambda_i z_i w_i\right|^2 \le \left(\sum_{i=1}^{n} |\lambda_i||z_i||w_i|\right)^2$$

$$\le \left(\max_i |\lambda_i|\right)^2 \sum_{i=1}^{n} |z_i|^2 \sum_{i=1}^{n} |w_i|^2 = (\max_i |\lambda_i|)^2 (\boldsymbol{x}, \boldsymbol{x})(\boldsymbol{y}, \boldsymbol{y})$$

故に (6.7) が示された． 証明終 ∎

系 6.4.1 (Parker) 定理 6.4 において $A = (a_{ij})$ とするとき

$$\frac{1}{2}(\max_i \lambda_i - \min_i \lambda_i) \ge \max_{\substack{i,j \\ i \ne j}} |a_{ij}|$$

【証明】 c を任意の実数とし，$\boldsymbol{x} = \boldsymbol{e}_j$, $\boldsymbol{y} = \boldsymbol{e}_i$ $(i \ne j)$ ($\boldsymbol{e}_j$, $\boldsymbol{e}_i$ は I_n の第 j 列と第 i 列) とすれば $\big((A - cI)\boldsymbol{x},\ \boldsymbol{y}\big) = a_{ij}$ かつ $A - cI$ の固有値は $\lambda_i - c$, $i = 1, 2, \ldots, n$ であるから (6.7) によって

$$|a_{ij}|^2 \le \max_i (\lambda_i - c)^2 \quad (\because\ (\boldsymbol{x}, \boldsymbol{x}) = (\boldsymbol{y}, \boldsymbol{y}) = 1)$$

ここで $\lambda_{\max} = \max_i \lambda_i$, $\lambda_{\min} = \min_i \lambda_i$, $c = \dfrac{1}{2}(\lambda_{\max} + \lambda_{\min})$ とおけば

$$|\lambda_i - c| \le \frac{1}{2}(\lambda_{\max} - \lambda_{\min}) \quad \forall i$$

よって

$$|a_{ij}| \le \frac{1}{2}(\lambda_{\max} - \lambda_{\min}) \quad i \ne j \quad \therefore\quad \max_{i \ne j} |a_{ij}| \le \frac{1}{2}(\lambda_{\max} - \lambda_{\min})$$

証明終 ∎

▒**注意** A を一般の n 次行列としその固有値を $\lambda_1, \ldots, \lambda_n$ とするとき $S(A) = \max_{i,j} |\lambda_i - \lambda_j|$ を A の**ひろがり (spread)** という．この定義は Mirsky(1956) によるものであるが，

$$S(A) \le \sqrt{2\|A\|_F^2 - \frac{2}{n}|\mathrm{tr}(A)|^2}$$

が成り立つことが知られている (証明は略す)．系 6.4.1 によって A が Hermite 行列ならば $S(A) \ge 2\max_{i \ne j} |a_{ij}|$ であり，この評価は最良 (best possible) である (等号が成り立つ例がある)．

定理 6.5 (基本不等式) A を n 次 Hermite 行列，その固有値を $\lambda_1 \le \cdots \le \lambda_n$，対応する正規直交固有ベクトルを $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ とする．このとき

$$S = \mathrm{span}\{\boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_l\} \quad (1 \le k \le l \le n)$$

とすれば

$$\lambda_k(\boldsymbol{x}, \boldsymbol{x}) \le (A\boldsymbol{x}, \boldsymbol{x}) \le \lambda_l(\boldsymbol{x}, \boldsymbol{x}) \quad \forall \boldsymbol{x} \in S$$

【証明】 $\boldsymbol{x} = \sum_{j=k}^{l} c_j \boldsymbol{u}_j$ ($c_k, \ldots, c_l$ は定数) と書けば

$$(A\boldsymbol{x}, \boldsymbol{x}) = \Big(\sum_{j=k}^{l} c_i A\boldsymbol{u}_i,\ \sum_{j=k}^{l} c_j \boldsymbol{u}_j\Big) = \Big(\sum_{j=k}^{l} c_i \lambda_i \boldsymbol{u}_i,\ \sum_{j=k}^{l} c_j \boldsymbol{u}_j\Big) = \sum_{i=k}^{l}\sum_{j=k}^{l} c_i \bar{c}_j \lambda_i (\boldsymbol{u}_i,\ \boldsymbol{u}_j)$$

$$= \sum_{i=k}^{l} |c_i|^2 \lambda_i \qquad \left(\because\ (\boldsymbol{u}_i,\ \boldsymbol{u}_j) = \delta_{ij} = \begin{cases} 1 & (i=j) \\ 0 & (i \neq j) \end{cases}\right)$$

$(\boldsymbol{x}, \boldsymbol{x}) = \sum_{i=k}^{l} |c_i|^2$ であるから定理が成り立つ. 証明終 ∎

定理 6.6 A, B は n 次 Hermite 行列で $A \preccurlyeq B$ (定義 2.1 参照) とする. A, B の固有値を $\lambda_1 \leq \cdots \leq \lambda_n$, $\mu_1 \leq \cdots \leq \mu_n$ とすれば

$$\lambda_i \leq \mu_i, \quad 1 \leq i \leq n$$

(特に $A \prec B$ ならば $\lambda_i < \mu_i$, $1 \leq i \leq n$ である.)

【証明】

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i, \quad (\boldsymbol{u}_i,\ \boldsymbol{u}_j) = \delta_{ij} \qquad B\boldsymbol{v}_i = \mu_i \boldsymbol{v}_i, \quad (\boldsymbol{v}_i,\ \boldsymbol{v}_j) = \delta_{ij}$$

$$S_1 = \mathrm{span}\{\boldsymbol{u}_i, \boldsymbol{u}_{i+1}, \ldots, \boldsymbol{u}_n\} \qquad S_2 = \mathrm{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_i\}$$

とすれば次元等式 (定理 1.8) によって

$$\dim(S_1 \cap S_2) = \dim S_1 + \dim S_2 - \dim(S_1 + S_2)$$
$$\geq (n - i + 1) + i - n = 1$$

よって $\boldsymbol{x} \in S_1 \cap S_2$, $(\boldsymbol{x},\ \boldsymbol{x}) = 1$ なる $\boldsymbol{x}$ がある. このとき定理 6.5 によって, $\boldsymbol{x} \in S_1$ より $\lambda_i \leq (A\boldsymbol{x},\ \boldsymbol{x})$ かつ $\boldsymbol{x} \in S_2$ より $(B\boldsymbol{x},\ \boldsymbol{x}) \leq \mu_i$. また仮定 $A \preccurlyeq B$ によって $(A\boldsymbol{x},\ \boldsymbol{x}) \leq (B\boldsymbol{x},\ \boldsymbol{x})$. よって $\lambda_i \leq \mu_i \quad \forall i$ を得る. (特に $A \prec B$ ならば $(A\boldsymbol{x},\ \boldsymbol{x}) < (B\boldsymbol{x},\ \boldsymbol{x})$ であるから $\lambda_i < \mu_i \quad \forall i$ である.) 証明終 ∎

定理 6.7 (Fan-Hoffman (ファン・ホフマン)) $A \in \mathbb{C}^{n\times n}$ かつ A の特異値を $\alpha_1 \leq \cdots \leq \alpha_n$, $\frac{1}{2}(A + A^*)$ の固有値を $\lambda_1 \leq \cdots \leq \lambda_n$ とすれば

$$\lambda_i \leq \alpha_i, \quad i = 1, 2, \ldots, n$$

【証明】 定理 6.5 を応用する. $\frac{1}{2}(A + A^*)$ は Hermite 行列であるから固有値 λ_i は実数である. $\lambda_1, \ldots, \lambda_n$ に対応する正規直交固有ベクトルを $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ とし, A^*A の固有値 α_i^2 ($\alpha_i \geq 0$) に対応する正規直交固有ベクトルを $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ とする. このとき

$$S_1 = \mathrm{span}\{\boldsymbol{v}_i, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}_n\} \qquad S_2 = \mathrm{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_i\}$$

とすれば定理 1.8 によって

$$\dim(S_1 \cap S_2) = \dim S_1 + \dim S_2 - \dim(S_1 + S_2)$$
$$\geq (n - i + 1) + i - n = 1$$

よって $(\boldsymbol{x},\ \boldsymbol{x}) = 1$ なる元 $\boldsymbol{x} \in S_1 \cap S_2$ がある. このとき $\boldsymbol{x} \in S_1$ より

$$\lambda_i \le \Big(\frac{1}{2}(A+A^*)\boldsymbol{x},\ \boldsymbol{x}\Big) = \frac{1}{2}\big\{(A\boldsymbol{x},\ \boldsymbol{x}) + (A^*\boldsymbol{x},\ \boldsymbol{x})\big\}$$
$$= \frac{1}{2}\big\{(A\boldsymbol{x},\ \boldsymbol{x}) + (\boldsymbol{x},\ A\boldsymbol{x})\big\} = \mathrm{Re}(A\boldsymbol{x},\ \boldsymbol{x})$$
$$\le |(A\boldsymbol{x},\ \boldsymbol{x})| \le \|A\boldsymbol{x}\|_2\|\boldsymbol{x}\|_2 = \|A\boldsymbol{x}\|_2 = \sqrt{(A^*A\boldsymbol{x},\ \boldsymbol{x})} \tag{6.9}$$

また $\boldsymbol{x} \in S_2$ より

$$(A^*A\boldsymbol{x},\ \boldsymbol{x}) \le \alpha_i^2(\boldsymbol{x},\ \boldsymbol{x}) = \alpha_i^2$$
$$\therefore\ \sqrt{(A^*A\boldsymbol{x},\ \boldsymbol{x})} \le \alpha_i \tag{6.10}$$

(6.9) と (6.10) より

$$\lambda_i \le \alpha_i \quad (1 \le i \le n)$$

証明終 ∎

6.3 ミニマックス定理

A を n 次 Hermite 行列とすればその固有値 $\lambda_1, \ldots, \lambda_n$ は実数であり，適当なユニタリ行列 U をえらんで (6.5) のように対角化できる．Hermite 行列の性質はすべてこの事実から導かれる．

定理 6.8 (Courant-Fischer (クーラン・フィッシャー) のミニマックス定理) n 次 Hermite 行列 A の固有値を $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ とし，S_k を $X = \mathbb{C}^n$ の任意の k 次元部分空間とするとき

$$\lambda_k = \min_{S_k} \max_{\boldsymbol{0} \ne \boldsymbol{x} \in S_k} \frac{(A\boldsymbol{x},\ \boldsymbol{x})}{(\boldsymbol{x},\ \boldsymbol{x})}, \quad 1 \le k \le n$$

【証明】 (6.5) をみたす U の各列 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ は $\lambda_1, \ldots, \lambda_n$ に対応する正規直交固有ベクトルである．いま

$$M_k = \mathrm{span}\{\boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_n\}$$

とおき，S_k を任意の k 次元部分空間とすれば

$$\dim(M_k \cap S_k) = \dim M_k + \dim S_k - \dim(M_k + S_k)$$
$$\ge (n-k+1) + k - n = 1$$

よって $(\boldsymbol{x},\ \boldsymbol{x}) = 1$ をみたす元 $\boldsymbol{x} \in M_k \cap S_k$ がある．$\boldsymbol{x} \in M_k$ より

$$\boldsymbol{x} = \sum_{i=k}^{n} c_i \boldsymbol{u}_i$$

と書けば $\{\boldsymbol{u}_i\}$ は正規直交固有ベクトルであるから

$$(A\boldsymbol{x},\ \boldsymbol{x}) = \Big(\sum_{i=k}^{n} c_i\lambda_i\boldsymbol{u}_i,\ \sum_{i=k}^{n} c_i\boldsymbol{u}_i\Big)$$
$$= \sum_{i=k}^{n} \lambda_i|c_i|^2 \ge \lambda_k \sum_{i=k}^{n} |c_i|^2 = \lambda_k \quad \Big(\because\ \sum_{i=k}^{n} |c_i|^2 = (\boldsymbol{x},\ \boldsymbol{x}) = 1\Big)$$

したがって

$$\max_{\substack{\boldsymbol{x} \in S_k \\ (\boldsymbol{x},\boldsymbol{x})=1}} (A\boldsymbol{x},\ \boldsymbol{x}) \ge \lambda_k$$
$$\therefore\ \min_{S_k} \max_{\substack{\boldsymbol{x} \in S_k \\ (\boldsymbol{x},\ \boldsymbol{x})=1}} (A\boldsymbol{x},\ \boldsymbol{x}) \ge \lambda_k \tag{6.11}$$

一方

$$S'_k = \mathrm{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$$

とすれば $(\boldsymbol{x},\ \boldsymbol{x}) = 1$ なる $\boldsymbol{x} \in S'_k$ に対して $\boldsymbol{x} = \sum_{i=1}^{k} c'_i \boldsymbol{u}_i$ とあらわすとき

$$(A\boldsymbol{x},\ \boldsymbol{x}) = \sum_{i=1}^{k} \lambda_i |c'_i|^2 \leq \lambda_k \sum_{i=1}^{k} |c'_i|^2 = \lambda_k \quad \therefore \quad \max_{\substack{\boldsymbol{x} \in S'_k \\ (\boldsymbol{x},\ \boldsymbol{x})=1}} (A\boldsymbol{x},\ \boldsymbol{x}) \leq \lambda_k$$

$$\therefore \quad \min_{S_k} \max_{\substack{\boldsymbol{x} \in S_k \\ (\boldsymbol{x},\ \boldsymbol{x})=1}} (A\boldsymbol{x},\ \boldsymbol{x}) \leq \lambda_k \tag{6.12}$$

(6.11) と (6.12) より

$$\lambda_k = \min_{S_k} \max_{\substack{\boldsymbol{x} \in S_k \\ (\boldsymbol{x},\boldsymbol{x})=1}} (A\boldsymbol{x},\ \boldsymbol{x}) = \min_{S_k} \max_{\boldsymbol{0} \neq \boldsymbol{x} \in S_k} \frac{(A\boldsymbol{x},\ \boldsymbol{x})}{(\boldsymbol{x},\ \boldsymbol{x})}$$

証明終 ∎

注意　A の固有値を $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ と並べるときには，定理 6.8 は

$$\lambda_k = \min_{S_{k-1}} \max_{\substack{\boldsymbol{x} \in S_{k-1}^{\perp} \\ (\boldsymbol{x},\boldsymbol{x})=1}} (A\boldsymbol{x},\ \boldsymbol{x}) = \min_{S_{k-1}} \max_{\boldsymbol{0} \neq \boldsymbol{x} \in S_{k-1}^{\perp}} \frac{(A\boldsymbol{x},\ \boldsymbol{x})}{(\boldsymbol{x},\ \boldsymbol{x})}$$

または

$$\lambda_k = \max_{S_k} \min_{\substack{\boldsymbol{x} \in S_k \\ (\boldsymbol{x},\boldsymbol{x})=1}} (A\boldsymbol{x},\ \boldsymbol{x}) = \max_{S_k} \min_{\boldsymbol{0} \neq \boldsymbol{x} \in S_k} \frac{(A\boldsymbol{x},\ \boldsymbol{x})}{(\boldsymbol{x},\ \boldsymbol{x})}$$

の形になる．

6.4　分離定理

A を n 次 Hermite 行列とし

$$A = \begin{pmatrix} H & K \\ K^* & L \end{pmatrix} \quad (H \text{ は } m \times m \text{ 行列},\ m < n)$$

と分割する．A の固有値を $\lambda_1 \leq \cdots \leq \lambda_n$，対応する正規直交固有ベクトルを $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ とする．また H の固有値を $\mu_1 \leq \cdots \leq \mu_m$，対応する正規直交固有ベクトルを $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ とする．(各 $\boldsymbol{v}_i$ は m 次元列ベクトルである．) さらに

$$\boldsymbol{w}_i = \begin{pmatrix} \boldsymbol{v}_i \\ \boldsymbol{0} \end{pmatrix} \quad (n \text{ 次元ベクトル})$$

とおくと $(A\boldsymbol{w}_i,\ \boldsymbol{w}_j) = \mu_i \delta_{ij}$，$1 \leq i \leq m$ である．$1 \leq k \leq m$ なる k に対し

$$M = \mathrm{span}\{\boldsymbol{u}_k, \ldots, \boldsymbol{u}_n\}, \quad S = \mathrm{span}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$$

とおけば

$$\begin{aligned} \dim(M \cap S) &= \dim M + \dim S - \dim(M + S) \\ &\geq (n - k + 1) + k - n = 1 \end{aligned}$$

よって $(\boldsymbol{x},\ \boldsymbol{x}) = 1$ をみたす $\boldsymbol{x} \in M \cap S$ が存在する．このとき $\boldsymbol{x} \in M$ より $\lambda_k \leq (A\boldsymbol{x},\ \boldsymbol{x})$ (定理 6.5) また $\boldsymbol{x} \in S$ より $(A\boldsymbol{x},\ \boldsymbol{x}) \leq \mu_k$ も容易にわかる．

$$\therefore \quad \lambda_k \le \mu_k, \quad k = 1, 2, \ldots, m \tag{6.13}$$

次に

$$M' = \mathrm{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k+n-m}\} \quad (k+n-m \text{ 次元空間})$$
$$S' = \mathrm{span}\{\boldsymbol{w}_k, \ldots, \boldsymbol{w}_m\} \qquad (m-k+1 \text{ 次元空間})$$

として同じ議論をくり返せば

$$\dim(M' \cap S') \ge (k+n-m) + (m-k+1) - n = 1$$

よって $(\boldsymbol{x}, \boldsymbol{x}) = 1$ なる $\boldsymbol{x} \in M' \cap S'$ が存在する．このとき $\boldsymbol{x} \in M'$ より $(A\boldsymbol{x}, \boldsymbol{x}) \le \lambda_{k+n-m}$
また $\boldsymbol{x} \in S'$ より $\mu_k \le (A\boldsymbol{x}, \boldsymbol{x})$

$$\therefore \quad \mu_k \le \lambda_{k+n-m}$$

これと (6.13) とを併せて $\lambda_k \le \mu_k \le \lambda_{k+n-m}$, $k = 1, 2, \ldots, m$. 特に $m = n-1$ ならば $\lambda_k \le \mu_k \le \lambda_{k+1}$.

結局次の定理が得られた．

定理 6.9 (分離定理 (separation theorem)) n 次 Hermite 行列 A の m 次首座行列 $A_m = A\begin{pmatrix} 1 & 2 & \cdots & m \\ 1 & 2 & \cdots & m \end{pmatrix}$ の固有値を

$$\lambda_1^{(m)} \le \lambda_2^{(m)} \le \cdots \le \lambda_m^{(m)}$$

とすれば

$$\lambda_k^{(n)} \le \lambda_k^{(m)} \le \lambda_{k+n-m}^{(n)}$$

特に $m = n-1$ のとき

$$\lambda_1^{(n)} \le \lambda_1^{(n-1)} \le \lambda_2^{(n)} \le \lambda_2^{(n-1)} \le \cdots \le \lambda_{n-1}^{(n)} \le \lambda_{n-1}^{(n-1)} \le \lambda_n^{(n)} \tag{6.14}$$

注意 1 Householder [11] によれば，(6.14) はすでに Cauchy(1829) が実対称行列に対して得ている由である．

注意 2 A と $A_m = A\begin{pmatrix} 1 & 2 & \cdots & m \\ 1 & 2 & \cdots & m \end{pmatrix}$ の固有値を

$$\lambda_1^{(n)} \ge \cdots \ge \lambda_n^{(n)}, \quad \lambda_1^{(m)} \ge \cdots \ge \lambda_m^{(m)}$$

と並べるとき，定理 6.9 は

$$\lambda_{k+n-m}^{(n)} \le \lambda_k^{(m)} \le \lambda_k^{(n)}$$

となる．証明は定理 6.9 の場合と同様である．すなわち

$$M = \mathrm{span}\{\boldsymbol{u}_k, \ldots, \boldsymbol{u}_n\}, \quad S = \mathrm{span}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$$

とおけば $\dim(M \cap S) \ge 1$ より $(\boldsymbol{x}, \boldsymbol{x}) = 1$ なる $\boldsymbol{x} \in M \cap S$ に対して

$$\lambda_k^{(m)} \le (A\boldsymbol{x},\ \boldsymbol{x}) \le \lambda_k^{(n)}$$

また

$$M' = \mathrm{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k+n-m}\}, \quad S' = \mathrm{span}\{\boldsymbol{w}_k, \ldots, \boldsymbol{w}_m\}$$

とおけば $\dim(M' \cap S') \geq 1$ より $(\boldsymbol{x}, \boldsymbol{x}) = 1$ なる $\boldsymbol{x} \in M' \cap S'$ に対し

$$\lambda_{k+n-m}^{(n)} \leq (A\boldsymbol{x},\ \boldsymbol{x}) \leq \lambda_k^{(m)}$$

よって $\lambda_{k+n-m} \leq \lambda_k^{(m)} \leq \lambda_k^{(n)}$ を得る.

■**注意 3** 必ずしも実対称でない行列に対する分離定理として Hall-Porshing (ホール・ポーシング) (1968) の結果がある. この定理は第 8 章において述べる.

6.5 摂動定理

A, B を n 次 Hermite 行列とし $C = A + B$ とおく. A, B, C の固有値をそれぞれ

$$\lambda_1 \leq \cdots \leq \lambda_n, \quad \mu_1 \leq \cdots \leq \mu_n, \quad \nu_1 \leq \cdots \leq \nu_n$$

とするとき次が成り立つ.

定理 6.10 (摂動定理 (perturbation theorem)) 上の記号と仮定の下に

$$\mu_1 \leq \nu_k - \lambda_k \leq \mu_n, \quad 1 \leq k \leq n$$

したがって

$$|\nu_k - \lambda_k| \leq \|B\|_2 \leq \|B\|_F$$

【証明 (池辺他 1987)】 $A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i, \quad B\boldsymbol{v}_i = \mu_i \boldsymbol{v}_i, \quad C\boldsymbol{w}_i = \nu_i \boldsymbol{w}_i$

$$\boldsymbol{u}_i^* \boldsymbol{u}_j = \boldsymbol{v}_i^* \boldsymbol{v}_j = \boldsymbol{w}_i^* \boldsymbol{w}_j = \delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}$$

とする.

$$S_1 = \mathrm{span}\{\boldsymbol{u}_k, \ldots, \boldsymbol{u}_n\}, \quad S_2 = \mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}, \quad S_3 = \mathrm{span}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$$

とおけば

$$\begin{aligned} \dim(S_1 \cap S_2 \cap S_3) &= \dim S_1 + \dim(S_2 \cap S_3) - \dim(S_1 + S_2 \cap S_3) \\ &= \dim S_1 + \dim S_2 + \dim S_3 - \dim(S_2 + S_3) - \dim(S_1 + S_2 \cap S_3) \\ &\geq (n - k + 1) + n + k - n - n = 1 \end{aligned}$$

よって $\boldsymbol{x}^* \boldsymbol{x} = 1$ なる $\boldsymbol{x} \in S_1 \cap S_2 \cap S_3$ が存在する. このとき

$$\begin{aligned} &\boldsymbol{x} \in S_1 \text{より} \quad \lambda_k \leq (A\boldsymbol{x},\ \boldsymbol{x}) \\ &\boldsymbol{x} \in S_2 \text{より} \quad \mu_1 \leq (B\boldsymbol{x},\ \boldsymbol{x}) \\ &\boldsymbol{x} \in S_3 \text{より} \quad (A\boldsymbol{x},\ \boldsymbol{x}) + (B\boldsymbol{x},\ \boldsymbol{x}) = (C\boldsymbol{x},\ \boldsymbol{x}) \leq \nu_k \end{aligned}$$

$$\therefore \quad \lambda_k + \mu_1 \leq \nu_k \tag{6.15}$$

次に

$$S_1' = \mathrm{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}, \quad S_2' = S_2, \quad S_3' = \mathrm{span}\{\boldsymbol{w}_k, \ldots, \boldsymbol{w}_n\}$$

として同じ議論をくり返す．すなわち

$$\begin{aligned}\dim(S_1' \cap S_2' \cap S_3') &= \dim S_1' + \dim S_2' + \dim S_3' - \dim(S_2' + S_3') - \dim(S_1' + S_2' \cap S_3') \\ &\geq k + n + (n-k+1) - 2n = 1\end{aligned}$$

より $\boldsymbol{x}^*\boldsymbol{x} = 1$ なる $\boldsymbol{x} \in S_1' \cap S_2' \cap S_3'$ をとれば $\boldsymbol{x} \in S_1'$ より $(A\boldsymbol{x},\ \boldsymbol{x}) \leq \lambda_k$．また $\boldsymbol{x} \in S_2'$ より $(B\boldsymbol{x},\ \boldsymbol{x}) \leq \mu_n$．さらに $\boldsymbol{x} \in S_3'$ より

$$\begin{aligned}\nu_k &\leq (C\boldsymbol{x},\ \boldsymbol{x}) = (A\boldsymbol{x},\ \boldsymbol{x}) + (B\boldsymbol{x},\ \boldsymbol{x}) \\ \therefore\quad \nu_k &\leq \lambda_k + \mu_n \qquad (6.16)\end{aligned}$$

(6.14) と (6.15) より　$\mu_1 \leq \nu_k - \lambda_k \leq \mu_n$

したがって $B = (b_{ij})$ とするとき

$$|\nu_k - \lambda_k| \leq \max(|\mu_1|,\ |\mu_n|) = \max_i |\mu_i| = \rho(B) = \|B\|_2 \leq \|B\|_F = \sqrt{\sum_{i,j=1}^{n} |b_{ij}|^2}$$

証明終 ∎

【付記】上記証明は池辺他 [26] による．この鮮やかな証明技法はこの種の定理を一挙に単純化する．実際 定理 6.6–6.9 の証明および後述の定理 6.11, 6.12 の証明もこの線に沿っている．

定理 6.11　定理 6.10 の記号と仮定の下で次が成り立つ．

(i)　$i+j \leq n+1$　ならば　$\lambda_i + \mu_j \leq \nu_{i+j-1}$

(ii)　$i+j \geq n+1$　ならば　$\nu_{i+j-n} \leq \lambda_i + \mu_j$

(iii)　$\lambda_i + \mu_1 \leq \nu_i \leq \lambda_i + \mu_n,\quad 1 \leq i \leq n$

【証明】　(i) $i+j \leq n+1$ のとき

$$S_1 = \mathrm{span}\{\boldsymbol{u}_i, \boldsymbol{u}_{i+1}, \ldots, \boldsymbol{u}_n\},\quad S_2 = \mathrm{span}\{\boldsymbol{v}_j, \ldots, \boldsymbol{v}_n\},\quad S_3 = \mathrm{span}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{i+j-1}\}$$

とおけば

$$\begin{aligned}\dim(S_1 \cap S_2 \cap S_3) &\geq \dim S_1 + \dim S_2 + \dim S_3 - 2n \\ &= (n-i+1) + (n-j+1) + (i+j-1) - 2n = 1\end{aligned}$$

よって $\boldsymbol{x}^*\boldsymbol{x} = 1$ なる $\boldsymbol{x} \in S_1 \cap S_2 \cap S_3$ をとれば $\boldsymbol{x} \in S_1$ より $\lambda_i \leq (A\boldsymbol{x},\ \boldsymbol{x})$，また $\boldsymbol{x} \in S_2$ より $\mu_j \leq (B\boldsymbol{x},\ \boldsymbol{x})$．さらに $\boldsymbol{x} \in S_3$ より

$$(A\boldsymbol{x},\ \boldsymbol{x}) + (B\boldsymbol{x},\ \boldsymbol{x}) = (C\boldsymbol{x},\ \boldsymbol{x}) \leq \nu_{i+j-1} \quad \therefore \quad \lambda_i + \mu_j \leq \nu_{i+j-1}$$

(ii) $i+j \geq n+1$ のとき

$$S_1' = \mathrm{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_i\},\quad S_2' = \mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j\},\quad S_3' = \mathrm{span}\{\boldsymbol{w}_{i+j-n}, \ldots, \boldsymbol{w}_n\}$$

とおけば

$$\begin{aligned}\dim(S_1' \cap S_2' \cap S_3') &\geq \dim S_1' + \dim S_2' + \dim S_3' - 2n \\ &= i + j + n - (i+j-n) + 1 - 2n = 1\end{aligned}$$

よって $\boldsymbol{x}^*\boldsymbol{x} = 1$ なる $\boldsymbol{x} \in S_1' \cap S_2' \cap S_3'$ に対し $\boldsymbol{x} \in S_1'$ より $(A\boldsymbol{x},\ \boldsymbol{x}) \leq \lambda_i$ また $\boldsymbol{x} \in S_2'$ より $(B\boldsymbol{x},\ \boldsymbol{x}) \leq \mu_j$ さらに $\boldsymbol{x} \in S_3'$ より

$$\nu_{i+j-n} \le (C\boldsymbol{x},\ \boldsymbol{x}) = (A\boldsymbol{x},\ \boldsymbol{x}) + (B\boldsymbol{x},\ \boldsymbol{x}) \quad \therefore\ \ \nu_{i+j-n} \le \lambda_i + \mu_j$$

(iii) これは定理 6.10 で済みであるが (i) で $j=1$, (ii) で $j=n$ をとれば得られる. 証明終 ∎

なお, A, B, C の固有値を

$$\lambda_1 \ge \cdots \ge \lambda_n, \quad \mu_1 \ge \cdots \ge \mu_n, \quad \nu_1 \ge \cdots \ge \nu_n$$

と並べるとき次の定理が成り立つ.

定理 6.12 (Weyl (ワイル)) 上の仮定の下で次が成り立つ.

(i) $i+j \le n+1$ のとき $\nu_{i+j-1} \le \lambda_i + \mu_j$
(ii) $i+j \ge n+1$ のとき $\mu_{i+j-n} \le \lambda_i + \mu_j$

【証明】 (i) は

$$S_1 = \mathrm{span}\{\boldsymbol{u}_i, \ldots, \boldsymbol{u}_n\}, \quad S_2 = \mathrm{span}\{\boldsymbol{v}_j, \ldots, \boldsymbol{v}_n\}, \quad S_3 = \mathrm{span}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{i+j-1}\}$$

(ii) は

$$S_1' = \mathrm{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_i\}, \quad S_2' = \mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j\}, \quad S_3' = \mathrm{span}\{\boldsymbol{w}_{i+j-n}, \ldots, \boldsymbol{w}_n\}$$

とおき定理 6.11 と同様な議論をくり返せばよい. 証明終 ∎

6.6 Hoffman-Wielandt の定理

この節で述べる Hoffman-Wielandt (ホフマン・ウィーラント) の定理 [25] は正規行列に対する摂動定理であって, 前節の証明技法とは全く異なる手法が適用される.

定理 6.13 (Hoffman-Wielandt) $A = (a_{ij}), B = (b_{ij})$ を n 次の正規行列, それらの固有値をそれぞれ $\lambda_1, \ldots, \lambda_n$, $\mu_1, \ldots, \mu_n$ とする. このときそれらの適当な並べかえにより

$$\sum_{i=1}^{n} |\lambda_i - \mu_i|^2 \le \|A - B\|_F^2$$

とできる. すなわち $1, 2, \ldots, n$ の置換 σ を適当にえらべば

$$\sum_{i=1}^{n} |\lambda_i - \mu_{\sigma(i)}|^2 \le \sum_{i,j=1}^{n} |a_{ij} - b_{ij}|^2$$

【証明】 適当なユニタリ行列 U, V により

$$U^*AU = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad (= A_0 \text{とおく})$$

$$V^*(U^*BU)V = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} \quad (= B_0 \text{とおく})$$

と対角化できる (U^*BU は正規行列であることに注意). このとき

$$A = UA_0U^*, \quad B = UVB_0V^*U^*$$

ここで任意の n 次行列 C とユニタリ行列 U に対して成り立つ等式

$$\begin{aligned}\|UCU^*\|_F^2 &= \mathrm{tr}\{(UCU^*)(UCU^*)^*\} \quad (\S 2.5 \text{ 参照}) \\ &= \mathrm{tr}(UCC^*U^*) = \mathrm{tr}(CC^*) = \mathrm{tr}(C^*C) = \|C\|_F^2\end{aligned}$$

を $C = A_0 - VB_0V^*$ に適用すれば

$$\begin{aligned}\|A-B\|_F^2 &= \|U(A_0 - VB_0V^*)U^*\|_F^2 = \|A_0 - VB_0V^*\|_F^2 \\ &= \mathrm{tr}\{(A_0 - VB_0V^*)(A_0 - VB_0V^*)^*\} = \mathrm{tr}\{(A_0 - VB_0V^*)(A_0^* - VB_0^*V^*)\} \\ &= \mathrm{tr}(A_0A_0^* + VB_0B_0^*V^*) + r(V) \\ &\qquad \left(\begin{array}{ll} r(V) & = -\mathrm{tr}(VB_0V^*A_0^* + A_0VB_0^*V^*) \\ & = -\mathrm{tr}(A_0^*VB_0V^* + A_0VB_0^*V^*) \end{array} \right) \\ &= \mathrm{tr}(A_0A_0^*) + \mathrm{tr}(VB_0B_0^*V^*) + r(V) = \mathrm{tr}(A_0A_0^*) + \mathrm{tr}(B_0B_0^*) + r(V) \\ &= \|A_0\|_F^2 + \|B_0\|_F^2 + r(V)\end{aligned}$$

$V = (v_{ij})$ とすれば

$$A_0V = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} (v_{ij}) = (\lambda_i v_{ij}),$$

$$\begin{aligned}B_0V^* &= \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} (v_{ij}^*) \quad (v_{ij}^* = \overline{v}_{ji}), \\ &= (\mu_i v_{ij}^*)\end{aligned}$$

$$\therefore \quad r(V) = -\mathrm{tr}\left\{(\overline{\lambda}_i v_{ij})(\mu_i v_{ij}^*) + \lambda_i v_{ij})(\overline{\mu}_i v_{ij}^*)\right\} = -\mathrm{tr}\left\{\left(\sum_{k=1}^n \overline{\lambda}_i v_{ik}\mu_k v_{kj}^* + \sum_{k=1}^n \lambda_i v_{ik}\overline{\mu}_k v_{kj}^*\right)\right\}$$

$$= -\sum_{i=1}^n\sum_{k=1}^n (\overline{\lambda}_i v_{ik}\mu_k v_{ki}^* + \lambda_i v_{ik}\overline{\mu}_k v_{ki}^*) = -\sum_{i=1}^n\sum_{k=1}^n (\overline{\lambda}_i\mu_k + \lambda_i\overline{\mu}_k) v_{ik}v_{ki}^*$$

$$= -\sum_{i=1}^n\sum_{k=1}^n \alpha_{ik}w_{ik}$$

$$\left(\begin{array}{ll} \alpha_{ik} & = -(\overline{\lambda}_i\mu_k + \lambda_i\overline{\mu}_k) \\ w_{ik} & = v_{ik}v_{ki}^* = v_{ik}\overline{v}_{ik} = |v_{ik}|^2 \end{array} \right)$$

$\left(= f(W), \quad W = (w_{ij}) \text{ とおく}\right)$

ここで

$$\mathcal{W} = \left\{ W = (w_{ij}) = (|v_{ij}|^2) \;\middle|\; V = (v_{ij}) : \text{ユニタリ行列} \right\}$$

とおけば $W \in \mathcal{W}$ は n 次の 2 重確率行列である. n 次の 2 重確率行列の全体を S_n とすれば

$$\min_{V:\text{ユニタリ}} r(V) = \min_{W\in\mathcal{W}} f(W) \geq \min_{X\in S_n} f(X)$$

定理 1.22 によって $X \in S_n$ は n 次の置換行列 P_i の凸 1 次結合として

$$X=\sum_{i=1}^{m}\theta_i P_i,\quad \theta_i\geq 0,\quad \sum_{i=1}^{m}\theta_i=1,\quad m=n!$$

と書ける．$f(W)=-\sum_{i,j=1}^{n}\alpha_{ij}w_{ij}$ は W の成分 w_{ij} の 1 次式であるから

$$f(X)=f\Big(\sum_{i=1}^{m}\theta_i P_i\Big)=\sum_{i=1}^{m}\theta_i f(P_i)\geq\Big(\min_i f(P_i)\Big)\sum_{i=1}^{m}\theta_i=\min_i f(P_i)=f(P)$$

とする．このとき $f(W)$ の最小値は $W=P$ のとき達せられるが，このとき $V=P$ でもあるから結局 $r(V)$ の最小値は $V=P$ のときとられる．行列 P に対応する置換を σ とすれば $A_0-PB_0P^{\mathrm{t}}=\mathrm{diag}(\lambda_1-\mu_{\sigma(1)},\dots,\lambda_n-\mu_{\sigma(n)})$ であるから

$$\|A-B\|_F^2=\|A_0-VB_0V^*\|_F^2\geq\|A_0-PB_0P^{\mathrm{t}}\|_F^2=\sum_{i=1}^{n}|\lambda_i-\mu_{\sigma(i)}|^2$$

となって定理が示された．　証明終 ∎

■**注意**　A,B が共に正規行列でなければ定理は成立しない．[25] において Hoffman-Wielandt は次の例を与えている．

$$n=2,\quad A=\begin{pmatrix}0&0\\0&4\end{pmatrix},\quad B=\begin{pmatrix}-1&-1\\1&1\end{pmatrix}$$

のとき A は実対称したがって正規であるが B は正規でない．$(\because\ B^{\mathrm{t}}B\neq BB^{\mathrm{t}})$ このとき $\lambda_1=0$, $\lambda_2=4$, $\mu_1=\mu_2=0$ (2 重根) で $\|A-B\|_F^2=12$．しかし任意の σ に対して $\sum_{i=1}^{2}|\lambda_i-\mu_{\sigma(i)}|^2=16>\|A-B\|_F^2$ である．

定理 6.14 (Hoffman-Wielandt)　A が Hermite 行列ならば定理 6.13 における最良の並べ方は

$$\lambda_1\leq\cdots\leq\lambda_n,\quad \mathrm{Re}\mu_1\leq\cdots\leq\mathrm{Re}\mu_n$$

である．したがって，A,B が Hermite 行列ならば $\lambda_1\leq\cdots\leq\lambda_n$, $\mu_1\leq\cdots\leq\mu_n$ のとき

$$\sum_{i=1}^{n}|\lambda_i-\mu_i|^2\leq\|A-B\|_F^2$$

【証明】　$\lambda_1\leq\cdots\leq\lambda_n$ を最良な並べ方とする．ここで $\mathrm{Re}\mu_1>\mathrm{Re}\mu_2$ ならば

$$\begin{aligned}
&|\lambda_1-\mu_1|^2+|\lambda_2-\mu_2|^2-(|\lambda_1-\mu_2|^2+|\lambda_2-\mu_1|^2)\\
&\quad=(\lambda_1-\mu_1)(\lambda_1-\overline{\mu}_1)+(\lambda_2-\mu_2)(\lambda_2-\overline{\mu}_2)\\
&\qquad-(\lambda_1-\mu_2)(\lambda_1-\overline{\mu}_2)-(\lambda_2-\mu_1)(\lambda_2-\overline{\mu}_1)\\
&\quad=\{\lambda_1^2+\lambda_2^2+|\mu_1|^2+|\mu_2|^2-(\mu_1+\overline{\mu}_1)\lambda_1-(\mu_2+\overline{\mu}_2)\lambda_2\}\\
&\qquad-\{\lambda_1^2+\lambda_2^2+|\mu_1|^2+|\mu_2|^2-(\mu_2+\overline{\mu}_2)\lambda_1-(\mu_1+\overline{\mu}_1)\lambda_2\}\\
&\quad=(\mu_1+\overline{\mu}_1)(\lambda_2-\lambda_1)+(\mu_2+\overline{\mu}_2)(\lambda_1-\lambda_2)\\
&\quad=2(\mathrm{Re}\mu_1-\mathrm{Re}\mu_2)(\lambda_2-\lambda_1)\geq 0\\
&\therefore\quad |\lambda_1-\mu_1|^2+|\lambda_2-\mu_2|^2\geq|\lambda_1-\mu_2|^2+|\lambda_2-\mu_1|^2
\end{aligned}$$

よって μ_1 と μ_2 をとりかえて $\mathrm{Re}\mu_1\leq\mathrm{Re}\mu_2$ としても $\delta=\sum_{i=1}^{n}|\lambda_i-\mu_i|^2$ の値は増加しない．以下これをくり返せば δ の値を増加させることなく

$$\mathrm{Re}\mu_1 \leq \cdots \leq \mathrm{Re}\mu_n$$

とできる．よって $\lambda_1 \leq \cdots \leq \lambda_n$ を最良な並べ方とすれば $\mathrm{Re}\mu_1 \leq \cdots \leq \mathrm{Re}\mu_n$ と並べるのも最良な並べ方である． 証明終 ∎

【付記】H. Wielandt(1910–2001) の指導教授は I. Schur と E.Schmidt である．Wielandt は 1950 年代前半まで行列の固有値に関する多くの切れ味鋭い結果を発表した後，群論 (特に巡回群) の研究に転じた．

尚，J. H. Wilkinson は彼の著書 The Algebraic Eigenvalue Problem, Oxford 1966 において定理 6.14 を Wielandt-Hoffman の定理と呼んでいる．しかし，本書では原論文 [25] の著者順を尊重し，敢えて Hoffman-Wielandt の定理と呼ぶ．最近ではこの呼び名の方が定着しているようである．

6.7 Schur 補元の固有値

n 次行列

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad (A_{11}\text{は } p \text{ 次}, \ A_{22}\text{は } n-p \text{ 次})$$

において A_{11} は正則とする．このとき A_{11} の **Schur 補元** $\widetilde{A_{11}} = A_{22} - A_{21}A_{11}^{-1}A_{12}$ が定義される (§1.10 参照)．

定理 6.15 A が半正定値 Hermite 行列 (すなわち $A_{11} \succcurlyeq O$) で A_{11} が正則ならば

$$\lambda_{\min}(A) \leq \lambda_{\min}(\widetilde{A_{11}}) \leq \lambda_{\min}(A_{22})$$

ただし $\lambda_{\min}(A)$ は A の最小固有値をあらわす．

【証明】 仮定 $A \succcurlyeq O$ より $A_{11} \succcurlyeq O$, $A_{22} \succcurlyeq O$ であるが，A_{11} は正則としているから $A_{11} \succ O$ である．したがって $B = A_{21}A_{11}^{-1}A_{12} = (A_{12})^*A_{11}^{-1}A_{12} \succcurlyeq O$ よって $A_{22} - \widetilde{A_{11}} = B \succcurlyeq O$ すなわち $A_{22} \succcurlyeq \widetilde{A_{11}}$

$$\therefore \quad \lambda_{\min}(A_{22}) \geq \lambda_{\min}(\widetilde{A_{11}}) \quad (\text{定理 6.6 による}) \tag{6.17}$$

次に残りの不等式を証明するために 2 つの場合に分けて考える．

(i) $\widetilde{A_{11}}$ が正則のとき．このとき定理 1.19 (ii) により A は正則で

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\widetilde{A_{11}}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\widetilde{A_{11}}^{-1} \\ -\widetilde{A_{11}}^{-1}A_{21}A_{11}^{-1} & \widetilde{A_{11}}^{-1} \end{pmatrix}$$

この行列に分離定理を適用すれば

$$\lambda_{\max}(A^{-1}) \geq \lambda_{\max}(\widetilde{A_{11}}^{-1}) \quad (\lambda_{\max}(A^{-1}) \text{ は } A^{-1}\text{の最大固有値})$$

$\lambda_{\max}(A^{-1}) = \dfrac{1}{\lambda_{\min}(A)}$, $\lambda_{\max}(\widetilde{A_{11}}^{-1}) = \dfrac{1}{\lambda_{\min}(\widetilde{A_{11}})}$ であるから

$$\lambda_{\min}(A) \leq \lambda_{\min}(\widetilde{A_{11}}) \tag{6.18}$$

(ii) $\widetilde{A_{11}}$ が正則でないときは ε を任意の正数として $\widetilde{A_{11}} + \varepsilon I$ は正則である．このとき $A \succcurlyeq O$ より $A + \varepsilon I$ も正則である．したがって $A + \varepsilon I$ における $A_{11} + \varepsilon I_p$ の Schur 補元

$$\widetilde{A_{11}}(\varepsilon) = A_{22} + \varepsilon I_{n-p} - A_{21}(A_{11} + \varepsilon I_p)^{-1}A_{12}$$

も正則である．故に (6.18) によって

$$\lambda_{\min}(A+\varepsilon I_n) \leq \lambda_{\min}(\widetilde{A_{11}}(\varepsilon))$$

ここで $\varepsilon \to 0$ とすれば $A+\varepsilon I_n \to A$ かつ $\widetilde{A_{11}}(\varepsilon) \to \widetilde{A_{11}}$ (A_{11} は正則と仮定していることに注意) であるから

$$\lambda_{\min}(A) \leq \lambda_{\min}(\widetilde{A_{11}}) \tag{6.19}$$

(6.17)–(6.19) により定理の不等式が得られる． 証明終 ∎

第7章　非負行列

1907年O.Perron (ペロン) は要素がすべて正の行列の固有値・固有ベクトルに関する1つの結果をドイツの数学専門誌 Mathematische Annalen に発表した．1908–1909年G.Frobeniusも相続いて同じ結果をドイツの異なる論文誌に発表し，さらに1912年それを既約非負行列に拡張した．その後1950年その結果はH.Wielandtにより一層精緻なものとされた．それらは現在一括りにして**非負行列に関するPerron-Frobeniusの理論**と呼ばれている．この章ではこの理論を詳述する．

7.1　非負行列

行列 A の要素がすべて非負のとき**非負行列 (nonnegative matrix)** といい $A \geq O$ とかく．また要素がすべて正のとき**正行列 (positive matrix)** といい $A > O$ とかく．同じ次元の行列 A, B に対し $A - B \geq O$ のとき $A \geq B$, $A - B > O$ のとき $A > B$ と記す．特にベクトル $\boldsymbol{x}$ に対しても，同様に**非負ベクトル** $\boldsymbol{x} \geq \boldsymbol{0}$ と**正ベクトル** $\boldsymbol{x} > \boldsymbol{0}$ が定義され，ベクトル $\boldsymbol{x}, \boldsymbol{y}$ に対して $\boldsymbol{x} - \boldsymbol{y} \geq \boldsymbol{0}$ のとき $\boldsymbol{x} \geq \boldsymbol{y}$, $\boldsymbol{x} - \boldsymbol{y} > \boldsymbol{0}$ のとき $\boldsymbol{x} > \boldsymbol{y}$ とかく．

定理 7.1　A を n 次の非負行列とするとき，次の条件は互いに同値である．

(i)　$\lim_{p\to\infty} A^p = O$
(ii)　$\rho(A) < 1$　($\rho(A)$ は A のスペクトル半径)
(iii)　$I - A$ は正則で　$(I - A)^{-1} \geq O$
(iv)　任意の n 次元列ベクトル $\boldsymbol{b} \geq \boldsymbol{0}$ に対し $(I - A)\boldsymbol{x} = \boldsymbol{b}$ は一意解 $\boldsymbol{x} \geq \boldsymbol{0}$ をもつ．
(v)　任意の n 次元列ベクトル $\boldsymbol{b} > \boldsymbol{0}$ に対し $(I - A)\boldsymbol{x} = \boldsymbol{b}$ は一意解 $\boldsymbol{x} > \boldsymbol{0}$ をもつ．
(vi)　$\boldsymbol{y} > A\boldsymbol{y}$ をみたす列ベクトル $\boldsymbol{y} > \boldsymbol{0}$ が存在する．

【証明】　(i) と (ii) の同値性は系 4.5.2 で示されている．
(ii)⇒(iii) $\rho(A) < 1$ のとき $I - A$ が正則であることは明らかである．いま

$$S_p = I + A + \cdots + A^p \quad (p = 1, 2, \ldots)$$

とおくと $A \geq O$ より $S_p \geq O$ であり，系 4.5.3 によって $(I - A)^{-1} = \lim_{p\to\infty} S_p \geq O$ である．
(iii)⇒(iv) $\boldsymbol{b} \geq \boldsymbol{0}$ のとき $\boldsymbol{x} = (I - A)^{-1}\boldsymbol{b} \geq \boldsymbol{0}$
(iv)⇒(v) $\boldsymbol{b} > \boldsymbol{0}$ のとき $(I - A)\boldsymbol{x} = \boldsymbol{b}$ なる解 $\boldsymbol{x} \geq \boldsymbol{0}$ をとれば，$A \geq O$ より

$$\boldsymbol{x} = A\boldsymbol{x} + \boldsymbol{b} \geq \boldsymbol{b} > \boldsymbol{0}$$

(v)⇒(vi) $\boldsymbol{e} = (1, \ldots, 1)^{\mathrm{t}}$ とおくと仮定によって方程式 $(I - A)\boldsymbol{y} = \boldsymbol{e}$ は解 $\boldsymbol{y} > \boldsymbol{0}$ をもつ．このとき $\boldsymbol{y} - A\boldsymbol{y} = \boldsymbol{e} > \boldsymbol{0}$ したがって $\boldsymbol{y} > A\boldsymbol{y}$
(vi)⇒(i) $\boldsymbol{y} > A\boldsymbol{y}$, $\boldsymbol{y} > \boldsymbol{0}$ とし，$\boldsymbol{v}^{(p)} = \boldsymbol{y} - A^p\boldsymbol{y}$ とおくと $\boldsymbol{v}^{(1)} > \boldsymbol{0}$
また

$$
\begin{aligned}
\boldsymbol{v}^{(p+1)} &= \boldsymbol{y} - A^{p+1}\boldsymbol{y} \\
&= (\boldsymbol{y} - A\boldsymbol{y}) + A(\boldsymbol{y} - A\boldsymbol{y}) + \cdots + A^{p-1}(\boldsymbol{y} - A\boldsymbol{y}) + A^p(\boldsymbol{y} - A\boldsymbol{y}) \\
&\geq (\boldsymbol{y} - A\boldsymbol{y}) + A(\boldsymbol{y} - A\boldsymbol{y}) + \cdots + A^{p-1}(\boldsymbol{y} - A\boldsymbol{y}) \quad (\because\ A^p \geq O,\ \boldsymbol{y} - A\boldsymbol{y} > \boldsymbol{0}) \\
&= \boldsymbol{y} - A^p\boldsymbol{y} = \boldsymbol{v}^{(p)} \\
\therefore\ \boldsymbol{v}^{(p)} &\leq \boldsymbol{v}^{(p+1)} = \boldsymbol{y} - A^{p+1}\boldsymbol{y} \leq \boldsymbol{y} \\
\therefore\ \boldsymbol{0} < \boldsymbol{v}^{(1)} &\leq \boldsymbol{v}^{(2)} \leq \cdots \leq \boldsymbol{v}^{(p)} \leq \cdots \leq \boldsymbol{y}
\end{aligned}
$$

よって $\{\boldsymbol{v}^{(p)}\}$ は上に有界な単調増加列ベクトルであるから収束する．収束先を $\boldsymbol{v}$ とすれば

$$
\lim_{p\to\infty} A^p(\boldsymbol{y} - A\boldsymbol{y}) = \lim_{p\to\infty}(\boldsymbol{v}^{(p+1)} - \boldsymbol{v}^{(p)}) = \boldsymbol{v} - \boldsymbol{v} = \boldsymbol{0}
$$

であるが，$A^p \geq O$ かつ $\boldsymbol{y} - A\boldsymbol{y} > O$ であるから $\lim_{p\to\infty} A^p = O$ でなければならない． 証明終 ∎

系 7.1.1 $A = (a_{ij})$ は n 次非負行列で $I - A$ は正則かつ $(I - A)^{-1} = (\alpha_{ij}) \geq O$ とすれば $a_{ii} < 1$ かつ $\alpha_{ii} \geq 1,\ 1 \leq i \leq n$.

【証明】 $(I - A)^{-1}(I - A) = I$ の (i, i) 要素を比較して

$$
\alpha_{ii}(1 - a_{ii}) - \sum_{\substack{k=1 \\ k\neq i}}^{n} \alpha_{ik}a_{ki} = 1
$$

$$
\therefore\quad \alpha_{ii}(1 - a_{ii}) = 1 + \sum_{k\neq 1} \alpha_{ik}a_{ki} \geq 1 \quad (\because\ \alpha_{ij} \geq 0,\ a_{ij} \geq 0 \quad \forall i, j) \tag{7.1}
$$

$$
\therefore\quad \alpha_{ii} \geq 1 + \alpha_{ii}a_{ii} \geq 1 \tag{7.2}
$$

故に (7.1) と (7.2) より

$$
1 - a_{ii} \geq \frac{1}{\alpha_{ii}} > 0 \quad \therefore\ a_{ii} < 1 \quad \forall i
$$

証明終 ∎

系 7.1.2 $A \geq O,\ \boldsymbol{y} \geq \boldsymbol{0},\ \boldsymbol{y} \neq \boldsymbol{0},\ A\boldsymbol{y} > \lambda\boldsymbol{y}$ かつ $\lambda \geq 0$ ならば

$$
\rho(A) > \lambda
$$

【証明】 $A\boldsymbol{y} \geq (\lambda + \varepsilon)\boldsymbol{y}$ をみたす $\varepsilon > 0$ をとることができる．このとき

$$
A^p\boldsymbol{y} \geq (\lambda + \varepsilon)^p\boldsymbol{y},\quad p = 1, 2, \ldots
$$

$$
\therefore\quad \left(\frac{1}{\lambda + \varepsilon}A\right)^p \boldsymbol{y} \geq \boldsymbol{y},\quad p = 1, 2, \ldots \tag{7.3}
$$

よって $\rho\left(\frac{1}{\lambda+\varepsilon}A\right) \geq 1$ $\left(\text{なぜならば仮に } \rho\left(\frac{1}{\lambda+\varepsilon}A\right) < 1 \text{ とすれば } \left(\frac{1}{\lambda+\varepsilon}A\right)^p \to 0\ (p \to \infty)\ (\text{系 } 4.5.2)\right.$ であるが (7.3) で $p \to \infty$ として $\boldsymbol{y} = \boldsymbol{0}$. これは矛盾である．$\left.\right)$ $\rho\left(\frac{1}{\lambda+\varepsilon}A\right) = \frac{1}{\lambda+\varepsilon}\rho(A)$ であるから $\rho(A) \geq \lambda + \varepsilon > \lambda$ を得る． 証明終 ∎

系 7.1.3 $A \geq O$ と $\lambda > 0$ に対し，$A\boldsymbol{y} < \lambda\boldsymbol{y}$ なる $\boldsymbol{y} > \boldsymbol{0}$ が存在すれば

$$
\rho(A) < \lambda
$$

【証明】 定理 7.1 (vi) と (ii) を $\frac{1}{\lambda}A$ に適用すれば $\rho\left(\frac{1}{\lambda}A\right) < 1$

$$
\therefore\quad \rho(A) < \lambda
$$

証明終 ∎

系 7.1.4 $A \geq O$ かつ $\alpha > 0$ のとき

$$\rho(A) < \alpha \Leftrightarrow (\alpha I - A)^{-1} \geq O$$

【証明】 $$\rho(A) < \alpha \Leftrightarrow \rho\Big(\frac{1}{\alpha}A\Big) < 1 \Leftrightarrow \Big(I - \frac{1}{\alpha}A\Big)^{-1} \geq 0 \quad \text{(定理 7.1 (iii))}$$
$$\Leftrightarrow (\alpha I - A)^{-1} \geq O$$
証明終 ∎

定理 7.2 n 次行列 $A = (a_{ij})$, $B = (b_{ij})$ に対し $|A| = (|a_{ij}|)$, $|B| = (|b_{ij}|)$ とおく．このとき $|A| \geq |B|$ ならば $\rho(|A|) \geq \rho(|B|) \geq \rho(B)$．特に $\rho(|A|) \geq \rho(A)$ である．

【証明】 $\rho(|A|) \geq \rho(B)$ を示せばよい．$\lambda_{|A|} = \rho(|A|)$, $\lambda_B = \rho(B)$ とおく．仮に $\lambda_B > \lambda_{|A|}$ とすれば $\lambda_B > c > \lambda_{|A|}$ なる正数 c が存在する．このとき $P = \frac{1}{c}|A|$, $Q = \frac{1}{c}B$ とおけば $\rho(P) = \frac{1}{c}\rho(|A|) < 1$ かつ $\rho(Q) = \frac{1}{c}\lambda_B > 1$

一方

$$P \geq |Q| = \frac{1}{c}|B| \geq Q$$

であるから $P^m \geq |Q|^m \geq Q^m \geq -|Q|^m \geq -P^m$

$$\rho(P) < 1 \text{ より} \quad P^m \to O \ (m \to \infty) \quad \therefore\ Q^m \to O \ (m \to \infty) \quad \therefore\ \rho(Q) < 1$$

これは $\rho(Q) > 1$ に反する． $\therefore\ \lambda_B \leq \lambda_{|A|}$.
証明終 ∎

系 7.2.1 (Frobenius) $A \geq B \geq O$ ならば $\rho(A) \geq \rho(B)$

定理 7.3 (Wielandt) A を n 次非負行列とする．A が既約であるための必要十分条件は $(I + A)^{n-1} > O$

【証明】 (i) 必要条件：A を既約として，任意の非負ベクトル $\boldsymbol{x} \neq \boldsymbol{0}$ に対して

$$(I + A)^{n-1}\boldsymbol{x} > \boldsymbol{0}$$

を示せばよい

$$\boldsymbol{x}^{(m)} = (I + A)^{m-1}\boldsymbol{x}, \quad m = 1, 2, \ldots$$

とおけば

$$\boldsymbol{x}^{(m+1)} = (I + A)\boldsymbol{x}^{(m)} = \boldsymbol{x}^{(m)} + A\boldsymbol{x}^{(m)}$$

$\boldsymbol{x} \neq \boldsymbol{0}$ であるから適当な置換行列 P をえらんで

$$P\boldsymbol{x} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{0} \end{pmatrix} \quad (\boldsymbol{u} \text{ は } l \text{ 次元正ベクトル})$$

とできる $(1 \leq l \leq n)$．このとき

$$P\boldsymbol{x}^{(2)} = P\boldsymbol{x}^{(1)} + PA\boldsymbol{x}^{(1)} = P\boldsymbol{x} + (PAP^{\mathrm{t}})P\boldsymbol{x}$$
$$= \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{0} \end{pmatrix} + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{u} + A_{11}\boldsymbol{u} \\ A_{21}\boldsymbol{u} \end{pmatrix}$$

A は既約であるから $A_{21} \neq O$．よって $n - l$ 次元ベクトル $A_{21}\boldsymbol{u}$ は少なくとも 1 つの正成分をもつ．し

たがって $P\boldsymbol{x}^{(2)}$ は少なくとも $l+1$ 個の正成分をもつから

$$P\boldsymbol{x}^{(n)} > \boldsymbol{0} \quad \therefore\ \boldsymbol{x}^{(n)} > \boldsymbol{0} \quad \therefore\ (I+A)^{n-1}\boldsymbol{x} > \boldsymbol{0}$$

$\boldsymbol{x} \neq \boldsymbol{0}$ は任意の非負ベクトルであるから $(I+A)^{n-1} > O$

(ii) 十分条件：$(I+A)^{n-1} > O$ とする．仮に A が既約でないとし，適当な置換行列 P をえらんで

$$PAP^{\mathrm{t}} = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix} \quad (A_{11} : l \text{次},\ \ l < n)$$

とすれば

$$\begin{aligned} P(I+A)^{n-1}P^{\mathrm{t}} &= \{P(I+A)P^{\mathrm{t}}\}\{P(I+A)P^{\mathrm{t}}\}\cdots\{P(I+A)P^{\mathrm{t}}\} \\ &= \{(I+PAP^{\mathrm{t}})\}^{n-1} \\ &= \begin{pmatrix} \widetilde{A}_{11} & A_{12} \\ O & \widetilde{A}_{22} \end{pmatrix}^{n-1} \qquad (\widetilde{A}_{11} = I_l + A_{11},\ \widetilde{A}_{22} = I_{n-l} + A_{22}) \\ &= \begin{pmatrix} \widetilde{A}_{11}^{n-1} & * \\ O & \widetilde{A}_{22}^{n-1} \end{pmatrix} \end{aligned}$$

一方 $(I+A)^{n-1} > O$ より $P(I+A)^{n-1}P^{\mathrm{t}} > O$ であるからこれは矛盾である．よって A は既約である．　　証明終 ∎

■**注意**　定理 7.3 は通常 **Wielandt の補題**と呼ばれている．

7.2 正行列に対する Perron-Frobenius の定理

定理 7.4 (Perron-Frobenius)　n 次正行列 A の固有値 λ_i を $|\lambda_1| \geq \cdots \geq |\lambda_n|$ と並べるとき，次が成り立つ．

(i)　$\lambda_A = \lambda_1 > 0$ かつ $\lambda_A > |\lambda_i|$　$(i \geq 2)$
(ii)　λ_A に対応する固有ベクトルとして正ベクトルがとれる．
(iii)　λ_A の幾何学的重複度は 1
(iv)　$A \geq B \geq O$ ならば $\lambda_A \geq \rho(B)$ で，等号は $A = B$ のときに限る．
(v)　λ_A は単純すなわち代数的重複度は 1

λ_A を **Perron 根**，対応する正固有ベクトルを **Perron ベクトル**という．

【証明】　(i), (ii) λ を $A = (a_{ij})$ の固有値として $A\boldsymbol{v} = \lambda\boldsymbol{v}$, $\boldsymbol{v} \neq \boldsymbol{0}$ とする．まず

$$|\lambda| = \lambda_A = \rho(A) \Rightarrow \lambda > 0$$

を示そう．$\boldsymbol{v} = (v_1, \ldots, v_n)^{\mathrm{t}}$ とすれば $\lambda v_i = \sum_{j=1}^{n} a_{ij}v_j$
いま，ある k と l に対し $v_k \neq 0$, $v_l \neq 0$ かつ $\arg v_k \neq \arg v_l$ ($\arg v$ は v の偏角をあらわす) とすれば

$$|a_{ik}v_k + a_{il}v_l| < a_{ik}|v_k| + a_{il}|v_l| \quad \therefore\ \ |\lambda| \cdot |v_i| < \sum_{j=1}^{n} a_{ij}|v_j|$$

よって $|\boldsymbol{v}| = (|v_1|, \ldots, |v_n|)^{\mathrm{t}}$ とおくとき　$|\lambda| \cdot |\boldsymbol{v}| < A|\boldsymbol{v}|$
したがって系 7.1.2 によって $|\lambda| < \rho(A) = \lambda_A$　これは仮定 $|\lambda| = \lambda_A$ に反する．故に $v_k \neq 0$ なる v_k

は $\arg v_k = \theta$ (一定) であり，$v_i \geq 0 \quad \forall i$ としてよい．

$\max_i v_i = v_k$, $\min v_i = v_m$ とすれば $v_k > 0$ で

$$\lambda v_m = \sum_{j=1}^{n} a_{mj} v_j \geq a_{mk} v_k > 0 \quad \therefore \quad \lambda > 0 \quad \text{かつ} \quad v_m > 0$$

これは $\lambda = \lambda_A > 0$ かつ $\boldsymbol{v} > \boldsymbol{0}$ を意味する．残る主張 $\lambda_A > |\lambda_i|$ $(i \geq 2)$ は後に示す (v) より従う．

(iii) $\dim \mathcal{N}(\lambda_A I - A) = 1$ を示す．仮に $\boldsymbol{v}_1 > \boldsymbol{0}$, $\boldsymbol{v}_2 > \boldsymbol{0}$ を λ_A に対応する 1 次独立な固有ベクトルとすれば $\boldsymbol{v} = \alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2$ (α, β：スカラー) も λ_A に対応する固有ベクトルであるが，α, β を適当にとれば $\boldsymbol{v}$ の成分の 1 つをゼロにできるから (ii) に矛盾する．

(iv) $A \geq B \geq O$ とする．$\rho(A^{\mathrm{t}}) = \rho(A)$ であるから，(i) により $A^{\mathrm{t}}\boldsymbol{x} = \lambda_A \boldsymbol{x}$ をみたす $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} > \boldsymbol{0}$ がある．μ を $B = (b_{ij})$ の固有値で $B\boldsymbol{y} = \mu \boldsymbol{y}$, $\boldsymbol{y} = (y_1, \ldots, y_n)^{\mathrm{t}} \neq \boldsymbol{0}$ とすれば

$$|\mu||y_i| \leq \sum_{j=1}^{n} b_{ij}|y_j|$$

$$\therefore \quad |\mu||\boldsymbol{y}| \leq B|\boldsymbol{y}|, \quad |\boldsymbol{y}| = (|y_1|, \ldots, |y_n|)^{\mathrm{t}}$$

$$\therefore \quad |\mu| \cdot \boldsymbol{x}^{\mathrm{t}}|\boldsymbol{y}| \leq \boldsymbol{x}^{\mathrm{t}} B|\boldsymbol{y}| \leq \boldsymbol{x}^{\mathrm{t}} A|\boldsymbol{y}| = \lambda_A \boldsymbol{x}^{\mathrm{t}}|\boldsymbol{y}| \tag{7.4}$$

また $\boldsymbol{x}^{\mathrm{t}}|\boldsymbol{y}| > 0$ ($\because$ $|\boldsymbol{y}| \neq \boldsymbol{0}$, $|\boldsymbol{y}| \geq \boldsymbol{0}$ かつ $\boldsymbol{x}^{\mathrm{t}} > \boldsymbol{0}$) であるから (7.4) より $|\mu| \leq \lambda_A$ である．特に $|\mu| = \lambda_A$ ならば (7.4) の不等号はすべて等号となり

$$|\mu| \boldsymbol{x}^{\mathrm{t}}|\boldsymbol{y}| = \boldsymbol{x}^{\mathrm{t}} B|\boldsymbol{y}| = \boldsymbol{x}^{\mathrm{t}} A|\boldsymbol{y}| = \lambda_A \boldsymbol{x}^{\mathrm{t}}|\boldsymbol{y}|$$

よって

$$\boldsymbol{x}^{\mathrm{t}}(|\mu||\boldsymbol{y}| - B|\boldsymbol{y}|) = \boldsymbol{0} \quad \text{かつ} \quad \boldsymbol{x}^{\mathrm{t}}(|\mu||\boldsymbol{y}| - A|\boldsymbol{y}|) = \boldsymbol{0}$$

しかし $|\mu||\boldsymbol{y}| - B|\boldsymbol{y}| \leq \boldsymbol{0}$ かつ $|\mu||\boldsymbol{y}| - A|\boldsymbol{y}| \leq \boldsymbol{0}$ であるから上の 2 式と $\boldsymbol{x}^{\mathrm{t}} > \boldsymbol{0}$ より

$$|\mu||\boldsymbol{y}| - B\boldsymbol{y} = \boldsymbol{0} \quad \text{かつ} \quad |\mu||\boldsymbol{y}| = A|\boldsymbol{y}|$$

$$\therefore \quad A|\boldsymbol{y}| = B|\boldsymbol{y}| = |\mu||\boldsymbol{y}| \tag{7.5}$$

このとき

$$(I + A)^{n-1}|\boldsymbol{y}| = (1 + |\mu|)^{n-1}|\boldsymbol{y}| \tag{7.6}$$

正行列は既約であるから定理 7.3 によって $(I + A)^{n-1} > O$

$$\therefore \quad (I + A)^{n-1}|\boldsymbol{y}| > O \tag{7.7}$$

(7.6) と (7.7) より $|\boldsymbol{y}| > \boldsymbol{0}$ よって (7.5) より $B = A$ を得る．

(v) A の第 i 行と第 i 列の要素をすべてゼロでおきかえて得られる行列を B_i とすれば $A \neq B_i$ かつ $A \geq B_i$ ($\because$ $A > O$) よって (iv) より

$$\rho(B_i) < \lambda_A \tag{7.8}$$

A の第 i 行と第 i 列を取り除いて得られる $n-1$ 次行列を A_{ii} とすれば

$$(B_i\text{の固有値}) = (A_{ii}\text{の固有値}) \cup \{0\} \tag{7.9}$$

$g_i(\lambda) = \det(\lambda I_{n-1} - A_{ii})$ (A_{ii} の特性多項式) とおけば (7.8) と (7.9) によって $\lambda \geq \lambda_A$ のとき

$g_i(\lambda) > 0$ である．

ここで $f(\lambda) = \det(\lambda I_n - A)$ とおけば §4.7 の公式によって

$$f'(\lambda) = \frac{d}{d\lambda}\begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ & & \cdots & \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & \cdots & 0 \\ -a_{21} & \lambda - a_{22} & & \\ & & \ddots & \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ & & \ddots & \\ 0 & 0 & \cdots & 1 \end{vmatrix}$$

$$= \begin{vmatrix} \lambda - a_{22} & -a_{23} & \cdots & -a_{2n} \\ -a_{32} & \lambda - a_{33} & \cdots & -a_{3n} \\ & & \ddots & \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1\,n-1} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2\,n-1} \\ & & \ddots & \\ -a_{n-1\,1} & -a_{n-1\,2} & \cdots & \lambda - a_{n-1\,n-1} \end{vmatrix}$$

$$= g_1(\lambda) + \cdots + g_n(\lambda)$$

よって

$$\lambda \geq \lambda_A \;\Rightarrow\; f'(\lambda) > 0 \quad \text{特に} \quad f'(\lambda_A) > 0$$

したがって λ_A は $f(\lambda) = 0$ の単根である． 証明終 ∎

7.3 非負行列に対する Frobenius の定理

定理 7.5 (Frobenius) A を n 次非負行列としその固有値 λ_i を $|\lambda_1| \geq \cdots \geq |\lambda_n|$ と並べるとき，次が成り立つ．

(i) $\lambda_1 = \lambda_A \geq 0$

(ii) λ_A に対応する固有値ベクトルとして非負ベクトルがとれる．

(iii) $A \geq B \geq O$ ならば $\lambda_A \geq \rho(B)$

この場合も λ_A を **Perron 根** (ときに **Perron-Frobenius 根**) という．

【証明】 (i), (ii) $A = (a_{ij})$ とする．k を自然数として

$$A^{(k)} = (a_{ij}^{(k)}), \quad a_{ij}^{(k)} = \begin{cases} a_{ij} & (a_{ij} \neq 0 \text{ のとき}) \\ \dfrac{1}{k} & (a_{ij} = 0 \text{ のとき}) \end{cases}$$

とおけば $A^{(k)}$ は正行列である．故に定理 7.4 により $A^{(k)}$ は Perron 根 $\lambda_{A^{(k)}} > 0$ と対応する固有ベクトル $\boldsymbol{x}^{(k)} > \boldsymbol{0}$ をもち，$A^{(k)}$ のその他の固有値 $\lambda(A^{(k)})$ は $|\lambda(A^{(k)})| < \lambda_{A^{(k)}}$ をみたす．ここで

$$\boldsymbol{x}^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})^{\mathrm{t}}, \quad \sum_{i=1}^{n} x_i^{(k)} = 1$$

と正規化すれば，$\{\boldsymbol{x}^{(k)}\}$ $(k = 1, 2, \ldots)$ は有界列であるから適当な収束部分列を抽き出すことができる．これを $\boldsymbol{x}^{(k_j)} \to \boldsymbol{x}$ $(j \to \infty)$, $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$, $k_1 < k_2 < \cdots$ とする．

$$A^{(k_j)} \geq A^{(k_{j+1})} \geq \cdots$$

であるから定理 7.4 (iv) によって $\lambda_{A^{(k_j)}} \geq \lambda_{A^{(k_j+1)}} \geq \cdots \geq 0$. よって

$$\lambda_A = \lim_{j\to\infty} \lambda_{A^{(k_j)}} \geq 0$$

また $A^{(k_j)}\boldsymbol{x}^{(k_j)} = \lambda_{A^{(k_j)}}\boldsymbol{x}^{(k_j)}$ で $j \to \infty$ とすれば $A\boldsymbol{x} = \lambda_A\boldsymbol{x}$. さらに $\sum_{i=1}^n x_i^{(k_j)} = 1$, $x_i^{(k)} > 0 \quad \forall i$ と $\boldsymbol{x}^{(k_j)} \to \boldsymbol{x}$ $(j \to \infty)$ より $\sum_{i=1}^n x_i = 1$, $x_i \geq 0$. よって $\boldsymbol{x} \geq \boldsymbol{0}$ かつ $\boldsymbol{x} \neq \boldsymbol{0}$

A の任意の固有値 λ_i は $A^{(k_j)}$ のある固有値 $\lambda_i^{(k_j)}$ の $j \to \infty$ としたときの極根であるから, $|\lambda_i^{(k_j)}| \leq \lambda_{A^{(k_j)}}$ より $|\lambda_i| \leq \lambda_A \quad \therefore\ \lambda_A = \rho(A)$

(iii) $A^{(k_j)} > B \geq O$ と系 7.2.1 によって $\rho(B) \leq \lambda_{A^{(k_j)}} \to \lambda_A$ $(j \to \infty)$

$$\therefore\quad \rho(B) \leq \rho(A)$$

証明終 ∎

系 7.5.1 $A = (a_{ij}) \geq O$, $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} > \boldsymbol{0}$ のとき

$$M_A(\boldsymbol{x}) = \max_i \frac{\sum_{j=1}^n a_{ij}x_j}{x_i}, \quad m_A(\boldsymbol{x}) = \min_i \frac{\sum_{j=1}^n a_{ij}x_j}{x_i}$$

とおけば

$$m_A(\boldsymbol{x}) \leq \lambda_A \leq M_A(\boldsymbol{x})$$

【証明】 $A^{\mathrm{t}}\boldsymbol{y} = \lambda_A\boldsymbol{y}$ なる非負ベクトル $\boldsymbol{y} \neq \boldsymbol{0}$ をとる. $M_A(\boldsymbol{x})$ と $m_A(\boldsymbol{x})$ の定義より

$$m_A(\boldsymbol{x})\boldsymbol{x} \leq A\boldsymbol{x} \leq M_A(\boldsymbol{x})\boldsymbol{x} \quad \therefore\quad m_A(\boldsymbol{x})\boldsymbol{y}^{\mathrm{t}}\boldsymbol{x} \leq \boldsymbol{y}^{\mathrm{t}}A\boldsymbol{x} \leq M_A(\boldsymbol{x})\boldsymbol{y}^{\mathrm{t}}\boldsymbol{x}$$

$\boldsymbol{y}^{\mathrm{t}}A\boldsymbol{x} = (A^{\mathrm{t}}\boldsymbol{y})^{\mathrm{t}}\boldsymbol{x} = \lambda_A\boldsymbol{y}^{\mathrm{t}}\boldsymbol{x}$ であるから上式より

$$m_A(\boldsymbol{x})\boldsymbol{y}^{\mathrm{t}}\boldsymbol{x} \leq \lambda_A\boldsymbol{y}^{\mathrm{t}}\boldsymbol{x} \leq M_A(\boldsymbol{x})\boldsymbol{y}^{\mathrm{t}}\boldsymbol{x}$$

ここで $\boldsymbol{y}^{\mathrm{t}}\boldsymbol{x} > 0$ ($\because\ \boldsymbol{y} \geq \boldsymbol{0}$, $\boldsymbol{y} \neq \boldsymbol{0}$ かつ $\boldsymbol{x} > \boldsymbol{0}$) であるから $m_A(\boldsymbol{x}) \leq \lambda_A \leq M_A(\boldsymbol{x})$ を得る. 証明終 ∎

系 7.5.2 $A = (a_{ij}) \geq O$ とし, $r_i = \sum_{j=1}^n a_{ij}$ とおくとき

$$\min_i r_i \leq \lambda_A \leq \max_i r_i$$

【証明】 系 7.5.1 において $\boldsymbol{x} = (1, 1, \ldots, 1)^{\mathrm{t}}$ とすればよい. 証明終 ∎

系 7.5.3 $A = (a_{ij}) \geq O$ ならば $\lambda_A \geq \max_i a_{ii}$

【証明】 定理 7.5 (iii) において $B = \mathrm{diag}(a_{11}, \ldots, a_{nn})$ ととればよい ($\rho(B) = \max_i a_{ii}$ であるから). 証明終 ∎

7.4 既約非負行列に対する Frobenius の定理

定理 7.6 (Frobenius) A を n 次既約非負行列としその固有値 λ_i を $|\lambda_1| \geq \cdots \geq |\lambda_n|$ と並べるとき, 次が成り立つ.

(i) $\lambda_A = \lambda_1 > 0$
(ii) λ_A に対応する固有ベクトルとして正ベクトルがとれる.
(iii) λ_A の幾何学的重複度は 1 である.
(iv) $A \geq B \geq O$ ならば $\lambda_A \geq \rho(B)$ かつ等号が成り立つのは $A = B$ のときに限る.
(v) λ_A は単純, すなわち代数的重複度は 1 である.

この場合も λ_A を **Perron 根** (ときに **Perron-Frobenius 根**) という.

【証明】 (i), (ii) $\lambda_A=\rho(A)$, $A^{(k)}=\frac{1}{\lambda_A+\frac{1}{k}}A$ とおくと

$$\rho(A^{(k)})=\frac{\lambda_A}{\lambda_A+\frac{1}{k}}<1$$

よって定理 7.1 (vi) によって $\boldsymbol{y}^{(k)}>A^{(k)}\boldsymbol{y}^{(k)}$ をみたす正ベクトル $\boldsymbol{y}^{(k)}$ が存在する.

$$\therefore\quad \Big(\lambda_A+\frac{1}{k}\Big)\boldsymbol{y}^{(k)}>A\boldsymbol{y}^{(k)} \tag{7.10}$$

ここで

$$\boldsymbol{y}^{(k)}=(y_1^{(k)},\ldots,y_n^{(k)})^{\mathrm{t}},\quad \sum_{i=1}^{n}y_i^{(k)}=1$$

と正規化する. $\{\boldsymbol{y}^{(k)}\}$ $(k=1,2,\ldots)$ は有界列であるから収束部分列 $\{\boldsymbol{y}^{(k_j)}\}$ をとり出せる. これをあらためて $\{\boldsymbol{y}^{(l)}\}$ とかき, 収束先を $\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}}$ とすると $\boldsymbol{y}^{(l)}\to\boldsymbol{x}$ $(l\to\infty)$. ここで $\sum_{i=1}^{n}y_j^{(l)}=1$ より $\sum_{i=1}^{n}x_i=1$ $\quad\therefore\ \boldsymbol{x}\neq\boldsymbol{0}$ かつ $\boldsymbol{x}\geq\boldsymbol{0}$ $(\because\ \boldsymbol{y}^{(l)}>\boldsymbol{0}\quad\forall i)$
さらに $l\to\infty$ のとき

$$\Big(\lambda_A+\frac{1}{l}\Big)\boldsymbol{y}^{(l)}\to\lambda_A\boldsymbol{x}\quad\text{かつ}\quad A\boldsymbol{y}^{(l)}\to A\boldsymbol{x}$$

よって (7.10) より $\lambda_A\boldsymbol{x}\geq A\boldsymbol{x}$

したがって $\boldsymbol{u}=\lambda_A\boldsymbol{x}-A\boldsymbol{x}$ とおくと $\boldsymbol{u}\geq\boldsymbol{0}$ かつ定理 7.3 より $\boldsymbol{v}=(I+A)^{n-1}\boldsymbol{x}>\boldsymbol{0}$ で

$$\begin{aligned}(I+A)^{n-1}\boldsymbol{u}&=(I+A)^{n-1}(\lambda_A\boldsymbol{x}-A\boldsymbol{x})=(I+A)^{n-1}(\lambda_AI-A)\boldsymbol{x}\\&=(\lambda_AI-A)\boldsymbol{v}=\lambda_A\boldsymbol{v}-A\boldsymbol{v}\end{aligned}\tag{7.11}$$

ここで $\boldsymbol{u}=\boldsymbol{0}$ でなければならない. (なぜならば仮に $\boldsymbol{u}\neq\boldsymbol{0}$ ならば定理 7.3 より $(I+A)^{n-1}\boldsymbol{u}>\boldsymbol{0}$ したがって (7.11) より

$$\lambda_A\boldsymbol{v}>A\boldsymbol{v} \tag{7.12}$$

A は既約で $\boldsymbol{v}>\boldsymbol{0}$ であるから $A\boldsymbol{v}>\boldsymbol{0}$. よって (7.12) より $\lambda_A>0$ となるが, このとき (7.12) と系 7.1.3 より $\lambda_A>\rho(A)$. これは矛盾である.) よって $\lambda_A\boldsymbol{x}=A\boldsymbol{x}$ を得るから

$$\boldsymbol{0}<\boldsymbol{v}=(I+A)^{n-1}\boldsymbol{x}=(1+\lambda_A)^{n-1}\boldsymbol{x}\quad\text{したがって}\quad\boldsymbol{x}>\boldsymbol{0}$$

(iii)〜(v) は定理 7.4 (iii)〜(v) の証明がそのまま通用する. 証明終 ∎

系 7.6.1 A を既約非負行列とする. λ が A の固有値で $\lambda\neq\lambda_A$ ならば λ に対応する固有ベクトル $\boldsymbol{y}$ は非負ベクトルではない.

【証明】 $A^{\mathrm{t}}\boldsymbol{x}=\lambda_A\boldsymbol{x}$, $\boldsymbol{x}>\boldsymbol{0}$ とすれば $A\boldsymbol{y}=\lambda\boldsymbol{y}$, $\boldsymbol{y}\neq\boldsymbol{0}$ より

$$\lambda\boldsymbol{x}^{\mathrm{t}}\boldsymbol{y}=\boldsymbol{x}^{\mathrm{t}}A\boldsymbol{y}=\lambda_A\boldsymbol{x}^{\mathrm{t}}\boldsymbol{y}\quad\therefore\quad(\lambda-\lambda_A)\boldsymbol{x}^{\mathrm{t}}\boldsymbol{y}=\boldsymbol{0}$$

$\lambda\neq\lambda_A$ であるから $\boldsymbol{x}^{\mathrm{t}}\boldsymbol{y}=\boldsymbol{0}$. しかし $\boldsymbol{x}>\boldsymbol{0}$ であるから $\boldsymbol{y}\neq\boldsymbol{0}$ は非負ベクトルではあり得ない.

証明終 ∎

系 7.6.2 A を既約非負行列とする. $A\boldsymbol{x}=\lambda\boldsymbol{x}$, $\boldsymbol{x}>\boldsymbol{0}$ とすれば $\lambda=\lambda_A$

【証明】 系 7.6.1 より明らかである. 証明終 ∎

定理 7.7 (Frobenius) A を n 次既約非負行列，$B=(b_{ij})$ を n 次の (実または複素) 行列，$|B|=(|b_{ij}|)\le A$ とするとき，次が成り立つ．

(i) β を B の任意の固有値とすれば $|\beta|\le\lambda_A=\rho(A)$

(ii) 上の不等式において等号が成り立つのは $|B|=A$ のときに限る．

(iii) (Wielandt) さらに精密に $\beta=\lambda_A e^{\sqrt{-1}\varphi}$ ($\sqrt{-1}$ は虚数単位) ならば

$$B=e^{\sqrt{-1}\varphi}DAD^{-1},\ D=\mathrm{diag}(d_1,\dots,d_n),\ |d_i|=1\quad\forall i$$

と書ける．

【証明】 (i), (ii) は定理 7.2 と定理 7.6 で済んでいる．実際，定理 7.2 によって

$$|\beta|\le\rho(B)\le\rho(|B|)\le\rho(A)=\lambda_A$$

さらに $|\beta|=\lambda_A$ ならば $\rho(|B|)=\lambda_A$ よって定理 7.6 (iv) により $|B|=A$.
(iii) この部分が定理 7.7 の核心である．$\beta=\lambda_A e^{\sqrt{-1}\varphi}$ に対応する固有ベクトルを $\boldsymbol{y}\neq\boldsymbol{0}$ とすれば $B\boldsymbol{y}=\beta\boldsymbol{y}$ より

$$|\beta||\boldsymbol{y}|\le|B|\cdot|\boldsymbol{y}|,\quad |\boldsymbol{y}|=(|y_1|,\dots,|y_n|)^{\mathrm{t}}$$

$\beta=\lambda_A e^{\sqrt{-1}\varphi}$ ならば $|\beta|=\lambda_A$．このとき定理 7.4 (iv) の証明と同様にして

$$|\beta||\boldsymbol{y}|=|B|\cdot|\boldsymbol{y}|=A|\boldsymbol{y}|,\quad |\boldsymbol{y}|>\boldsymbol{0} \tag{7.13}$$

を得る．いま $y_i=e^{\sqrt{-1}\theta_i}|y_i|=d_i|y_i|$, $d_i=e^{\sqrt{-1}\theta_i}$, $D=\mathrm{diag}(d_1,\dots,d_n)$ とおけば $\boldsymbol{y}=D|\boldsymbol{y}|$ と書けるから $B\boldsymbol{y}=\beta\boldsymbol{y}$ より $BD|\boldsymbol{y}|=\beta D|\boldsymbol{y}|=|\beta|e^{\sqrt{-1}\varphi}D|\boldsymbol{y}|$

$$\therefore\quad e^{-\sqrt{-1}\varphi}D^{-1}BD|\boldsymbol{y}|=|\beta||\boldsymbol{y}|\quad(=A|\boldsymbol{y}|,\ (7.13)\text{ 参照})$$

$C=e^{-\sqrt{-1}\varphi}D^{-1}BD$ とおくと

$$C|\boldsymbol{y}|=|\beta||\boldsymbol{y}| \tag{7.14}$$

$C=(c_{ij})$ に対し $\overline{C}=(\overline{c}_{ij})$ とおくと (7.14) の共役をとって $\overline{C}|\boldsymbol{y}|=|\beta||\boldsymbol{y}|$

$$\therefore\quad (C+\overline{C})|\boldsymbol{y}|=2|\beta||\boldsymbol{y}|=2A|\boldsymbol{y}| \tag{7.15}$$

$|D^{-1}|=|D|=I_n$ であるから $|C|=|B|=A$．故に (7.14) と (7.15) より

$$\big((C+\overline{C})-2|C|\big)|\boldsymbol{y}|=\boldsymbol{0}$$

ここで $C+\overline{C}\le 2|C|$ ($\because\ c_{ij}+\overline{c}_{ij}=2\mathrm{Re}c_{ij}\le 2|c_{ij}|$) であるから $|\boldsymbol{y}|>\boldsymbol{0}$ と併せて

$$C+\overline{C}-2|C|=O\quad\text{すなわち}\quad C=|C|=A$$

$$\therefore\quad e^{-\sqrt{-1}\varphi}D^{-1}BD=A$$

$$\therefore\quad B=e^{\sqrt{-1}\varphi}DAD^{-1},\quad D=\mathrm{diag}(d_1,\dots,d_n),\quad |d_i|=1\quad\forall i$$

証明終 ∎

■**注意** 固有ベクトルは定数倍してもよいから上記証明において $\theta_1=0$ (したがって $d_1=1$) としてもよい．

7.5 巡回指数 $k > 1$ の既約非負行列に対する Frobenius の定理

A を既約非負行列とし $\rho(A) = |\lambda|$ なる A の固有値 λ の個数を代数的重複度を込めて k とするとき，$k = 1$ ならば A は**原始的 (primitive)** であるといい，$k > 1$ ならば指数 k の**巡回行列 (cyclic matrix of index k)** という．定理 7.4 によって正行列は原始的であるが一般にはそうではない．この節では $k > 1$ の場合における Frobenius の結果を述べる．

定理 7.8 (Frobenius) n 次既約非負行列 A が指数 $k > 1$ の巡回行列であるとし，A の固有値 $\lambda_1, \ldots, \lambda_n$ を次のように並べる．

$$\lambda_A = \rho(A) = |\lambda_1| = \cdots = |\lambda_k| > |\lambda_{k+1}| \geq \cdots \geq |\lambda_n|$$

このとき n 個の固有値 λ_i は原点のまわりに角 $\frac{2\pi}{k}$ の回転を行っても不変であるが，$\frac{2\pi}{k}$ より小さい角で回転すれば変化する．さらに適当な置換行列 P をえらんで

$$PAP^{\mathrm{t}} = \begin{pmatrix} O & A_{12} & O & \cdots & O \\ \vdots & O & A_{23} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & O \\ O & \vdots & & O & A_{k-1\ k} \\ A_{k1} & O & \cdots & \cdots & O \end{pmatrix}$$

とできる．ただしブロック対角要素の O は正方行列をあらわす．

【証明】 [Wielandt [41] による] 以下 λ_A を λ とかくことにし，定理 7.6 (v) を考慮して

$$\lambda_\kappa = \lambda e^{\sqrt{-1}\varphi_\kappa},\ \kappa = 1, 2, \ldots, k\ (0 = \varphi_1 < \varphi_2 \leq \cdots \leq \varphi_k < 2\pi)$$

とおく．定理 7.7 を $B = A$ に適用すれば適当な対角行列 D_κ をえらんで

$$A = e^{\sqrt{-1}\varphi_\kappa} D_\kappa A D_\kappa^{-1},\ \kappa = 1, 2, \ldots, k \tag{7.16}$$

と書ける．定理 7.6 により λ は単純であるから，$\lambda e^{\sqrt{-1}\varphi_\kappa}$ は $e^{\sqrt{-1}\varphi_\kappa} D_\kappa A D_\kappa^{-1}$ の単純固有値である．一方 $\lambda e^{\sqrt{-1}\varphi_\kappa}$ は A の固有値であるから結局 $\lambda_\kappa = \lambda e^{\sqrt{-1}\varphi_\kappa}$ は A の単純固有値である．すなわち $\varphi_1, \ldots, \varphi_k$ は互いに相異なる． (証明つづく)

ここで証明を中断し，2 つの補題を証明する．

補題 7.1 $\varphi_1, \ldots, \varphi_k$ は $\mathrm{mod} 2\pi$ で加群をなす．

【証明】 (7.16) において $\kappa = \mu$ として $A = e^{\sqrt{-1}\varphi_\mu}(D_\mu A D_\mu^{-1})$．これを (7.16) の右辺に代入して

$$\begin{aligned} A &= e^{\sqrt{-1}\varphi_\kappa} D_\kappa (e^{\sqrt{-1}\varphi_\mu} D_\mu A D_\mu^{-1}) D_\kappa^{-1} = e^{\sqrt{-1}(\varphi_\kappa + \varphi_\mu)} D_\kappa D_\mu A D_\mu^{-1} D_\kappa^{-1} \\ &= e^{\sqrt{-1}(\varphi_\kappa + \varphi_\mu)} D_\mu D_\kappa A D_\kappa^{-1} D_\mu^{-1} \end{aligned} \tag{7.17}$$

また (7.16) より

$$e^{-\sqrt{-1}\varphi_\kappa} A = D_\kappa A D_\kappa^{-1}$$

$$\therefore\quad A = e^{-\sqrt{-1}\varphi_\kappa} D_\kappa^{-1} A D_\kappa$$

ここで再び $\kappa = \mu$ として $A = e^{-\sqrt{-1}\varphi_\mu} D_\mu^{-1} A D_\mu$ を (7.16) の右辺に代入すれば

$$A = e^{\sqrt{-1}\varphi_\kappa} D_\kappa (e^{-\sqrt{-1}\varphi_\mu} D_\mu^{-1} A D_\mu) D_\kappa^{-1} = e^{\sqrt{-1}(\varphi_\kappa - \varphi_\mu)} D_\mu^{-1} D_\kappa A D_\kappa^{-1} D_\mu \tag{7.18}$$

ここで $T = D_\mu^{-1} D_\kappa$ とおけば

$$A = e^{\sqrt{-1}(\varphi_\kappa - \varphi_\mu)} TAT^{-1} \tag{7.19}$$

また $S = D_\mu D_\kappa$ とおけば (7.17) より

$$A = e^{\sqrt{-1}(\varphi_\kappa + \varphi_\mu)} SAS^{-1} \tag{7.20}$$

(7.19) と (7.20) によって A と $e^{\sqrt{-1}(\varphi_\kappa \pm \varphi_\mu)} A$ の固有値は一致する.

よって $\lambda e^{\sqrt{-1}(\varphi_\kappa \pm \varphi_\mu)}$ は A の固有値であり, 適当な φ_ν により

$$\lambda e^{\sqrt{-1}(\varphi_\kappa \pm \varphi_\mu)} = \lambda e^{i\varphi_\nu}$$

と書ける. $\therefore\ \varphi_\kappa \pm \varphi_\mu = \varphi_\nu \ (\mathrm{mod} 2\pi)$ 証明終 ∎

補題 7.2 $\varphi_\kappa = (\kappa - 1)\dfrac{2\pi}{k},\ \kappa = 1, 2, \ldots, k$

【証明】 $0 = \varphi_1 < \varphi_2 < \cdots < \varphi_k < 2\pi$ であるが補題 7.1 によって

$$\varphi_2 + \varphi_2 = \varphi_m \quad (\mathrm{mod} 2\pi)$$

をみたす φ_m がある. ここで $m > 3$ ならば $\varphi_2 + \varphi_2 > \varphi_3 > \varphi_2$

$$\therefore\quad \varphi_2 > \varphi_3 - \varphi_2 > 0 = \varphi_1$$

補題 7.1 によって $\varphi_3 - \varphi_2 = \varphi_l$ となる l があるがこれは不可能であるから $m = 3$ でなければならない. したがって $2\varphi_2 = \varphi_3$

同様にある m につき $3\varphi_2 = \varphi_m \ (\mathrm{mod} 2\pi)$ であるが, 仮に $m > 4$ とすれば $\varphi_m = 3\varphi_2 > \varphi_4 > \varphi_3 = 2\varphi_2 \quad \therefore\ \varphi_2 > \varphi_4 - \varphi_3 > 0 = \varphi_1$. これは不可能であるから $m \leq 4 \quad \therefore\ m = 4,\ 3\varphi_2 = \varphi_4$

以下これを続けて $(\kappa - 1)\varphi_2 = \varphi_\kappa,\ \kappa = 1, 2, \ldots, k$ となるから $\{\varphi_\kappa\}\ (1 \leq \kappa \leq k)$ は φ_2 を生成元とする巡回群で $k\varphi_2 = \varphi_1 = 0 \ (\mathrm{mod} 2\pi)$

$$\therefore\quad \varphi_2 = \frac{2\pi}{k}$$

結局 $\varphi_\kappa = (\kappa - 1)\dfrac{2\pi}{k},\ \kappa = 1, 2, \ldots, k$ を得る. 証明終 ∎

以上の準備の下に定理 7.8 の証明に戻ろう.

【定理 7.8 の証明続き】 補題 7.2 によって $\lambda_1, \ldots, \lambda_k$ は原点を中心として半径 $\lambda = \lambda_A$ の円周上の等分点であり, 正 k 角形の頂点をなす. また残りの固有値も $\frac{2\pi}{k}$ の回転により不変である. なぜならば A の任意の固有値を $re^{\sqrt{-1}\theta}$ とすれば $A = e^{\sqrt{-1}\varphi_2} D_2 A D_2^{-1}$ より

$$re^{\sqrt{-1}\theta + \sqrt{-1}\varphi_2} = re^{\sqrt{-1}(\theta + \varphi_2)}$$

も A の固有値であるからである.

次に (7.16) の分解において $\lambda_2 = \lambda e^{\sqrt{-1}\varphi_2}\ (\lambda = \lambda_A)$ に属する固有ベクトル $\boldsymbol{y}_2$ を適当な対角行列 D_2 を用いて $\boldsymbol{y}_2 = D_2 |\boldsymbol{y}_2|$ とあらわす (定理 7.7 の証明参照). $\boldsymbol{y}_2$ の成分のうち偏角の等しいもの毎に集めれば適当な置換行列 P により

$$PD_2P^{\mathrm{t}} = \begin{pmatrix} e^{\sqrt{-1}\delta_1} I_{n_1} & & \\ & \ddots & \\ & & e^{\sqrt{-1}\delta_s} I_{n_s} \end{pmatrix} \quad \left(\sum_{j=1}^{s} n_j = n\right) \tag{7.21}$$

とできる. $\boldsymbol{y}_2$ の第 1 成分を 1 と正規化して $\delta_1 = 0$ としてよい. $\lambda_2 \neq \lambda_A$ であるから系 7.6.1 により $s > 1$ である. このとき (7.16) より

$$PAP^{\mathrm{t}} = e^{\sqrt{-1}\varphi_2}(PD_2P^{\mathrm{t}})(PAP^{\mathrm{t}})(PD_2P^{\mathrm{t}})^{-1} \quad \left(\varphi_2 = \frac{2\pi}{k}\right) \tag{7.22}$$

(7.21) に対応して PAP^{t} を s^2 個の部分行列 $A_{\nu\mu}$ $(1 \leq \nu,\ \mu \leq s)$ に分割すれば，(7.21)，(7.22) より

$$A_{\nu\mu} = e^{\sqrt{-1}(\varphi_2+\delta_\nu-\delta_\mu)}A_{\nu\mu} \quad (1 \leq \nu,\ \mu \leq s) \tag{7.23}$$

したがって $\frac{2\pi}{k} + \delta_\nu \not\equiv \delta_\mu \pmod{2\pi}$ ならば $A_{\nu\mu} = O$.

各 ν につき $\frac{2\pi}{k} + \delta_\nu \equiv \delta_\mu \pmod{2\pi}$ なる μ は高々 1 つしかないから $A_{\nu\mu} \neq O$ なる $A_{\nu\mu}$ は各 ν につき高々 1 つ存在する．しかし A は既約であるから各 ν につき $A_{\nu\mu} \neq O$ なる μ は少なくとも 1 つは存在する．よって各 ν につき $A_{\nu\mu} \neq O$ なる $A_{\nu\mu}$ はちょうど 1 つ存在する．したがって各 ν につき合同式 $\delta_\mu \equiv \frac{2\pi}{k} + \delta_\nu \pmod{2\pi}$ は一意解 $\delta_\mu = \delta_{\mu(\nu)}$ をもつ．(同様に各 μ につき $A_{\nu\mu} \neq O$ なる $A_{\nu\mu}$ もちょうど 1 つ存在し，各 μ につき $\delta_\nu \equiv \delta_\mu - \frac{2\pi}{k} \pmod{2\pi}$ は一意解 $\delta_\nu = \delta_{\nu(\mu)}$ をもつ．)

このとき

$$\begin{aligned}
\delta_1 &= 0 \\
\delta_{\mu(1)} &= \frac{2\pi}{k} + \delta_1 \\
\delta_{\mu(2)} &= \frac{2\pi}{k} + \delta_2 \\
&\vdots \\
\delta_{\mu(s-1)} &= \frac{2\pi}{k} + \delta_{s-1}
\end{aligned}$$

よって PAP^{t}, PD_2P^{t} の行と列を同時に並べかえて $\delta_{\mu(\kappa)} = \delta_{\kappa+1}$ $(1 \leq \kappa \leq s-1)$ とすれば

$$\begin{aligned}
\delta_2 &= \delta_{\mu(1)} = \frac{2\pi}{k} \\
\delta_3 &= \delta_{\mu(2)} = \frac{2\pi}{k} + \delta_2 = 2 \times \frac{2\pi}{k} \\
&\vdots \\
\delta_\nu &= \delta_{\mu(\nu-1)} = \frac{2\pi}{k} + \delta_{\nu-1} = (\nu-1)\frac{2\pi}{k} \quad (1 \leq \nu \leq \min(k,\ s))
\end{aligned} \tag{7.24}$$

次に $s = k$ を示す．

(i)　まず $s \geq k$ をいう．仮に $1 < s < k$ とすれば $k > 2$ である．(7.24) より

$1 \leq \mu < s-1$ のとき　　$\frac{2\pi}{k} + \delta_s - \delta_\mu = \frac{2\pi}{k} + (s-\mu)\frac{2\pi}{k}$

$1 \leq s-\mu < s < k$ より　$1 \leq s-\mu \leq k-2$　$\therefore$　$2 \leq 1+s-\mu \leq k-1$

よって

$$\frac{2\pi}{k} + (s-\mu)\frac{2\pi}{k} \not\equiv 0 \pmod{2\pi} \quad \therefore \quad A_{s\mu} = O \quad (\mu \leq s-1))$$

これは A が可約であることになり仮定に反する．よって $s \geq k$ である．

(ii)　次に $s \leq k$ をいう．仮に $s > k$ として矛盾を導く．(7.24) より

$1 \leq \nu \leq k$, $k < \mu \leq s$ のとき

$$\frac{2\pi}{k} + \delta_\nu - \delta_\mu = \frac{2\pi}{k} + (\nu-1)\frac{2\pi}{k} - (\mu-1)\frac{2\pi}{k} = \frac{2\pi}{k}(1+\nu-\mu)$$

ここで $1+\nu-\mu < 1$ に注意すれば

$$\frac{2\pi}{k} + \delta_\nu - \delta_\mu \not\equiv 0 \pmod{2\pi} \quad \therefore \quad A_{\nu\mu} = O \quad (1 \leq \nu \leq k,\ k < \mu \leq s)$$

$$\therefore\quad PAP^{\mathrm{t}}=\begin{pmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{s1} & \cdots & A_{ss} \end{pmatrix}=\left(\begin{array}{ccc|ccc} A_{11} & \cdots & A_{1k} & & & \\ \vdots & & \vdots & & \mathbf{0} & \\ A_{k1} & \cdots & A_{kk} & & & \\ \hline A_{k+1\ 1} & \cdots & A_{k+1\ k} & A_{k+1\ k+1} & \cdots & A_{k+1\ s} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{s1} & \cdots & A_{sk} & A_{s\ k+1} & \cdots & A_{ss} \end{array}\right)$$

故に A は可約となって仮定に反す．　$\therefore s\leq k$

よって (i), (ii) より $s=k$ を得る．したがって $A_{\kappa\,\kappa+1}\neq O$ $(1\leq\kappa\leq k-1)$.

結局適当な置換行列 P をえらんで

$$PAP^{\mathrm{t}}=e^{\sqrt{-1}\varphi_2}(PD_2P^{\mathrm{t}})(PAP^{\mathrm{t}})(PD_2P^{\mathrm{t}})^{-1}$$

$$=\begin{pmatrix} O & A_{12} & & & \mathbf{0} \\ O & O & A_{23} & & \\ \vdots & & \ddots & \ddots & \\ \vdots & & & \ddots & A_{k-1\ k} \\ A_{k1} & O & \cdots & \cdots & O \end{pmatrix}$$

となる．ただし最後のブロック行は各行各列に $A_{\nu\mu}\neq O$ がただ 1 つ存在することより $A_{k1}\neq O$ かつ $A_{kj}=O$ $(2\leq j\leq k)$ がわかる．　証明終 ∎

注意　$A_{k1}\neq O$ は

$$\delta_{\nu(1)}\equiv\delta_1-\frac{2\pi}{k}\ (\mathrm{mod}2\pi)\equiv-\frac{2\pi}{k}\ (\mathrm{mod}2\pi)\equiv(k-1)\frac{2\pi}{k}\ (\mathrm{mod}2\pi)\quad\therefore\quad\nu(1)=k$$

として導くこともできる．

【付記】定理 7.8 は行列論における珠玉の定理であって，既約非負行列の固有値は原点を中心とする同心円の周上にまるで結晶格子のように整然と並んでいることを主張している．45 年前，この定理にはじめて出会ったときの感動を今あらためて思い出す．

第8章　Z 行列と M 行列

この章では，行列とベクトルはすべて実行列と実ベクトルに限定し，すべての非対角要素が負または零の行列 A と $A=\alpha I-B$ $(\alpha>0,\ B\geq O$ かつ $\rho(B)<\alpha)$ の形にあらわされる行列 A を中心にしてその性質を論じる．前者は **Z 行列**，後者は **M 行列**と呼ばれ数理経済，回路理論，数値解析等いろいろな分野で重要な役割を果す．

8.1　Z 行列

実行列 $A=(a_{ij})$ が $a_{ij}\leq 0$ $(i\neq j)$ をみたすとき A を **Z 行列**という．n 次 Z 行列の全体を $\mathcal{Z}_n$ または単に $\mathcal{Z}$ であらわす．$A\in\mathcal{Z}_n$ に対し

$$\alpha=\max_i a_{ii},\quad B=\alpha I-A$$

とおけば B は非負行列であるから次が成り立つ．

$$A\in\mathcal{Z}_n \Leftrightarrow \text{適当な正数}\alpha\text{と非負行列 } B \text{ により } A=\alpha I_n-B \quad \text{と書ける．}$$

定理 3.1～3.3 と密接に関連する結果として次が成り立つ．

定理 8.1 (LU 分解)　$A\in\mathcal{Z}_n$ かつ $A=LU$ $(L=(l_{ij})$ は下 3 角で $l_{ii}>0\ \forall i,\ U=(u_{ij})$ は上 3 角で $u_{ii}>0\ \forall i)$ と分解できるならば $L,U\in\mathcal{Z}_n$

【証明】　$A=(a_{ij}),\ a_{ij}\leq 0$ $(i\neq j)$ とおき，$l_{ii}>0,\ u_{ii}>0\ \forall i$ として，

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{pmatrix}$$

の両辺の要素を順次比較すればよい．

まず $a_{1j}=l_{11}u_{1j}\quad (j\geq 2)$ より

$$u_{1j}=\frac{a_{1j}}{l_{11}}\leq 0\quad (j\geq 2)\quad (\because\ a_{1j}\leq 0,\ l_{11}>0)$$

次に $a_{i1}=l_{i1}u_{11}\quad (i\geq 2)$ より

$$l_{i1}=\frac{a_{i1}}{u_{11}}\leq 0\quad (i\geq 2)\quad (\because\ a_{i1}\leq 0,\ u_{11}>0)$$

さらに $a_{2j}=l_{21}u_{1j}+l_{22}u_{2j}\quad (j\geq 3)$ より

$$u_{2j}=\frac{a_{2j}-l_{21}u_{1j}}{l_{22}}\leq 0\quad (j\geq 3)\quad (\because\ a_{2j}\leq 0,\ l_{21}\leq 0,\ u_{1j}\leq 0,\ l_{22}>0)$$

同様に $a_{i2}=l_{i1}u_{12}+l_{i2}u_{22}\quad (i\geq 3)$ より

$$l_{i2} = \frac{a_{i2} - l_{i1}u_{12}}{u_{22}} \leq 0 \quad (i \geq 3) \quad (\because\ a_{i2} \leq 0,\ l_{i1} \leq 0,\ u_{12} \leq 0,\ u_{22} > 0)$$

等々. 証明終 ∎

8.2 M 行列

n 次実行列 A が任意の $\boldsymbol{x} \in \mathbb{R}^n$ に対して

$$A\boldsymbol{x} \geq \boldsymbol{0} \ \Rightarrow\ \boldsymbol{x} \geq \boldsymbol{0}$$

をみたすとき A を**単調行列 (monotone matrix)** という.

定理 8.2 A が単調行列 $\Leftrightarrow$ A が正則で $A^{-1} \geq O$

【証明】 $\Rightarrow$ A を単調とし $A\boldsymbol{x} = \boldsymbol{0}$ とすれば $\boldsymbol{x} \geq \boldsymbol{0}$

一方このとき $A(-\boldsymbol{x}) = \boldsymbol{0}$ でもあるから $-\boldsymbol{x} \geq \boldsymbol{0}$ よって $\boldsymbol{x} = \boldsymbol{0}$ が従う. すなわち $A\boldsymbol{x} = \boldsymbol{0}$ は自明解 $\boldsymbol{x} = \boldsymbol{0}$ しかもたないから A は正則である.

さらに $\boldsymbol{e}_i$ を n 次の第 i 基本ベクトル (n 次単位行列の第 i 列) とすれば $\boldsymbol{e}_i \geq \boldsymbol{0}$ で $\boldsymbol{e}_i = A(A^{-1}\boldsymbol{e}_i) \geq \boldsymbol{0}$ であるから A の単調性により $A^{-1}\boldsymbol{e}_i \geq \boldsymbol{0}\ \forall i$

$A^{-1}\boldsymbol{e}_i$ は A^{-1} の第 i 列をあらわすから, これより $A^{-1} \geq O$ が従う.

$\Leftarrow$ $A\boldsymbol{x} \geq \boldsymbol{0}$ ならば両辺に左から $A^{-1} \geq O$ を乗じて $\boldsymbol{x} \geq \boldsymbol{0}$. よって A は単調である. 証明終 ∎

定義 8.1 (M 行列) 単調な $\mathcal{Z}$ 行列を **M 行列**という.

定理 8.3 $A = (a_{ij})$ が M 行列ならば $a_{ii} > 0\ \forall i$

【証明】 M 行列は $\mathcal{Z}$ 行列であるから $a_{ij} \leq 0\ (i \neq j)$ である.
いま $A^{-1} = (\alpha_{ij})$ とおけば $\alpha_{ij} \geq 0\ (\forall i, j)$. よって $A^{-1}A = I_n$ の (i,i) 要素を比較して

$$\alpha_{ii}a_{ii} + \sum_{j \neq i} \alpha_{ij}a_{ji} = 1$$

$$\therefore\quad \alpha_{ii}a_{ii} = 1 - \sum_{j \neq i} \alpha_{ij}a_{ji} \geq 1 \tag{8.1}$$

$\alpha_{ii} \geq 0$ であるから (8.1) より $\alpha_{ii} > 0$ かつ $a_{ii} > 0$ を得る. 証明終 ∎

定理 8.4 次の条件は互いに同値である.

(i) A が n 次 M 行列である. (すなわち $A \in \mathcal{Z}_n$ かつ A は正則で $A^{-1} \geq O$)
(ii) $A = \alpha I_n - B$ かつ $\rho(B) < \alpha$ をみたす正数 α と非負行列 B が存在する.
(iii) $A \in \mathcal{Z}_n$ かつ A の固有値の実部はすべて正である.
(iv) $A \in \mathcal{Z}_n$ かつ A の実固有値はすべて正である.

【証明】 (i)⇒(ii) $\alpha = \max_i a_{ii}$ とおけば定理 8.3 によって $\alpha > 0$. ここで $B = \alpha I_n - A$ とおけば $B \geq O$ で $A = \alpha I - B = \alpha(I - \frac{1}{\alpha}B)$. よって A が正則かつ $A^{-1} \geq O$ ならば $I - \frac{1}{\alpha}B$ が正則かつ

$(I-\frac{1}{\alpha}B)^{-1} \geq O$ である．故に定理7.1 (ii) と (iii) の同値性を $\frac{1}{\alpha}B$ に適用して $\rho(\frac{1}{\alpha}B) < 1 \quad \therefore\ \rho(B) < \alpha$

(ii)⇒(i)　再び定理 7.1 を $A = \alpha(I - \frac{1}{\alpha}B)$ に適用すればよい．よって (i) と (ii) は同値である．

(ii)⇒(iii)　λ を A の任意の固有値とし λ_B を B の Perron 根とすれば $B = \alpha I - A$ であるから

$$|\alpha - \lambda| \leq \rho(B) = \lambda_B < \alpha$$

$|\alpha - \mathrm{Re}\lambda| \leq |\alpha - \lambda|$ であるから，上式より $\mathrm{Re}\lambda > 0$ が従う．

(iii)⇒(iv)　明らかである．

(iv)⇒(ii)　$\alpha > \max_i a_{ii}$ をみたす正数 α をとると，$A \in \mathcal{Z}_n$ であるから

$$B = \alpha I - A \geq O$$

B の任意の固有値は $\alpha - \lambda(A)$ $\big(\lambda(A)$ は A の固有値$\big)$ と書けるから，特に

$$\rho(B) = \lambda_B = \alpha - \lambda_0 \ (\lambda_0 \text{は } A \text{ の実固有値})$$

と書ける．すると仮定によって $\lambda_0 > 0$ であるから

$$\rho(B) = \alpha - \lambda_0 < \alpha \quad \text{かつ} \quad A = \alpha I - B,\ B \geq O$$

となって (iv)⇒(ii) が示された．　証明終 ∎

定理 8.5　A が n 次 M 行列で $A \leq B$ (すなわち $B - A$ が非負行列) かつ $B \in \mathcal{Z}_n$ とすれば，次が成り立つ．

(i)　B も M 行列で $A^{-1} \geq B^{-1} \geq O$
(ii)　$\det B \geq \det A > O$
(iii)　$A^{-1}B \geq I$ かつ $BA^{-1} \geq I$
(iv)　$B^{-1}A$ および AB^{-1} は M 行列で $B^{-1}A \leq I$ かつ $AB^{-1} \leq I$
(v)　$\rho(I - B^{-1}A) = \rho(I - AB^{-1}) < 1$

【証明】　(i) $B \in \mathcal{Z}$ であるから ε を十分小さくとれば $U = I - \varepsilon B \geq O$．このとき $V = I - \varepsilon A \geq I - \varepsilon B = U \geq O$

V の Perron 根を λ_V とすれば

$$\det\big[(1 - \lambda_V)I - \varepsilon A\big] = \det(V - \lambda_V I) = 0$$

よって $\frac{1-\lambda_V}{\varepsilon}$ は A の実固有値であり，A は M 行列であるから，定理 8.4 (iv) によって

$$\frac{1 - \lambda_V}{\varepsilon} > 0 \quad \therefore\ \lambda_V < 1$$

したがって $I - V$ は正則で

$$(\varepsilon A)^{-1} = (I - V)^{-1} = I + V + V^2 + \cdots \geq O$$

次に $O \leq U \leq V$ より　$\rho(U) \leq \rho(V) = \lambda_V < 1$ (系 7.2.1)．よって $\varepsilon B = I - U$ も正則で

$$(\varepsilon B)^{-1} = (I - U)^{-1} = I + U + U^2 + \cdots \geq O$$

また $U^k \leq V^k$ $(k \geq 0)$ より

$$(I-U)^{-1}=\sum_{k=0}^{\infty}U^k\le\sum_{k=0}^{\infty}V^k=(I-V)^{-1}$$
$$\therefore\quad O\le(\varepsilon B)^{-1}\le(\varepsilon A)^{-1}$$
$$\therefore\quad O\le B^{-1}\le A^{-1}\tag{8.2}$$

(ii) n に関する帰納法による．$n=1$ のときは明らかであるから $n>1$ とし，$n-1$ のとき成り立つと仮定する．このとき A,B の $n-1$ 次首座行列 $A\begin{pmatrix}1&2&\cdots&n-1\\1&2&\cdots&n-1\end{pmatrix}$, $B\begin{pmatrix}1&2&\cdots&n-1\\1&2&\cdots&n-1\end{pmatrix}$ を A_1,B_1 とすれば $A_1,B_1\in\mathcal{Z}$ かつ $A_1\le B_1$ である．いま

$$\widetilde{A}=\begin{pmatrix}A_1&O\\O&a_{nn}\end{pmatrix}$$

とおけば $A\le\widetilde{A}$ かつ $\widetilde{A}\in\mathcal{Z}$ であるから (i) によって $\widetilde{A}$ も M 行列である．すなわち $\widetilde{A}$ は正則で

$$\widetilde{A}^{-1}=\begin{pmatrix}A_1^{-1}&\\&a_{nn}^{-1}\end{pmatrix}\ge O$$

したがって A_1 は $n-1$ 次 M 行列である．また $B_1\in\mathcal{Z}_{n-1}$ であるから $A_1\le B_1$ に (i) を適用して B_1 も M 行列である．よって帰納法の仮定によって

$$\det B_1\ge\det A_1\tag{8.3}$$

一方，定理 8.4 により A の固有値 $\lambda_j(A)=\alpha_j\pm\sqrt{-1}\beta_j$ の実部はすべて正であるから

$$\det A=\prod_j\lambda_j(A)=\Big(\prod_{\beta_j=0}\alpha_j\Big)\Big\{\prod_{\beta_j\neq0}(\alpha_j^2+\beta_j^2)\Big\}>0$$

同様に $\det B>0$ である．

最後に (8.2) の (n,n) 要素を比較すれば定理 1.3 (ii) によって

$$\frac{\det A_1}{\det A}\ge\frac{\det B_1}{\det B}\tag{8.4}$$

(8.3) と (8.4) より

$$\det B\ge\Big(\frac{\det B_1}{\det A_1}\Big)\det A\ge\det A>0$$

を得て帰納法による証明が完了する．

(iii) $A^{-1}\ge O$ かつ $B-A\ge O$ より $A^{-1}(B-A)\ge O\quad\therefore\ A^{-1}B\ge I$

同様に $(B-A)A^{-1}\ge O$ より $BA^{-1}\ge I$

(iv) $B^{-1}\ge O$ かつ $B-A\ge O$ より $B^{-1}(B-A)\ge O\quad\therefore\ B^{-1}A\le I$

よって $B^{-1}A\in\mathcal{Z}$ かつ $(B^{-1}A)^{-1}=A^{-1}B\ge O$ ((iii) による)

故に $B^{-1}A$ は M 行列である．同様に $(B-A)B^{-1}\ge O$ より $AB^{-1}\le I$

したがって $AB^{-1}\in\mathcal{Z}$ かつ $(AB^{-1})^{-1}=BA^{-1}\ge I\ge O$ よって AB^{-1} も M 行列である．

(v) 定理 5.1 によって $B^{-1}A$ と AB^{-1} の固有値は等しく

$$\rho(I-AB^{-1})=\rho(I-B^{-1}A)$$

したがって $\rho(I-B^{-1}A)<1$ を示せばよい．(iv) によって $C=I-B^{-1}A\ge O$ であるから C の Perron 根を λ_C であらわせば

$$\rho(C) = \lambda_C = 1 - \lambda(B^{-1}A)\ (= 1 - \mu \text{ とおく})$$

ただし $\mu = \lambda(B^{-1}A)$ は $B^{-1}A$ の適当な実固有値である．(iv) によって $B^{-1}A$ は M 行列であるから $\mu > 0$ (定理 8.4 (iv)) $\therefore\ \rho(C) < 1$ 証明終 ∎

定理 8.6 A が既約 M 行列ならば A^{-1} は正行列である．

【証明】 A の次数を n とする．仮に A^{-1} の第 i 列 $\boldsymbol{u}$ の要素にゼロがあるとすれば適当な置換行列 P を定めて

$$P\boldsymbol{u} = \begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{0} \end{pmatrix} \quad (\boldsymbol{v} \text{ は } m(<n) \text{ 次元正ベクトル})$$

とできる．このとき

$$A\boldsymbol{u} = \boldsymbol{e}_i \quad (n \text{ 次単位行列の第 } i \text{ 列}),\ P^{\mathrm{t}}P = I_n$$

より

$$(PAP^{\mathrm{t}})P\boldsymbol{u} = P\boldsymbol{e}_i \geq \boldsymbol{0}$$

ここで

$$PAP^{\mathrm{t}} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad (A_{11} : m \text{ 次行列})$$

と書けば

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{0} \end{pmatrix} \geq \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix} \quad \therefore\ A_{21}\boldsymbol{v} \geq \boldsymbol{0}$$

しかし $A \in \mathcal{Z}$ であるから $A_{21} \leq O$．したがって $\boldsymbol{v} > \boldsymbol{0}$ と併せて $A_{21}\boldsymbol{v} \leq \boldsymbol{0}$

$$\therefore\ \ A_{21}\boldsymbol{v} = \boldsymbol{0} \quad \therefore\ \ A_{21} = O$$

故に A は可約となって仮定に反する．よって A^{-1} の各列はゼロを含まず $A^{-1} > O$ である．証明終 ∎

定理 8.7 A を M 行列とするとき，次が成り立つ．

(i) A の正固有値 $\lambda_0 = \lambda_0(A)$ で A の任意の固有値 $\lambda(A)$ に対し $\mathrm{Re}\lambda(A) \geq \lambda_0$ となるものがある．

(ii) $B \in \mathcal{Z},\ B \geq A \Rightarrow\ \lambda_0(B) \geq \lambda_0(A)$

(iii) $A = (a_{ij})$ とするとき $\lambda_0(A) \leq \min_i a_{ii}$

【証明】 (i) 定理 8.4 によって $A = \alpha I - B,\ \alpha > 0,\ B \geq O,\ \rho(B) < \alpha$ をみたす $\alpha > 0$ と非負行列 B がある．このとき $\lambda_0 = \alpha - \rho(B)$ が求めるものである．実際 B の Perron 根 $\lambda_B = \rho(B)$ は B の実固有値で $\lambda_0 = \alpha - \lambda_B > 0$．よって λ_0 は A の正固有値である．また A の任意の固有値を $\lambda = \lambda(A)$ とすれば $\alpha - \lambda$ は $B = \alpha I - A$ の固有値である．

$$\therefore\ \ |\alpha - \lambda| \leq \lambda_B = \alpha - \lambda_0$$

λ は複素平面上の点 $(\alpha, 0)$ を中心とし半径 $\alpha - \lambda_0$ の閉円板内にあるから上式より $\mathrm{Re}\lambda \geq \lambda_0$

(ii) $B \in \mathcal{Z},\ B \geq A$ とすれば定理 8.5 によって B は M 行列である．さて $\lambda_0 = \lambda_0(A) > c > 0$ な

る c をとれば $A-cI \in \mathcal{Z}$ かつ (i) によって A の任意の実固有値 $\lambda(A)$ に対し $\lambda(A) \geq \lambda_0$ であるから $\lambda(A)-c \geq \lambda_0 - c > 0$. 故に $A-cI$ の実固有値はすべて正である.

故に定理 8.4 により $A-cI$ は M 行列である. また $A-cI \leq B-cI \in \mathcal{Z}$ であるから, 定理 8.5 によって $B-cI$ も M 行列である. よって B の任意の固有値 $\lambda(B)$ に対し

$$\therefore \quad \mathrm{Re}\lambda(B-cI) = \mathrm{Re}\lambda(B) - c > 0$$
$$\therefore \quad \mathrm{Re}\lambda(B) > c \tag{8.5}$$

$\lambda(B)$ は M 行列 B の任意の固有値であるから, $\lambda(B)$ として $\lambda_0(B)$ をとれば, (8.5) より

$$\lambda_0(B) > c \tag{8.6}$$

c は $\lambda_0(A) > c > 0$ をみたす任意の正数であったから

$$\lambda_0(B) \geq \lambda_0(A)$$

(仮に $\lambda_0(B) < \lambda_0(A)$ ならば c として $\lambda_0(B)$ をとることができて (8.6) と矛盾する.)

(iii) n 次 M 行列 $A=(a_{ij})$ に対し $B=\mathrm{diag}(a_{11},\ldots,a_{nn})$ とすれば $B \in \mathcal{Z}$ かつ $B \geq A$ であるから (ii) によって $\lambda_0(B) \geq \lambda_0(A)$.

B の固有値は a_{ii} $(1 \leq i \leq n)$ であるから $\lambda_0(B) = \min_i a_{ii}$

故に $\quad \lambda_0(A) \leq \min_i a_{ii}$ 証明終 ∎

【付記】単調行列は 1952 年 L. Collatz により導入された概念である. 一方, M 行列は 1937 年 A. M. Ostrowski により導入された概念であって, "M" はこのタイプの行列を 19 世紀末に他に先駆けて研究した Minkowski および 20 世紀はじめに既約性を加味して研究した Markoff の頭文字を意味するものとされる. この行列は一見, いかにも数学者が空想でつくりあげた概念のように思えるが, 1997 年九州大学 (現早稲田大学) 西哲生教授から, M 行列は後述する P 行列や P_0 行列と共に電気回路理論における重要な概念であると聞き一驚した想い出がある.

なお, Ostrowski は数学者 Hadamard (アダマール) の名に因んだ H 行列の概念も導入している:

行列 $A=(a_{ij})$ が H 行列であるとは

$$\alpha_{ij} = \begin{cases} |a_{ii}| & (i=j) \\ -|a_{ij}| & (i \neq j) \end{cases}$$

により定義される A の比較行列 $\langle A \rangle = (\alpha_{ij})$ が M 行列のときをいう.

この概念は数値解析において有用であるが本書では立ち入らない.

なお, 既約強優対角でかつすべての対角要素が正の実対称行列を ***S* 行列**といい, 実対称正定値な Z 行列を **Stieltjes (スチルチェス) 行列**ということも付記しておく.

8.3 M 行列であるための条件

前節において M 行列に関する基本的な結果を述べた. この節では Fiedler-Pták [23] にしたがって与えられた $\mathcal{Z}$ 行列が M 行列であるための 12 個の必要十分条件を定理として掲げておく. さらに詳細な結果は原論文を参照されたい.

以下において対角要素がすべて正の対角行列 D を**正対角行列**といい $D > O$ とかく. また対角要素がすべて非負の対角行列 D を**非負対角行列**といい $D \geq O$ とかく.

定理 8.8 (Fiedler-Pták) $A=(a_{ij})\in\mathcal{Z}_n$ のとき次の条件は互いに同値である.

(i) A が正則で $A^{-1}\geq O$ (すなわち A は単調行列)

(ii) 適当な正数 α と非負行列 B をえらんで $A=\alpha I-B$, $\rho(B)<\alpha$ と書ける.

(iii) $A\boldsymbol{x}>\boldsymbol{0}$ なる非負ベクトル $\boldsymbol{x}$ が存在する.

(iv) $A\boldsymbol{x}>\boldsymbol{0}$ なる正ベクトル $\boldsymbol{x}$ が存在する.

(v) $\boldsymbol{e}=(1,\ldots,1)^{\mathrm{t}}\in\mathbb{R}^n$ とするとき, $AD\boldsymbol{e}>\boldsymbol{0}$ となる正対角行列 D が存在する.

(vi) 正対角行列 D を適当にえらんで $W=AD$ が狭義優対角かつ W の対角要素がすべて正とできる.

(vii) R が $R\geq A$ なる対角行列ならば R は正則で $Q=\mathrm{diag}(a_{11},\ldots,a_{nn})$ とおくとき $\rho(R^{-1}(Q-A))<1$

(viii) $B\in Z_n$ かつ $B\geq A$ ならば B は正則である.

(ix) A の実固有値は正である.

(x) A のすべての主小行列式が正である. すなわち

$$\det A\begin{pmatrix} i_1 & \cdots & i_p \\ i_1 & \cdots & i_p \end{pmatrix}>0 \quad (1\leq i_1<\cdots<i_p\leq n,\ 1\leq p\leq n)$$

(xi) 適当な置換行列 P をとり

$$PAP^{\mathrm{t}}=LU,\quad L,U\in\mathcal{Z}$$

L は対角要素が正の下3角, U は対角要素が正の上3角

とできる.

(xii) A の固有値の実部が正である.

【証明】

(i)⇔(ii)⇔(ix)⇔(xii) は定理 8.4 で済んでいる.

(ii)⇒(iii) $\boldsymbol{e}=(1,\ldots,1)^{\mathrm{t}}$, $\boldsymbol{x}=A^{-1}\boldsymbol{e}$ とおけば $A^{-1}\geq O$ であるから

$$\boldsymbol{x}\geq\boldsymbol{0}\quad\text{かつ}\quad A\boldsymbol{x}=\boldsymbol{e}>\boldsymbol{0}$$

(iii)⇒(iv) $A\boldsymbol{x}>\boldsymbol{0}$ なる $\boldsymbol{x}\geq\boldsymbol{0}$ をとる. 正数 ε を十分小さくとれば $\boldsymbol{e}=(1,\ldots,1)^{\mathrm{t}}$ に対し $\boldsymbol{x}+\varepsilon\boldsymbol{e}>\boldsymbol{0}$ で

$$A(\boldsymbol{x}+\varepsilon\boldsymbol{e})=A\boldsymbol{x}+\varepsilon A\boldsymbol{e}>\boldsymbol{0}$$

(iv)⇒(v) $A\boldsymbol{x}>\boldsymbol{0}$ なる $\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}}>\boldsymbol{0}$ に対し $D=\mathrm{diag}(x_1,\ldots,x_n)$ とおけば D は正対角行列で $AD\boldsymbol{e}=A\boldsymbol{x}>\boldsymbol{0}$

(v)⇒(vi) $D=\mathrm{diag}(d_1,\ldots,d_n)$, $d_i>0\ \forall i$, $AD\boldsymbol{e}>\boldsymbol{0}$ なる正対角行列 D に対し $W=AD=(w_{ij})$ とおけば

$$(AD\boldsymbol{e})_i=\sum_{j=1}^{n}w_{ij}>0\quad\therefore\ w_{ii}>-\sum_{j\neq i}w_{ij}\ \forall i$$

$A\in\mathcal{Z}$ かつ $d_i>0\quad\therefore\ w_{ij}=a_{ij}d_j\leq 0\ (i\neq j)$. よって上式は

$$w_{ii}>\sum_{j\neq i}|w_{ij}|\geq 0\quad\forall i$$

を意味し，W は対角要素がすべて正の狭義優対角行列である．

(vi)⇒(vii) $W = AD = (w_{ij})$ (D は正対角) が狭義優対角で $w_{ii} > 0\ \forall i$ ならば $\boldsymbol{e} = (1, \ldots, 1)^{\mathrm{t}}$ に対し

$$(W\boldsymbol{e})_i = \sum_{j=1}^{n} w_{ij} = w_{ii} + \sum_{j \neq i} w_{ij} = w_{ii} - \sum_{j \neq i} |w_{ij}| > 0$$

$D = \mathrm{diag}(d_1, \ldots, d_n)$ とおけば，$w_{ii} = a_{ii} d_i > 0$ である．故に $d_i > 0$ と併せて $a_{ii} > 0\ \forall i$

このとき $R = \mathrm{diag}(r_1, \ldots, r_n)$，$R \geq A$ ならば $r_i \geq a_{ii} > 0$ より R は正則である．

さらに $H = \mathrm{diag}(w_{11}, \ldots, w_{nn})$，$Q = \mathrm{diag}(a_{11}, \ldots, a_{nn})$ とおけば $H = QD$ ($\because\ W = AD$)

また $r_i \geq a_{ii} > 0\ \forall i$ より $R^{-1} \leq Q^{-1}$．よって $Q - A \geq O$ に注意して

$$\begin{aligned} \rho\big(R^{-1}(Q-A)\big) &\leq \rho\big(Q^{-1}(Q-A)\big) = \rho(I - Q^{-1}A) \\ &= \rho\big(D^{-1}(I - Q^{-1}A)D\big) = \rho\big(I - (QD)^{-1}AD\big) \\ &= \rho(I - H^{-1}W) < 1 \quad (\text{系 } 6.1.2 \text{ を狭義優対角 } W \text{ に適用する．}) \end{aligned}$$

(vii)⇒(viii) $B = (b_{ij}) \in \mathcal{Z}_n$，$B \geq A$ のとき $b_{ii} \geq a_{ii} > 0\ \forall i$

$$R = \mathrm{diag}(b_{11}, \ldots, b_{nn}), \quad Q = \mathrm{diag}(a_{11}, \ldots, a_{nn})$$

とおけば

$$O \leq R - B \leq Q - A \quad \therefore\ O \leq R^{-1}(R-B) \leq R^{-1}(Q-A)$$

$$\therefore\ \ \rho(I - R^{-1}B) = \rho\big(R^{-1}(R-B)\big) \leq \rho\big(R^{-1}(Q-A)\big) < 1$$

よって $C = I - R^{-1}B$ とおくと $R^{-1}B = I - C$ かつ $\rho(C) < 1$．故に $B = R(I-C)$ は正則で $B^{-1} = (I-C)^{-1}R^{-1}$

(viii)⇒(ix) 実数 $\mu \leq 0$ を任意にとれば $B = A - \mu I \in \mathcal{Z}$ かつ $B \geq A$．よって B は正則である．故に A の実固有値は (存在するならば) 正である．

(ix)⇒(x) $M_p = \{i_1, \ldots, i_p\} \subset \{1, 2, \ldots, n\}$，$i_1 < \cdots < i_p$ に対し

$$b_{ij} = \begin{cases} a_{ij} & (i = j \text{ または } i, j \in M_p) \\ 0 & (i \neq j \text{ かつ } i \text{ または } j \notin M_p) \end{cases}$$

とおくと $B = (b_{ij}) \geq A$ かつ $B \in \mathcal{Z}$

定理 8.4 によって A は M 行列であるから B も M 行列 (定理 8.5) であり，再び定理 8.4 により B の実固有値は正である．

さて

$$A(M_p) = A\begin{pmatrix} i_1 & \cdots & i_p \\ i_1 & \cdots & i_p \end{pmatrix},\ \{i_1, \ldots, i_p\} \cup \{j_1, \ldots, j_q\} = \{1, 2, \ldots, n\}$$

とすれば B の定義によって適当な置換行列 P により

$$PBP^{\mathrm{t}} = \begin{pmatrix} A(M_p) & & & \\ & a_{j_1 j_1} & & \\ & & \ddots & \\ & & & a_{j_q j_q} \end{pmatrix} \tag{8.7}$$

となる．B は M 行列であるから定理 8.5 によって $\det B > 0$.

一方 (8.7) より

$$\det B = \det(PBP^{\mathrm{t}}) = \big(\det A(M_p)\big) \prod_{k=1}^{q} a_{j_k j_k}, \quad a_{j_k j_k} > 0$$

$$\therefore \quad \det A(M_p) = \frac{\det B}{\prod_{k=1}^{q} a_{j_k j_k}} > 0, \qquad p = 1, 2, \ldots, n$$

(x)⇒(xi) 再び $M_p = \{i_1, \ldots, i_p\}$, $i_1 < \cdots < i_p$, $p = 1, 2, \ldots, n$ とおく．

置換

$$\sigma = \begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

に対応する置換行列を P とするとき

$$(PAP^{\mathrm{t}}) \begin{pmatrix} 1 & 2 & \cdots & p \\ 1 & 2 & \cdots & p \end{pmatrix} = A(M_p), \quad \det A(M_p) > 0$$

よって定理 3.3 により

$$PAP^{\mathrm{t}} = LU \quad (L \text{ は対角要素が正の下 3 角行列，} U \text{ は対角要素が正の上 3 角行列})$$

とあらわされる．このとき $PAP^{\mathrm{t}} \in \mathcal{Z}$ であるから定理 8.1 により $L, U \in \mathcal{Z}$

(xi)⇒(i) $PAP^{\mathrm{t}} = LU (L = (l_{ij}) : \text{下 3 角}, \ U = (u_{ij}) : \text{上 3 角}, \ l_{ii} > 0, \ u_{ii} > 0 \ \forall i \text{ かつ } L, U \in \mathcal{Z})$ とすれば，L, U は正則で $A = P^{\mathrm{t}} LUP$ も正則である．このとき

$$A^{-1} = (P^{\mathrm{t}} LUP)^{-1} = P^{-1} U^{-1} L^{-1} P^{\mathrm{t}^{-1}} = P^{\mathrm{t}} U^{-1} L^{-1} P$$

$D = \mathrm{diag}(l_{11}, \ldots, l_{nn})$, $L_0 = D - L$ とおけば $L_0 \geq 0$ かつ

$$L = D - L_0 = D(I - D^{-1} L_0)$$

$D^{-1} L_0$ は対角要素がゼロの下 3 角非負行列であるから $(D^{-1} L_0)^k = 0 \ (k \geq n)$

故に

$$L^{-1} = (I - D^{-1} L_0)^{-1} D^{-1} = \{I + (D^{-1} L_0) + (D^{-1} L_0)^2 + \cdots + (D^{-1} L_0)^{n-1}\} D^{-1}$$
$$\geq O$$

同様に $U^{-1} \geq O$ も示すことができて $A^{-1} = P^{\mathrm{t}} U^{-1} L^{-1} P \geq O$ 証明終 ∎

■**注意**　稀ではあるが，定理 8.8 の同値条件 (ii) を緩めて

$$A = \alpha I - B, \ \alpha > 0, \ B \geq 0, \ \rho(B) \leq \alpha \tag{8.8}$$

とあらわされる行列 $A \in \mathcal{Z}$ を**特異 M 行列 (singular M-matrix)** ということがある．その場合には条件 (ii) をみたす行列 $A \in \mathcal{Z}$ を**正則 M 行列 (nonsingular M-matrix)** と呼んで (8.8) との違いを強調する．

8.4 L 行列

定義 8.2 (L 行列) 対角要素がすべて正の Z 行列を **L 行列**という.

定理 8.9 狭義優対角な L 行列は M 行列である.

【証明】 $A=(a_{ij})$ を n 次の狭義優対角 L 行列として

$$A=D+L+U$$

$$D=\mathrm{diag}(a_{11},\ldots,a_{nn}),\ L=\begin{pmatrix}0 & & 0\\ & \ddots & \\ a_{ij} & & 0\end{pmatrix},\ U=\begin{pmatrix}0 & & a_{ij}\\ & \ddots & \\ 0 & & 0\end{pmatrix}$$

と分解する. このとき $B=I-D^{-1}A=(b_{ij})$ は非負行列で

$$\sum_{j=1}^{n} b_{ij}=\frac{1}{a_{ii}}\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|<1 \quad (a_{ii}>0\ \forall i \text{ に注意})$$

$$\therefore\quad \rho(B)\leq ||B||_{\infty}=\max_{i}\sum_{j\neq i} b_{ij}<1$$

したがって $A^{-1}=(I-B)^{-1}D^{-1}\geq 0$ となり $A\in\mathcal{Z}$ は M 行列である. 証明終 ∎

定理 8.10 既約強優対角 L 行列 A は M 行列で A^{-1} は正行列である.

【証明】 前定理の証明に用いた行列 B を考えると $b_{ii}=0\ \forall i$ かつ

$$\rho(B)\leq ||B||_{\infty}\leq 1$$

A は既約であるから B も既約である. このとき $\rho(B)=1$ は起こりえない. なぜならば, 仮に $\rho(B)=1$ とすれば既約非負行列 B は Perron 根 $\lambda_B=1$ と対応する正固有ベクトル $\boldsymbol{x}$ をもつが, このとき

$$B\boldsymbol{x}=\boldsymbol{x}\Leftrightarrow D^{-1}A\boldsymbol{x}=\boldsymbol{0}\Leftrightarrow A\boldsymbol{x}=\boldsymbol{0}$$

となって A は正則でない. A は既約強優対角であるからこれは定理 1.21 に矛盾する.

故に $\rho(B)<1$ で定理 7.1 (iii) によって $(D^{-1}A)^{-1}=(I-B)^{-1}\geq O$

$$\therefore\quad A^{-1}=(I-B)^{-1}D^{-1}\geq O$$

よって A は M 行列であるが, A は既約であるから定理 8.6 によって $A^{-1}>O$ を得る.

証明終 ∎

8.5 非負行列に対する分離定理

M 行列の応用として, 一般な非負行列に対する次の分離定理が知られている.

定理 8.11 (Hall-Porsching) A を n 次非負行列，その固有値 $\{\lambda_i(A)\}$ を

$$\lambda_1(A) = \rho(A) \geq |\lambda_2(A)| \geq \cdots \geq |\lambda_n(A)|$$

と並べる．行列 A の第 i 行，第 i 列を取り除いて得られる $n-1$ 次非負行列を A_i とし，その固有値 $\{\lambda_k(A_i)\}$, $k = 1, 2, \ldots, n-1$ を

$$\lambda_1(A_i) = \rho(A_i) \geq |\lambda_2(A_i)| \geq \cdots \geq |\lambda_{n-1}(A_i)|$$

と並べる．$\lambda_p(A)$ を A の実固有値で $\lambda_p(A) \neq \lambda_1(A)$ とすれば

$$\lambda_1(A) \geq \lambda_1(A_i) \geq \lambda_p(A), \ i = 1, 2, \ldots, p \tag{8.9}$$

【証明】 必要ならば行と列の同時入れかえを行えばよいから A_1 $(i = 1)$ に対して定理を証明すればよい．非負行列の理論によって $\lambda_1(A) \geq \lambda_1(A_1)$ であり，A が既約ならば $\lambda_1(A) > \lambda_1(A_1)$ であるから，(8.9) の右側の不等式を示せばよい．

仮に $\lambda_1(A_1) < \lambda_p(A)$ として矛盾を導く．この仮定の下で

$$B_j = \lambda_j(A) I_{n-1} - A_1$$

は $j = 1$ と $j = p$ のとき M 行列である．(実際 $j = 1$ のとき $\alpha = \lambda_1(A)$ とおいて $B_1 = \alpha I - A_1$, $A_1 \geq O$ かつ $\rho(A_1) = \lambda_1(A_1) < \lambda_p(A) \leq \lambda_1(A) = \alpha$．よって定理 8.4 (ii) より B_1 は M 行列である．また $j = p$ のとき $\tilde{\alpha} = \lambda_p(A)$ とおけば $B_p = \tilde{\alpha} I_{n-1} - A_1$, $A_1 \geq O$ で $\rho(A_1) = \lambda_1(A_1) < \lambda_p(A) = \tilde{\alpha}$ より B_p は M 行列である (再び定理 8.4 による)．)

よって

$$B_j^{-1} \geq O, \ j = 1, p$$

いま $A = (a_{ij})$ とし $\boldsymbol{x}^{\mathrm{t}} = -(a_{12}, \ldots, a_{1n})$, $y = -(a_{21}, \ldots, a_{n1})^{\mathrm{t}}$ とおけば

$$\lambda_j(A) I_n - A = \left(\begin{array}{c|c} \lambda_j(A) - a_{11} & \boldsymbol{x}^{\mathrm{t}} \\ \hline \boldsymbol{y} & B_j \end{array} \right)$$

ここで

$$C_j = \begin{pmatrix} 1 & 0 \\ 0 & B_j^{-1} \end{pmatrix}$$

とおけば

$$T_j \equiv (\lambda_j(A) I_n - A) C_j = \left(\begin{array}{c|c} \lambda_j(A) - a_{11} & \boldsymbol{x}^{\mathrm{t}} B_j^{-1} \\ \hline \boldsymbol{y} & I_{n-1} \end{array} \right)$$

さらに

$$E = \left(\begin{array}{c|c} 1 & -\boldsymbol{x}^{\mathrm{t}} B_j^{-1} \\ \hline 0 & \\ \vdots & I_{n-1} \\ 0 & \end{array} \right)$$

とおけば

$$ET_j = \left(\begin{array}{c|ccc} \lambda_j(A) - a_{11} - \boldsymbol{x}^{\mathrm{t}} B_j^{-1} \boldsymbol{y} & 0 & \cdots & 0 \\ \hline \boldsymbol{y} & & I_{n-1} & \end{array}\right)$$

$\lambda_j(A)$ は A の固有値であるから T_j は正則でなく，したがって ET_j も正則でない．よって $\lambda_j(A) - a_{11} - \boldsymbol{x}^{\mathrm{t}} B_j^{-1} \boldsymbol{y} = 0$ でなければならない．

この式において $j = 1, p$ とおけば

$$a_{11} = \lambda_1(A) - \boldsymbol{x}^{\mathrm{t}} B_1^{-1} \boldsymbol{y} = \lambda_p(A) - \boldsymbol{x}^{\mathrm{t}} B_p^{-1} \boldsymbol{y} \tag{8.10}$$

すでにみたように B_1 と B_p は M 行列で $B_1 \geq B_p$ であるから $B_p^{-1} \geq B_1^{-1} \geq O$

$$\therefore \quad \boldsymbol{x}^{\mathrm{t}} B_p^{-1} \boldsymbol{y} \geq \boldsymbol{x}^{\mathrm{t}} B_1^{-1} \boldsymbol{y} \quad (\because\ \boldsymbol{x} \leq \boldsymbol{0},\ \boldsymbol{y} \leq \boldsymbol{0})$$

よって (8.10) と仮定 $\lambda_p(A) \neq \lambda_1(A)$ (したがって $\lambda_1(A) > \lambda_p(A)$) より

$$a_{11} = \lambda_1(A) - \boldsymbol{x}^{\mathrm{t}} B_1^{-1} \boldsymbol{y} > \lambda_p(A) - \boldsymbol{x}^{\mathrm{t}} B_p^{-1} \boldsymbol{y} = a_{11}$$

これは矛盾であるから $\lambda_1(A_1) < \lambda_p(A)$ は起こりえない．$\therefore\ \lambda_1(A_1) \geq \lambda_p(A)$ 証明終 ∎

第9章　P行列とP_0行列

正定値実対称行列と半正定値実対称行列の非対称実行列への拡張としてP行列とP_0行列が知られている．本章では，前章に続いて行列とベクトルは実行列，実ベクトルに限定し，Fiedler-Pták (フィードラー・プタック) と Sandberg-Willson, Jr. (サンドベルグ・ウィルソン) により得られているP行列とP_0行列の性質を中心に述べる．

9.1　P行列

定義 9.1 (P 行列)　n次行列A自身を含めてAのすべての主小行列の行列式が正，すなわち

$$\det A\begin{pmatrix} i_1 & \cdots & i_r \\ i_1 & \cdots & i_r \end{pmatrix} > 0, \quad 1 \le i_1 < \cdots < i_r \le n, \quad 1 \le r \le n$$

$$\{i_1, i_2, \ldots, i_n\} = \{1, 2, \ldots, n\}$$

のとき，Aを**P行列 (P-matrix)**という．n次P行列の全体を$\mathcal{P}_n$または$\mathcal{P}$であらわす．

したがって$A \in \mathcal{P}$ならば$A^{\mathrm{t}} \in \mathcal{P}$である．明らかに$P$行列は正定値実対称行列の非対称実行列への自然な一般化であり，その基本的な性質はZ行列，M行列の性質とともにFiedler-Pták [23]により詳しく論じられている．次の定理はその一部である．

定理 9.1　n次行列Aに対し次の性質は互いに同値である．

(i)　$A \in \mathcal{P}$

(ii)　任意のベクトル$\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} \neq \boldsymbol{0}$に対し，$\boldsymbol{y} = A\boldsymbol{x} = (y_1, \ldots, y_n)^{\mathrm{t}}$とおけば$x_k y_k > 0$となる$k$がある．

(iii)　任意のn次元ベクトル$\boldsymbol{x} \neq \boldsymbol{0}$に対し適当な正対角行列$D_{\boldsymbol{x}}$をえらんで

$$(A\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x}) > 0$$

(iv)　任意のn次元ベクトル$\boldsymbol{x} \neq \boldsymbol{0}$に対し適当な非負対角行列$H_{\boldsymbol{x}}$をえらんで

$$(A\boldsymbol{x},\ H_{\boldsymbol{x}}\boldsymbol{x}) > 0$$

(v)　Aを含むAのすべての主小行列の実固有値は正である．

【証明】　(i)⇒(ii) 仮に$x_i y_i \le 0\ \forall i$とし

$$M = \{i \mid x_i \neq 0\} = \{i_1, \ldots, i_m\}$$

とする．$\boldsymbol{x} \neq \boldsymbol{0}$であるから$M \neq \emptyset$である．いま

$$A(M) = A\begin{pmatrix} i_1 & \cdots & i_m \\ i_1 & \cdots & i_m \end{pmatrix}$$

$$\boldsymbol{x}(M) = (x_{i_1}, \ldots, x_{i_m})^{\mathrm{t}}, \qquad \boldsymbol{z} = \boldsymbol{y}(M) = (y_{i_1}, \ldots, y_{i_m})^{\mathrm{t}}$$

とすれば，$\boldsymbol{z} = A(M)\boldsymbol{x}(M)$

ここで

$$u_k = -\frac{y_{i_k}}{x_{i_k}} \quad (1 \le k \le m), \quad U = \mathrm{diag}(u_1, \ldots, u_m)$$

とおけば $x_{i_k} y_{i_k} \le 0$ より $u_1, \ldots, u_m \ge 0$ かつ

$$\big(A(M) + U\big)\boldsymbol{x}(M) = \boldsymbol{0}, \quad \boldsymbol{x}(M) \ne \boldsymbol{0} \tag{9.1}$$

よって m 次行列 $A(M)+U$ は正則でない．しかし仮定によって $A(M)$ の主小行列式はすべて正で $U \ge O$ であるから

$$\det\big(A(M) + U\big) > 0 \tag{9.2}$$

が示される．実際

$$U_1 = \mathrm{diag}(u_1, 0, \ldots, 0), \ldots, U_m = \mathrm{diag}(0, \ldots, 0, u_m)$$

とおけば $U = U_1 + \cdots + U_m$ で

$$\begin{aligned}
\det\big(A(M) + U\big) &= \det\big(A(M) + U_1 + \cdots + U_m\big) \\
&= \det\big(A(M) + U_2 + \cdots + U_m\big) \\
&\qquad + u_1 \det\left(\big(A(M) + U_2 + \cdots + U_m\big)\begin{pmatrix} i_2 & \cdots & i_m \\ i_2 & \cdots & i_m \end{pmatrix}\right) \\
&\ge \det\big(A(M) + U_2 + \cdots + U_m\big) \\
&\left(\because\ u_1 \ge 0 \text{ かつ } \det\left(\big(A(M) + U_2 + \cdots + U_m\big)\begin{pmatrix} i_2 & \cdots & i_m \\ i_2 & \cdots & i_m \end{pmatrix}\right) \ge 0\right) \\
&\ge \quad \cdots \\
&\ge \det A(M) > 0 \quad \text{(厳密には } m \text{ に関する帰納法による.)}
\end{aligned}$$

よって (9.2) が成り立ち $A(M) + U$ は正則であるがこれは (9.1) と矛盾する．よってある k につき $x_k y_k > 0$.

(ii)⇒(iii) ある k につき $x_k y_k > 0$ とすれば $\varepsilon > 0$ を十分小さくえらんで

$$x_k y_k + \varepsilon \sum_{j \ne k} x_j y_j > 0$$

このとき $D_{\boldsymbol{x}} = \mathrm{diag}(\varepsilon, \varepsilon, \ldots, \varepsilon, \overset{k}{\overset{\vee}{1}}, \varepsilon, \ldots, \varepsilon)$ とおけば，$D_{\boldsymbol{x}}$ は正対角行列で，$(A\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x}) > 0$

(iii)⇒(iv) 明らか．(非負対角行列 $H_{\boldsymbol{x}}$ として正対角行列 $D_{\boldsymbol{x}}$ をとればよい．)

(iv)⇒(v) $A(M)$ を A の任意の m 次主小行列 (A 自身も含む) とし，λ と $\boldsymbol{v}$ をその実固有値と固有ベクトルとする：

$$A(M)\boldsymbol{v} = \lambda\boldsymbol{v}, \quad \boldsymbol{v} = (v_1, \ldots, v_m)^{\mathrm{t}} \ne \boldsymbol{0}$$

このとき

$$\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}}, \quad \text{ただし } x_i = \begin{cases} v_i & (i \in M) \\ 0 & (i \notin M) \end{cases}$$

とおくと $\boldsymbol{x} \neq \boldsymbol{0}$ であるから，仮定により非負対角行列

$$H_{\boldsymbol{x}} = \operatorname{diag}(h_1, \ldots, h_n), \quad h_i \geq 0 \ \forall i$$

が存在して

$$(A\boldsymbol{x},\ H_{\boldsymbol{x}}\boldsymbol{x}) > 0$$

このとき

$$\begin{aligned}(A\boldsymbol{x},\ H_{\boldsymbol{x}}\boldsymbol{x}) &= \big(A(M)\boldsymbol{x}(M),\ H_{\boldsymbol{x}}(M)\boldsymbol{x}(M)\big) \\ &= \big(\lambda\boldsymbol{x}(M),\ H_{\boldsymbol{x}}(M)\boldsymbol{x}(M)\big) = \lambda(\boldsymbol{x},\ H_{\boldsymbol{x}}\boldsymbol{x}) \\ &= \lambda\sum_{i=1}^{n} h_i x_i^2 \quad \text{かつ} \quad \sum_{i=1}^{n} h_i x_i^2 \geq 0\end{aligned}$$

であるから，$\sum_{i=1}^{n} h_i x_i^2 > 0$ かつ $\lambda > 0$ である．
(v)⇒(i) $M = \{i_1, \ldots, i_r\}$ とすれば

$$\det A(M) = \prod_{j=1}^{r} \lambda_j(A(M))$$

$\lambda_j = \lambda_j(A(M)) = \alpha_j + \sqrt{-1}\beta_j$ が複素根で $\beta_j \neq 0$ ならば $\overline{\lambda}_j = \alpha_j - \sqrt{-1}\beta_j$ も固有値で $\lambda_j\overline{\lambda}_j = \alpha_j^2 + \beta_j^2 > 0$
また $A(M)$ の実固有値は正であるから $\beta_j = 0$ のとき $\alpha_j > 0$．よって

$$\prod_{j=1}^{r} \lambda_j(A(M)) = \Big(\prod_{\beta_j=0} \alpha_j\Big)\Big(\prod_{\beta_j\neq 0} (\alpha_j^2 + \beta_j^2)\Big) > 0, \quad 1 \leq r \leq n$$

$$\therefore\ A \in \mathcal{P}$$

証明終 ∎

系 9.1.1 (Sandberg-Willson, Jr.) $A \in \mathcal{P}$ ならば $A^{-1} \in \mathcal{P}$

【証明】 $A \in \mathcal{P}$ ならば A は正則である．$\boldsymbol{y} = A^{-1}\boldsymbol{x}$ とおけば $\boldsymbol{x} = A\boldsymbol{y}$ で $\boldsymbol{x} \neq \boldsymbol{0} \Leftrightarrow \boldsymbol{y} \neq \boldsymbol{0}$.
定理 9.1 (ii) によって任意の $\boldsymbol{y} = (y_1, \ldots, y_n)^{\mathrm{t}} \neq \boldsymbol{0}$ に対し $y_k x_k = y_k(A\boldsymbol{y})_k > 0$ となる k がある．これは任意の $\boldsymbol{x} \neq \boldsymbol{0}$ に対し $x_k(A^{-1}\boldsymbol{x})_k > 0$ となる k が存在することであるから，定理 9.1 の性質 (ii) が A^{-1} に対してみたされて $A^{-1} \in \mathcal{P}$ を意味する．
証明終 ∎

系 9.1.2 $A \in \mathcal{P},\ M = \{i_1, \ldots, i_m\} \subseteq \{1, 2, \ldots, n\}$ ならば $A(M) \in \mathcal{P}$

【証明】 $A(M)$ の主小行列は A の主小行列であるから明らかである．
証明終 ∎

系 9.1.3 (Fiedler-Pták) $A \in \mathcal{P}$ で D が正対角行列 (すなわち $D > O$) ならば $DA, AD \in \mathcal{P}$

【証明】 任意の $M = \{i_1, \ldots, i_m\}$ に対して

$$(DA)(M) = D(M)A(M)$$

であるから

$$\det(DA)(M) = \det D(M) \det A(M) > 0 \quad \therefore\ DA \in \mathcal{P}$$

AD についても同様である．
証明終 ∎

系 9.1.4 $A \in \mathcal{P}$ かつ D が非負対角行列 (すなわち $D \geq O$) ならば $A + D \in \mathcal{P}$. 特に

$$\det(A + D) > 0$$

【証明】 (9.2) の証明をくり返せばよいが次のようにしてもよい.

$$D = \mathrm{diag}(d_1, \ldots, d_n), \quad d_i \geq 0 \ \forall i$$

$$D_i = \mathrm{diag}(0, \ldots, 0, \overset{i}{\check{d}_i}, 0, \ldots, 0)$$

$$M = \{i_1, \ldots, i_m\} \subseteq \{1, 2, \ldots, n\}$$

とするとき

$$\det(A + D_i)(M) = \begin{cases} \det A(M) & (i \notin M) \\ \det A(M) + d_i \det A(M \setminus \{i\}) & (i \in M) \end{cases}$$

$$\geq \det A(M) > 0$$

$$\therefore \ A + D_i \in \mathcal{P}$$

よって $A + D_1 \in \mathcal{P}$, $\therefore\ (A + D_1) + D_2 \in \mathcal{P}$, $\therefore\ (A + D_1 + D_2) + D_3 \in \mathcal{P}, \ldots$
以下これを続けて

$$A + D = (A + D_1 + D_2 + \cdots + D_{n-1}) + D_n \in \mathcal{P}$$

証明終 ∎

系 9.1.5 $A + A^{\mathrm{t}}$ が正定値ならば $A \in \mathcal{P}$

【証明】 $((A + A^{\mathrm{t}})\boldsymbol{x},\ \boldsymbol{x}) = (A\boldsymbol{x},\ \boldsymbol{x}) + (A^{\mathrm{t}}\boldsymbol{x},\ \boldsymbol{x}) = 2(A\boldsymbol{x},\ \boldsymbol{x})$

$$\therefore\ (A\boldsymbol{x},\ \boldsymbol{x}) > 0 \quad \forall \boldsymbol{x} \neq \boldsymbol{0}$$

よって $D_{\boldsymbol{x}} = I$ は定理 9.1 (iii) をみたすから $A \in \mathcal{P}$ である.

証明終 ∎

9.2 P_0 行列

定義 9.2 (P_0 行列) n 次行列 A のすべての主小行列式の値が非負, すなわち

$$\det A \begin{pmatrix} i_1 & \cdots & i_m \\ i_1 & \cdots & i_m \end{pmatrix} \geq 0,\ 1 \leq i_1 < \cdots < i_m \leq n,\ 1 \leq m \leq n$$

$$\{i_1, i_2, \ldots, i_n\} = \{1, 2, \ldots, n\}$$

のとき, A を **P_0 行列 (P_0-matrix)** という. P_0 行列の全体を $\mathcal{P}_0$ であらわす.

明らかに $A \in \mathcal{P}_0$ ならば $A^{\mathrm{t}} \in \mathcal{P}_0$ である. P 行列と同様に P_0 行列は半正定値実対称行列の一般化である. その基本的な性質は Fiedler-Pták [24] に述べられているが, その後 Sandberg-Willson, Jr. [36], [37] による重要な貢献がある.

本節および次節においてこれらの結果の一部を解説する.

定理 9.2 (Sandberg-Willson, Jr.) $A \in \mathcal{P}_0$ とし，D を対角行列とすると次が成り立つ．

(i) $D \geq O$ ならば $A+D \in \mathcal{P}_0$

(ii) $D > O$ ならば $A+D \in \mathcal{P}$

【証明】 (i) 定理 9.1 の (i)⇒(ii) の証明 または系 9.1.4 の証明に用いた論法が使える．
(ii) も同様であるが，念のため証明を記す．A の次数 n に関する帰納法による．
$n=1$ のときは明らかであるから $n>1$ とし $m<n$ のとき成り立つとする．

$$M=\{i_1,\dots,i_m\} \subset N=\{1,2,\dots,n\}, \quad i_1<\cdots<i_m$$

$$D(M)=D_1+D_2, \quad D_1=\operatorname{diag}(d_{i_1},0,\dots,0), \quad D_2=\operatorname{diag}(0,d_{i_2},\dots,d_{i_m})$$

とすれば

$$\begin{aligned}\det\big((A+D)(M)\big) &= \det\big(A(M)+D(M)\big)=\det\big(A(M)+D_1+D_2\big)\\ &= \det\big(A(M)+D_2\big)+d_{i_1}\det\big((A(M)+D_2)(\widetilde{M})\big)\\ &\qquad\qquad (\text{ただし}\widetilde{M}=M\setminus\{i_1\})\\ &= \det\big(A(M)+D_2\big)+d_{i_1}\det\big(A(\widetilde{M})+D(\widetilde{M})\big) \qquad (9.3)\end{aligned}$$

$D_2 \geq O$ かつ $A(M) \in \mathcal{P}_0$ であるから (i) によって $A(M)+D_2 \in \mathcal{P}_0$

$$\therefore\ \det(A(M)+D_2) \geq 0$$

また $A(\widetilde{M}) \in \mathcal{P}_0$ かつ $D(\widetilde{M})>O$ でそれらの行列の次数は $m-1<n$ であるから帰納法の仮定によって

$$A(\widetilde{M})+D(\widetilde{M}) \in \mathcal{P} \quad \therefore\ \det\big(A(\widetilde{M})+D(\widetilde{M})\big)>0 \quad \therefore\ d_{i_1}\det\big(A(\widetilde{M})+D(\widetilde{M})\big)>0$$

よって (9.3) より $\det\big((A+D)(M)\big)>0 \quad \forall M=\{i_1,\dots,i_m\},\ 1\leq i_1<\cdots<i_m\leq n$

$$\therefore\quad A+D \in \mathcal{P}$$

証明終 ∎

定理 9.3 n 次行列 A に対し次の性質は互いに同値である．

(i) $A \in \mathcal{P}_0$

(ii) 任意の $\boldsymbol{x}=(x_1,\dots,x_n)^{\mathrm{t}} \neq \boldsymbol{0}$ に対し $\boldsymbol{y}=A\boldsymbol{x}=(y_1,\dots,y_n)^{\mathrm{t}}$ とおけば $x_k \neq 0$ かつ $x_k y_k \geq 0$ となる k が存在する．

(iii) 任意の $\boldsymbol{x} \neq \boldsymbol{0}$ に対して非負対角行列 $D_{\boldsymbol{x}}$ で

$$(\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x})>0 \quad \text{かつ} \quad (A\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x}) \geq 0$$

となるものが存在する．

(iv) A を含む A のすべての主小行列の実固有値は非負である．

(v) 任意の正数 ε に対して $A+\varepsilon I_n \in \mathcal{P}$

(vi) 任意の正対角行列 D に対して $A+D \in \mathcal{P}$

(vii) 任意の正対角行列 D に対して $\det(A+D)>0$

(viii) 任意の正対角行列 D に対して $\det(A+D) \neq 0$

【証明】 (i)⇒(ii) $M=\{i \mid x_i \neq 0\}=\{i_1,\ldots,i_m\}$ とすれば $\boldsymbol{x}\neq\boldsymbol{0}$ より $M\neq\emptyset$ である．

仮に $x_iy_i<0\ \forall i\in M$ として矛盾を導けばよい．$A(M)$, $\boldsymbol{x}(M)$, $\boldsymbol{y}(M)$, U を定理 9.1 (i) ⇒(ii) の証明のように定義すれば

$$\boldsymbol{y}(M)=A(M)\boldsymbol{x}(M)=-U\boldsymbol{x}(M)\quad\therefore\ \big(A(M)+U\big)\boldsymbol{x}(M)=\boldsymbol{0}$$

$\boldsymbol{x}(M)\neq\boldsymbol{0}$ であるから $A(M)+U$ は正則でない．しかし仮定 $x_iy_i<0\ \forall i\in M$ によって U は正対角行列であるから定理 9.2 (ii) によって $A(M)+U\in\mathcal{P}$．したがって $\det\big(A(M)+U\big)>0$ でなければならず矛盾が引き起こされた．

(ii)⇒(iii) $\boldsymbol{x}=(x_1,\ldots,x_n)\neq\boldsymbol{0}$ に対して $\boldsymbol{y}=A\boldsymbol{x}=(y_1,\ldots,y_n)^{\mathrm{t}}$ とし

$$x_k\neq 0,\quad x_ky_k\geq 0$$

なる k をとる．このとき

$$D_{\boldsymbol{x}}=\mathrm{diag}(d_1,\ldots,d_n),\quad d_k=1,\ d_j=0\ (j\neq k)$$

とおけば

$$(\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x})=x_k^2>0\quad\text{かつ}\quad(A\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x})=y_kx_k\geq 0$$

(iii)⇒(iv) $M=\{i_1,\ldots,i_m\}\subseteq N=\{1,2,\ldots,n\}$, $M\neq\emptyset$ に対し

$$A(M)\boldsymbol{u}=\lambda\boldsymbol{u},\quad \boldsymbol{u}=(u_1,\ldots,u_m)\neq\boldsymbol{0},\quad \lambda\text{は実数}$$

とする．このとき $\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}}$ を

$$x_i=\begin{cases}u_i & (i\in M)\\ 0 & (i\notin M)\end{cases}$$

により定義すれば $\boldsymbol{x}\neq\boldsymbol{0}$．よって仮定により非負対角行列 $D_{\boldsymbol{x}}$ をとり

$$(\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x})>0\quad\text{かつ}\quad(A\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x})\geq 0 \tag{9.4}$$

とできて

$$\begin{aligned}(A\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x})&=\big(A(M)\boldsymbol{x}(M),\ D_{\boldsymbol{x}}(M)\boldsymbol{x}(M)\big)=\big(A(M)\boldsymbol{u},\ D_{\boldsymbol{x}}(M)\boldsymbol{u}\big)\quad(\because\ \boldsymbol{x}(M)=\boldsymbol{u})\\&=(\lambda\boldsymbol{u},\ D_{\boldsymbol{x}}(M)\boldsymbol{u})=\lambda(\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x})\\ \therefore\quad\lambda&=\frac{(A\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x})}{(\boldsymbol{x},\ D_{\boldsymbol{x}}\boldsymbol{x})}\geq 0\end{aligned} \tag{9.5}$$

(iv)⇒(v) $A(M)$ の実固有値を $\lambda(A(M))$ であらわせば $(A+\varepsilon I)(M)$ の実固有値は $\lambda(A(M))+\varepsilon$ であるから $\lambda(A(M))\geq 0$ より $\lambda(A(M))+\varepsilon>0$．故に定理 9.1 (v) を $A+\varepsilon I$ に適用すれば $A+\varepsilon I\in\mathcal{P}$

(v)⇒(vi) D を任意の正対角行列とすれば，正数 ε を十分小さくとって $D_\varepsilon=D-\varepsilon I$ を正対角行列にできる．このとき $A+\varepsilon I\in\mathcal{P}$ であるから，定理 9.2 (ii) によって

$$A+D=(A+\varepsilon I)+D_\varepsilon\in\mathcal{P}$$

(vi)⇒(vii) 明らか．

(vii)⇒(viii) 明らか．

(viii)⇒(i) 仮に $A\notin\mathcal{P}_0$ とすれば，$\det(A+D)=0$ なる正対角行列 D が存在することを示す．そのために $\det A<0$ と $\det A\geq 0$ の 2 つの場合に分けて考える．

(a) $\det A < 0$ のとき．　正数 ξ を十分小さくとれば $\det(\xi I + A) < 0$
一方 ξ を十分大きくとれば $\det(\xi I + A) = \xi^n \det(I + \frac{1}{\xi} A) > 0$
$\det(\xi I + A)$ は ξ の連続関数であるから $\det(\xi I + A) = 0$ なる ξ が存在する．この ξ に対し $D = \xi I$ とおけば，D は正対角行列で，$\det(A + D) = 0$

(b) $\det A \geq 0$ のとき．　$A \notin \mathcal{P}_0$ と仮定しているから，ある $k < n$ につき A の k 次主小行列式が負となる．一般性を失うことなくこれを $A_1 = A \begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}$ とし

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, \quad \det A_1 < 0$$

とブロック表示すれば，十分小なる $\xi > 0$ につき $\det(\xi I_k + A_1) < 0$
いま

$$D^{(1)} = \begin{pmatrix} \xi I_k & \\ & \eta I_{n-k} \end{pmatrix}, \quad D^{(2)} = \zeta I_n$$

とおけば，η が十分大きいとき

$$\begin{aligned} \det(A + D^{(1)}) &= \det \begin{pmatrix} A_1 + \xi I_k & A_2 \\ A_3 & A_4 + \eta I_{n-k} \end{pmatrix} \\ &= \eta^{n-k} \det \begin{pmatrix} A_1 + \xi I_k & A_2 \\ \frac{1}{\eta} A_3 & \frac{1}{\eta} A_4 + I_{n-k} \end{pmatrix} < 0 \end{aligned}$$

また ζ が十分大きいとき

$$\det(A + D^{(2)}) = \zeta^n \det(I + \frac{1}{\zeta} A) > 0$$

よって

$$D(\varepsilon) = \varepsilon D^{(1)} + (1 - \varepsilon) D^{(2)}, \quad 0 < \varepsilon < 1$$

とおくとき，$D(\varepsilon)$ は正対角行列で，$\det(A + D(\varepsilon)) = 0$ をみたす $\varepsilon \in (0, 1)$ が存在する．

よって $A \notin \mathcal{P}_0$ と仮定すれば，(a), (b) いずれの場合でも $\det(A + D) = 0$ なる正対角行列 D が存在し矛盾が生じるから，$A \in \mathcal{P}_0$ でなければならない．　証明終 ∎

■**注意**　定理 9.3 において (i)～(v) の同値性は Fiedler-Pták [24] により与えられた．(vi) と (vii) は児玉・須田 [3] にある．また (i) と (viii) の同値性は Sandberg-Willson, Jr. [37] により示された．

系 9.3.1　$A + A^{\mathrm{t}}$ が半正定値ならば $A \in \mathcal{P}_0$

【証明】　$(A\boldsymbol{x},\ \boldsymbol{x}) = \frac{1}{2}\big((A + A^{\mathrm{t}})\boldsymbol{x},\ \boldsymbol{x}\big) \geq 0\ \forall \boldsymbol{x} \neq \boldsymbol{0}$ であるから $D_{\boldsymbol{x}} = I_n$ は定理 9.3 の条件 (iii) をみたす．　証明終 ∎

系 9.3.2 (Sandberg-Willson, Jr.)　$A \in \mathcal{P}_0$ かつ $\det A \neq 0$ であるための必要十分条件は任意の非負対角行列 D につき

$$\det(A + D) > 0$$

が成り立つことである．

【証明】 (i) 必要条件 ($\Rightarrow$) 定理 9.2 によって $A \in \mathcal{P}_0$ かつ D が非負対角ならば $A+D \in \mathcal{P}_0$ である．また $A \in \mathcal{P}_0$ より $\det A \geq 0$ であるが仮定 $\det A \neq 0$ によって $\det A > 0$ である．

さて $D = \mathrm{diag}(d_1, \ldots, d_n)$, $d_i \geq 0\ \forall i$, $D_i = \mathrm{diag}(0, \ldots, 0, d_i, 0, \ldots, 0)$ とする．A の第 i 行第 i 列を取り除いて得られる $n-1$ 次小行列を A_i とすれば $\det A_i \geq 0$ である．

$$\therefore \quad \det(A + D_i) = \det A + d_i \det A_i \geq \det A > 0$$

$i = 1$ として

$$\det(A + D_1) > 0$$

$$\therefore \quad \det(A + D_1 + D_2) = \det(A + D_1) + d_2 \det A_2 \geq \det(A + D_1) > 0$$

以下同様にして

$$\begin{aligned} \det(A + D) &= \det(A + D_1 + \cdots + D_{n-1}) + d_n \det A_n \\ &\geq \det(A + D_1 + \cdots + D_{n-1}) > 0 \end{aligned}$$

(ii) 十分条件 ($\Leftarrow$) 任意の非負対角行列 D に対して $\det(A+D) > 0$ ならば，正対角行列に対しても $\det(A+D) > 0$．よって定理 9.3 の条件 (vii) がみたされて $A \in \mathcal{P}_0$ となる．特に $D = O$ ととれば $\det A = \det(A+D) > 0$ よって $\det A \neq 0$ である． 証明終 ∎

系 9.3.3 $A \in \mathcal{P}_0$ かつ $\det A \neq 0$ ならば $A^{-1} \in \mathcal{P}_0$

【証明】 D を正対角行列とすると

$$\det(A^{-1} + D) = \det\bigl(A^{-1}(D^{-1} + A)D\bigr) = \det A^{-1} \det(D^{-1} + A) \det D$$

系 9.3.2 によって $\det(D^{-1} + A) > 0$ である．

また $A \in \mathcal{P}_0$ より $\det A \geq 0$ であるが $\det A \neq 0$ と仮定しているから $\det A > 0$

$$\therefore \quad \det A^{-1} = \frac{1}{\det A} > 0 \quad \therefore \quad \det(A^{-1} + D) > 0$$

故に A^{-1} は定理 9.3 の条件 (vii) をみたす．よって $A^{-1} \in \mathcal{P}_0$ を得る． 証明終 ∎

定理 9.4 (Sandberg-Willson, Jr.) n 次行列 $A = (a_{ij})$ は $a_{ii} > 0\ \forall i$ をみたし，かつ列に関して狭義優対角

$$a_{ii} > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ji}| \quad \forall i$$

とする．n 次行列 $B = (b_{ij})$ が $b_{ii} \geq 0\ \forall i$ かつ列につき優対角

$$b_{ii} \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |b_{ji}| \quad \forall i$$

ならば $A^{-1}B \in \mathcal{P}_0$，また列につき狭義優対角ならば $A^{-1}B \in \mathcal{P}$

【証明】 仮に $A^{-1}B \notin \mathcal{P}_0$ とすれば定理 9.3 (viii) により $\det(A^{-1}B + D) = 0$ をみたす正対角行列 $D = \mathrm{diag}(d_1, \ldots, d_n)$ が存在する．このとき

$$\det(A^{-1}B + D) = \det A^{-1} \det(B + AD), \quad \det A^{-1} \neq 0$$

より

$$\det(B+AD)=0 \tag{9.6}$$

B が $b_{ii}\geq 0\ \forall i$ かつ列につき優対角ならば，$C=B+AD=(c_{ij})$ は列に関して狭義優対角である．実際

$$\sum_{\substack{j=1\\ j\neq i}}^{n}|c_{ji}|=\sum_{\substack{j=1\\ j\neq i}}^{n}|b_{ji}+a_{ji}d_i|\leq\sum_{j\neq i}|b_{ji}|+d_i\sum_{j\neq i}|a_{ji}|<b_{ii}+d_ia_{ii}\ (\because\ d_i>0)\ =c_{ii}$$

故に定理 1.21 によって C は正則であるからこれは (9.6) と矛盾する．$\therefore\ A^{-1}B\in\mathcal{P}_0$

さらに B が列に関して狭義優対角ならば適当な正数 δ をえらんで $B-\delta A$ も狭義優対角にできる．このとき任意の正対角行列 D に対し $B-\delta A+AD$ は列に関して狭義優対角，したがって正則である．

$$\therefore\ \det(A^{-1}B-\delta I+D)=\det A^{-1}\det(B-\delta A+AD)\neq 0\quad\forall D>O$$

よって定理 9.3 (viii) の条件が $A^{-1}B-\delta I$ に対してみたされて $A^{-1}B-\delta I\in\mathcal{P}_0$．故に定理 9.2 (ii) によって $A^{-1}B=(A^{-1}B-\delta I)+\delta I\in\mathcal{P}$ ($\because\ D=\delta I$ は正対角行列)．　証明終 ∎

9.3 Sandberg-Willson, Jr. の定理

$A=(a_{ij})\in\mathbb{R}^{n\times n}$ とし，$\boldsymbol{x}=(x_1,\ldots,x_n)^{\mathrm{t}}$，$\boldsymbol{b}=(b_1,\ldots,b_n)^{\mathrm{t}}\in\mathbb{R}^n$ とおく．

$$\mathcal{F}=\left\{f\in C^1(\mathbb{R})\ \middle|\ f(\mathbb{R})=\mathbb{R},\ f'(x)>0\ \forall x\in\mathbb{R}\right\}$$

$$\mathcal{F}^n=\overbrace{\mathcal{F}\times\cdots\times\mathcal{F}}^{n}=\left\{F(\boldsymbol{x})=(f_1(x_1),\ldots,f_n(x_n))^{\mathrm{t}}\ \middle|\ f_i\in\mathcal{F},\ 1\leq i\leq n\right\}$$

$$F\in\mathcal{F}^n$$

として n 元連立準線形方程式

$$A\boldsymbol{x}+F(\boldsymbol{x})=\boldsymbol{b} \tag{9.7}$$

を考える．このよう方程式はある種の非線形抵抗回路の方程式をあらわすが，2 次元境界値問題

$$-\left(\frac{\partial^2u}{\partial x^2}+\frac{\partial^2u}{\partial y^2}\right)+f(x,\ y,\ u)=0,\quad (x,y)\in\Omega\ (\text{有界領域}) \tag{9.8}$$

$$u=g(x,y),\quad (x,y)\in\partial\Omega\ (\Omega\text{の境界}) \tag{9.9}$$

を離散近似する場合にもあらわれる．

次の定理は回路理論の分野ではよく知られているが，数学関係者には殆んど知られていない．

定理 9.5 (Sandberg-Willson, Jr. の定理 [36])　方程式 (9.7) が任意の $F\in\mathcal{F}^n$ と任意の $\boldsymbol{b}\in\mathbb{R}^n$ につき一意解をもつための必要十分条件は $A\in\mathcal{P}_0$．このとき解は $\boldsymbol{b}$ の連続関数である．

【証明】　以下の証明は原論文 [36] に与えられたものである．やや冗長であるが初等的であるので敢えてこれを記す．多変数関数に関する陰関数定理を使えばこの証明は大幅に短縮され，解は $\boldsymbol{b}$ の成分につき微分可能であることまで主張できる．

(i) 十分条件．$A\in\mathcal{P}_0$ のとき (9.7) は任意の $F\in\mathcal{F}^n$ と任意の $\boldsymbol{b}$ につき一意解をもち，解は $\boldsymbol{b}$ の連続関数であることを示そう．

(a) 解の一意性： $\boldsymbol{x}$ と $\boldsymbol{y}$ を (9.7) の 2 つの解とすれば

$$f_i(x_i) - f_i(y_i) = \frac{\partial f_i}{\partial x}\big(y_i + \theta_i(x_i - y_i)\big)(x_i - y_i), \quad 0 < \theta_i < 1$$

と書ける．

$$d_i = \frac{\partial f}{\partial x}\big(y_i + \theta_i(x_i - y_i)\big), \quad D = \mathrm{diag}(d_1, \ldots, d_n)$$

とおけば $D > O$ で

$$F(\boldsymbol{x}) - F(\boldsymbol{y}) = D \cdot (\boldsymbol{x} - \boldsymbol{y}) \tag{9.10}$$

よって (9.7) より $(A+D)(\boldsymbol{x}-\boldsymbol{y}) = \boldsymbol{0}$ を得るが，定理 9.3 より $A+D$ は正則であるから $\boldsymbol{x}-\boldsymbol{y} = \boldsymbol{0}$.

(b) 解の存在と $\boldsymbol{b}$ に関する連続性：

$$F_k(\boldsymbol{x}) = (f_1(x_1), \ldots, f_k(x_k))^{\mathrm{t}}, \quad A_k = A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}, \quad \boldsymbol{b}^{(k)} = (b_1, \ldots, b_k)^{\mathrm{t}}$$

とおくと $F_k \in \mathcal{F}^k$, $A_k \in \mathcal{P}_0$, $\boldsymbol{b}^{(k)} \in \mathbb{R}^k$, $k = 1, 2, \ldots, n$

以下 (9.7) の解の (一意) 存在を k に関する帰納法により示す．

$k = 1$ のとき

$$F_1(x) + A_1 x = f_1(x) + a_{11}x$$

$g(x) = f_1(x) + a_{11}x$ は $g'(x) = f_1'(x) + a_{11} > 0$ をみたす ($\because\ a_{11} \geq 0$)．よって $F_1(x) + A_1 x = b_1$ は任意の $F_1 \in \mathcal{F}^1 = \mathcal{F}$, $b_1 \in \mathbb{R}$ に対して一意解 x をもち，x は b_1 の連続関数である．

$k < n$ のとき　各 $F_k \in \mathcal{F}^k$ と $\boldsymbol{b}^{(k)} \in \mathbb{R}^k$ に対し

$$F_k(\boldsymbol{x}) + A_k\boldsymbol{x} = \boldsymbol{b}^{(k)}$$

は一意解 $\boldsymbol{x} \in \mathbb{R}^k$ をもち，$\boldsymbol{b}^{(k)}$ がスカラーパラメータ η につき連続的に変化するとき解は η の連続関数であると仮定する．

$$A_{k,k+1} = \begin{pmatrix} a_{1\,k+1} \\ \vdots \\ a_{k\,k+1} \end{pmatrix}, \quad A_{k+1\ k} = (a_{k+1\,1}, \ldots, a_{k+1\,k})$$

とおけば任意の $x_{k+1} \in \mathbb{R}$ に対して

$$F_k(\boldsymbol{x}) + A_k\boldsymbol{x} + A_{k\ k+1}x_{k+1} - \boldsymbol{b}^{(k)} = \boldsymbol{0} \tag{9.11}$$

は一意解 $\boldsymbol{x} = (x_1, \ldots, x_k)^{\mathrm{t}}$ をもち，その解は x_{k+1} と η の連続関数である．

$$x_i = m_i(x_{k+1},\ \eta), \quad 1 \leq i \leq k$$

$$M_k = M_k(x_{k+1},\ \eta) = \begin{pmatrix} m_1(x_{k+1},\ \eta) \\ \vdots \\ m_k(x_{k+1},\ \eta) \end{pmatrix}$$

$$\varphi_{k+1}(x_{k+1},\ \eta) = A_{k+1\,k}M_k(x_{k+1},\ \eta) + a_{k+1\,k+1}x_{k+1} - b_{k+1}(\eta)$$

とおくとき，次の補題が成り立つ．　(証明つづく)

補題 9.1　$\varphi_{k+1}(x_{k+1},\ \eta)$ は x_{k+1} につき広義単調増加関数である．

【証明】 $x_{k+1}^1, x_{k+1}^2 \in \mathbb{R}$, $x_{k+1}^1 < x_{k+1}^2$ とし $x_{k+1} = x_{k+1}^i$ $(i=1,2)$ に対する (9.11) の解を

$$M^i = M(x_{k+1}^i,\ \eta) = \begin{pmatrix} m_1(x_{k+1}^i,\ \eta) \\ \vdots \\ m_k(x_{k+1}^i,\ \eta) \end{pmatrix}$$

とすれば

$$F_k(M^1) + A_k M^1 + A_{k\,k+1} x_{k+1}^1 = \boldsymbol{b}^{(k)}$$
$$F_k(M^2) + A_k M^2 + A_{k\,k+1} x_{k+1}^2 = \boldsymbol{b}^{(k)}$$
$$\therefore\ F_k(M^2) - F_k(M^1) + A_k(M^2 - M^1) + A_{k\,k+1}(x_{k+1}^2 - x_{k+1}^1) = \boldsymbol{0}$$

(9.10) によって k 次正対角行列 Δ_k が存在して

$$F_k(M^2) - F_k(M^1) = \Delta_k(M^2 - M^1)$$

とかくことができる.

$$\therefore\ (\Delta_k + A_k)(M^2 - M^1) + A_{k\,k+1}(x_{k+1}^2 - x_{k+1}^1) = \boldsymbol{0}$$
$$\therefore\ M^2 - M^1 = -(x_{k+1}^2 - x_{k+1}^1)(A_k + \Delta_k)^{-1} A_{k\,k+1}$$
$$\begin{aligned} \therefore\ \varphi_{k+1}(x_{k+1}^2,\ \eta) - \varphi_{k+1}(x_{k+1}^1,\ \eta) &= A_{k+1\,k}(M^2 - M^1) + a_{k+1\,k+1}(x_{k+1}^2 - x_{k+1}^1) \\ &= \big(a_{k+1\,k+1} - A_{k+1\,k}(A_k + \Delta_k)^{-1} A_{k\,k+1}\big)(x_{k+1}^2 - x_{k+1}^1) \\ &= \sigma(x_{k+1}^2 - x_{k+1}^1) \end{aligned} \tag{9.12}$$
$$\text{ただし}\quad \sigma = a_{k+1\,k+1} - A_{k+1\,k}(A_k + \Delta_k)^{-1} A_{k\,k+1}$$

等式

$$\begin{pmatrix} A_k + \Delta_k & A_{k\,k+1} \\ A_{k+1\,k} & a_{k+1\,k+1} \end{pmatrix} \begin{pmatrix} I_k & -(A_k + \Delta_k)^{-1} A_{k\,k+1} \\ O & 1 \end{pmatrix} = \begin{pmatrix} A_k + \Delta_k & O \\ A_{k+1\,k} & \sigma \end{pmatrix}$$

の両辺の行列式をとれば, $D = \begin{pmatrix} \Delta_k & \\ & 0 \end{pmatrix}$ ($k+1$ 次非負対角行列) として

$$\det(A_{k+1} + D) = \det(A_k + \Delta_k)\sigma \tag{9.13}$$

$A_k \in \mathcal{P}_0$ かつ $\Delta_k > O$ より $A_k + \Delta_k \in \mathcal{P}$ (定理 9.2 (ii))

$$\therefore\quad \det(A_k + \Delta_k) > 0 \tag{9.14}$$

また $A_{k+1} = A \begin{pmatrix} 1 & 2 & \cdots & k+1 \\ 1 & 2 & \cdots & k+1 \end{pmatrix} \in \mathcal{P}_0$ かつ $D \geq O$ より $A_{k+1} + D \in \mathcal{P}_0$

$$\therefore\quad \det(A_{k+1} + D) \geq 0 \tag{9.15}$$

(9.13)–(9.15) より $\sigma \geq 0$

したがって (9.12) より

$$x_{k+1}^1 < x_{k+2}^2 \quad \text{ならば} \quad \varphi_{k+1}(x_{k+1}^1,\ \eta) \leq \varphi_{k+1}(x_{k+1}^2,\ \eta)$$

補題 9.1 の証明終 ∎

(定理 9.5 の証明つづき) さて十分条件の証明に戻る.

$\varphi_{k+1}(x_{k+1},\ \eta)$ は x_{k+1} につき連続であるから $f_{k+1}(x_{k+1})+\varphi_{k+1}(x_{k+1},\ \eta)$ は x_{k+1} の狭義単調増加な連続関数で $\mathbb{R}^1$ から $\mathbb{R}^1$ の上への写像である. よって方程式

$$f_{k+1}(x_{k+1})+\varphi_{k+1}(x_{k+1},\ \eta)=0 \tag{9.16}$$

は一意解 $x_{k+1}=x_{k+1}(\eta)$ をもつ. このとき

$$\boldsymbol{x}=(m_1(x_{k+1},\ \eta),\ldots,m_k(x_{k+1},\ \eta),x_{k+1})^{\mathrm{t}}$$

は $k+1$ 元連立方程式

$$F_{k+1}(\boldsymbol{x})+A_{k+1}\boldsymbol{x}=\boldsymbol{b}^{(k+1)}$$

$$\left(F_{k+1}=\begin{pmatrix} f_1(x_1)\\ \vdots\\ f_{k+1}(x_{k+1})\end{pmatrix},\ A_{k+1}=A\begin{pmatrix}1 & 2 & \cdots & k+1\\ 1 & 2 & \cdots & k+1\end{pmatrix},\ \boldsymbol{b}^{(k+1)}=\begin{pmatrix} b_1\\ \vdots\\ b_{k+1}\end{pmatrix}\right)$$

の (一意) 解である.

この解は η の連続関数であることを次の補題で示そう. これが示されれば帰納法が完了する. (証明つづく)

補題 9.2 $x_{k+1}=x_{k+1}(\eta)$ は η の連続関数である.

【証明】 $\eta=\eta^0$ に対する (9.16) の解を $x^0_{k+1}=x^0_{k+1}(\eta^0)$ とおく. したがって

$$f_{k+1}(x^0_{k+1})+\varphi_{k+1}(x^0_{k+1},\ \eta^0)=0$$

f_{k+1} は $\mathbb{R}$ から $\mathbb{R}$ の上への狭義単調増加な連続関数であるから逆関数 f^{-1}_{k+1} も存在して $\mathbb{R}$ から $\mathbb{R}$ の上への狭義単調増加な連続関数である. よって任意に与えられた $\varepsilon>0$ に対し適当な $\delta'>0$ を定めて

$$\left|f_{k+1}(x_{k+1})-f_{k+1}(x^0_{k+1})\right|<\delta'\ \Rightarrow\ |x_{k+1}-x^0_{k+1}|<\varepsilon \tag{9.17}$$

とできる. 一方 φ_{k+1} は η の連続関数であるから適当な $\delta>0$ を定めて

$$|\eta-\eta^0|<\delta\ \Rightarrow\ \left|\varphi_{k+1}(x^0_{k+1},\ \eta)-\varphi_{k+1}(x^0_{k+1},\ \eta^0)\right|<\delta' \tag{9.18}$$

とできる. このとき

$$|\eta-\eta^0|<\delta\quad\text{かつ}\quad f_{k+1}(x_{k+1})+\varphi_{k+1}(x_{k+1},\ \eta)=0$$

ならば

$$\begin{aligned}&f_{k+1}(x_{k+1})-f_{k+1}(x^0_{k+1})+\varphi_{k+1}(x_{k+1},\ \eta)-\varphi_{k+1}(x^0_{k+1},\ \eta)\\ &\quad=-\{\varphi_{k+1}(x_{k+1},\ \eta)-\varphi_{k+1}(x^0_{k+1},\ \eta^0)\}+\varphi_{k+1}(x_{k+1},\ \eta)-\varphi_{k+1}(x^0_{k+1},\ \eta)\\ &\quad=\varphi_{k+1}(x^0_{k+1},\ \eta^0)-\varphi_{k+1}(x^0_{k+1},\ \eta)\end{aligned}$$

さらに

$$\begin{aligned}&(x_{k+1}-x^0_{k+1})\big(f_{k+1}(x_{k+1})-f_{k+1}(x^0_{k+1})\big)\ge 0\\ &(x_{k+1}-x^0_{k+1})\big(\varphi_{k+1}(x_{k+1},\ \eta)-\varphi_{k+1}(x^0_{k+1},\ \eta)\big)\ge 0\ \text{(補題 9.1 による)}\end{aligned}$$

$$\begin{aligned}\therefore\ &(x_{k+1}-x^0_{k+1})\big(\varphi_{k+1}(x^0_{k+1},\ \eta^0)-\varphi_{k+1}(x^0_{k+1},\ \eta)\big)\\ &\quad=(x_{k+1}-x^0_{k+1})\big\{\big(f_{k+1}(x_{k+1})-f_{k+1}(x^0_{k+1})\big)+(\varphi_{k+1}(x_{k+1},\ \eta)-\varphi_{k+1}(x^0_{k+1},\ \eta)\big\}\end{aligned}$$

$$\geq (x_{k+1}-x_{k+1}^0)\big(f_{k+1}(x_{k+1})-f_{k+1}(x_{k+1}^0)\big)\geq 0$$

$$\therefore\ |x_{k+1}-x_{k+1}^0|\,\big|\varphi_{k+1}(x_{k+1}^0,\ \eta^0)-\varphi_{k+1}(x_{k+1}^0,\ \eta)\big|$$
$$\geq |x_{k+1}-x_{k+1}^0|\,\big|f_{k+1}(x_{k+1})-f_{k+1}(x_{k+1}^0)\big|$$

したがって $x_{k+1}-x_{k+1}^0\neq 0$ ならば

$$\big|f_{k+1}(x_{k+1})-f_{k+1}(x_{k+1}^0)\big|\leq\big|\varphi_{k+1}(x_{k+1}^0,\ \eta)-\varphi_{k+1}(x_{k+1}^0,\ \eta^0)\big|$$

この不等式は $x_{k+1}-x_{k+1}^0=0$ のときも成り立つ.

したがって (9.18) より

$$|\eta-\eta_0|<\delta\ \Rightarrow\ \big|f_{k+1}(x_{k+1})-f_{k+1}(x_{k+1}^0)\big|<\delta'$$

よって (9.17) より

$$|\eta-\eta_0|<\delta\ \Rightarrow\ |x_{k+1}-x_{k+1}^0|<\varepsilon$$

これは $x_{k+1}=x_{k+1}(\eta)$ が η につき連続であることを示している. 補題 9.2 の証明終 ∎

(定理 9.5 の証明つづき) 結局 補題 9.1 と 9.2 によって

$$F_{k+1}(\boldsymbol{x})+A_{k+1}\boldsymbol{x}=\boldsymbol{b}^{(k+1)}$$

は η につき連続な解 $\boldsymbol{x}\in\mathbb{R}^{k+1}$ をただ 1 つもつ. 故に帰納法が完了し, $A\in\mathcal{P}_0$ ならば (9.7) は一意解をもち, それは $\boldsymbol{b}$ の連続関数であることが示された.

(ii) 必要条件. (9.7) が任意の $F\in\mathcal{F}^n$ と任意の $\boldsymbol{b}$ につき一意解をもつと仮定する.

$D=\mathrm{diag}(d_1,\ldots,d_n)$ を正対角行列として $F(\boldsymbol{x})=D\boldsymbol{x}$ (したがって $f_i(\boldsymbol{x})=f_i(x_i)=d_ix_i$) とすれば (9.7) は

$$(A+D)\boldsymbol{x}=\boldsymbol{b}$$

となりこの方程式は一意解をもつ. $\therefore\ \det(A+D)\neq 0$

よって定理 9.3 の条件 (viii) がみたされるから $A\in\mathcal{P}_0$ である. 定理 9.5 の証明終 ∎

■**注意 1** $\mathcal{F}$ の定義において $f(\mathbb{R})=\mathbb{R}$ なる仮定は必要である (Sandberg-Willson, Jr. [36]). 例えば

$$n=2,\quad f_1(x)=f_2(x)=e^x,\quad A=\begin{pmatrix}1&-1\\-1&1\end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix}1\\-2\end{pmatrix}$$

とすれば, (9.7) は

$$e^{x_1}+x_1-x_2=1,\quad e^{x_2}-x_1+x_2=-2$$

となるが, この連立方程式は解をもたない. 実際, 上の 2 つの式を辺々加えれば

$$e^{x_1}+e^{x_2}=-1$$

となって不合理である. この場合 $A\in\mathcal{P}_0$ であるが $f(x)=e^x$ は $f(\mathbb{R})=\mathbb{R}$ をみたさないからこの例では $F\notin\mathcal{F}^2$ である.

■**注意 2** 定理 9.5 において A を $\mathcal{P}$ 行列でおきかえることはできない ([36]).

実際 $f_1,f_2\in\mathcal{F}$, $A=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$ とすれば $A\notin\mathcal{P}$ であるが,

$$f_1(x_1)-x_2=b_1,\quad f_2(x_2)+x_1=b_2$$

は任意の $b_1, b_2 \in \mathbb{R}$ に対して一意解をもつ．なぜならば第1式を第2式に代入すれば

$$f_2(f_1(x_1) - b_1) + x_1 = b_2$$

左辺の関数は $\mathbb{R}$ を $\mathbb{R}$ の上へ写す狭義単調増加関数であるから，与えられた b_1, b_2 に対して x_1 は一意に定まり，それを第1式に代入して x_2 が一意に定まるからである．

注意 3 $\mathcal{F}$ のクラスを拡げて

$$\widehat{\mathcal{F}} = \left\{ f \in C^1(\mathbb{R}) \;\middle|\; f(\mathbb{R}) = \mathbb{R},\ f'(x) \geq 0 \quad \forall x \in \mathbb{R} \right\}$$

$$\widehat{\mathcal{F}}^n = \overbrace{\widehat{\mathcal{F}} \times \cdots \times \widehat{\mathcal{F}}}^{n}$$

とするとき次の定理が成り立つ．

定理 9.6 $A \in \mathcal{P}$ ならば方程式 (9.7) は任意の $F \in \widehat{\mathcal{F}}^n$ と任意の $\boldsymbol{b} \in \mathbb{R}^n$ につき一意解 $\boldsymbol{x}$ をもつ．特に A が正定値対称行列ならば

$$||\boldsymbol{x}||_2 \leq ||A^{-1}||_2 \, ||F(\boldsymbol{0}) - \boldsymbol{b}||_2 = \frac{1}{\min \lambda(A)} ||F(\boldsymbol{0}) - \boldsymbol{b}||_2$$

ただし $\lambda(A)$ は A の固有値をあらわす．

【証明】 前半は定理 9.5 の十分条件の証明と同様である (ただし系 9.1.4 を用いる)．後半を示そう．A を正定値対称行列，$\boldsymbol{u} \in \mathbb{R}^n$, $G(\boldsymbol{u}) = F(\boldsymbol{u}) - \boldsymbol{b}$, $F \in \widehat{\mathcal{F}}^n$ とおくとき，任意の $\boldsymbol{u} = (u_1, \ldots, u_n)^{\mathrm{t}}$, $\boldsymbol{v} = (v_1, \ldots, v_n)^{\mathrm{t}} \in \mathbb{R}^n$ に対して

$$(\boldsymbol{u} - \boldsymbol{v})^{\mathrm{t}} \big(G(\boldsymbol{u}) - G(\boldsymbol{v})\big) = \sum_{i=1}^{n} (u_i - v_i)\big(f_i(u_i) - f_i(v_i)\big) \geq 0$$

$$\left(\because\ (u_i - v_i)\big(f_i(u_i) - f_i(v_i)\big) \geq 0 \quad \forall i\right)$$

特に $\boldsymbol{u} = \boldsymbol{x}$, $\boldsymbol{v} = \boldsymbol{0}$ とおけば

$$\boldsymbol{x}^{\mathrm{t}} G(\boldsymbol{x}) \geq \boldsymbol{x}^{\mathrm{t}} G(\boldsymbol{0})$$

$$\therefore \quad \boldsymbol{x}^{\mathrm{t}} A \boldsymbol{x} = -\boldsymbol{x}^{\mathrm{t}} G(\boldsymbol{x}) \leq -\boldsymbol{x}^{\mathrm{t}} G(\boldsymbol{0}) \leq ||\boldsymbol{x}||_2 ||G(\boldsymbol{0})||_2$$

$$\therefore \quad \min \lambda(A) \boldsymbol{x}^{\mathrm{t}} \boldsymbol{x} \leq ||\boldsymbol{x}||_2 ||G(\boldsymbol{0})||_2$$

$\boldsymbol{x} \neq \boldsymbol{0}$ のとき両辺を $||\boldsymbol{x}||_2 > 0$ で割って

$$\min \lambda(A) ||\boldsymbol{x}||_2 \leq ||G(\boldsymbol{0})||_2$$

$$\therefore \quad ||\boldsymbol{x}||_2 \leq \frac{1}{\min \lambda(A)} ||G(\boldsymbol{0})||_2 = ||A^{-1}||_2 \, ||F(\boldsymbol{0}) - \boldsymbol{b}||_2$$

この不等式は $\boldsymbol{x} = \boldsymbol{0}$ のときも成り立つ． 証明終 ∎

9.4 Ω 行列

定義 9.3 (Ω 行列) 行列 A が **Ω 行列**であるとは，$A + D$ が正則となるような任意の正対角行列 D に対し $(A + D)^{-1} = (b_{ij})$ が符号条件

$$b_{ii} < 0 \ \Rightarrow\ b_{ij} \leq 0 \quad (j \neq i)$$

をみたすときをいう．

Ω 行列は P, P_0 行列の一般化として 1991 年西哲生教授 (当時，九州大学，現早稲田大学) により導入された．

前節の $\mathcal{F}$ と $\mathcal{F}^n$ を

$$\widetilde{\mathcal{F}} = \left\{ f \in C^2(\mathbb{R}) \;\middle|\; f(\mathbb{R}) = \mathbb{R},\ f'(x) > 0,\ f''(x) \geq 0 \quad \forall x \in \mathbb{R} \right\}$$

$$\widetilde{\mathcal{F}}^n = \overbrace{\widetilde{\mathcal{F}} \times \cdots \times \widetilde{\mathcal{F}}}^{n}$$

でおきかえるとき，次の結果が知られている．

定理 9.7 (西 [32]) 方程式 (9.7) が任意の $F \in \widetilde{\mathcal{F}}^n$ と任意の $\boldsymbol{b} \in \mathbb{R}^n$ につき高々 2^n 個の解をもつための必要十分条件は A が Ω 行列であることである．

【付記】著者の独断によれば，そこに記された証明には細部になお若干詰めねばならない箇所が見受けられる．今後この理論の進展を期待したい．細部の詰めが完了すればこの定理は大定理となる．

第10章　テンソル積

行列論におけるやや進んだ概念としてテンソル積とGrassmann積および複合行列がある．この章ではテンソル積を定義しその基本的性質と応用について述べる．

10.1　テンソル積

定義 10.1 (テンソル積 $\mathbf{A}\otimes\mathbf{B}$)　$m\times n$ 行列 $A=(a_{ij})$ と $p\times q$ 行列 $B=(b_{ij})$ に対し

$$(mp)\times(nq)\text{ 行列 }\begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ \vdots & & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix} \tag{10.1}$$

を A と B のテンソル積といい $A\otimes B$ であらわす．

以下にこの行列の基本的性質を述べる．

定理 10.1　A,B,X,Y をそれぞれ $m\times n$, $p\times q$, $n\times p$, $m\times q$ 行列とする．

(i)　線形写像 $f:X\to Y=AXB$ に対応する行列は $B^{\mathrm{t}}\otimes A$ に等しい．

(ii)　$m=n$, $p=q$ のとき線形写像 $g:X\to Y=AX+XB$ に対応する行列は $I_p\otimes A+B^{\mathrm{t}}\otimes I_n$ に等しい．

【証明】　(i) $X=(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p)$ ($\boldsymbol{x}_i$ は X の第 i 列), $B=(b_{ij})$, $Y=(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_q)$ ($\boldsymbol{y}_i$ は Y の第 i 列) とすれば

$$Y=AXB=(A\boldsymbol{x}_1,\ldots,A\boldsymbol{x}_p)(b_{ij})=\Big(\sum_{i=1}^{p}b_{i1}A\boldsymbol{x}_i,\ldots,\sum_{i=1}^{p}b_{iq}A\boldsymbol{x}_i\Big)$$

$$\therefore\quad \begin{pmatrix}\boldsymbol{y}_1\\ \vdots\\ \boldsymbol{y}_q\end{pmatrix}=\begin{pmatrix} b_{11}A & \cdots & b_{p1}A \\ \vdots & & \vdots \\ b_{1q}A & \cdots & b_{pq}A\end{pmatrix}\begin{pmatrix}\boldsymbol{x}_1\\ \vdots\\ \boldsymbol{x}_p\end{pmatrix}=(B^{\mathrm{t}}\otimes A)\begin{pmatrix}\boldsymbol{x}_1\\ \vdots\\ \boldsymbol{x}_p\end{pmatrix}$$

(ii) 同様に

$$AX=(A\boldsymbol{x}_1,\ldots,A\boldsymbol{x}_p),\ XB=\Big(\sum_{i=1}^{p}b_{i1}\boldsymbol{x}_i,\ldots,\sum_{i=1}^{p}b_{ip}\boldsymbol{x}_i\Big)$$

より

$$Y=AX+XB=(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_p)$$

とするとき

$$\begin{pmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_p \end{pmatrix} = \left\{ \begin{pmatrix} A & & \\ & \ddots & \\ & & A \end{pmatrix} + \begin{pmatrix} b_{11}I_n & \cdots & b_{p1}I_n \\ \vdots & & \vdots \\ b_{1p}I_n & \cdots & b_{pp}I_n \end{pmatrix} \right\} \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_p \end{pmatrix}$$

$$= (I_p \otimes A + B^{\mathrm{t}} \otimes I_n) \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_p \end{pmatrix}$$

証明終 ∎

定理 10.2 $A, \widetilde{A}, B, \widetilde{B}$ をそれぞれ $m \times n$, $n \times l$, $p \times q$, $q \times r$ 行列とするとき

$$(A \otimes B)(\widetilde{A} \otimes \widetilde{B}) = (A\widetilde{A}) \otimes (B\widetilde{B})$$

【証明】 $A = (a_{ij})$, $\widetilde{A} = (\widetilde{a}_{ij})$, $A\widetilde{A} = (c_{ij})$ とおけば

$$(A \otimes B)(\widetilde{A} \otimes \widetilde{B}) = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix} \begin{pmatrix} \widetilde{a}_{11}\widetilde{B} & \cdots & \widetilde{a}_{1l}\widetilde{B} \\ \vdots & & \vdots \\ \widetilde{a}_{n1}\widetilde{B} & \cdots & \widetilde{a}_{nl}\widetilde{B} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{n} a_{1j}\widetilde{a}_{j1}B\widetilde{B} & \cdots & \sum_{j=1}^{n} a_{1j}\widetilde{a}_{jl}B\widetilde{B} \\ \vdots & & \vdots \\ \sum_{j=1}^{n} a_{mj}\widetilde{a}_{j1}B\widetilde{B} & \cdots & \sum_{j=1}^{n} a_{mj}\widetilde{a}_{jl}B\widetilde{B} \end{pmatrix} = \begin{pmatrix} c_{11}B\widetilde{B} & \cdots & c_{1l}B\widetilde{B} \\ \vdots & & \vdots \\ c_{m1}B\widetilde{B} & \cdots & c_{ml}B\widetilde{B} \end{pmatrix}$$

$$= (A\widetilde{A}) \otimes (B\widetilde{B})$$

証明終 ∎

定理 10.3 A, B をそれぞれ $m \times n$, $p \times q$ 行列とするとき

$$(A \otimes B)^* = A^* \otimes B^*, \quad (A \otimes B)^{\mathrm{t}} = A^{\mathrm{t}} \otimes B^{\mathrm{t}}$$

【証明】

$$(A \otimes B)^* = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}^* = \begin{pmatrix} (a_{11}B)^* & \cdots & (a_{m1}B)^* \\ \vdots & & \vdots \\ (a_{1n}B)^* & \cdots & (a_{mn}B)^* \end{pmatrix}$$

$$= \begin{pmatrix} \overline{a}_{11}B^* & \cdots & \overline{a}_{m1}B^* \\ \vdots & & \vdots \\ \overline{a}_{1n}B^* & \cdots & \overline{a}_{mn}B^* \end{pmatrix} = A^* \otimes B^*$$

証明終 ∎

系 10.3.1 A, B をそれぞれ m 次, n 次の行列とする.

(i) A, B が上 3 角行列ならば $A \otimes B$ も上 3 角行列である.
また A, B が共に下 3 角行列ならば $A \otimes B$ も下 3 角行列である.

特に A, B が対角行列ならば $A \otimes B$ も対角行列である.

(ii) A, B が Hermite 行列ならば $A \otimes B$ も Hermite 行列である.

(iii) A, B がユニタリ行列ならば $A \otimes B$ もユニタリ行列である.

【証明】 明らかである. 証明終 ∎

定理 10.4 A, B を m 次, n 次の正方行列とする. それらの固有値を $\lambda_1, \ldots, \lambda_m$ および $\mu_1, \ldots, \mu_n$ とすれば $A \otimes B$ の固有値は

$$\lambda_i \mu_j \quad (1 \leq i \leq m,\ 1 \leq j \leq n)$$

である.

【証明】 Schur の定理によって適当なユニタリ行列 U, V により

$$U^*AU = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_m \end{pmatrix} = S, \quad V^*BV = \begin{pmatrix} \mu_1 & & * \\ & \ddots & \\ & & \mu_n \end{pmatrix} = T$$

とできる. このとき定理 10.2 と 10.3 によって

$$\begin{aligned} A \otimes B &= (USU^*) \otimes (VTV^*) = (U \otimes V)\big((SU^*) \otimes (TV^*)\big) = (U \otimes V)(S \otimes T)(U^* \otimes V^*) \\ &= (U \otimes V)(S \otimes T)(U \otimes V)^* \\ &= (U \otimes V) \begin{pmatrix} \lambda_1 \begin{pmatrix} \mu_1 & & * \\ & \ddots & \\ & & \mu_n \end{pmatrix} & & * \\ & \ddots & \\ * & & \lambda_m \begin{pmatrix} \mu_1 & & * \\ & \ddots & \\ & & \mu_n \end{pmatrix} \end{pmatrix} (U \otimes V)^* \\ &= (U \otimes V) \begin{pmatrix} \lambda_1\mu_1 & & & & * \\ & \ddots & & & \\ & & \lambda_1\mu_n & & \\ & & & \ddots & \\ * & & & & \lambda_m\mu_n \end{pmatrix} (U \otimes V)^* \end{aligned}$$

系 10.3.1 (iii) によって $U \otimes V$ はユニタリ行列であるから $A \otimes B$ の固有値は $\lambda_i \mu_j$ $(1 \leq i \leq m,\ 1 \leq j \leq n)$ である. 証明終 ∎

系 10.4.1 A, B が正定値 (半正定値) Hermite 行列ならば $A \otimes B$ も正定値 (半正定値) Hermite 行列である.

【証明】 系 10.3.1 (ii) と定理 10.4 による. 証明終 ∎

定理 10.5 A, B を m 次, n 次の行列とし, それらの固有値をそれぞれ $\lambda_1, \ldots, \lambda_m$ および $\mu_1, \ldots, \mu_n$ とすれば $A \otimes I_n + I_m \otimes B$ の固有値は

$$\lambda_i + \mu_j \quad (1 \leq i \leq m,\ 1 \leq j \leq n)$$

である.

【証明】 定理 10.4 の証明に用いたユニタリ行列 U と V をとれば

$$\begin{aligned} A \otimes I_n + I_m \otimes B &= (I_m + A) \otimes (I_n + B) - I_m \otimes I_n - A \otimes B \\ &= \big(U(I_m + S)U^*\big) \otimes \big(V(I_n + T)V^*\big) - I_m \otimes I_n \\ &\qquad - (USU^*) \otimes (VTV^*) \\ &= (U \otimes V)\big[(I_m + S) \otimes (I_n + T) - I_m \otimes I_n - S \otimes T\big](U \otimes V)^* \end{aligned}$$

ただし, 上の式変形に定理 10.2 を用いた. [ ] 内の行列は上 3 角行列でその対角成分は

$$(1 + \lambda_i)(1 + \mu_j) - 1 - \lambda_i \mu_j = \lambda_i + \mu_j, \quad 1 \leq i \leq m,\ 1 \leq j \leq n$$

である. これらが $A \otimes I_n + I_m \otimes B$ の固有値である. 証明終 ∎

■**注意** (10.1) の代わりに $(mp) \times (nq)$ 行列

$$\begin{pmatrix} Ab_{11} & \cdots & Ab_{1q} \\ \vdots & & \vdots \\ Ab_{p1} & \cdots & Ab_{pq} \end{pmatrix}$$

を A と B のテンソル積または **Kronecker** (クロネッカー) 積ということもある.

定義 10.2 (Hadamard (アダマール) 積) n 次正方行列 $A = (a_{ij})$ と $B = (b_{ij})$ に対して $A \circ B = (a_{ij} b_{ij})$ とおき, これを A と B の **Hadamard 積**という.

定理 10.6 A, B が n 次半正定値 Hermite 行列ならば $A \circ B$ も半正定値 Hermite 行列である. さらに A の固有値を $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, B の固有値を $\mu_1 \geq \cdots \geq \mu_n \geq 0$ とすれば $A \circ B$ の任意の固有値 λ は

$$\lambda_1 \mu_1 \geq \lambda \geq \lambda_n \mu_n$$

をみたす.

【証明】 系 10.4.1 によって n 次 Hermite 行列 A, B に対し

$$A \succcurlyeq O,\ B \succcurlyeq O \ \Rightarrow\ A \otimes B \succcurlyeq O$$

$A \circ B$ は n^2 次行列 $A \otimes B$ の n 次主小行列であるからこれもまた半正定値行列である. また定理 10.4 により $A \otimes B$ の固有値は閉区間 $[\lambda_n \mu_n,\ \lambda_1 \mu_1]$ に属するから分離定理 (定理 6.9) によって $A \circ B$ の固有値はその区間内にある. 証明終 ∎

10.2 行列方程式 $X - AXB = C$

テンソル積を用いて行列方程式 $X - AXB = C$ に対する解の存在と一意性を議論することができる.

定理 10.7 (Newman) A,B,C,X を n 次行列とする．$\rho(A)\rho(B)<1$ ならば行列方程式

$$X-AXB=C \tag{10.2}$$

をみたす X はただ 1 つ存在し

$$X=\sum_{k=0}^{\infty}A^kCB^k=C+ACB+A^2CB^2+\cdots \tag{10.3}$$

とあらわされる．

【証明】 $X=(X_1,\ldots,X_n)$, $C=(C_1,\ldots,C_n)$ とする．ここに X_i は X の第 i 列で C_i は C の第 i 列である．すると (10.2) は

$$(I_n\otimes I_n-{}^tB\otimes A)\begin{pmatrix}X_1\\ \vdots\\ X_n\end{pmatrix}=\begin{pmatrix}C_1\\ \vdots\\ C_n\end{pmatrix} \tag{10.4}$$

と書ける．A,B の固有値をそれぞれ $\lambda_1,\ldots,\lambda_n$; $\mu_1,\ldots,\mu_n$ とするとき $I_n\otimes I_n=I_{n^2}$ に注意して, n^2 次行列 $I_n\otimes I_n-{}^tB\otimes A$ の固有値は $1-\mu_i\lambda_j$ $(1\le i,j\le n)$ である．仮定によって $|\mu_i\lambda_j|\le\rho(A)\rho(B)<1$ であるから $I_n\otimes I_n-{}^tB\otimes A$ は正則で (10.4) は一意解をもつ．さらに

$$\begin{pmatrix}X_1\\ \vdots\\ X_n\end{pmatrix}=(I_n\otimes I_n-{}^tB\otimes A)^{-1}\begin{pmatrix}C_1\\ \vdots\\ C_n\end{pmatrix}=\left(I_n\otimes I_n+\sum_{k=1}^{\infty}({}^tB\otimes A)^k\right)\begin{pmatrix}C_1\\ \vdots\\ C_n\end{pmatrix}$$

この無限級数は収束する．

ここで $({}^tB\otimes A)^k=({}^tB)^k\otimes A^k=(B^k)^t\otimes A^k$ であり，$B^k=(b_{ij}^{(k)})$ の第 j 列を $B_j^{(k)}$, すなわち $B^k=(B_1^{(k)},\ldots,B_n^{(k)})$ とおけば

$$({}^tB)^k\otimes A^k\begin{pmatrix}C_1\\ \vdots\\ C_n\end{pmatrix}=\begin{pmatrix}\sum_{j=1}^{n}b_{j1}^{(k)}A^kC_j\\ \vdots\\ \sum_{j=1}^{n}b_{jn}^{(k)}A^kC_j\end{pmatrix}=\begin{pmatrix}A^k\sum_{j=1}^{n}b_{j1}^{(k)}C_j\\ \vdots\\ A^k\sum_{j=1}^{n}b_{jn}^{(k)}C_j\end{pmatrix}=\begin{pmatrix}A^kCB_1^{(k)}\\ \vdots\\ A^kCB_n^{(k)}\end{pmatrix}$$

$$\therefore\quad\begin{pmatrix}X_1\\ \vdots\\ X_n\end{pmatrix}=\left(I_n\otimes I_n+\sum_{k=1}^{\infty}({}^tB\otimes A)^k\right)\begin{pmatrix}C_1\\ \vdots\\ C_n\end{pmatrix}=\begin{pmatrix}C_1\\ \vdots\\ C_n\end{pmatrix}+\begin{pmatrix}\sum_{k=1}^{\infty}A^kCB_1^{(k)}\\ \vdots\\ \sum_{k=1}^{\infty}A^kCB_n^{(k)}\end{pmatrix}$$

$$\therefore\quad X=(X_1,\ldots,X_n)=(C_1,\ldots,C_n)+\left(\sum_{k=1}^{\infty}A^kCB_1^{(k)},\ldots,\sum_{k=1}^{\infty}A^kCB_n^{(k)}\right)$$

$$=\sum_{k=0}^{\infty}A^kCB^k$$

証明終 ∎

系 10.7.1 (Stein (スタイン) の定理) $H - A^*HA$ が正定値 Hermite 行列となるような正定値 Hermite 行列 H が存在するための必要十分条件は $\rho(A) < 1$.

【証明】 (i) 必要条件：H を正定値 Hermite 行列で $C = H - A^*HA \succ O$ とすれば，A の任意の固有値 λ と対応する固有ベクトル $\boldsymbol{x} \neq \boldsymbol{0}$ に対して $A\boldsymbol{x} = \lambda\boldsymbol{x}$ かつ

$$\boldsymbol{x}^*(H - A^*HA)\boldsymbol{x} = \boldsymbol{x}^*C\boldsymbol{x} > 0$$

一方 $\boldsymbol{x}^*A^* = (A\boldsymbol{x})^* = \overline{\lambda}\boldsymbol{x}^*$ であるから

$$\boldsymbol{x}^*C\boldsymbol{x} = (1 - |\lambda|^2)\boldsymbol{x}^*H\boldsymbol{x} \quad \text{かつ} \quad \boldsymbol{x}^*H\boldsymbol{x} > 0 \quad \therefore \quad 1 - |\lambda|^2 > 0$$

λ は A の任意の固有値であったから $\rho(A) < 1$.
(ii) 十分条件：$\rho(A) < 1$ ならば定理 10.7 により方程式 $H - A^*HA = I_n$ は一意解

$$H = I_n + A^*A + (A^*)^2A^2 + \cdots$$

をもつ．$(A^*)^kA^k = (A^k)^*A^k \succcurlyeq O$ であるから

$$H = I_n + \sum_{k=1}^{\infty}(A^*)^kA^k \succcurlyeq I_n \succ O$$

証明終 ∎

10.3 行列方程式 $AX + XB = C$

定理 10.8 (Hahn(ハーン)) A, B, X をn次行列とし A, B の固有値をそれぞれ $\lambda_1, \ldots, \lambda_n$; $\mu_1, \ldots, \mu_n$ とする．任意の n 次行列 C に対し，行列方程式

$$AX + XB = C \tag{10.5}$$

が一意解 X をもつための必要十分条件は

$$\lambda_i + \mu_j \neq 0 \quad \forall i, j \tag{10.6}$$

【証明】 $X = (X_1, \ldots, X_n)$, $C = (C_1, \ldots, C_n)$ (X_i, C_i は X, C の第 i 列) とすれば定理 10.1 (ii) によって (10.5) は

$$(I_n \otimes A + B^{\mathrm{t}} \otimes I_n)\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix}$$

と書ける．定理 10.5 によって係数行列の固有値は $\lambda_i + \mu_j$ $(1 \leq i, j \leq n)$ であるから定理が従う．

証明終 ∎

系 10.8.1 任意の n 次 Hermite 行列 C に対して方程式

$$AX + XA^* = C$$

が Hermite 行列の解 X をただ 1 つもつための必要十分条件は

$$\lambda_i + \overline{\lambda}_j \neq 0 \quad \forall i, j$$

【証明】 定理 10.8 により $\lambda_i+\overline{\lambda}_j\neq 0\ \forall i,j$ のとき，かつそのときに限り方程式 $AX+XA^*=C$ は一意解 X をもつ．両辺の共役転置をとれば

$$X^*A^*+AX^*=C$$

よって X^* も $AX+XA^*=C$ の解であるが，解はただ 1 つであるから $X=X^*$．すなわち一意解 X は必然的に Hermite 行列である． 証明終 ∎

注意 A,B,C を n 次行列とする．n 次行列値関数 $X=X(t)$ に関する 1 階線形常微分方程式の初期値問題

$$\frac{dX}{dt}=AX+XB,\quad X(0)=C$$

の解は (10.6) とは無関係に $X=e^{tA}Ce^{tB}$ で与えられる．

【証明】 $X_1,\dots,X_n,\ C_1,\dots,C_n$ を定理 10.8 の証明のように定義して

$$\boldsymbol{x}=\begin{pmatrix}X_1\\ \vdots\\ X_n\end{pmatrix},\quad \boldsymbol{c}=\begin{pmatrix}C_1\\ \vdots\\ C_n\end{pmatrix}\quad (n^2\text{次元ベクトル})$$

とおけば，与えられた初期値問題は次の n^2 次元連立初期値問題となる．

$$\frac{d\boldsymbol{x}}{dt}=(I_n\otimes A+B^{\mathrm{t}}\otimes I_n)\boldsymbol{x},\ \boldsymbol{x}(0)=\boldsymbol{c}$$

この初期値問題は一意解をもつ (§4.5 例 2 参照) から解 X は一意に定まる．さらに $X=e^{tA}Ce^{tB}$ とおけば $X(0)=C$ かつ

$$\frac{dX}{dt}=Ae^{tB}Ce^{tB}+e^{tA}Ce^{tB}B=AX+XB$$

である． 証明終 ∎

定理 10.9 n 次行列 A の固有値の実部がすべて負のとき A は**安定な行列**であるという．A,B を安定な n 次行列とすれば

$$X=-\int_0^\infty e^{tA}Ce^{tB}dt \tag{10.7}$$

は (10.5) のただ 1 つの解である．

【証明】 A,B が安定な行列ならば定理 10.8 の条件 (10.6) がみたされて (10.5) は一意解をもつ．いま $Y=e^{tA}Ce^{tB}$ とおくと

$$\frac{d}{dt}e^{tA}=Ae^{tA},\quad \frac{d}{dt}e^{tB}=Be^{tB}$$

に注意して

$$\frac{dY}{dt}=Ae^{tA}Ce^{tB}+e^{tA}C(Be^{tB})=AY+YB\quad (\because\ Be^{tB}=e^{tB}B)$$

$$\therefore\quad \int_0^\infty (AY+YB)dt=[Y]_0^\infty=-C$$

$(\because\ \mathrm{Re}\lambda(A)<0,\ \mathrm{Re}\lambda(B)<0$ であるから $e^{tA}Ce^{tB}\to 0\ (t\to\infty))$

$$\therefore\quad C=-\int_0^\infty (AY+YB)dt=A\Big(-\int_0^\infty Ydt\Big)+\Big(-\int_0^\infty Ydt\Big)B$$

よって $X=-\int_0^\infty Ydt$ は (10.5) の一意解である. 証明終 ∎

■**注意** 上記証明において，実は (10.7) の右辺の積分の存在をまず確認しておく必要がある．これは A の Jordan 分解を用いて示される．(後述の定理 10.10 の証明を参照されたい．)

系 10.9.1 A,B を n 次正定値 Hermite 行列とすれば (10.5) の解は

$$X=\int_0^\infty e^{-tA}Ce^{-tB}dt$$

と書ける.

【証明】 (10.5) を $(-A)(-X)+(-X)(-B)=C$ と書けば $-A,-B$ は安定な行列であるから，定理 10.9 により

$$-X=-\int_0^\infty e^{-tA}Ce^{-tB}dt\quad\therefore\quad X=\int_0^\infty e^{-tA}Ce^{-tB}dt$$ 証明終 ∎

定理 10.10 (Liapunov (リアプノフ)) A,C,X は n 次行列で C は正定値 Hermite 行列であるとする．このとき方程式

$$AX+XA^*=-C \tag{10.8}$$

が正定値 Hermite 行列の解 X をもつための必要十分条件は

$$\mathrm{Re}\lambda(A)<0\quad(\text{すなわち } A \text{ が安定})$$

ただし $\lambda(A)$ は A の任意の固有値をあらわす.

【証明】 (i) 必要条件：A の任意の固有値 $\lambda=\lambda(A)$ をとる．$\overline{\lambda}$ は A^* の固有値であるから $A^*\boldsymbol{y}=\overline{\lambda}\boldsymbol{y}$ なる $\boldsymbol{y}\neq\boldsymbol{0}$ がある．両辺の共役転置をとれば $\boldsymbol{y}^*A=\lambda\boldsymbol{y}^*$

$$\therefore\quad \boldsymbol{y}^*(AX+XA^*)\boldsymbol{y}=\lambda\boldsymbol{y}^*X\boldsymbol{y}+\boldsymbol{y}^*X(\overline{\lambda}\boldsymbol{y})=(\lambda+\overline{\lambda})\boldsymbol{y}^*X\boldsymbol{y} \tag{10.9}$$

一方 C は正定値 Hermite 行列であるから

$$\boldsymbol{y}^*(AX+XA^*)\boldsymbol{y}=\boldsymbol{y}^*(-C)\boldsymbol{y}=-\boldsymbol{y}^*C\boldsymbol{y}<\boldsymbol{0} \tag{10.10}$$

また X は正定値 Hermite 行列であるから

$$\boldsymbol{y}^*X\boldsymbol{y}>0 \tag{10.11}$$

(10.9)–(10.11) より $\lambda+\overline{\lambda}<0\quad\therefore\ \mathrm{Re}\lambda<0$

(ii) 十分条件：A が安定ならば A の固有値 λ_i は $\lambda_i+\overline{\lambda}_j\neq 0\ \forall i,j$ をみたすから系 10.8.1 により (10.8) は一意解 X をもち X は Hermite 行列である．定理 10.9 によってこの X は

$$X=\int_0^\infty e^{tA}Ce^{tA^*}dt \tag{10.12}$$

とあらわされるが，ここで右辺の積分が存在することを示そう．§4.5 例 1 により

$$e^{tA} = V \begin{pmatrix} F_1(t) & & \\ & \ddots & \\ & & F_m(t) \end{pmatrix} V^{-1} \quad (V\ :\text{正則行列})$$

$$F_i(t) = e^{\lambda_i t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{n_i-1}}{(n_i-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2!} \\ & & & \ddots & t \\ & & & & 1 \end{pmatrix}$$

と書けるから (10.12) の被積分関数 (n 次行列) $Y = e^{tA}Ce^{tA^*} = (y_{ij}(t))$ の各要素は

$$(t\text{ の多項式}) \times e^{\lambda_i t}, \quad (t\text{ の多項式}) \times e^{\overline{\lambda}_j t}$$

の形の項の積和である．そして，A が安定ならば $\mathrm{Re}\lambda_i \leq -\sigma\ \forall i$ なる正数 σ が存在する．任意の非負整数 k につき $\displaystyle\int_0^\infty t^k e^{-\sigma t}dt$ は存在するから $\displaystyle\int_0^\infty y_{ij}(t)dt$ は存在し，結局積分

$$X = \int_0^\infty e^{tA}Ce^{tA^*}dt = \Big(\int_0^\infty y_{ij}(t)dt\Big) \tag{10.13}$$

は存在する．

最後に Hermite 行列 X の正定値性を示す．C が正定値ならば Sylvester の慣性則により $Y = e^{tA}Ce^{tA^*} = \big(y_{ij}(t)\big)$ も正定値であるから任意の $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} \neq \boldsymbol{0}$ に対して

$$\begin{aligned} \boldsymbol{x}^*\Big(\int_0^\infty Ydt\Big)\boldsymbol{x} &= \boldsymbol{x}^*\Big(\int_0^\infty y_{ij}(t)dt\Big)\boldsymbol{x} \\ &\qquad \Big(\because \int_0^\infty Ydt = \Big(\int_0^\infty y_{ij}(t)dt\Big)\Big) \\ &= \sum_{i=1}^n \overline{x}_i \sum_{j=1}^n \Big(\int_0^\infty y_{ij}(t)dt\Big)x_j = \sum_{i,j=1}^n \int_0^\infty \overline{x}_i y_{ij}(t)x_j dt \\ &= \int_0^\infty \boldsymbol{x}^*Y\boldsymbol{x}dt > 0 \end{aligned}$$

よって $\displaystyle X = \int_0^\infty Ydt$ は正定値である． 証明終 ∎

10.4 差分近似行列

長方形領域 $\Omega = (a,\ b) \times (c,\ d) \subset \mathbb{R}^2$ における境界値問題

$$-\Delta u = f(x,\ y), \quad (x,\ y) \in \Omega \tag{10.14}$$

$$u = g(x,\ y), \quad (x,\ y) \in \partial\Omega\ (\Omega\text{の境界}) \tag{10.15}$$

を有限差分法を用いて数値的に解く．

$$h = \frac{b-a}{m+1}, \quad k = \frac{d-c}{n+1}$$

$$x_i = a + ih,\ (i = 0, 1, 2, \ldots, m+1), \qquad y_j = c + jk,\ (j = 0, 1, 2, \ldots, n+1)$$

$$u_{ij} = u(x_i,\ y_j), \quad f_{ij} = f(x_i,\ y_j), \quad g_{ij} = g(x_i,\ y_j)$$

とするとき $u \in C^4(\Omega)$ ならば

$$\frac{\partial^2 u(x,y)}{\partial x^2} = \frac{1}{h^2}\{u(x+h,\ y) - 2u(x,\ y) + u(x-h,\ y)\} + \frac{\partial^4 u}{\partial x^4}(x+\theta_x h,\ y) \quad (0<\theta_x<1)$$

$$\frac{\partial^2 u(x,y)}{\partial y^2} = \frac{1}{k^2}\{u(x,\ y+k) - 2u(x,\ y) + u(x,\ y-k)\} + \frac{\partial^4 u}{\partial y^4}(x,\ y+\theta_y k) \quad (0<\theta_y<1)$$

が成り立つから，$u_{ij} = u(x_i,\ y_j)$ の近似値 U_{ij} を次の mn 元連立 1 次方程式の解として定める．

$$-\frac{1}{h^2}\{U_{i+1\,j} - 2U_{ij} + U_{i-1\,j}\} - \frac{1}{k^2}\{U_{i\,j+1} - 2U_{ij} + U_{i\,j-1}\} = f_{ij} \quad 1 \le i \le m,\ 1 \le j \le n$$

ただし上式において $(x_i,\ y_j) \in \partial\Omega$ のときは $U_{ij} = g_{ij}$ とおく．すなわち

$$U_{i0} = g_{i0}, \quad U_{i\,n+1} = g_{i\,n+1}, \quad i = 0,1,2,\ldots,m+1$$
$$U_{0j} = g_{0j}, \quad U_{m+1\,j} = g_{m+1\,j}, \quad j = 0,1,2,\ldots,n+1$$

ここで

$$\boldsymbol{U} = (U_{11},\ U_{21},\ldots,\ U_{m1},\ U_{12},\ U_{22},\ldots,\ U_{m2},\ldots,\ U_{1n},\ U_{2n},\ldots,\ U_{mn})^{\mathrm{t}}$$

$$\boldsymbol{f} = \left(f_{11} + \frac{g_{01}}{h^2} + \frac{g_{10}}{k^2},\ f_{21} + \frac{g_{20}}{k^2},\ldots,\ f_{m-1\,1} + \frac{g_{m-1\,0}}{k^2},\ f_{mn} + \frac{g_{m+1\,1}}{h^2} + \frac{g_{m0}}{k^2},\right.$$
$$f_{12} + \frac{g_{02}}{h^2},\ f_{22},\ldots,\ f_{m-1\,2},\ f_{m2} + \frac{g_{m+1\,2}}{h^2} + \frac{g_{m0}}{k^2},\ldots,\ f_{1n} + \frac{g_{0n}}{h^2} + \frac{g_{1\,n+1}}{k^2},$$
$$\left. f_{2n} + \frac{g_{2\,n+1}}{k^2},\ldots,\ f_{m-1\,n} + \frac{g_{m-1\,n+1}}{k^2},\ f_{mn} + \frac{g_{m+1\,n}}{h^2} + \frac{g_{m\,n+1}}{k^2}\right)^{\mathrm{t}}$$

$$A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix} \quad (m\text{ 次})$$

$$B = \frac{1}{k^2}\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix} \quad (n\text{ 次})$$

とおけば上記連立 1 次方程式は

$$(I_n \otimes A + B \otimes I_m)\boldsymbol{U} = \boldsymbol{f} \tag{10.16}$$

と書ける (各自検証されたい)．これを (10.14), (10.15) に対する**有限差分方程式**という．

A の固有値は $\lambda_i = \frac{4}{h^2}\sin^2\frac{i\pi}{2(m+1)}$ $(1 \le i \le m)$, B の固有値は $\mu_j = \frac{4}{k^2}\sin^2\frac{j\pi}{2(n+1)}$ $(1 \le j \le n)$ である ([6] 参照) から，定理 10.5 によって $I_n \otimes A + B \otimes I_m$ の固有値は $\lambda_i + \mu_j > 0$ $\forall i, j$. したがって (10.16) の係数行列は正定値対称行列で解 $\boldsymbol{U}$ はただ 1 通りに定まる.

最も簡単な収束定理として次が知られている.

定理 10.11 $u = u(x, y)$ を (10.14), (10.15) の解とする. $u \in C^4(\Omega)$ ならば h, k に無関係な正定数 C をえらんで次が成り立つ.

$$|u_{ij} - U_{ij}| \le C(h^2 + k^2) \quad \forall\, i, j$$

この定理の簡単な証明は山本 [6] を参照されたい.

【付記】(10.16) において，$n = m = 10^2$ のとき $nm = 10^4$ (1 万)，$n = m = 10^3$ のとき $nm = 10^6$ (100 万) となるから (10.14), (10.15) を実際に離散近似して解く場合，大型連立 1 次方程式を解くことになる. 1940 年代後半に Neumann 型コンピュータが出現して，従来不可能であった計算が可能になり，さらに 100 元を超える連立 1 次方程式が解けるようになると，連立 1 次方程式の効率のよい数値解法を含む行列演算への関心が高まり，現在数値線形代数 (numerical linear algebra) と呼ばれる分野が急成長するに至った.

なお，Varga の名著 [19] には，著者の実体験として，1960 年にはすでにアメリカの原子力研究所において，2 万元の連立 1 次方程式が反復法により解かれていたと記されている. 我が国でこのような大型方程式が解かれ始めたのはさらに後のことである. 以来，人間の欲望には際限がなく，科学技術の発展とコンピュータの性能向上と相俟って現在では天文学的な次元の方程式が解かれている. 現在の環境問題と同様に，このようにしてえられる数値解の信頼性が問題となる時代がやがて到来するかも知れない.

第11章　Grassmann積と複合行列

この章ではまずGrassmann (グラスマン) 積とその基本的性質を述べ, それに基づき複合行列を定義し基本的性質を導く. 次にその応用として次の結果を証明する：n次行列Aの固有値λ_iを$|\lambda_1| \geq \cdots \geq |\lambda_n|$, A^sの特異値を$\alpha_1^{(s)} \geq \cdots \geq \alpha_n^{(s)}$と並べるとき$\lim_{s\to\infty} \alpha_i^{(s)\frac{1}{s}} = |\lambda_i|$, $1 \leq i \leq n$.

11.1　Grassmann積

$1, 2, \ldots, n$からp個をとり出し$1 \leq i_1 < \cdots < i_p \leq n$と並べた組$(i_1, i_2, \ldots, i_p)$の全体

$$Q\binom{n}{p} = \left\{ (i_1, i_2, \ldots, i_p) \;\middle|\; 1 \leq i_1 < i_2 < \cdots < i_p \leq n \right\}$$

$$\left(\text{要素の数は} \binom{n}{p} = \frac{n!}{p!(n-p)!} \text{個ある.} \right)$$

に辞書式順序を導入して

$$\alpha = (i_1, \ldots, i_p), \quad \beta = (j_1, \ldots, j_p) \quad \in Q\binom{n}{p}$$

$$i_1 = j_1, \ldots, i_{m-1} = j_{m-1}, \quad i_m < j_m, \quad 1 \leq m \leq p$$

のとき$\alpha < \beta$と定める.

$$\begin{aligned}
\boldsymbol{x}_j &= (x_{1j}, \ldots, x_{nj})^{\mathrm{t}} \in \mathbb{F}^n \ (\mathbb{R}^n \text{または} \mathbb{C}^n), \quad 1 \leq j \leq p \\
X &= (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p) \quad (\text{列ベクトル } \boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \text{を並べてつくった } n \times p \text{ 行列}) \\
&= (x_{ij}) \\
\omega &= (i_1, \ldots, i_p) \in Q\binom{n}{p} = \{\omega_1, \ldots, \omega_{\binom{n}{p}}\} \\
\Delta_\omega &= \begin{vmatrix} x_{i_1 1} & \cdots & x_{i_1 p} \\ \vdots & & \vdots \\ x_{i_p 1} & \cdots & x_{i_p p} \end{vmatrix} = \det X \begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ 1 & 2 & \cdots & p \end{pmatrix} \quad (p \text{ 次小行列式})
\end{aligned}$$

とおく. このとき写像$f : \overbrace{\mathbb{F}^n \times \cdots \times \mathbb{F}^n}^{p} \to \mathbb{F}^{\binom{n}{p}}$を

$$f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p) = (\Delta_{\omega_1}, \ldots, \Delta_{w_{\binom{n}{p}}})^{\mathrm{t}}$$

により定義する.

fは次の性質をもつ (行列式の性質から明らかである).

(i)　fは各成分につき線形

(ii)　$f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p) = \mathrm{sgn}(\sigma) \cdot f(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(p)})$

$$\text{ただし} \quad \sigma = \begin{pmatrix} 1 & 2 & \cdots & p \\ \sigma(1) & \sigma(2) & \cdots & \sigma(p) \end{pmatrix} \quad (1, 2, \ldots, p \text{ の置換})$$

定義 11.1 (Grassmann 積) $f(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p)$ を $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p$ の **Grassmann 積**といい $\boldsymbol{x}_1\wedge\cdots\wedge\boldsymbol{x}_p$ であらわす．すなわち

$$\boldsymbol{x}_1\wedge\cdots\wedge\boldsymbol{x}_p=(\Delta_{\omega_1},\ldots,\Delta_{\omega_{\binom{n}{p}}})^{\mathrm{t}}\ \in\mathbb{F}^{\binom{n}{p}}$$

例 1 $n=4,\ p=3$ のとき $\binom{n}{p}=4$

$$\boldsymbol{x}_1=\begin{pmatrix}x_{11}\\x_{21}\\x_{31}\\x_{41}\end{pmatrix},\quad \boldsymbol{x}_2=\begin{pmatrix}x_{12}\\x_{22}\\x_{32}\\x_{42}\end{pmatrix},\quad \boldsymbol{x}_3=\begin{pmatrix}x_{13}\\x_{23}\\x_{33}\\x_{43}\end{pmatrix},\quad X=(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)$$

$$Q\binom{n}{p}=\{\omega_1=(1,2,3),\ w_2=(1,2,4),\ \omega_3=(1,3,4),\ \omega_4=(2,3,4)\}$$

$$\Delta_{\omega_1}=\begin{vmatrix}x_{11}&x_{12}&x_{13}\\x_{21}&x_{22}&x_{23}\\x_{31}&x_{32}&x_{33}\end{vmatrix},\quad \Delta_{\omega_2}=\begin{vmatrix}x_{11}&x_{12}&x_{13}\\x_{21}&x_{22}&x_{23}\\x_{41}&x_{42}&x_{43}\end{vmatrix},$$

$$\Delta_{\omega_3}=\begin{vmatrix}x_{11}&x_{12}&x_{13}\\x_{31}&x_{32}&x_{33}\\x_{41}&x_{42}&x_{43}\end{vmatrix},\quad \Delta_{\omega_4}=\begin{vmatrix}x_{21}&x_{22}&x_{23}\\x_{31}&x_{32}&x_{33}\\x_{41}&x_{42}&x_{43}\end{vmatrix}$$

$$\therefore\quad \boldsymbol{x}_1\wedge\boldsymbol{x}_2\wedge\boldsymbol{x}_3=(\Delta_{\omega_1},\Delta_{\omega_2},\Delta_{\omega_3},\Delta_{\omega_4})^{\mathrm{t}}\ \in\mathbb{F}^4$$

定理 11.1 $\boldsymbol{x}_1\wedge\cdots\wedge\boldsymbol{x}_p=\boldsymbol{0}\ \Leftrightarrow\ \boldsymbol{x}_1,\ldots,\boldsymbol{x}_p$ ：1 次従属

【証明】 $\boldsymbol{x}_1\wedge\cdots\wedge\boldsymbol{x}_p=\boldsymbol{0}\ \Leftrightarrow\ (\Delta_{\omega_1},\ldots,\Delta_{\omega\binom{n}{p}})^{\mathrm{t}}=\boldsymbol{0}$

$$\Leftrightarrow\ \begin{vmatrix}x_{i_11}&\cdots&x_{i_1p}\\\vdots&&\vdots\\x_{i_p1}&\cdots&x_{i_pp}\end{vmatrix}=0\quad\forall\omega=(i_1,\ldots,i_p)\in Q\binom{n}{p}$$

$$\Leftrightarrow\ \mathrm{rank}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p)<p$$

$$\Leftrightarrow\ \boldsymbol{x}_1,\ldots,\boldsymbol{x}_p\ :\ 1\text{ 次従属}$$

証明終 ∎

定理 11.2 $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_p\in\mathbb{F}^n,\ \boldsymbol{y}_i=\sum_{j=1}^n y_{ji}\boldsymbol{x}_j\quad(1\le i\le p)$ ならば

$$\boldsymbol{y}_1\wedge\cdots\wedge\boldsymbol{y}_p=\sum_{\omega=(j_1,\ldots,j_p)\in Q\binom{n}{p}}\det\begin{pmatrix}y_{j_11}&\cdots&y_{j_1p}\\\vdots&&\vdots\\y_{j_p1}&\cdots&y_{j_pp}\end{pmatrix}\boldsymbol{x}_{j_1}\wedge\cdots\wedge\boldsymbol{x}_{j_p}$$

【証明】

$$\boldsymbol{y}_1\wedge\cdots\wedge\boldsymbol{y}_p=\Big(\sum_{j=1}^n y_{j1}\boldsymbol{x}_j\Big)\wedge\cdots\wedge\Big(\sum_{j=1}^n y_{jp}\boldsymbol{x}_j\Big)$$

$$=\sum_{j_1}y_{j_11}\Big(\boldsymbol{x}_{j_1}\wedge\Big(\sum_{j=1}^n y_{j2}\boldsymbol{x}_j\Big)\wedge\cdots\wedge\Big(\sum_{j=1}^n y_{jp}\boldsymbol{x}_j\Big)\Big)$$

$$= \cdots$$

$$= \sum_{j_1, j_2, \ldots, j_p} y_{j_1 1} y_{j_2 2} \cdots y_{j_p p}\, \boldsymbol{x}_{j_1} \wedge \cdots \wedge \boldsymbol{x}_{j_p} \tag{11.1}$$

ここで，ある s と t $(s \neq t)$ につき $j_s = j_t$ ならば定理 11.1 によって $\boldsymbol{x}_{j_1} \wedge \cdots \wedge \boldsymbol{x}_{j_p} = \boldsymbol{0}$ となるから (11.1) の $j_1, \ldots, j_p$ は相異なるもののみ残る．

$$\therefore \quad \boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p = \sum_{\omega=(j_1,\ldots,j_p)\in Q\binom{n}{p}} \left(\sum_{\sigma} \operatorname{sgn}(\sigma) y_{\sigma(1)1} \cdots y_{\sigma(p)p} \right) \boldsymbol{x}_{j_1} \wedge \cdots \wedge \boldsymbol{x}_{j_p}$$

$$= \sum_{\omega=(j_1,\ldots,j_p)\in Q\binom{n}{p}} \det \begin{pmatrix} y_{j_1 1} & \cdots & y_{j_1 p} \\ \vdots & & \vdots \\ y_{j_p 1} & \cdots & y_{j_p p} \end{pmatrix} \boldsymbol{x}_{j_1} \wedge \cdots \wedge \boldsymbol{x}_{j_p}$$

証明終 ∎

定理 11.3 $(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p,\ \boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p) = \det\{(\boldsymbol{x}_i,\ \boldsymbol{y}_j)\} \quad 1 \leq i, j \leq p$

【証明】 $\boldsymbol{x}_i = (x_{1i}, \ldots, x_{ni})^{\mathrm{t}}$, $\boldsymbol{y}_j = (y_{1j}, \ldots, y_{nj})^{\mathrm{t}}$, $i, j = 1, 2, \ldots, p$ とすれば $(\boldsymbol{x}_i,\ \boldsymbol{y}_j) = \boldsymbol{y}_j^* \boldsymbol{x}_i$ であるから

$$\det\{(\boldsymbol{x}_i,\ \boldsymbol{y}_j)\} = \det\{(\boldsymbol{x}_j,\ \boldsymbol{y}_i)\} \quad (\because\ \det A = \det A^{\mathrm{t}})$$

$$= \det \left[\begin{pmatrix} \boldsymbol{y}_1^* \\ \vdots \\ \boldsymbol{y}_p^* \end{pmatrix} (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p) \right] = \det \left[\begin{pmatrix} \overline{y}_{11} & \cdots & \overline{y}_{n1} \\ \vdots & & \vdots \\ \overline{y}_{1p} & \cdots & \overline{y}_{np} \end{pmatrix} \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix} \right]$$

Binet-Cauchy の定理によって

$$= \sum_{\omega=(j_1,\ldots,j_p)\in Q\binom{n}{p}} \det \begin{pmatrix} \overline{y}_{j_1 1} & \cdots & \overline{y}_{j_p 1} \\ \vdots & & \vdots \\ \overline{y}_{j_1 p} & \cdots & \overline{y}_{j_p p} \end{pmatrix} \det \begin{pmatrix} x_{j_1 1} & \cdots & x_{j_1 p} \\ \vdots & & \vdots \\ x_{j_p 1} & \cdots & x_{j_p p} \end{pmatrix} \tag{11.2}$$

Grassmann 積の定義によって

$$\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p = (\Delta_{\omega_1}, \ldots, \Delta_{\omega_{\binom{n}{p}}})^{\mathrm{t}}$$

$$\boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p = (\widetilde{\Delta}_{\omega_1}, \ldots, \widetilde{\Delta}_{\omega_{\binom{n}{p}}})^{\mathrm{t}}$$

$$\left(\begin{array}{l} \text{ただし} \quad \omega = (j_1, \ldots, j_p) \text{ のとき} \\ \\ \Delta_\omega = \det \begin{pmatrix} x_{j_1 1} & \cdots & x_{j_1 p} \\ \vdots & & \vdots \\ x_{j_p 1} & \cdots & x_{j_p p} \end{pmatrix}, \quad \widetilde{\Delta}_\omega = \det \begin{pmatrix} y_{j_1 1} & \cdots & y_{j_1 p} \\ \vdots & & \vdots \\ y_{j_p 1} & \cdots & y_{j_p p} \end{pmatrix} \end{array} \right)$$

であるから，(11.2) は 2 つの $\binom{n}{p}$ 次元ベクトル $\boldsymbol{x} \wedge \cdots \wedge \boldsymbol{x}_p$ と $\boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p$ の複素内積

$$\overline{\widetilde{\Delta}}_{\omega_1} \cdot \Delta_{\omega_1} + \cdots + \overline{\widetilde{\Delta}}_{\omega_{\binom{n}{p}}} \cdot \Delta_{\omega_{\binom{n}{p}}}$$

に等しい．

$$\therefore \quad \det\{(\boldsymbol{x}_i,\ \boldsymbol{y}_j)\} = (\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p,\ \boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p)$$

証明終 ∎

定理 11.4 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n$ が n 次元線形空間 $\mathbb{F}^n$ における正規直交系で $1 \le p \le n$ ならば $\widehat{\boldsymbol{x}}_\omega = \boldsymbol{x}_{j_1} \wedge \cdots \wedge \boldsymbol{x}_{j_p}$ $(\omega = (j_1,\dots,j_p) \in Q(\binom{n}{p}))$ は $\binom{n}{p}$ 次元空間 $\mathbb{F}^{\binom{n}{p}}$ の正規直交系をなす.

【証明】 殆ど明らかであるが念のため証明を記す.
$\omega = (i_1,\dots,i_p)$, $\omega' = (j_1,\dots,j_p) \in Q(\binom{n}{p})$ に対し

$$\begin{aligned}(\widehat{\boldsymbol{x}}_\omega,\ \widehat{\boldsymbol{x}}_{\omega'}) &= (\boldsymbol{x}_{i_1} \wedge \cdots \wedge \boldsymbol{x}_{i_p},\ \boldsymbol{x}_{j_1} \wedge \cdots \wedge \boldsymbol{x}_{j_p}) \\ &= \det\bigl\{(\boldsymbol{x}_{i_k},\ \boldsymbol{x}_{j_l})\bigr\} \quad (1 \le k,l \le p) \quad (\text{定理 11.3 による})\end{aligned}$$

したがって $\omega = \omega'$ のとき

$$(\widehat{\boldsymbol{x}}_\omega,\ \widehat{\boldsymbol{x}}_{\omega'}) = \det\bigl\{(\boldsymbol{x}_{i_k},\ \boldsymbol{x}_{i_l})\bigr\} = \det I_p = 1$$

$\omega \neq \omega'$ のときは一般性を失うことなく $\omega < \omega'$ としてよく, このときある m につき

$$i_1 = j_1,\quad \dots,\quad i_{m-1} = j_{m-1},\quad i_m < j_m$$

$$\therefore\quad \det\bigl\{(\boldsymbol{x}_{i_k},\ \boldsymbol{x}_{j_l})\bigr\} = \det\begin{pmatrix} & & & 0 & & & \\ & I_{m-1} & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ & & & * & & & \end{pmatrix} = 0$$

$$\therefore\quad (\widehat{\boldsymbol{x}}_\omega,\ \widehat{\boldsymbol{x}}_{\omega'}) = 0$$

よって $\{\widehat{\boldsymbol{x}}_\omega\}_{\omega \in Q(\binom{n}{p})}$ は $\mathbb{F}^{\binom{n}{p}}$ の正規直交系をなす. 証明終 ∎

定理 11.5 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n$ が $\mathbb{F}^n$ の基底ならば

$$\widehat{\boldsymbol{x}}_\omega = \boldsymbol{x}_{j_1} \wedge \cdots \wedge \boldsymbol{x}_{j_p},\quad \omega = (j_1,\dots,j_p) \in Q(\binom{n}{p})$$

は $\mathbb{F}^{\binom{n}{p}}$ の基底をなす.

【証明】 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n$ に Gram-Schmidt の直交化法を施して $\mathbb{F}^n$ の正規直交系

$$\boldsymbol{y}_i = \sum_{j=1}^{i} y_{ji}\boldsymbol{x}_j\ (y_{ji} \in \mathbb{F}),\quad i = 1,2,\dots,n$$

をつくる. このとき定理 11.4 により $\binom{n}{p}$ 個のベクトル

$$\widehat{\boldsymbol{y}}_\omega = \boldsymbol{y}_{i_1} \wedge \cdots \wedge \boldsymbol{y}_{i_p},\quad \omega = (i_1,\dots,i_p) \in Q(\binom{n}{p})$$

は $\mathbb{F}^{\binom{n}{p}}$ の正規直交基底である. 定理 11.2 によって

$$\boldsymbol{y}_{i_1} \wedge \cdots \wedge \boldsymbol{y}_{i_p} = \sum_{\omega=(j_1,\dots,j_p) \in Q(\binom{n}{p})} \det\begin{pmatrix} y_{j_1 i_1} & \cdots & y_{j_1 i_p} \\ \cdots & \cdots & \cdots \\ y_{j_p i_1} & \cdots & y_{j_p i_p} \end{pmatrix} \boldsymbol{x}_{j_1} \wedge \cdots \wedge \boldsymbol{x}_{j_p}$$

と書けるから

$$Q\binom{n}{p} = \{\omega_1, \ldots, \omega_m\}, \quad m = \binom{n}{p}$$

に対して

$$Y = (\widehat{\boldsymbol{y}}_{\omega_1}, \ldots, \widehat{\boldsymbol{y}}_{\omega_m}) \quad (\widehat{\boldsymbol{y}}_{\omega_1}, \ldots, \widehat{\boldsymbol{y}}_{\omega_m} \text{を第} 1\sim m \text{列とする} m \times m \text{行列})$$

$$X = (\widehat{\boldsymbol{x}}_{\omega_1}, \ldots, \widehat{\boldsymbol{x}}_{\omega_m}) \quad (\widehat{\boldsymbol{x}}_{\omega_1}, \ldots, \widehat{\boldsymbol{x}}_{\omega_m} \text{を第} 1\sim m \text{列とする} m \times m \text{行列})$$

$$C = (c_{ij}), \quad c_{ij} = \det \begin{pmatrix} y_{j_1 i_1} & \cdots & y_{j_1 i_p} \\ \vdots & & \vdots \\ y_{j_p i_1} & \cdots & y_{j_p i_p} \end{pmatrix}$$

とおけば Y, X, C は $m = \binom{n}{p}$ 次行列で $Y = XC$ と書ける．Y の各列は正規直交系であるから $\det Y = 1$．したがって両辺の行列式をとれば $\det X \neq 0$．よって X の各列 $\widehat{\boldsymbol{x}}_{\omega_1}, \ldots, \widehat{\boldsymbol{x}}_{\omega_m}$ は 1 次独立で $\mathbb{F}^m = \mathbb{F}^{\binom{n}{p}}$ の基底をなす． 証明終 ∎

11.2 複合行列

$\boldsymbol{e}_j$ を n 次単位行列の第 j 列とすれば定理 11.4 によって

$$\boldsymbol{e}_{j_1} \wedge \cdots \wedge \boldsymbol{e}_{j_p}, \quad (j_1, \ldots, j_p) \in Q\binom{n}{p}$$

は $\mathbb{F}^{\binom{n}{p}}$ の正規直交基底である．

定義 11.2 (複合行列) n 次行列 A に対し

$$C_p(A)\boldsymbol{e}_{j_1} \wedge \cdots \wedge \boldsymbol{e}_{j_p} = A\boldsymbol{e}_{j_1} \wedge \cdots \wedge A\boldsymbol{e}_{j_p}, \quad (j_1, \ldots, j_p) \in Q\binom{n}{p}$$

により定義される $\binom{n}{p}$ 次行列 $C_p(A)$ を **p 次複合行列 (p-th compound matrix)** という．

$C_p(A)$ の具体的な形をみるために

$$Q\binom{n}{p} = \left\{(i_1, \ldots, i_p) \;\middle|\; 1 \leq i_1 < \cdots < i_p \leq n\right\} = \{\omega_1, \ldots, \omega_m\}, \quad m = \binom{n}{p}$$

とし，$\omega_i = (i_1, \ldots, i_p)$, $\omega_j = (j_1, \ldots, j_p)$ のとき

$$\Delta_{ij} = \det A \begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix} = \det A\binom{\omega_i}{\omega_j}$$

とおけば Grassmann 積の定義によって

$$A\boldsymbol{e}_{j_1} \wedge \cdots \wedge A\boldsymbol{e}_{j_p} = \begin{pmatrix} \Delta_{1j} \\ \vdots \\ \Delta_{mj} \end{pmatrix}$$

同様に $m = \binom{n}{p}$ 次単位行列の第 j 列を $\widetilde{\boldsymbol{e}}_j$ とすれば容易に確かめられるように

$$\boldsymbol{e}_{j_1} \wedge \cdots \wedge \boldsymbol{e}_{j_p} = \widetilde{\boldsymbol{e}}_j \quad (\text{ただし} \omega_j = (j_1, \ldots, j_p) \in Q\binom{n}{p})$$

よって $C_p(A)\boldsymbol{e}_j \wedge \cdots \wedge \boldsymbol{e}_{j_p} = C_p(A)\widetilde{\boldsymbol{e}}_j$ は $C_p(A)$ の第 j 列をあらわす．

$$\therefore \; C_p(A)\boldsymbol{e}_{j_1} \wedge \cdots \wedge \boldsymbol{e}_{j_p} = \begin{pmatrix} \Delta_{1j} \\ \vdots \\ \Delta_{mj} \end{pmatrix} \quad \forall \omega_j = (j_1, \ldots, j_p) \in Q\binom{n}{p},\ m = \binom{n}{p}$$

$$\therefore \; C_p(A) = \begin{pmatrix} \Delta_{11} & \cdots & \Delta_{1m} \\ \vdots & & \vdots \\ \Delta_{m1} & \cdots & \Delta_{mm} \end{pmatrix}, \quad \Delta_{ij} = \det A\binom{\omega_i}{\omega_j},\ \omega_i, \omega_j \in Q\binom{n}{p}$$

これが $C_p(A)$ の具体的な形である.

例 2 $n=3,\ p=2,\ A=(a_{ij})$ のとき $m=\binom{n}{p}=3$

$$Q\binom{n}{p}=\{\omega_1=(1,\ 2),\ \omega_2=(1,\ 3),\ \omega_3=(2,\ 3)\}$$

$$\therefore\quad C_2(A)=\begin{pmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \end{pmatrix}$$

注意 任意の $j_1,\dots,j_p\ (1\le j_1,\dots,j_p\le n)$ につき

$$C_p(A)\boldsymbol{e}_{j_1}\wedge\cdots\wedge\boldsymbol{e}_{j_p}=A\boldsymbol{e}_{j_1}\wedge\cdots\wedge A\boldsymbol{e}_{j_p}\tag{11.3}$$

が成り立つ ($(j_1,\dots,j_p)\in Q\binom{n}{p}$ でなくてもよい). 実際, ある k と l につき $j_k=j_l$ ならば定理 11.1 により $\boldsymbol{e}_{j_1}\wedge\cdots\wedge\boldsymbol{e}_{j_p}=\boldsymbol{0}$ かつ $A\boldsymbol{e}_{j_1}\wedge\cdots\wedge A\boldsymbol{e}_{j_p}=\boldsymbol{0}$ となって (11.3) が成り立つ. また $j_k\ne j_l\ (k\ne l),\ k,l=1,2,\dots,p$ のときも $\sigma(j_1)<\cdots<\sigma(j_p)$ となる置換 σ をとれば $\big(\sigma(j_1),\dots,\sigma(j_p)\big)\in Q\binom{n}{p}$ かつ

$$\begin{aligned}\boldsymbol{e}_{j_1}\wedge\cdots\wedge\boldsymbol{e}_{j_p}&=(\mathrm{sgn}\sigma)\boldsymbol{e}_{\sigma(j_1)}\wedge\cdots\wedge\boldsymbol{e}_{\sigma(j_p)}\\ A\boldsymbol{e}_{j_1}\wedge\cdots\wedge A\boldsymbol{e}_{j_p}&=(\mathrm{sgn}\sigma)A\boldsymbol{e}_{\sigma(1)}\wedge\cdots\wedge A\boldsymbol{e}_{\sigma(j_p)}\\ C_p(A)\boldsymbol{e}_{j_1}\wedge\cdots\wedge\boldsymbol{e}_{j_p}&=C_p(A)\big[(\mathrm{sgn}\sigma)\boldsymbol{e}_{\sigma(j_1)}\wedge\cdots\wedge\boldsymbol{e}_{\sigma(j_p)}\big]\\ &=(\mathrm{sgn}\sigma)C_p(A)\boldsymbol{e}_{\sigma(j_1)}\wedge\cdots\wedge\boldsymbol{e}_{\sigma(j_p)}\quad\big(\because\ \sigma(j_1),\dots,\sigma(j_p)\in Q\tbinom{n}{p}\big)\\ &=(\mathrm{sgn}\sigma)A\boldsymbol{e}_{\sigma(j_1)}\wedge\cdots\wedge\boldsymbol{e}_{\sigma(j_p)}=A\boldsymbol{e}_{j_1}\wedge\cdots\wedge A\boldsymbol{e}_{j_p}\end{aligned}$$

となって (11.3) が成り立つ.

定理 11.6 任意の $\boldsymbol{y}_1,\dots,\boldsymbol{y}_p\in\mathbb{F}^n$ に対し

$$C_p(A)\boldsymbol{y}_1\wedge\cdots\wedge\boldsymbol{y}_p=A\boldsymbol{y}_1\wedge\cdots\wedge A\boldsymbol{y}_p$$

【証明】 $\boldsymbol{y}_i=\sum_{j=1}^n y_{ji}\boldsymbol{e}_j,\ 1\le i\le p$ とおくと

$$\begin{aligned}A\boldsymbol{y}_1\wedge\cdots\wedge A\boldsymbol{y}_p&=\Big(A\sum_{j=1}^n y_{j1}\boldsymbol{e}_j\Big)\wedge\cdots\wedge\Big(A\sum_{j=1}^n y_{jp}\boldsymbol{e}_j\Big)\\ &=\Big(\sum_{j=1}^n y_{j1}A\boldsymbol{e}_j\Big)\wedge\cdots\wedge\Big(\sum_{j=1}^n y_{jp}A\boldsymbol{e}_j\Big)\\ &=\sum_{j_1,\dots,j_p}(y_{j_1\,1}\cdots y_{j_p\,p})A\boldsymbol{e}_{j_1}\wedge\cdots\wedge A\boldsymbol{e}_{j_p}\\ &=\sum_{j_1,\dots,j_p}(y_{j_1\,1}\cdots y_{j_p\,p})C_p(A)\boldsymbol{e}_{j_1}\wedge\cdots\wedge\boldsymbol{e}_{j_p}\qquad\big((11.3)\text{ による}\big)\end{aligned}$$

$$= C_p(A) \sum_{j_1,\dots,j_p} (y_{j_1 1} \cdots y_{j_p p}) \boldsymbol{e}_{j_1} \wedge \cdots \wedge \boldsymbol{e}_{j_p}$$

$$= C_p(A) \Big(\sum_{j=1}^{n} y_{j1} \boldsymbol{e}_j \Big) \wedge \cdots \wedge \Big(\sum_{j=1}^{n} y_{jp} \boldsymbol{e}_j \Big) = C_p(A) \boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p \qquad \text{証明終} \ \blacksquare$$

定理 11.7 A, B を n 次行列とする．$1 \leq p \leq n$ のとき次が成り立つ．

(i) $C_p(AB) = C_p(A) C_p(B)$

(ii) $C_p(A)^* = C_p(A^*)$

(iii) A が正則ならば $C_p(A)$ も正則で $C_p(A)^{-1} = C_p(A^{-1})$

(iv) A がユニタリ行列ならば $C_p(A)$ もユニタリ行列である．

(v) A の固有値を $\lambda_1, \dots, \lambda_n$ とするとき $C_p(A)$ の固有値は

$$\lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_p}, \ (i_1, i_2, \dots, i_p) \in Q(_p^n)$$

である．

(vi) A が正規行列ならば $C_p(A)$ も正規行列である．

(vii) A が Hermite 行列ならば $C_p(A)$ も Hermite 行列である．

(viii) A が正定値 (半正定値) Hermite 行列ならば $C_p(A)$ も正定値 (半正定値) Hermite 行列である．

【証明】 (i) $\boldsymbol{x}_1, \dots, \boldsymbol{x}_p$ を $\mathbb{F}^n$ の基底とすれば定理 11.5 によって

$$\widehat{\boldsymbol{x}}_\omega = \boldsymbol{x}_{i_1} \wedge \cdots \wedge \boldsymbol{x}_{i_p}, \ \omega = (i_1, \dots, i_p) \in Q(_p^n)$$

は $\mathbb{F}^{\binom{n}{p}}$ の基底であり

$$\begin{aligned} C_p(AB) \boldsymbol{x}_{i_1} \wedge \cdots \wedge \boldsymbol{x}_{i_p} &= \big((AB) \boldsymbol{x}_{i_1} \big) \wedge \cdots \wedge \big((AB) \boldsymbol{x}_{i_p} \big) \\ &= C_p(A) (B \boldsymbol{x}_{i_1}) \wedge \cdots \wedge (B \boldsymbol{x}_{i_p}) \\ &= C_p(A) C_p(B) \boldsymbol{x}_{i_1} \wedge \cdots \wedge \boldsymbol{x}_{i_p} \quad \forall (i_1, \dots, i_p) \in Q(_p^n) \end{aligned}$$

$$\therefore \quad C_p(AB) = C_p(A) C_p(B)$$

(ii)
$$\begin{aligned} &\big(C_p(A) \boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p, \ \boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p \big) \\ &\quad = (A \boldsymbol{x}_1 \wedge \cdots \wedge A \boldsymbol{x}_p, \ \boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p) \\ &\quad = \det \big\{ (A \boldsymbol{x}_i, \ \boldsymbol{y}_j) \big\} \quad 1 \leq i, j \leq p \quad \text{(定理 11.3)} \\ &\quad = \det \big\{ (\boldsymbol{x}_i, \ A^* \boldsymbol{y}_j) \big\} \quad 1 \leq i, j \leq p \\ &\quad = \big(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p, \ (A^* \boldsymbol{y}_1) \wedge \cdots \wedge (A^* \boldsymbol{y}_p) \big) \\ &\quad = \big(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p, \ C_p(A^*) \boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p \big) \\ &\qquad\qquad \forall \boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p, \ \boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_p \in \mathbb{F}^{\binom{n}{p}} \end{aligned}$$

$$\therefore \quad C_p(A)^* = C_p(A^*)$$

(iii) (i) によって $C_p(A) C_p(A^{-1}) = C_p(A \cdot A^{-1}) = C_p(I_n) = I_{\binom{n}{p}}$

$$\therefore \quad C_p(A)^{-1} = C_p(A^{-1})$$

(iv) A がユニタリ行列ならば $AA^* = A^* A = I_n$ で

$$C_p(AA^*) = C_p(A^*A) = C_p(I_n) = I_m, \quad m = \binom{n}{p}$$

(i) と (ii) によって上式は $C_p(A)C_p(A)^* = C_p(A)^*C_p(A) = I_m$ を意味する．よって $C_p(A)$ はユニタリ行列である．

(v) 適当なユニタリ行列により

$$U^*AU = \begin{pmatrix} \lambda_1 & & \\ & \ddots & * \\ & & \lambda_n \end{pmatrix} \quad (= \Lambda \text{とおく})$$

と上 3 角行列に変換される (Schur の定理)．(iv) によって $C_p(U^*) = C_p(U)^{-1}$ であるから上式より

$$C_p(U)^{-1}C_p(A)C_p(U) = C_p(\Lambda)$$

故に $C_p(A)$ の固有値は $C_p(\Lambda)$ の固有値に等しい．Λ は上 3 角行列であるから $C_p(\Lambda)$ も上 3 角行列でその対角要素は

$$\det \Lambda \begin{pmatrix} i_1 & \cdots & i_p \\ i_1 & \cdots & i_p \end{pmatrix} = \lambda_{i_1} \cdots \lambda_{i_p}, \quad (i_1, \ldots, i_p) \in Q(\binom{n}{p})$$

である．これらが $C_p(A)$ の固有値である．

(vi) A が正規ならば $AA^* = A^*A$ であるから

$$C_p(A)C_p(A^*) = C_p(AA^*) = C_p(A^*A) = C_p(A^*)C_p(A)$$

(ii) によって $C_p(A^*) = C_p(A)^*$ であるから上式より $C_p(A)C_p(A)^* = C_p(A)^*C_p(A)$
よって $C_p(A)$ は正規である．

(vii) A が Hermite 行列ならば $A = A^*$ より $C_p(A) = C_p(A^*) = C_p(A)^*$
したがって $C_p(A)$ も Hermite 行列である．

(viii) $\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p \neq \boldsymbol{0}$ ならば定理 11.1 によって $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p$ は 1 次独立で

$$\begin{aligned} &\big(C_p(A)\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p,\ \boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p\big) \\ &\qquad = \big((A\boldsymbol{x}_1) \wedge \cdots \wedge (A\boldsymbol{x}_p),\ \boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p\big) \\ &\qquad = \det\big\{(A\boldsymbol{x}_i,\ \boldsymbol{x}_j)\big\} \quad 1 \leq i, j \leq p \quad (\text{定理 } 11.3) \end{aligned} \tag{11.4}$$

$g_{ij} = (A\boldsymbol{x}_i,\ \boldsymbol{x}_j)$ を (i,j) 要素とする p 次行列 $G = (g_{ij})$ は A が正定値 Hermite 行列のとき正定値 Hermite 行列である．なぜならば $G^* = G$ は明らかであり，$\boldsymbol{u} = (u_1, \ldots, u_p) \neq \boldsymbol{0}$ のとき

$$(G\boldsymbol{u},\ \boldsymbol{u}) = \sum_{i,j=1}^{p} g_{ij}\overline{u}_i u_j = \sum_{i,j=1}^{p} (A\boldsymbol{x}_i,\ \boldsymbol{x}_j)\overline{u}_i u_j = \Big(A\sum_{i=1}^{p} \boldsymbol{x}_i\overline{u}_i,\ \sum_{j=1}^{p} \overline{u}_j\boldsymbol{x}_j\Big)$$

$$= (A\boldsymbol{y},\ \boldsymbol{y}) \quad \Big(\boldsymbol{y} = \sum_{i=1}^{p} \overline{u}_i\boldsymbol{x}_i \text{とおく}\Big)$$

$$> 0 \quad (\because\ \boldsymbol{x}_1, \ldots, \boldsymbol{x}_p \text{は 1 次独立で } \boldsymbol{u} \neq \boldsymbol{0} \text{ より } \boldsymbol{y} \neq \boldsymbol{0} \text{ かつ } A \text{ は正定値})$$

故に G は正定値で $\det G > 0$

したがって (11.4) より，任意の $\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p \neq \boldsymbol{0}$ に対して

$$\big(C_p(A)\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p,\ \boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_p\big) > 0$$

これは $C_p(A)$ が正定値 Hermite 行列であることを意味する．

最後に，A が半正定値行列ならば上記証明において $(A\boldsymbol{y},\ \boldsymbol{y}) > 0$, $\det G > 0$ は $(A\boldsymbol{y},\ \boldsymbol{y}) \geq 0$, $\det G \geq 0$ でおきかえられて $C_p(A)$ は半正定値 Hermite 行列である．　証明終 ∎

11.3 複合行列の応用

A を n 次行列としその固有値 λ_i と特異値 α_i を

$$|\lambda_1| \geq \cdots \geq |\lambda_n|, \quad \alpha_1 \geq \cdots \geq \alpha_n \ (\geq 0)$$

と並べるとき系 4.4.1 (iii) によって $|\lambda_1| \leq \alpha_1$ である．これを複合行列 $C_p(A)$ に適用すると，$C_p(A)$ の絶対値最大な固有値は $\lambda_1\lambda_2\cdots\lambda_p$ で，最大特異値は $\alpha_1\alpha_2\cdots\alpha_p$ であるから

$$|\lambda_1\lambda_2\cdots\lambda_p| \leq \alpha_1\alpha_2\cdots\alpha_p \quad (1 \leq p \leq n)$$

この評価は Weyl[40] に述べられている．Weyl は同じ論文の中で

$$|\lambda_1|^s + \cdots + |\lambda_n|^s \leq \alpha_1^s + \cdots + \alpha_n^s \quad \forall s > 0$$

を示し，これを複合行列 $C_2(A), C_3(A)$ に適用して，任意の正数 s につき，

$$\sum_{i_1<i_2} \left|\lambda_{i_1}^s \lambda_{i_2}^s\right| \leq \sum_{i_1<i_2} \alpha_{i_1}^s \alpha_{i_2}^s$$

$$\sum_{i_1<i_2<i_3} \left|\lambda_{i_1}^s \lambda_{i_2}^s \lambda_{i_3}^s\right| \leq \sum_{i_1<i_2<i_3} \alpha_{i_1}^s \alpha_{i_2}^s \alpha_{i_3}^s$$

を示した．この節では，s を自然数とし，同じアイデアを $\rho(A) = \lim_{s\to\infty} ||A^s||^{\frac{1}{s}}$ に適用して A^s の特異値を $\alpha_i^{(s)}(\alpha_1^{(s)} \geq \cdots \geq \alpha_n^{(s)})$ とするとき

$$\lim_{s\to\infty} {\alpha_i^{(s)}}^{\frac{1}{s}} = |\lambda_i| \quad (1 \leq i \leq n)$$

を示す．

まず次の定理の証明からはじめよう．

定理 11.8 s を自然数とする．与えられたベクトルノルム $||\boldsymbol{x}||$ と両立する任意の行列ノルム $||A||$ に対し

$$\rho(A) \leq ||A^s||^{\frac{1}{s}} \leq ||A|| \quad \text{かつ} \quad \rho(A) = \lim_{s\to\infty} ||A^s||^{\frac{1}{s}}$$

【証明】 これは Banach 代数でよく知られた結果であるが，以下 [43] に従い行列論による初等的証明を与える．

λ を A の任意の固有値として $A\boldsymbol{x} = \lambda\boldsymbol{x}$, $\boldsymbol{x} \neq \boldsymbol{0}$ とすれば $A^s\boldsymbol{x} = \lambda^s\boldsymbol{x}$

よって

$$|\lambda|^s||\boldsymbol{x}|| = ||\lambda^s\boldsymbol{x}|| = ||A^s\boldsymbol{x}|| \leq ||A^s||\ ||\boldsymbol{x}||, \quad ||\boldsymbol{x}|| > 0$$

$$\therefore \quad |\lambda| \leq ||A^s||^{\frac{1}{s}} \leq ||A|| \quad \therefore \quad \rho(A) \leq ||A^s||^{\frac{1}{s}} \leq ||A||$$

$\left\{||A^s||^{\frac{1}{s}}\right\}$, $s = 1, 2, \ldots$ は有界閉区間 $[\rho(A),\ ||A||]$ に属し，少なくとも 1 つの集積点をもつ．集積点の 1 つを α とし $\rho(A) < \alpha$ とすれば

$$1 \leq s_1 < s_2 < \cdots \quad \text{かつ} \quad ||A^{s_i}||^{\frac{1}{s_i}} \to \alpha \quad (i \to \infty)$$

となる部分列 $\left\{||A^{s_i}||^{\frac{1}{s_i}}\right\}$ がある．このとき $\rho(A) < \alpha' < \alpha$ なる α' をとれば

$$\rho\left(\frac{1}{\alpha'}A\right) = \frac{1}{\alpha'}\rho(A) < 1$$

したがって Oldenburger の定理 (系 4.5.2) によって

$$\lim_{i\to\infty}\left(\frac{1}{\alpha'}A\right)^{s_i}=0$$

よって任意に与えられた正数 ε に対し自然数 $N(\varepsilon)$ を定めて

$$\left\|\left(\frac{1}{\alpha'}A\right)^{s_i}\right\|<\varepsilon,\quad \forall i>N(\varepsilon)$$

とできる．すると

$$1<\frac{\alpha}{\alpha'}=\lim_{i\to\infty}\left\|\left(\frac{1}{\alpha'}A\right)^{s_i}\right\|^{\frac{1}{s_i}}\le\lim_{i\to\infty}\varepsilon^{\frac{1}{s_i}}=1$$

これは不合理であるから $\rho(A)=\alpha$ でなければならない．これは $s\to\infty$ のとき $||A^s||^{\frac{1}{s}}$ が収束して

$$\lim_{s\to\infty}||A^s||^{\frac{1}{s}}=\rho(A)$$

であることを意味する． 証明終 ∎

系 11.8.1 $1<s_1<s_2<\cdots$ かつ s_{i+1} が s_i で割り切れるならば $\left\{||A^{s_i}||^{\frac{1}{s_i}}\right\}$ は単調減少で $\rho(A)$ に収束する．

【証明】 $s_{i+1}=m_is_i$ (m_i は自然数) とすれば

$$||A^{s_{i+1}}||^{\frac{1}{s_{i+1}}}=||A^{m_is_i}||^{\frac{1}{m_is_i}}\le||A^{s_i}||^{\frac{1}{s_i}},\quad i=1,2,\ldots$$

である． 証明終 ∎

さて次の定理を示すことが本節の目的であった．

定理 11.9 ([43]) n 次行列 A の固有値 λ_i と A^s の特異値 $\alpha_i^{(s)}$ を

$$|\lambda_1|\ge\cdots\ge|\lambda_n|,\quad \alpha_1^{(s)}\ge\cdots\ge\alpha_n^{(s)}$$

と並べるとき

$$\lim_{s\to\infty}\alpha_i^{(s)\frac{1}{s}}=|\lambda_i|,\quad i=1,2,\ldots,n$$

【証明】 ベクトルノルムとして Euclid ノルム $||\cdot||_2$ をとれば，このベクトルノルムに従属する行列ノルムはスペクトルノルム $||A||_2=\sqrt{\rho(A^*A)}$ であるから

$$||A^s||_2=\sqrt{\rho\big((A^s)^*A^s\big)}=\alpha_1^{(s)}$$

故に定理 11.8 によって

$$\lim_{s\to\infty}\alpha_1^{(s)\frac{1}{s}}=|\lambda_1|$$

これを複合行列 $C_p(A^s)$ に適用すれば

$$\lim_{s\to\infty}\prod_{i=1}^{p}\alpha_i^{(s)\frac{1}{s}}=\prod_{i=1}^{p}|\lambda_i|$$

ここで $\lambda_1=0$ ならば定理は自明であるから

$$|\lambda_1| \geq \cdots \geq |\lambda_k| > 0 = |\lambda_{k+1}| = \cdots = |\lambda_n|$$

と仮定する．このとき $\alpha_k^{(s)} > 0\ \forall s$ である．なぜならば仮にある s に対して $\alpha_k^{(s)} = 0$ とすれば定理 1.18 によって

$$\mathrm{rank}(A^{s+1})^* A^{s+1} = \mathrm{rank} A^{s+1} \leq \mathrm{rank} A^s = \mathrm{rank}(A^s)^* A^s$$

故に $m \geq s$ のとき $\alpha_k^{(m)} = 0$ したがって $\prod_{i=1}^k \alpha_i^{(m)\frac{1}{m}} = 0$

しかし，再び $C_k(A^m)$ に前定理を適用すれば

$$\lim_{m\to\infty} \prod_{i=1}^k \alpha_i^{(m)\frac{1}{m}} = \prod_{i=1}^k |\lambda_i| > 0$$

であるから矛盾を生じる．よって任意の自然数 s に対して $\alpha_i^{(s)} > 0\ (1 \leq i \leq k)$ であり，$j \leq k+1$ のとき

$$\lim_{s\to\infty} \alpha_j^{(s)\frac{1}{s}} = \lim_{s\to\infty} \frac{\left(\prod_{i=1}^j \alpha_i^{(s)}\right)^{\frac{1}{s}}}{\left(\prod_{i=1}^{j-1} \alpha_i^{(s)}\right)^{\frac{1}{s}}} = \frac{\prod_{i=1}^j |\lambda_i|}{\prod_{i=1}^{j-1} |\lambda_i|} = |\lambda_j|$$

すなわち

$$\lim_{s\to\infty} \alpha_j^{(s)\frac{1}{s}} = |\lambda_j| \quad (1 \leq j \leq k+1) \tag{11.5}$$

一方 $j > k+1$ ならば

$$\alpha_{k+1}^{(s)\frac{1}{s}} \geq \alpha_j^{(s)\frac{1}{s}} \geq 0 \tag{11.6}$$

また (11.5) で $j = k+1$ とおくと $\lim_{s\to\infty} \alpha_{k+1}^{(s)\frac{1}{s}} = |\lambda_{k+1}| = 0$

よって (11.5) と (11.6) より

$$\lim_{s\to\infty} \alpha_j^{(s)\frac{1}{s}} = 0 = |\lambda_j| \quad (j > k+1)$$

結局

$$\lim_{s\to\infty} \alpha_j^{(s)\frac{1}{s}} = |\lambda_j| \quad (1 \leq j \leq n)$$

を得る．　　証明終 ∎

【付記】定理 11.9 の別証明が [10], [27], [30], [42] にある．また Hilbert 空間におけるコンパクト作用素への拡張については [22], [29] を参照されたい．なお，最近半単純 Lie 群への翻訳が Tam-Huang [38] によりなされている．

第12章　線形反復法

大型の連立1次方程式に対する高速解法の開発は数値線形代数の主要な部分を占め，現在でもいろいろな解法が提案されている．この最終章ではそれらの解法の基礎となる基本反復解法としてJacobi法，緩和Jacobi法，Gauss-Seildel法，SOR法，CG法をとりあげ，それらの収束証明を与える．最後に与えられた反復停止条件で反復を打ち切ったときの数値解の誤差の分布に関するKrasnosel'skii-Kreinの定理を紹介する．

12.1　Jacobi法と緩和Jacobi法

$A=(a_{ij})$ を $a_{ii}\neq 0\ \forall i$ なる n 次正則行列とし，n 元連立1次方程式

$$A\boldsymbol{x}=\boldsymbol{b} \tag{12.1}$$

を考える．A を $A=M-N$ (M：正則行列) と分解して (12.1) を $M\boldsymbol{x}=N\boldsymbol{x}+\boldsymbol{b}$ と書き直し，$\boldsymbol{x}^{(0)}$ を初期値とする反復

$$M\boldsymbol{x}^{(k+1)}=N\boldsymbol{x}^{(k)}+\boldsymbol{b},\quad k=0,1,2,\ldots \tag{12.2}$$

あるいは

$$\boldsymbol{x}^{(k+1)}=H\boldsymbol{x}^{(k)}+\boldsymbol{c},\quad H=M^{-1}N,\quad \boldsymbol{c}=M^{-1}\boldsymbol{b} \tag{12.2$'$}$$

をつくれば，$\{\boldsymbol{x}^{(k)}\}$ が任意の $\boldsymbol{x}^{(0)}$ に対して (12.1) の一意解 $\widehat{\boldsymbol{x}}$ に収束するための必要十分条件は

$$\rho(H)=\rho(M^{-1}N)<1 \tag{12.3}$$

である．実際，(12.2) あるいは (12.2$'$) と $M\widehat{\boldsymbol{x}}^{(k)}=N\widehat{\boldsymbol{x}}+\boldsymbol{b}$ より

$$\boldsymbol{x}^{(k)}-\widehat{\boldsymbol{x}}=H(\boldsymbol{x}^{(k-1)}-\widehat{\boldsymbol{x}})=H^k(\boldsymbol{x}^{(0)}-\widehat{\boldsymbol{x}}) \tag{12.4}$$

よって $\rho(H)<1$ ならば $H^k\to 0\ (k\to\infty)$ で

$$||\boldsymbol{x}^{(k)}-\widehat{\boldsymbol{x}}||\leq||H^k||\cdot||\boldsymbol{x}^{(0)}-\widehat{\boldsymbol{x}}||\to 0\ (k\to\infty)$$

逆に $\boldsymbol{x}^{(k)}\to\widehat{\boldsymbol{x}}\ (k\to\infty)$ ならば $\boldsymbol{x}^{(0)}$ として $\boldsymbol{x}^{(0)}-\widehat{\boldsymbol{x}}=\boldsymbol{e}^{(j)}$ (n 次単位行列の第 j 列) なるものをえらべば (12.4) より

$$H^k\boldsymbol{e}^{(j)}\to 0\quad (k\to\infty)$$

$H^k\boldsymbol{e}^{(j)}$ は H^k の第 j 列をあらわすから $j=1,2,\ldots,n$ として H^k の各列は $k\to\infty$ のときゼロベクトルに収束する．故に $H^k\to O\ (k\to\infty)$．したがって系 4.5.2 によって $\rho(H)<1$ である．これで主張が示された．

特に $M=D=\mathrm{diag}(a_{11},\ldots,a_{nn})$ のとき反復 (12.2) すなわち

$$\begin{aligned}\boldsymbol{x}^{(k+1)}&=D^{-1}(N\boldsymbol{x}^{(k)}+\boldsymbol{b})\\&=B\boldsymbol{x}^{(k)}+D^{-1}\boldsymbol{b}\quad\big(\text{ただし } B=D^{-1}N=D^{-1}(D-A)\big)\\&=\boldsymbol{x}^{(k)}+D^{-1}(\boldsymbol{b}-A\boldsymbol{x}^{(k)})\end{aligned} \tag{12.5}$$

を **Jacobi 法**といい，B を **Jacobi 行列**という．

また緩和パラメータ $\omega > 0$ を導入して，反復

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \omega D^{-1}(\boldsymbol{b} - A\boldsymbol{x}^{(k)}) \tag{12.6}$$

を考え，Jacobi 法を加速することも考えられる．これを**緩和 Jacobi 法 (the Jacobi method with relaxation)** という．この方法の収束については次の定理が成り立つ．

定理 12.1 $\rho(B) < 1$ すなわち Jacobi 反復 (12.5) が収束すると仮定する．さらに Jacobi 行列 B の固有値 λ_i が実数で $1 > \lambda_1 \geq \cdots \geq \lambda_n > -1$ とするとき，緩和 Jacobi 法 (12.6) のパラメータ ω の最適値は次で与えられる．

$$\omega_{\text{opt}} = \frac{2}{2 - \lambda_1 - \lambda_n}$$

【証明】 (12.6) は

$$\boldsymbol{x}^{(k+1)} = (I - \omega D^{-1}A)\boldsymbol{x}^{(k)} + \omega D^{-1}\boldsymbol{b}$$

と書けるから反復行列は

$$\begin{aligned} B_\omega = I - \omega D^{-1}A &= I - \omega D^{-1}(D - N) = I - \omega(I - D^{-1}N) \\ &= (1 - \omega)I + \omega B, \quad \omega > 0 \end{aligned}$$

よって λ を B の固有値とすれば

$$B\boldsymbol{u} = \lambda\boldsymbol{u},\ \boldsymbol{u} \neq \boldsymbol{0} \ \Leftrightarrow\ B_\omega\boldsymbol{u} = (1 - \omega + \omega\lambda)\boldsymbol{u},\ \boldsymbol{u} \neq \boldsymbol{0}$$

よって $\rho(B_\omega) = \max(|1 - \omega + \omega\lambda_1|,\ |1 - \omega + \omega\lambda_n|)$ であり，$\rho(B_\omega)$ のグラフは図 12.1 のようになる．

$\rho(B_\omega)$ を最小にする $\omega > 0$ は図 12.1 より $-(1 - \omega + \omega\lambda_n) = 1 - \omega + \omega\lambda_1$ のときで

$$\omega_{\text{opt}} = \frac{2}{2 - \lambda_1 - \lambda_n}$$

このとき

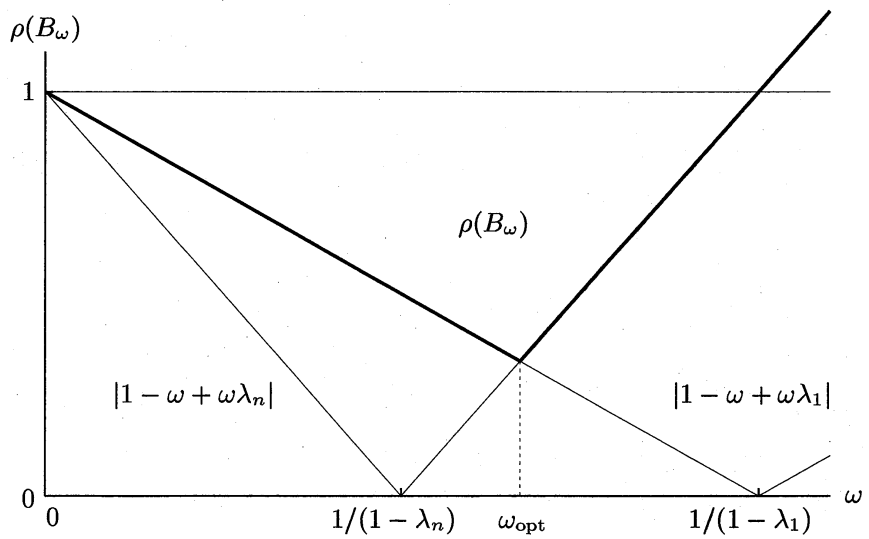

図 12.1 $\rho(B_\omega) = \max(|1 - \omega + \omega\lambda_1|,\ |1 - \omega + \omega\lambda_n|)$

$$\rho(B_{\omega_{\text{opt}}}) = 1 + \omega_{\text{opt}}(\lambda_1 - 1)$$
$$= \frac{\lambda_1 - \lambda_n}{2 - \lambda_1 - \lambda_n}$$

証明終 ∎

12.2 Gauss-Seidel 法

Jacobi 法は

$$\boldsymbol{x}_i^{(k+1)} = \frac{1}{a_{ii}}\Big(b_i - \sum_{j\neq i} a_{ij}x_j^{(k)}\Big), \quad 1 \le i \le n, \quad k \ge 0$$

と書ける．このとき $x_1^{(k)}, \ldots, x_{i-1}^{(k)}$ をすでに求めた最新の値 $x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}$ でおきかえて

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\Big(b_i - \sum_{j<i} a_{ij}x_j^{(k+1)} - \sum_{j>i} a_{ij}x_j^{(k)}\Big), \quad 1 \le i \le n, \quad k \ge 0 \tag{12.7}$$

とする反復を **Gauss-Seidel** (ガウス・ザイデル) **法**という．

$$L = \begin{pmatrix} 0 & & & \\ a_{21} & 0 & & \\ \vdots & & 0 & \\ a_{n1} & \cdots & a_{n\,n-1} & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & 0 & \ddots & \vdots \\ & & 0 & a_{n-1\,n} \\ & & & 0 \end{pmatrix},$$

$$D = \begin{pmatrix} a_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & a_{n\,n} \end{pmatrix}$$

とおけば (12.7) は

$$D\boldsymbol{x}^{(k+1)} = \boldsymbol{b} - L\boldsymbol{x}^{(k+1)} - U\boldsymbol{x}^{(k)}$$
$$\therefore \quad (D+L)\boldsymbol{x}^{(k+1)} = \boldsymbol{b} - U\boldsymbol{x}^{(k)}$$
$$\therefore \quad \boldsymbol{x}^{(k+1)} = -(D+L)^{-1}U\boldsymbol{x}^{(k)} + (D+L)^{-1}\boldsymbol{b}$$

12.3 SOR 法 (Gauss-Seidel 法の緩和)

ω を緩和パラメータとして

$$\widetilde{\boldsymbol{x}}^{(k+1)} = D^{-1}\big[\boldsymbol{b} - L\boldsymbol{x}^{(k+1)} - U\boldsymbol{x}^{(k)}\big] \tag{12.8}$$
$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \omega(\widetilde{\boldsymbol{x}}^{(k+1)} - \boldsymbol{x}^{(k)}), \quad k \ge 0 \tag{12.9}$$

とする反復を **SOR 法 (successive over-relaxation method)** という．

$\omega = 1$ のときはこの反復は Gauss-Seidel 法であるから (12.8), (12.9) は Gaus-Seidel 法の緩和 (加速) とみなせる．(12.8) を (12.9) に代入して両辺に左から D をかければ

$$D\boldsymbol{x}^{(k+1)} = D\boldsymbol{x}^{(k)} + \omega\big[(\boldsymbol{b} - L\boldsymbol{x}^{(k+1)} - U\boldsymbol{x}^{(k)}) - D\boldsymbol{x}^{(k)}\big]$$
$$\therefore \quad (D+\omega L)\boldsymbol{x}^{(k+1)} = \big\{(1-\omega)D - \omega U\big\}\boldsymbol{x}^{(k)} + \omega\boldsymbol{b}$$

これは (12.2) において

$$M=\frac{1}{\omega}(D+\omega L),\quad N=\frac{1}{\omega}\{(1-\omega)D-\omega U\},\quad \omega\neq 0$$

としたものである．したがって SOR 法の反復行列は

$$H_\omega=(D+\omega L)^{-1}\{(1-\omega)D-\omega U\}$$

故に SOR 法の収束条件は $\rho(H_\omega)<1$ である．

定理 12.2 (Kahan) SOR 法が任意の初期値 $\boldsymbol{x}^{(0)}$ に対して収束するならば $0<\omega<2$

【証明】 H_ω の固有値を $\lambda_1,\dots,\lambda_n$ とすれば

$$\begin{aligned}\lambda_1\lambda_2\cdots\lambda_n=\det H_\omega&=\det(D+\omega L)^{-1}\det\{(1-\omega)D-\omega U\}\\&=\frac{1}{a_{11}\cdots a_{nn}}(1-\omega)^n a_{11}\cdots a_{nn}=(1-\omega)^n\end{aligned}$$

よって任意の $\boldsymbol{x}^{(0)}$ に対して SOR 法が収束する $\Rightarrow\ \rho(H_\omega)<1\ \Rightarrow\ |\lambda_1\cdots\lambda_n|<1$

$$\Rightarrow\ |1-\omega|^n<1\ \Rightarrow\ 0<\omega<2$$

証明終 ∎

定理 12.3 (Ostrowski) A が正定値 Hermite 行列ならば

$$\rho(H_\omega)<1\ \Leftrightarrow\ 0<\omega<2$$

【証明】 定理 12.2 によって $\Leftarrow$ のみ証明すればよい．

$S=A-(M^{-1}N)^*A(M^{-1}N)$ とおくと

$$\begin{aligned}S&=A-\big(M^{-1}(M-A)\big)^*A\big(M^{-1}(M-A)\big)=A-(I-M^{-1}A)^*A(I-M^{-1}A)\\&=A-\big(A-(M^{-1}A)^*A\big)(I-M^{-1}A)=(M^{-1}A)^*A+AM^{-1}A-(M^{-1}A)^*A(M^{-1}A)\\&=(M^{-1}A)^*(M+M^*-A)(M^{-1}A)=(M^{-1}A)^*(M^*+N)(M^{-1}A)\\&=\frac{2-\omega}{\omega}(M^{-1}A)^*D(M^{-1}A)\end{aligned}\tag{12.10}$$

A は正定値 Hermite 行列であるから D も正定値行列 (正対角行列) で，$0<\omega<2$ ならば (12.10) は正定値 Hermite 行列である．故に Stein の定理 (系 10.7.1) によって $\rho(M^{-1}N)<1$．　証明終 ∎

■**注意**　SOR 法の緩和パラメータ ω は $0<\omega<1$ のとき**過小緩和 (under-relaxation) パラメータ**，$1<\omega<2$ のとき**過大緩和 (over-relaxation) パラメータ**と呼ばれる．しかし $0<\omega<2$ のとき両者を区別せず **SOR 法 (逐次過大緩和法)** と呼ぶのが慣例である．それは，実計算では通常 $\omega>1$ ととるからである (後述の定理 12.5 を参照)．

12.4 整合順序行列と Young の定理

SOR 法はブロック行列

$$A=\begin{pmatrix}A_{11}&A_{12}&\cdots&A_{1m}\\\vdots&&\vdots&\\\vdots&&\vdots&\\A_{m1}&A_{m2}&\cdots&A_{mm}\end{pmatrix}\quad(A_{ii}\text{は}n_i\text{次行列},\ \sum_{i=1}^{m}n_i=n)\tag{12.11}$$

に対してもそのまま定義される．すなわち A_{ii} $(1 \le i \le m)$ は正則と仮定し

$$D = \begin{pmatrix} A_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & A_{mm} \end{pmatrix}, \quad L = \begin{pmatrix} O & & & \\ A_{21} & O & & \\ \vdots & \ddots & \ddots & \\ A_{m1} & \cdots & A_{m\,m-1} & O \end{pmatrix},$$

$$U = \begin{pmatrix} O & A_{12} & \cdots & A_{1m} \\ & \ddots & \ddots & \vdots \\ & & \ddots & A_{m-1\,m} \\ & & & O \end{pmatrix},$$

$$M_\omega = \frac{1}{\omega}(D + \omega L), \quad N_\omega = \frac{1}{\omega}\{(1-\omega)D - \omega U\}, \quad H_\omega = M_\omega^{-1} N_\omega$$

とおきブロック SOR 反復を

$$\boldsymbol{x}^{(k+1)} = H_\omega \boldsymbol{x}^{(k)} + M_\omega^{-1}\boldsymbol{b}, \quad k \ge 0 \tag{12.12}$$

により定義する．$A_{ij} = a_{ij}$ の場合は通常の SOR 反復となる．なお，(12.12) に対応してブロック Jacobi 反復

$$\boldsymbol{x}^{(k+1)} = B\boldsymbol{x}^{(k)} + D^{-1}\boldsymbol{b}, \quad k \ge 0 \quad \text{ただし } B = D^{-1}(D - A)$$

も定義される．

定義 12.1 行列 A が**整合順序行列 (consistently ordered matrix)** であるとは，A の上記分解 $A = D + L + U$ に対し

$$C(\alpha) = \alpha D^{-1}L + \frac{1}{\alpha}D^{-1}U \quad (\alpha\text{は実数または複素数で}\alpha \ne 0)$$

の固有値が α に依存しないときをいう．

ブロック 3 重対角行列

$$A = \begin{pmatrix} A_{11} & A_{12} & & \\ A_{21} & A_{22} & A_{23} & \\ & \ddots & \ddots & \ddots \\ & & A_{m\,m-1} & A_{mm} \end{pmatrix}$$

は整合順序行列である．実際

$$S(\alpha) = \mathrm{diag}(I_{n_1}, \alpha I_{n_2}, \ldots, \alpha^{m-1} I_{n_m})$$

とおけば容易に確かめられるように $S(\alpha)C(1)S(\alpha)^{-1} = C(\alpha)$, $\alpha \ne 0$ が成り立つ．故に $C(\alpha)$ の固有値は $C(1) = D^{-1}(L+U)$ の固有値に等しい．

さて $0 < \omega < 2$ において $\rho(H_\omega)$ を最小にする ω の値 ω_{opt} はいろいろな行列に対し Young その他により詳しく調べられている．([19], [20] 参照) この節では A が整合順序行列のとき ω_{opt} を $\rho(B)$ の関数として求めてみよう．

そのための準備としてまず次の定理を示す．

定理 12.4 (Ostrowski) A を正定値 Hermite 行列として $B = D^{-1}(D - A)$ (ブロック Jacobi 行列) とおく．A が整合順序ならば $\rho(B) < 1$ であり，ブロック Jacobi 法は任意の初期値に対して収束する．

【証明】 $A=(a_{ij})$ が正定値 Hermite 行列ならば D も正定値である．定理 5.1 によって $D^{-1}A=\sqrt{D}^{-1}(\sqrt{D}^{-1}A)$ と $(\sqrt{D}^{-1}A\sqrt{D}^{-1})$ の固有値は等しいが，後者は慣性則 (系 2.15.1) によって正定値である．よって $B=I-D^{-1}A$ の任意の固有値 β は $\beta=1-\alpha$ (α は $D^{-1}A$ の固有値で $\alpha>0$) の形であるから $\beta<1$ である．また A は整合順序行列であるから $B=C(-1)$ の固有値と $C(1)=-B$ の固有値は等しい．よって $-\beta$ も B の固有値で $-\beta<1$．したがって $\rho(B)<1$ であり任意の初期値に対してブロック Jacobi 法は収束する． 証明終 ∎

定理 12.5 (Young (ヤング)) A を n 次整合順序行列とし，Jacobi 行列 $B=D^{-1}(D-A)$ の固有値はすべて実数で $0<\rho(B)<1$ を仮定する．このとき次が成り立つ．

(i) $0<\omega<2$ のときブロック SOR 反復 (12.12) は収束する．

(ii) $\omega_{\mathrm{opt}}=\frac{2}{1+\sqrt{1-\rho(B)^2}}$ とおけば $\omega_{\mathrm{opt}}>1$ かつ

$$\min_{0<\omega<2}\rho(H_\omega)=\rho(H_{\omega_{\mathrm{opt}}})=\omega_{\mathrm{opt}}-1=\frac{1-\sqrt{1-\rho(B)^2}}{1+\sqrt{1-\rho(B)^2}}$$

【証明】

$$\begin{aligned}
H_\omega&=(D+\omega L)^{-1}\{(1-\omega)D-\omega U\}\\
&=(I+\omega D^{-1}L)^{-1}\{(1-\omega)I-\omega D^{-1}U\}\\
\therefore\quad (I+\omega D^{-1}L)(\mu I-H_\omega)&=\mu(I+\omega D^{-1}L)-\{(1-\omega)I-\omega D^{-1}U\}\\
&=(\mu+\omega-1)I+\sqrt{\mu}\omega(\sqrt{\mu}D^{-1}L+\frac{1}{\sqrt{\mu}}D^{-1}U)\\
&=\sqrt{\mu}\omega\left\{\frac{\mu+\omega-1}{\sqrt{\mu}\omega}I+\sqrt{\mu}D^{-1}L+\frac{1}{\sqrt{\mu}}D^{-1}U\right\}\\
&=\sqrt{\mu}\omega\left\{\frac{\mu+\omega-1}{\sqrt{\mu}\omega}I+C(\sqrt{\mu})\right\}\\
&\qquad\qquad\left(\text{ただし } C(\alpha)=\alpha D^{-1}L+\frac{1}{\alpha}D^{-1}U\right)
\end{aligned}$$

よって μ が H_ω の非零固有値 $\Leftrightarrow$ $\lambda=\frac{\mu+\omega-1}{\sqrt{\mu}\omega}$ が $C(-1)=-D^{-1}(L+U)=B$ の固有値 ($\because$ A は整合順序)

さらに

$$\lambda=\frac{\mu+\omega-1}{\sqrt{\mu}\omega}\Leftrightarrow\ \mu-\lambda\omega\sqrt{\mu}+\omega-1=0 \tag{12.13}$$

$$\Leftrightarrow\ \sqrt{\mu}=\frac{\omega\lambda\pm\sqrt{(\omega\lambda)^2-4(\omega-1)}}{2}$$

$$\Leftrightarrow\ \mu=\left(\frac{\omega\lambda\pm\sqrt{(\omega\lambda)^2-4(\omega-1)}}{2}\right)^2 \tag{12.14}$$

整合順序行列の性質によって λ が $B=C(-1)$ の固有値ならば $-\lambda$ も B の固有値である (定理 12.4 の証明参照)．

したがって $d(\lambda,\ \omega)=(\omega\lambda)^2-4(\omega-1)$ とおけば，μ が実数 ($d(\lambda,\ \omega)\geq 0$) のとき (12.14) より

$$|\mu|=\left(\frac{\omega|\lambda|+\sqrt{d(\lambda,\ \omega)}}{2}\right)^2 \tag{12.15}$$

μ が複素数 ($d(\lambda,\ \omega)<0$) のとき (12.13) より根と係数の関係によって

$$|\mu| = \omega - 1 \tag{12.16}$$

容易にわかるように $d(\lambda, \omega)$ を ω の2次式とみれば $f(\omega) = \lambda^2\omega^2 - 4\omega + 4$ は $0 < \omega < 2$ においてただ1つの根 $\omega_\lambda = \dfrac{2}{1+\sqrt{1-\lambda^2}} > 1$ をもつ ($\because$ $f(0) = 4 > 0$, $f(2) = 4\lambda^2 - 4 < 0$ ($\because$ $0 < |\lambda| \leq \rho(B) < 1$) かつ $f(\omega)$ は $\omega = \frac{2}{\lambda^2}$ (> 2) のとき最小値をとる). よって

$$|\mu| = \begin{cases} \left(\dfrac{\omega|\lambda| + \sqrt{d(\lambda,\ \omega)}}{2}\right)^2 & (0 < \omega \leq \omega_\lambda) \\ \omega - 1 & (\omega_\lambda < \omega < 2) \end{cases} \tag{12.17}$$

特に λ を $|\lambda| = \rho(B)$ をみたすものとすれば $\omega_\lambda = \dfrac{2}{1+\sqrt{1-\rho(B)^2}} = \omega_{\text{opt}} > 1$ かつ (12.17) より

$$\rho(H_\omega) = \begin{cases} \left(\dfrac{\omega\rho(B) + \sqrt{d(\rho(B),\ \omega)}}{2}\right)^2 & (0 < \omega \leq \omega_{\text{opt}}) \\ \omega - 1 & (\omega_{\text{opt}} < \omega < 2) \end{cases}$$

さらに関数

$$\varphi(\omega) = \left(\frac{\omega\rho + \sqrt{\rho^2\omega^2 - 4\omega + 4}}{2}\right)^2 \quad (\rho = \rho(B) \text{ と略記する})$$

は次の性質をもつ.

(a) $\varphi(0) = 1$

(b) $\varphi'(\omega) < 0 \quad (0 < \omega < \omega_{\text{opt}})$

$$\because\ \varphi'(\omega) = 2\left(\frac{\omega\rho + \sqrt{\rho^2\omega^2 - 4\omega + 4}}{2}\right) \cdot \left(\frac{\rho}{2} + \frac{\rho^2\omega - 2}{2\sqrt{\rho^2\omega^2 - 4\omega + 4}}\right) \tag{12.18}$$

$$\text{かつ} \quad \frac{\rho}{2} + \frac{\rho^2\omega - 2}{2\sqrt{\rho^2\omega^2 - 4\omega + 4}} = \frac{\rho\sqrt{\rho^2\omega^2 - 4\omega + 4} + \rho^2\omega - 2}{2\sqrt{\rho^2\omega^2 - 4\omega + 4}}$$

$$< \frac{\sqrt{\omega^2 - 4\omega + 4} + \omega - 2}{2\sqrt{\rho^2\omega^2 - 4\omega + 4}} = \frac{(2-\omega) + (\omega - 2)}{2\sqrt{\rho^2\omega^2 - 4\omega + 4}} = 0$$

(c) $\lim\limits_{\omega \to \omega_{\text{opt}}-} \varphi'(\omega) = -\infty$ ((12.18) より明らか)

また

$$\lim_{\omega\to 0} \rho(H_\omega) = \lim_{\omega\to 0} \varphi(\omega) = \varphi(0) = 1 \text{ かつ } \lim_{\omega\to 2} \rho(H_\omega) = \lim_{\omega\to 2}(\omega - 1) = 1$$

よって $\rho(H_\omega)$ のグラフは図 12.2 のようになり, $0 < \omega < 2$ のとき $\rho(H_\omega) < 1$

したがってブロック SOR 法は $0 < \omega < 2$ のとき収束し

$$\min_{0<\omega<2} \rho(H_\omega) = \rho(H_{\omega_{\text{opt}}}) = \omega_{\text{opt}} - 1 = \frac{1 - \sqrt{1-\rho(B)^2}}{1 + \sqrt{1-\rho(B)^2}}$$

証明終 ∎

注意 定理 12.5 は D. M. Young (1950) により得られた古典的な結果であるが, さらに一般な結果が Varga [19] にある.

系 12.5.1 ブロック3重正定値対称行列 A に対するブロック SOR 法は $0 < \omega < 2$ のとき任意の初期値に対して収束する.

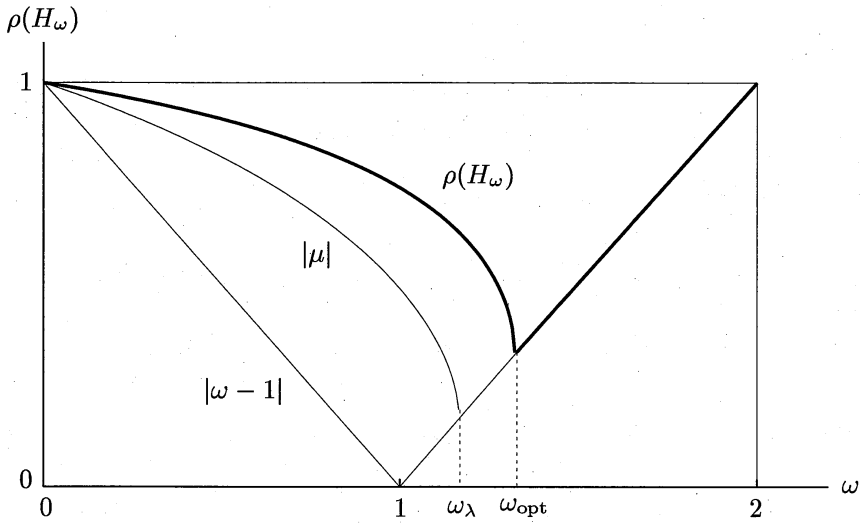

図 12.2　$\rho(H_\omega)$ と $|\mu|$ のグラフ ($\rho(B)=0.86,\ \lambda=0.7$ のとき)

【証明】　ブロック 3 重対角行列は整合順序であるから，定理 12.4 と定理 12.5 による．　証明終 ∎

系 12.5.2　$\rho(H_1)=\rho(B)^2$．したがって Gauss-Seidel 法は Jacobi 法より 2 倍速く収束する．

【証明】　$\omega=1$ のとき

$$\rho(H_1)=\left(\frac{\rho(B)+\sqrt{d(\rho(B),\ 1)}}{2}\right)^2=\left(\frac{\rho(B)+\rho(B)}{2}\right)^2=\rho(B)^2$$

であるから．　証明終 ∎

12.5　最急降下法

$A=(a_{ij})$ を n 次の実正定値対称行列とし，$\widehat{\boldsymbol{x}}$ を連立 1 次方程式 $A\boldsymbol{x}=\boldsymbol{b}$ の一意解とすれば，汎関数

$$f(\boldsymbol{x})=\frac{1}{2}\big(A(\boldsymbol{x}-\widehat{\boldsymbol{x}}),\ \boldsymbol{x}-\widehat{\boldsymbol{x}}\big)$$

は $\boldsymbol{x}=\widehat{\boldsymbol{x}}$ のとき最小 (最小値 0) となる．したがって

$$F(\boldsymbol{x})=f(\boldsymbol{x})-f(\boldsymbol{0})$$

も $\boldsymbol{x}=\widehat{\boldsymbol{x}}$ のとき最小で

$$\begin{aligned}F(\boldsymbol{x})&=\frac{1}{2}\big(A(\boldsymbol{x}-\widehat{\boldsymbol{x}}),\ \boldsymbol{x}-\widehat{\boldsymbol{x}}\big)-\frac{1}{2}(A\widehat{\boldsymbol{x}},\ \widehat{\boldsymbol{x}})=\frac{1}{2}(A\boldsymbol{x},\ \boldsymbol{x})-(A\widehat{\boldsymbol{x}},\ \boldsymbol{x})\\&=\frac{1}{2}(A\boldsymbol{x},\ \boldsymbol{x})-(\boldsymbol{b},\ \boldsymbol{x})=\frac{1}{2}\sum_{i,j=1}^{n}a_{ij}x_ix_j-\sum_{i=1}^{n}b_ix_i\end{aligned}$$

ここで $F(\boldsymbol{x})$ の勾配 grad$F$ は

$$\begin{aligned}\boldsymbol{g}(\boldsymbol{x})&=\left(\frac{\partial F(\boldsymbol{x})}{\partial x_1},\dots,\frac{\partial F(\boldsymbol{x})}{\partial x_n}\right)^{\mathrm{t}}\\&=A\boldsymbol{x}-\boldsymbol{b}=-\boldsymbol{r}\quad(\boldsymbol{r}=\boldsymbol{b}-A\boldsymbol{x}\ \text{は 残差ベクトルと呼ばれる})\end{aligned}\tag{12.19}$$

である．いま $\boldsymbol{v}=(v_1,\dots,v_n)^{\mathrm{t}}\in\mathbb{R}^n$ を長さ 1 のベクトル ($||\boldsymbol{v}||_2^2=\sum_{i=1}^n v_i^2=1$) とすれば，$\boldsymbol{x}$ における F の $\boldsymbol{v}$ 方向の微分 (方向微分) は

$$F^{(1)}(\boldsymbol{x},\ \boldsymbol{v}) = \frac{\partial}{\partial t}F(\boldsymbol{x}+t\boldsymbol{v})\ \Big|_{t=0} = \sum_{j=1}^{n}\frac{\partial F(\boldsymbol{x})}{\partial x_j}v_j = (\boldsymbol{g}(\boldsymbol{x}),\ \boldsymbol{v})$$

$$\therefore\quad |F^{(1)}(\boldsymbol{x},\ \boldsymbol{v})| = |(\boldsymbol{g}(\boldsymbol{x}),\ \boldsymbol{v})| \le ||\boldsymbol{g}(\boldsymbol{x})||_2 \cdot ||\boldsymbol{v}||_2 = ||\boldsymbol{g}(\boldsymbol{x})||_2 \tag{12.20}$$

特に $\boldsymbol{v} = -\boldsymbol{g}(\boldsymbol{x})/||\boldsymbol{g}(\boldsymbol{x})||_2$ とれば上式は等号となり

$$F^{(1)}(\boldsymbol{x},\ \boldsymbol{v}) = -||\boldsymbol{g}(\boldsymbol{x})||_2 \quad \text{かつ} \quad -\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{r} \quad ((12.19)\text{による})$$

よって (12.20) より $F(\boldsymbol{x})$ の最も減少する方向は $\boldsymbol{r}$ 方向である．したがって

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{r}^{(k)},\quad k \ge 0$$
$$\boldsymbol{r}^{(k)} = \boldsymbol{b} - A\boldsymbol{x}^{(k)} \quad (\text{残差ベクトル})$$

とおき，パラメータ α_k を $F(\boldsymbol{x}^{(k+1)})$ を最小にするように定める反復が考えられる．これを**最急降下法 (method of steepest descent)** という．

$$\begin{aligned} F(\boldsymbol{x}+\alpha\boldsymbol{r}) &= \frac{1}{2}\big(A(\boldsymbol{x}+\alpha\boldsymbol{r}),\ \boldsymbol{x}+\alpha\boldsymbol{r}\big) - (\boldsymbol{b},\ \boldsymbol{x}+\alpha\boldsymbol{r}) \\ &= \frac{1}{2}(A\boldsymbol{x},\ \boldsymbol{x}) - (\boldsymbol{b},\ \boldsymbol{x}) + \alpha(A\boldsymbol{x}-\boldsymbol{b},\ \boldsymbol{r}) + \frac{1}{2}\alpha^2(A\boldsymbol{r},\ \boldsymbol{r}) \\ &= F(\boldsymbol{x}) - \alpha(\boldsymbol{r},\ \boldsymbol{r}) + \frac{1}{2}\alpha^2(A\boldsymbol{r},\ \boldsymbol{r}) \end{aligned}$$

であるから α の 2 次式 $F(\boldsymbol{x}+\alpha\boldsymbol{r})$ は

$$\alpha = \frac{(\boldsymbol{r},\ \boldsymbol{r})}{(A\boldsymbol{r},\ \boldsymbol{r})} = \frac{||\boldsymbol{r}||_2^2}{(A\boldsymbol{r},\ \boldsymbol{r})}$$

のとき最小となり，最小値は

$$F(\boldsymbol{x}) - \frac{||\boldsymbol{r}||_2^4}{2(A\boldsymbol{r},\ \boldsymbol{r})}$$

故に最急降下法は次式で定義される．

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{r}^{(k)} \tag{12.21}$$
$$\boldsymbol{r}^{(k+1)} = \boldsymbol{r}^{(k)} - \alpha_k A\boldsymbol{r}^{(k)} \tag{12.22}$$
$$\alpha_k = \frac{||\boldsymbol{r}^{(k)}||_2^2}{(A\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)})},\quad k \ge 0 \tag{12.23}$$

したがって

$$\begin{aligned} (\boldsymbol{r}^{(k+1)},\ \boldsymbol{r}^{(k)}) &= (\boldsymbol{r}^{(k)} - \alpha_k A\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) \\ &= (\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) - \alpha_k(A\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) = 0 \quad (k \ge 0) \end{aligned}$$

定理 12.6 (収束定理)　A が正定値対称行列ならば，最急降下法は任意の $\boldsymbol{x}^{(0)}$ に対して解 $\widehat{\boldsymbol{x}}$ に収束する．

【証明】　A の固有値を $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$ とすれば任意の $\boldsymbol{x} \in \mathbb{R}^n$ に対して

$$(\boldsymbol{x},\ A^{-1}\boldsymbol{x}) \le \frac{1}{\lambda_n}(\boldsymbol{x},\ \boldsymbol{x}) \tag{12.24}$$

および

$$(\boldsymbol{x},\ A\boldsymbol{x}) \le \lambda_1(\boldsymbol{x},\ \boldsymbol{x}) \tag{12.25}$$

さて

$$\begin{aligned}(\boldsymbol{r}^{(k+1)},\ A^{-1}\boldsymbol{r}^{(k+1)}) &= (\boldsymbol{r}^{(k)} - \alpha_k A\boldsymbol{r}^{(k)},\ A^{-1}\boldsymbol{r}^{(k)} - \alpha_k \boldsymbol{r}^{(k)}) \\ &= (\boldsymbol{r}^{(k)},\ A^{-1}\boldsymbol{r}^{(k)}) - \alpha_k(A\boldsymbol{r}^{(k)},\ A^{-1}\boldsymbol{r}^{(k)}) \\ &\qquad - \alpha_k(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) + \alpha_k^2(A\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) \\ &= (\boldsymbol{r}^{(k)},\ A^{-1}\boldsymbol{r}^{(k)}) - \alpha_k(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) \qquad (12.26)\end{aligned}$$

(12.24)–(12.26) より

$$\alpha_k \geq \frac{||\boldsymbol{r}^{(k)}||_2^2}{\lambda_1(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)})} \geq \frac{\lambda_n}{\lambda_1}\,\frac{(\boldsymbol{r}^{(k)},\ A^{-1}\boldsymbol{r}^{(k)})}{(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)})}$$

$$\therefore\quad (\boldsymbol{r}^{(k+1)},\ A^{-1}\boldsymbol{r}^{(k+1)}) \leq \left(1 - \frac{\lambda_n}{\lambda_1}\right)(\boldsymbol{r}^{(k)},\ A^{-1}\boldsymbol{r}^{(k)})$$

$$\leq \left(1 - \frac{\lambda_n}{\lambda_1}\right)^{k+1}(\boldsymbol{r}^{(0)},\ A^{-1}\boldsymbol{r}^{(0)}) \to 0\ (k \to \infty) \qquad (12.27)$$

($\lambda_1 = \lambda_n$の場合も (12.27) に含まれる.)

よって $\boldsymbol{r}^{(k+1)} \to 0\ (k \to \infty)$ を得る. したがって $k \to \infty$ のとき

$$\boldsymbol{x}^{(k)} = A^{-1}(\boldsymbol{b} - \boldsymbol{r}^{(k)}) \to A^{-1}\boldsymbol{b} = \widehat{\boldsymbol{x}}$$

証明終 ∎

定理 12.7 (誤差評価) A を正定値対称行列として $||\boldsymbol{x}||_A = \sqrt{(A\boldsymbol{x},\ \boldsymbol{x})}$ (エネルギーノルム), $\kappa(A) = ||A^{-1}||_2\,||A||_2$ (スペクトル条件数) とおくとき, $\kappa(A) \geq 1$ かつ

$$||\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}||_A \leq \left(\frac{\kappa(A)-1}{\kappa(A)+1}\right)^k ||\boldsymbol{x}^{(0)} - \widehat{\boldsymbol{x}}||_A, \quad k \geq 0$$

【証明】 明らかに $\kappa(A) = ||A^{-1}||_2||A||_2 \geq ||A^{-1}A||_2 = ||I||_2 = 1$ である.

さて

$$\begin{aligned}2f(\boldsymbol{x}^{(k)}) &= (\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}})^{\mathrm{t}}A(\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}) = (\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}})^{\mathrm{t}}A \cdot A^{-1}A(\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}) \\ &= \big(A(\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}})\big)^{\mathrm{t}}A^{-1}A(\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}) = \boldsymbol{r}^{(k)^{\mathrm{t}}}A^{-1}\boldsymbol{r}^{(k)}\end{aligned}$$

$$\begin{aligned}\therefore\quad 2f(\boldsymbol{x}^{(k+1)}) &= (A^{-1}\boldsymbol{r}^{(k+1)},\ \boldsymbol{r}^{(k+1)}) \\ &= (A^{-1}\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) - \alpha_k(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) \quad ((12.26)\text{ による}) \\ &= \left\{1 - \frac{(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)})}{(A\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)})} \cdot \frac{(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)})}{(A^{-1}\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)})}\right\}(A^{-1}\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) \\ &\leq \left\{1 - \frac{4\lambda_1\lambda_n}{(\lambda_1+\lambda_n)^2}\right\}(A^{-1}\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) \quad (\text{Kantorovich の不等式による}) \\ &= \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 \big(2f(\boldsymbol{x}^{(k)})\big) \\ &= \left(\frac{\kappa(A)-1}{\kappa(A)+1}\right)^2 \big(2f(\boldsymbol{x}^{(k)})\big) \qquad (12.28)\end{aligned}$$

$2f(\boldsymbol{x}^{(k)}) = \big(A(\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}),\ \boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}\big) = ||\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}||_A^2$ であるから (12.28) より

$$||\boldsymbol{x}^{(k+1)}-\widehat{\boldsymbol{x}}||_A \le \left(\frac{\kappa(A)-1}{\kappa(A)+1}\right)||\boldsymbol{x}^{(k)}-\widehat{\boldsymbol{x}}||_A$$
$$\le \left(\frac{\kappa(A)-1}{\kappa(A)+1}\right)^{k+1}||\boldsymbol{x}^{(0)}-\widehat{\boldsymbol{x}}||_A, \quad k\ge 0$$

これは求める不等式である．　　証明終 ∎

12.6 共役勾配法 (最急降下法の加速)

最急降下法の加速として**共役勾配法** (**conjugate gradient method** 略して **CG 法**) と呼ばれる反復法がある．この反復は

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\alpha_k\{\boldsymbol{r}^{(k)}+\gamma_k(\boldsymbol{x}^{(k)}-\boldsymbol{x}^{(k-1)})\} \tag{12.29}$$

の形であり，パラメータ α_k と γ_k は次のように定める．まず

$$\boldsymbol{p}^{(k)}=\boldsymbol{r}^{(k)}+\gamma_k(\boldsymbol{x}^{(k)}-\boldsymbol{x}^{(k-1)})$$

とおけば (12.29) は

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)}$$

と書けるから

$$\boldsymbol{p}^{(k)}=\boldsymbol{r}^{(k)}+\gamma_k(\alpha_{k-1}\boldsymbol{p}^{(k-1)})=\boldsymbol{r}^{(k)}+\beta_{k-1}\boldsymbol{p}^{(k-1)} \quad (\beta_{k-1}=\alpha_{k-1}\gamma_k)$$

よって

$$\begin{aligned}\boldsymbol{x}^{(k+1)}&=\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)}\\ \boldsymbol{r}^{(k+1)}&=\boldsymbol{b}-A\boldsymbol{x}^{(k+1)}=\boldsymbol{b}-A(\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)})\\ &=\boldsymbol{r}^{(k)}-\alpha_kA\boldsymbol{p}^{(k)}\\ \boldsymbol{p}^{(k+1)}&=\boldsymbol{r}^{(k+1)}+\beta_k\boldsymbol{p}^{(k)}\end{aligned}$$

最急降下法と同様に α_k は $F(\boldsymbol{x}^{(k+1)})=F(\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)})$ を最小にするように定める．

$$\begin{aligned}F(\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)})&=\frac{1}{2}(A(\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)}),\ \boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)})-(\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)},\ \boldsymbol{b})\\ &=\frac{1}{2}(A\boldsymbol{x}^{(k)},\ \boldsymbol{x}^{(k)})-(\boldsymbol{x}^{(k)},\ \boldsymbol{b})+\alpha_k(A\boldsymbol{p}^{(k)},\ \boldsymbol{x}^{(k)})\\ &\qquad+\frac{1}{2}\alpha_k^2(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)})-\alpha_k(\boldsymbol{p}^{(k)},\ \boldsymbol{b})\\ &=F(\boldsymbol{x}^{(k)})-\alpha_k(\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)})+\frac{1}{2}\alpha_k^2(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)})\end{aligned}$$

であるから，$F(\boldsymbol{x}^{(k+1)})$ は

$$\alpha_k=\frac{(\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)})}{(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)})} \tag{12.30}$$

のとき最小となり最小値

$$F(\boldsymbol{x}^{(k+1)})=F(\boldsymbol{x}^{(k)})-\frac{1}{2}\frac{(\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)})^2}{(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)})} \tag{12.31}$$

をとる．このとき

$$(\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k+1)})=(\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)}-\alpha_kA\boldsymbol{p}^{(k)})=(\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)})-\alpha_k(\boldsymbol{p}^{(k)},\ A\boldsymbol{p}^{(k)})=0$$

よって

$$(\boldsymbol{p}^{(k+1)},\ \boldsymbol{r}^{(k+1)}) = (\boldsymbol{r}^{(k+1)} + \beta_k \boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k+1)}) = (\boldsymbol{r}^{(k+1)},\ \boldsymbol{r}^{(k+1)}) \quad (k \geq 0)$$

この関係式が $k+1=0$ のときも成り立つように $\boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)}$ と定めれば

$$(\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)}) = (\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}), \quad k \geq 0 \tag{12.32}$$

となって (12.31) より

$$F(\boldsymbol{x}^{(k+1)}) = F(\boldsymbol{x}^{(k)}) - \frac{1}{2}\frac{(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)})^2}{(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)})}$$

よって再び $F(\boldsymbol{x}^{(k+1)})$ を最小にするために $(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)})$ を最小にする β_k を定める.

$$\begin{aligned}(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)}) &= \big(A(\boldsymbol{r}^{(k)} + \beta_{k-1}\boldsymbol{p}^{(k-1)}),\ \boldsymbol{r}^{(k)} + \beta_{k-1}\boldsymbol{p}^{(k-1)}\big) \\ &= (A\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) + 2\beta_{k-1}(A\boldsymbol{p}^{(k-1)},\ \boldsymbol{r}^{(k)}) + \beta_{k-1}^2(A\boldsymbol{p}^{(k-1)},\ \boldsymbol{p}^{(k-1)})\end{aligned}$$

であるから

$$\beta_{k-1} = -\frac{(A\boldsymbol{p}^{(k-1)},\ \boldsymbol{r}^{(k)})}{(A\boldsymbol{p}^{(k-1)},\ \boldsymbol{p}^{(k-1)})} \quad (k \geq 1)$$

すなわち

$$\beta_k = -\frac{(A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k+1)})}{(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)})} \qquad (k \geq 0) \tag{12.33}$$

すると

$$\begin{aligned}(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k+1)}) &= (A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k+1)} + \beta_k \boldsymbol{p}^{(k)}) \\ &= (A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k+1)}) + \beta_k(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)}) = 0 \qquad (12.34)\end{aligned}$$

($\boldsymbol{p}^{(k)}$ と $\boldsymbol{p}^{(k+1)}$ は**共役**であるという.)

したがって

$$\begin{aligned}(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)}) &= (A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)} + \beta_{k-1}\boldsymbol{p}^{(k-1)}) = (A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)}) + \beta_{k-1}(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k-1)}) \\ &= (A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)})\end{aligned}$$

この等式と (12.30) および (12.32) を用いて

$$\begin{aligned}(\boldsymbol{r}^{(k+1)},\ \boldsymbol{r}^{(k)}) &= (\boldsymbol{r}^{(k)} - \alpha_k A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)}) \\ &= (\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) - \alpha_k(A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)}) \\ &= (\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) - \alpha_k(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)}) \\ &= (\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)}) - (\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)}) = 0 \qquad (12.35)\end{aligned}$$

$$\begin{aligned}\therefore\quad (\boldsymbol{r}^{(k+1)},\ \boldsymbol{r}^{(k+1)}) &= (\boldsymbol{r}^{(k)} - \alpha_k A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k+1)}) \\ &= (\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k+1)}) - \alpha_k(A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k+1)}) \\ &= -\alpha_k(A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k+1)})\end{aligned}$$

$$\therefore\quad (A\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k+1)}) = -\frac{1}{\alpha_k}(\boldsymbol{r}^{(k+1)},\ \boldsymbol{r}^{(k+1)})$$

これを (12.33) に代入すれば

$$\beta_k = \frac{1}{\alpha_k} \cdot \frac{(\boldsymbol{r}^{(k+1)},\ \boldsymbol{r}^{(k+1)})}{(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)})} = \frac{(\boldsymbol{r}^{(k+1)},\ \boldsymbol{r}^{(k+1)})}{(\boldsymbol{p}^{(k)},\ \boldsymbol{r}^{(k)})}$$

$$= \frac{(\boldsymbol{r}^{(k+1)},\ \boldsymbol{r}^{(k+1)})}{(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(k)})} \quad ((12.32)\text{ による})$$

以上をまとめて CG 法のアルゴリズムは次のようになる.

[CG 法]

$$\boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)} = \boldsymbol{b} - A\boldsymbol{x}^{(0)} \tag{12.36a}$$

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}^{(k)} \tag{12.36b}$$

$$\boldsymbol{r}^{(k+1)} = \boldsymbol{r}^{(k)} - \alpha_k A\boldsymbol{p}^{(k)} \tag{12.36c}$$

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{r}^{(k+1)} + \beta_k \boldsymbol{p}^{(k)} \tag{12.36d}$$

$$\alpha_k = \frac{||\boldsymbol{r}^{(k)}||_2^2}{(A\boldsymbol{p}^{(k)},\ \boldsymbol{p}^{(k)})} \tag{12.36e}$$

$$\beta_k = \frac{||\boldsymbol{r}^{(k+1)}||_2^2}{||\boldsymbol{r}^{(k)}||_2^2}, \quad k \geq 0 \tag{12.36f}$$

定理 12.8 CG 法では次が成り立つ.

$$(\boldsymbol{r}^{(k)},\ \boldsymbol{r}^{(l)}) = (\boldsymbol{p}^{(k)},\ A\boldsymbol{p}^{(l)}) = 0 \quad (k \neq l) \tag{12.37}$$

この定理を証明するために，少し回り道をして，CG 反復 $\{\boldsymbol{x}^{(k)}\}$ と一見異なる反復 $\{\widetilde{\boldsymbol{x}}^{(k)}\}$ を定義した後,

$$\widetilde{\boldsymbol{x}}^{(0)} = \boldsymbol{x}^{(0)} \ \Rightarrow\ \widetilde{\boldsymbol{x}}^{(k)} = \boldsymbol{x}^{(k)}, \quad k \geq 1$$

を示すことにより (12.37) を示そう．$l < k$ として示せばよい.

以下 Axelsson-Barker[8] に従い，$\mathbb{R}^n$ の内積 $(\ ,\ )_{A^{-1}}$ とこの内積から導かれるノルム $||\cdot||_{A^{-1}}$ を

$$(\boldsymbol{x},\ \boldsymbol{y})_{A^{-1}} = \boldsymbol{y}^{\mathrm{t}} A^{-1} \boldsymbol{x} = \boldsymbol{x}^{\mathrm{t}} A^{-1} \boldsymbol{y}, \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$$

$$||\boldsymbol{x}||_{A^{-1}} = \sqrt{(\boldsymbol{x},\ \boldsymbol{x})_{A^{-1}}} = (\boldsymbol{x}^{\mathrm{t}} A^{-1} \boldsymbol{x})^{1/2}, \quad \boldsymbol{x} \in \mathbb{R}^n$$

により定義する．このとき次の補題をまず示す.

補題 12.1 S を $\mathbb{R}^n$ の線形部分空間とし，$\boldsymbol{r}_0 \in S$ とする．内積 $(\ ,\ )_{A^{-1}}$ に関する S の直交補空間を $S^{\perp}$ とするとき次が成り立つ.

(i) $\boldsymbol{r}_0 + \boldsymbol{u}_0 \in S^{\perp}$ なる $\boldsymbol{u}_0 \in S$ はただ 1 つ存在する.

(ii) $\boldsymbol{u} \in S$ かつ $\boldsymbol{u} \neq \boldsymbol{u}_0$ ならば $||\boldsymbol{r}_0 + \boldsymbol{u}||_{A^{-1}} > ||\boldsymbol{r}_0 + \boldsymbol{u}_0||_{A^{-1}}$

【証明】 (i) $\dim S = m$ とし，$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ を S の基底とする．S の元 $\boldsymbol{u}$ を $\boldsymbol{u} = \sum_{j=1}^{m} c_j \boldsymbol{u}_j$ とあらわすとき

$$\boldsymbol{r}_0 + \boldsymbol{u} \in S^{\perp} \quad \Leftrightarrow \quad (\boldsymbol{u}_i,\ \boldsymbol{r}_0 + \boldsymbol{u})_{A^{-1}} = 0,\ 1 \leq i \leq m$$

$$\Leftrightarrow \quad \sum_{j=1}^{m} (\boldsymbol{u}_i,\ \boldsymbol{u}_j)_{A^{-1}} c_j = -(\boldsymbol{u}_i,\ \boldsymbol{r}_0)_{A^{-1}} \tag{12.38}$$

$c_1, \ldots, c_m$ に関する連立 1 次方程式 (12.38) の係数行列は $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ のつくる Gram 行列であり，$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ は 1 次独立であるから定理 2.14 により正則である.

よって (12.38) は一意解 $\boldsymbol{c} = (c_1, \ldots, c_m)^{\mathrm{t}}$ をもち，$\boldsymbol{r}_0 + \boldsymbol{u} \in S^{\perp}$ となる $\boldsymbol{u} \in S$ はただ 1 つ確定する.

これを $\boldsymbol{u}_0$ であらわす.

(ii) $\boldsymbol{u} \in S$ かつ $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{u}_0$ とおけば $\boldsymbol{v} \in S$ で $(\boldsymbol{r}_0 + \boldsymbol{u}_0,\ \boldsymbol{v})_{A^{-1}} = 0$

$$\begin{aligned}\therefore\quad ||\boldsymbol{r}_0 + \boldsymbol{u}||^2_{A^{-1}} &= (\boldsymbol{r}_0 + \boldsymbol{u}_0 + \boldsymbol{v},\ \boldsymbol{r}_0 + \boldsymbol{u}_0 + \boldsymbol{v})_{A^{-1}} \\ &= ||\boldsymbol{r}_0 + \boldsymbol{u}_0||^2_{A^{-1}} + 2(\boldsymbol{r}_0 + \boldsymbol{u}_0,\ \boldsymbol{v})_{A^{-1}} + ||\boldsymbol{v}||^2_{A^{-1}} \\ &= ||\boldsymbol{r}_0 + \boldsymbol{u}_0||^2_{A^{-1}} + ||\boldsymbol{v}||^2_{A^{-1}} \\ &\geq ||\boldsymbol{r}_0 + \boldsymbol{u}_0||^2_{A^{-1}}\end{aligned}$$

等号が成り立つのは $\boldsymbol{v} = \boldsymbol{0}$ すなわち $\boldsymbol{u} = \boldsymbol{u}_0$ のときに限る. 証明終 ∎

■**注意 1** 補題 12.1 は通常の Euclid 内積 (,) でも成り立つ.

【定理 12.8 の証明】 (12.36a)–(12.36f) に類似な反復 $\{\widetilde{\boldsymbol{x}}^{(k)}\}$, $k \geq 0$ を次により定義する.

$$\widetilde{\boldsymbol{p}}^{(0)} = \widetilde{\boldsymbol{r}}^{(0)} = \boldsymbol{b} - A\widetilde{\boldsymbol{x}}^{(0)} \tag{12.36a′}$$

$$\widetilde{\boldsymbol{x}}^{(k+1)} = \widetilde{\boldsymbol{x}}^{(k)} + \widetilde{\alpha}_k \widetilde{\boldsymbol{p}}^{(k)} \tag{12.36b′}$$

$$\widetilde{\boldsymbol{r}}^{(k+1)} = \boldsymbol{b} - A\widetilde{\boldsymbol{x}}^{(k+1)} = \widetilde{\boldsymbol{r}}^{(k)} - \widetilde{\alpha}_k A\widetilde{\boldsymbol{p}}^{(k)} \tag{12.36c′}$$

$$\widetilde{\boldsymbol{p}}^{(k+1)} = \widetilde{\boldsymbol{r}}^{(k+1)} + \widetilde{\beta}_k \widetilde{\boldsymbol{p}}^{(k)} \tag{12.36d′}$$

ただし $\widetilde{\alpha}_k$, $\widetilde{\beta}_k$ は実数で以下のように定める.

1^0. $\widetilde{\alpha}_k$ の決定：条件

$$\widetilde{\boldsymbol{p}}^{(k)^{\mathrm{t}}} \widetilde{\boldsymbol{r}}^{(k+1)} = 0 \quad (k \geq 0) \tag{12.39}$$

をみたすように定める. すなわち

$$\widetilde{\boldsymbol{p}}^{(k)^{\mathrm{t}}} \widetilde{\boldsymbol{r}}^{(k+1)} = \widetilde{\boldsymbol{p}}^{(k)^{\mathrm{t}}} (\widetilde{\boldsymbol{r}}^{(k)} - \widetilde{\alpha}_k A\widetilde{\boldsymbol{p}}^{(k)}) = 0$$

より

$$\widetilde{\alpha}_k = \frac{(\widetilde{\boldsymbol{p}}^{(k)},\ \widetilde{\boldsymbol{r}}^{(k)})}{(A\widetilde{\boldsymbol{p}}^{(k)},\ \widetilde{\boldsymbol{p}}^{(k)})} \quad (k \geq 0) \tag{12.40}$$

2^0. $\widetilde{\beta}_k$ の決定：(12.36c′) より

$$\widetilde{\boldsymbol{r}}^{(1)} = \widetilde{\boldsymbol{r}}^{(0)} - \widetilde{\alpha}_0 A\widetilde{\boldsymbol{p}}^{(0)} = \widetilde{\boldsymbol{r}}^{(0)} - \widetilde{\alpha}_0 A\widetilde{\boldsymbol{r}}^{(0)} \in \boldsymbol{r}^{(0)} + \mathrm{span}\{A\widetilde{\boldsymbol{r}}^{(0)}\}$$

$$\begin{aligned}\widetilde{\boldsymbol{r}}^{(2)} &= \widetilde{\boldsymbol{r}}^{(1)} - \widetilde{\alpha}_1 A\widetilde{\boldsymbol{p}}^{(1)} = \widetilde{\boldsymbol{r}}^{(0)} - \widetilde{\alpha}_0 A\widetilde{\boldsymbol{r}}^{(0)} - \widetilde{\alpha}_1 A\{\widetilde{\boldsymbol{r}}^{(1)} + \widetilde{\beta}_0 \widetilde{\boldsymbol{p}}^{(0)}\} \\ &= \widetilde{\boldsymbol{r}}^{(0)} - \widetilde{\alpha}_0 A\widetilde{\boldsymbol{r}}^{(0)} - \widetilde{\alpha}_1 A\{\widetilde{\boldsymbol{r}}^{(0)} - \widetilde{\alpha}_0 A\widetilde{\boldsymbol{p}}^{(0)} + \widetilde{\beta}_0 \widetilde{\boldsymbol{p}}^{(0)}\} \in \widetilde{\boldsymbol{r}}^{(0)} + \mathrm{span}\{A\widetilde{\boldsymbol{r}}^{(0)},\ A^2\widetilde{\boldsymbol{r}}^{(0)}\}\end{aligned}$$

以下これを続けて

$$\widetilde{\boldsymbol{r}}^{(k)} \in \widetilde{\boldsymbol{r}}^{(0)} + \mathrm{span}\{A\widetilde{\boldsymbol{r}}^{(0)},\ A^2\widetilde{\boldsymbol{r}}^{(0)}, \ldots, A^k\widetilde{\boldsymbol{r}}^{(0)}\}$$

よって

$$S_k = \mathrm{span}\{A\widetilde{\boldsymbol{r}}^{(0)},\ A^2\widetilde{\boldsymbol{r}}^{(0)}, \ldots, A^k\widetilde{\boldsymbol{r}}^{(0)}\}$$

$$V_k = \widetilde{\boldsymbol{r}}^{(0)} + S_k = \{\boldsymbol{v} = \widetilde{\boldsymbol{r}}^{(0)} + \boldsymbol{s} \mid \boldsymbol{s} \in S_k\}$$

とおくとき $\widetilde{\boldsymbol{r}}^{(k)} \in V_k$ $(k \geq 1)$.

ここで内積 (,)$_{A^{-1}}$ に関する S_k の直交補空間を $S_k^{\perp}$ であらわすとき, $\widetilde{\boldsymbol{r}}^{(k)} \in S_k^{\perp}$ となるように $\widetilde{\beta}_{k-1}$ を定めると (補題 12.1 によってこれは $\min_{\boldsymbol{v} \in V_k} ||\boldsymbol{v}||_{A^{-1}} = ||\widetilde{\boldsymbol{r}}^{(k)}||_{A^{-1}}$ を意味する)

$$(\widetilde{\boldsymbol{r}}^{(k)},\ \boldsymbol{s})_{A^{-1}} = 0 \quad \forall \boldsymbol{s} \in S_k$$

いま $\boldsymbol{w} = A^l\widetilde{\boldsymbol{r}}^{(0)}$ $(0 \le l \le k-1)$ とおけば $\boldsymbol{s} = A\boldsymbol{w} = A^{l+1}\widetilde{\boldsymbol{r}}^{(0)} \in S_k$ であるから

$$\widetilde{\boldsymbol{r}}^{(k)\mathrm{t}} A^l \widetilde{\boldsymbol{r}}^{(0)} = (\widetilde{\boldsymbol{r}}^{(k)},\ \boldsymbol{s})_{A^{-1}} = 0 \quad (0 \le l \le k-1) \tag{12.41}$$

$\widetilde{\boldsymbol{r}}^{(l)}$ は $\widetilde{\boldsymbol{r}}^{(0)},\ A\widetilde{\boldsymbol{r}}^{(0)},\ldots,A^l\widetilde{\boldsymbol{r}}^{(0)}$ の 1 次統合 ($\widetilde{\boldsymbol{r}}^{(0)}$ の係数は 1) として書けるから，$\widetilde{\boldsymbol{r}}^{(l)} = \sum_{j=0}^{l} c_j A^j \widetilde{\boldsymbol{r}}^{(0)}$ $(c_0 = 1)$ とおけば (12.41) より

$$(\widetilde{\boldsymbol{r}}^{(k)},\ \widetilde{\boldsymbol{r}}^{(l)}) = (\widetilde{\boldsymbol{r}}^{(k)},\ \sum_{j=0}^{l} c_j A^j \widetilde{\boldsymbol{r}}^{(0)}) = \sum_{j=0}^{l} c_j (\widetilde{\boldsymbol{r}}^{(k)},\ A^j \widetilde{\boldsymbol{r}}^{(0)}) = 0$$

$$\therefore \quad (\widetilde{\boldsymbol{r}}^{(k)},\ \widetilde{\boldsymbol{r}}^{(l)}) = 0 \quad (0 \le l \le k-1) \tag{12.42}$$

次に $l < k$ のとき

$$\begin{aligned}
\widetilde{\boldsymbol{p}}^{(k)^{\mathrm{t}}} A\widetilde{\boldsymbol{p}}^{(l)} &= (A\widetilde{\boldsymbol{p}}^{(k)})^{\mathrm{t}}\ \widetilde{\boldsymbol{p}}^{(l)} \\
&= \frac{1}{\widetilde{\alpha}_k}(\widetilde{\boldsymbol{r}}^{(k)} - \widetilde{\boldsymbol{r}}^{(k+1)})^{\mathrm{t}}\ (\widetilde{\boldsymbol{r}}^{(l)} + \widetilde{\beta}_{l-1}\widetilde{\boldsymbol{p}}^{(l-1)}) \\
&= \frac{1}{\widetilde{\alpha}_k}\widetilde{\beta}_{l-1}(\widetilde{\boldsymbol{r}}^{(k)} - \widetilde{\boldsymbol{r}}^{(k+1)})^{\mathrm{t}}\ \widetilde{\boldsymbol{p}}^{(l-1)} \qquad ((12.42)\text{ による}) \\
&= \frac{1}{\widetilde{\alpha}_k}\widetilde{\beta}_{l-1}(\widetilde{\boldsymbol{r}}^{(k)} - \widetilde{\boldsymbol{r}}^{(k+1)})^{\mathrm{t}}\ (\widetilde{\boldsymbol{r}}^{(l-1)} + \widetilde{\beta}_{l-2}\widetilde{\boldsymbol{p}}^{(l-2)}) \\
&= \frac{1}{\widetilde{\alpha}_k}\widetilde{\beta}_{l-1}\widetilde{\beta}_{l-2}(\widetilde{\boldsymbol{r}}^{(k)} - \widetilde{\boldsymbol{r}}^{(k+1)})\ \widetilde{\boldsymbol{p}}^{(l-2)} \\
&= \cdots \\
&= \frac{1}{\widetilde{\alpha}_k}(\widetilde{\beta}_{l-1}\widetilde{\beta}_{l-2}\cdots\widetilde{\beta}_0)(\widetilde{\boldsymbol{r}}^{(k)} - \widetilde{\boldsymbol{r}}^{(k+1)})\widetilde{\boldsymbol{p}}^{(0)} \\
&= 0 \quad (\because\ \widetilde{\boldsymbol{p}}^{(0)} = \widetilde{\boldsymbol{r}}^{(0)}\text{ と }(12.42)\text{ による})
\end{aligned}$$

$$\therefore \quad (\widetilde{\boldsymbol{p}}^{(k)},\ A\widetilde{\boldsymbol{p}}^{(l)}) = 0 \quad (0 \le l \le k-1) \tag{12.43}$$

(12.43) において k を $k+1$, l を k でおきかえれば

$$0 = \widetilde{\boldsymbol{p}}^{(k+1)^{\mathrm{t}}} A\widetilde{\boldsymbol{p}}^{(k)} = (\widetilde{\boldsymbol{r}}^{(k+1)} + \widetilde{\beta}_k\widetilde{\boldsymbol{p}}^{(k)})^{\mathrm{t}} A\widetilde{\boldsymbol{p}}^{(k)} \quad (k \ge 0)$$

$$\therefore \quad \widetilde{\beta}_k = -\frac{\widetilde{\boldsymbol{r}}^{(k+1)^{\mathrm{t}}} A\widetilde{\boldsymbol{p}}^{(k)}}{\widetilde{\boldsymbol{p}}^{(k)^{\mathrm{t}}} A\widetilde{\boldsymbol{p}}^{(k)}} = -\frac{(A\widetilde{\boldsymbol{p}}^{k},\ \widetilde{\boldsymbol{r}}^{(k+1)})}{(A\widetilde{\boldsymbol{p}}^{(k)},\ \widetilde{\boldsymbol{p}}^{(k)})} \tag{12.44}$$

ここで $k \ge 1$ のとき

$$\begin{aligned}
\widetilde{\boldsymbol{p}}^{(k)} &= \widetilde{\boldsymbol{r}}^{(k)} + \widetilde{\beta}_{k-1}\ \widetilde{\boldsymbol{p}}^{(k-1)} = \widetilde{\boldsymbol{r}}^{(k)} + \widetilde{\beta}_{k-1}(\widetilde{\boldsymbol{r}}^{(k-1)} + \widetilde{\beta}_{k-2}\ \widetilde{\boldsymbol{p}}^{(k-2)}) \\
&= \widetilde{\boldsymbol{r}}^{(k)} + \widetilde{\beta}_{k-1}\ \widetilde{\boldsymbol{r}}^{(k-1)} + \widetilde{\beta}_{k-1}\widetilde{\beta}_{k-2}(\widetilde{\boldsymbol{r}}^{(k-2)} + \widetilde{\beta}_{k-3}\ \widetilde{\boldsymbol{p}}^{(k-3)}) \\
&= \cdots
\end{aligned} \tag{12.45}$$

と展開すれば

$$\begin{aligned}
\widetilde{\boldsymbol{p}}^{(k)} &\in \widetilde{\boldsymbol{r}}^{(k)} + \mathrm{span}\{\widetilde{\boldsymbol{r}}^{(k-1)},\ \widetilde{\boldsymbol{r}}^{(k-2)},\ldots,\widetilde{\boldsymbol{r}}^{(1)},\ \widetilde{\boldsymbol{p}}^{(0)}\} \\
&= \widetilde{\boldsymbol{r}}^{(k)} + \mathrm{span}\{\widetilde{\boldsymbol{r}}^{(k-1)},\ \widetilde{\boldsymbol{r}}^{(k-2)},\ldots,\widetilde{\boldsymbol{r}}^{(1)},\ \widetilde{\boldsymbol{r}}^{(0)}\}
\end{aligned}$$

よって (12.39) と (12.42),(12.45) を用いて，$k \ge 1$ のとき

$$(A\widetilde{\boldsymbol{p}}^{(k)},\ \widetilde{\boldsymbol{p}}^{(k)}) = (\frac{1}{\widetilde{\alpha}_k}(\widetilde{\boldsymbol{r}}^{(k)} - \widetilde{\boldsymbol{r}}^{(k+1)}),\ \widetilde{\boldsymbol{p}}^{(k)}) = \frac{1}{\widetilde{\alpha}_k}(\widetilde{\boldsymbol{r}}^{(k)},\ \widetilde{\boldsymbol{r}}^{(k)})$$

上式は $k=0$ のときも成り立つ ((12.40) で $k=0$ としたものである).

さらに $k \geq 0$ のとき

$$(A\widetilde{\boldsymbol{p}}^{(k)},\ \widetilde{\boldsymbol{r}}^{(k+1)}) = \Big(\frac{1}{\widetilde{\alpha}_k}(\widetilde{\boldsymbol{r}}^{(k)} - \widetilde{\boldsymbol{r}}^{(k+1)}),\ \widetilde{\boldsymbol{r}}^{(k+1)}\Big)$$
$$= -\frac{1}{\widetilde{\alpha}_k}(\widetilde{\boldsymbol{r}}^{(k+1)},\ \widetilde{\boldsymbol{r}}^{(k+1)})$$

よって (12.44) より

$$\widetilde{\beta}_k = \frac{(\widetilde{\boldsymbol{r}}^{(k+1)},\ \widetilde{\boldsymbol{r}}^{(k+1)})}{(\widetilde{\boldsymbol{r}}^{(k)},\ \widetilde{\boldsymbol{r}}^{(k)})} \quad (k \geq 0) \tag{12.36f'}$$

また (12.42) と (12.45) によって

$$(\widetilde{\boldsymbol{p}}^{(k)},\ \widetilde{\boldsymbol{r}}^{(k)}) = (\widetilde{\boldsymbol{r}}^{(k)},\ \widetilde{\boldsymbol{r}}^{(k)}) \quad (k \geq 1)$$

であるが, $\widetilde{\boldsymbol{p}}^{(0)} = \widetilde{\boldsymbol{r}}^{(0)}$ であるから上式は $k=0$ のときも成り立つ. よって (12.40) より

$$\widetilde{\alpha}_k = \frac{(\widetilde{\boldsymbol{r}}^{(k)},\ \widetilde{\boldsymbol{r}}^{(k)})}{(A\widetilde{\boldsymbol{p}}^{(k)},\ \widetilde{\boldsymbol{p}}^{(k)})} \quad (k \geq 0) \tag{12.36e'}$$

結局 (12.36a)–(12.36f) と (12.36a')–(12.36f') を比較して

$$\widetilde{\boldsymbol{x}}^{(0)} = \boldsymbol{x}^{(0)} \Rightarrow \quad \widetilde{\boldsymbol{x}}^{(k)} = \boldsymbol{x}^{(k)},\quad \widetilde{\boldsymbol{r}}^{(k)} = \boldsymbol{r}^{(k)},\quad \widetilde{\boldsymbol{p}}^{(k)} = \boldsymbol{p}^{(k)},$$
$$\widetilde{\alpha}_k = \alpha_k,\quad \widetilde{\beta}_k = \beta_k,\quad k \geq 0$$

を得る. したがって (12.42) と (12.43) は (12.37) を意味し, 定理 12.8 が示された. 証明終 ∎

■**注意 2** 上記証明では暗黙のうちに α_k, β_k がすべて定義できるものと仮定した.
もしある k につき $\widetilde{\boldsymbol{p}}^{(k)} = 0$ かつ $\widetilde{\boldsymbol{p}}^{(j)} \neq 0$ $(j < k)$ ならば (12.36e') と (12.44) によって $\widetilde{\alpha}_j$, $\widetilde{\beta}_j$ $(j \leq k-1)$ は定義され

$$0 = \widetilde{\boldsymbol{r}}^{(k)}\widetilde{\boldsymbol{p}}^{(k)} = \widetilde{\boldsymbol{r}}^{(k)}(\widetilde{\boldsymbol{r}}^{(k)} + \widetilde{\beta}_{k-1}\widetilde{\boldsymbol{p}}^{(k-1)}) = \widetilde{\boldsymbol{r}}^{(k)}\widetilde{\boldsymbol{r}}^{(k)}$$
$$(\because\ (12.39) \text{ によって } \widetilde{\boldsymbol{r}}^{(k)^{\mathrm{t}}}\widetilde{\boldsymbol{p}}^{(k-1)} = 0)$$
$$\therefore\quad \widetilde{\boldsymbol{r}}^{(k)} = 0$$

これは $\widetilde{\boldsymbol{x}}^{(k)} = \widehat{\boldsymbol{x}}$ であることを意味し, 反復はここで停止する.

■**注意 3** 実は CG 法を最初から (12.36a')–(12.36f') により定義してもよかったのであるが, 見かけの異なる 2 通りの反復 $\{\boldsymbol{x}^{(k)}\}$ と $\{\widetilde{\boldsymbol{x}}^{(k)}\}$ が同一のものであるという事実は CG 法の奥行きが見かけより深いことを示唆している. なお, 現在出版されている書物の中には, 定理 12.8 に対する怪しい証明が記されているものがある. たとえ外国の有名な書物といえども信用してはいけない.

定理 12.9 (収束定理) A を n 次の正定値実対称行列とする.

(i) CG 法は高々 n 回のステップで $A\boldsymbol{x} = \boldsymbol{b}$ の真の解 $\widehat{\boldsymbol{x}}$ に収束する.

(ii) A の固有値を $\lambda_1 \geq \cdots \geq \lambda_n$ とすればエネルギーノルム $||\boldsymbol{x}||_A = \sqrt{(A\boldsymbol{x},\ \boldsymbol{x})}$ に関して

$$||\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}||_A \leq 2\left(\frac{\sqrt{\lambda_1} - \sqrt{\lambda_n}}{\sqrt{\lambda_1} + \sqrt{\lambda_n}}\right)^k ||\boldsymbol{x}^{(0)} - \widehat{\boldsymbol{x}}||_A$$

【証明】 (i) もしある k $(0 \leq k \leq n-1)$ につき $\boldsymbol{r}^{(k)} = \boldsymbol{0}$ ならば $\boldsymbol{x}^{(k)} = \widehat{\boldsymbol{x}}$ である. またもし $\boldsymbol{r}^{(k)} \neq \boldsymbol{0}$ $(0 \leq k \leq n-1)$ ならば定理 12.8 によって

$$\frac{\boldsymbol{r}^{(0)}}{||\boldsymbol{r}^{(0)}||_2},\ \frac{\boldsymbol{r}^{(1)}}{||\boldsymbol{r}^{(1)}||_2},\cdots,\ \frac{\boldsymbol{r}^{(n-1)}}{||\boldsymbol{r}^{(n-1)}||_2}$$

は n 次元空間 $\mathbb{R}^n$ の正規直交系をなすから，これらのすべてに直交するベクトル $\boldsymbol{r}^{(n)}$ は零ベクトルでなければならない．すなわち $\boldsymbol{x}^{(n)} = \widehat{\boldsymbol{x}}$.

(ii) 略す．(Strikwerda[16] を参照されたい．) 証明終 ∎

■**注意 4** 適当な正則行列 S をえらんで方程式 $A\boldsymbol{x} = \boldsymbol{b}$ を $SAS^{\mathrm{t}}\boldsymbol{y} = S\boldsymbol{b}$ $(\boldsymbol{x} = S^{\mathrm{t}}\boldsymbol{y})$ と書き直し $\widetilde{A} = SAS^{\mathrm{t}}$ の固有値を $\widetilde{\lambda}_1 \geq \cdots \geq \widetilde{\lambda}_n$ とする．このとき $\frac{\widetilde{\lambda}_1}{\widetilde{\lambda}_n} < \frac{\lambda_1}{\lambda_n}$ ならば

$$\frac{\sqrt{\widetilde{\lambda}_1} - \sqrt{\widetilde{\lambda}_n}}{\sqrt{\widetilde{\lambda}_1} + \sqrt{\widetilde{\lambda}_n}} < \frac{\sqrt{\lambda_1} - \sqrt{\lambda_n}}{\sqrt{\lambda_1} + \sqrt{\lambda_n}} \tag{12.46}$$

となるから，定理 12.9 (ii) によって収束は速くなる．S を**前処理行列 (preconditioning matrix)** という．このような S の選択は簡単ではないが，$A = LL^{\mathrm{t}}$ (L：下 3 角) と分解するとき，S を L^{-1} に近くえらべば $\widetilde{A}$ は単位行列に近くなって (12.46) が実現すると考えられる．

12.7 線形反復解の誤差分布

n 元連立方程式 (12.1) を解く反復 (12.2) または (12.2′) が，任意の初期値 $\boldsymbol{x}^{(0)}$ に対して，$k \to \infty$ のとき一意解 $\widehat{\boldsymbol{x}}$ に収束するための必要十分条件は

$$\rho(H) < 1$$

である．$\delta_k = ||\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}||_2$，$\varepsilon_k = ||\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}||_2$ とおけば $\boldsymbol{x}^{(k+1)} - \widehat{\boldsymbol{x}} = H(\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}})$ であるから

$$\delta_k \leq ||\boldsymbol{x}^{(k)} - \widehat{\boldsymbol{x}}||_2 + ||\widehat{\boldsymbol{x}} - \boldsymbol{x}^{(k+1)}||_2 \leq (1 + ||H||_2)\varepsilon_k$$

かつ

$$\varepsilon_k \leq ||\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k+1)}||_2 + ||\boldsymbol{x}^{(k+1)} - \widehat{\boldsymbol{x}}||_2 \leq \delta_k + ||H||_2\varepsilon_k$$

よって仮定 $||H||_2 < 1$ の下で

$$\frac{\delta_k}{1 + ||H||_2} \leq \varepsilon_k \leq \frac{\delta_k}{1 - ||H||_2}$$

したがって $\alpha > 0$ を与えられた定数 (反復停止条件数) として，

$$||\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}||_2 > \alpha\ (k \leq p-1) \quad \text{かつ} \quad ||\boldsymbol{x}^{(p+1)} - \boldsymbol{x}^{(p)}||_2 \leq \alpha$$

のとき反復を打ち切るものとすれば，最終誤差 $\varepsilon = \varepsilon_p$ は閉区間

$$\left[\ 0,\ \frac{\alpha}{1 - ||H||_2}\ \right] \tag{12.47}$$

の中に落ちる．初期値 $\boldsymbol{x}^{(0)}$ をいろいろ取りかえるとき ε はこの区間内にどのように分布するのであろうか．

H が正定値対称行列でその固有値を $1 > \lambda_1 > \cdots > \lambda_n > 0$ とすれば，$||H||_2 = \lambda_1$ であるから，(12.47) は $\left[\ 0,\ \frac{\alpha}{1-\lambda_1}\ \right]$ となる．1952 年 Krasnosel'skii-Krein (クラスノセルスキー・クライン) は次の定理を証明した ([12] 参照).

定理 12.10 (Krasnosel'skii-Krein) 反復 (12.2′) において H を正定値対称行列，その固有値を $1 > \lambda_1 > \cdots > \lambda_n > 0$ とする．また θ は $0 < \theta < 1$ として任意に与える．このとき $\boldsymbol{\varepsilon}^{(0)} = \widehat{\boldsymbol{x}} - \boldsymbol{x}^{(0)}$ を閉球 $\overline{S}(0,r) = \left\{ \boldsymbol{x} \in \mathbb{R}^n \;\middle|\; ||\boldsymbol{x}||_2 \leq r \right\}$ の中から一様にえらぶとき，最終誤差 $\boldsymbol{\varepsilon} = \widehat{\boldsymbol{x}} - \boldsymbol{x}^{(p)}$ が

$$\frac{\theta\lambda_1\alpha}{1-\lambda_1} \leq ||\boldsymbol{\varepsilon}||_2 \leq \frac{\alpha}{1-\lambda_1}$$

をみたす確率 P は $r \to \infty$ のとき 1 に近づく：

$$\lim_{r\to\infty} P\left(\frac{\theta_1\lambda_1\alpha}{1-\lambda_1} \leq ||\boldsymbol{\varepsilon}||_2 \leq \frac{\alpha}{1-\lambda_1}\right) = 1$$

$\frac{\lambda_1\alpha}{1-\lambda_1} = \frac{\alpha}{1-\lambda_1} - \alpha$ であるから上記定理は最終誤差 $\boldsymbol{\varepsilon}$ がほぼ確実に幅 α の閉区間 $\left[\ \frac{\alpha}{1-\lambda_1} - \alpha,\ \frac{\alpha}{1-\lambda_1}\ \right]$ の中に属することを主張する．したがって，遅い反復 ($\lambda_1 \fallingdotseq 1$) よりも早い反復 ($\lambda_1 \fallingdotseq 0$) のほうが得られる解の精度もよいのである．その後 Peradge(ペラージェ) (1968) はこの定理を H が実正則行列の場合に対して拡張した [34]．H が任意の実行列に対する拡張は [45] にある．また $||\cdot||_2$ を $||\cdot||_\infty$ にとりかえた場合の結果は [46] に示されている．

参考文献

和書

[1] 池辺 八洲彦・池辺淑子・浅井信吉・宮崎佳典：現代線形代数：共立出版 2009

[2] 伊理正夫：一般線形代数，岩波書店 2003

[3] 児玉慎三・須田信英：システム制御のためのマトリックス理論，コロナ社 1981

[4] 佐武一郎：線型代数学，裳華房 1974

[5] 大関信雄・青柳雅計：不等式，慎書店 1967

[6] 山本哲朗：数値解析入門 [増訂版]，サイエンス社 2003

[7] 山本哲朗：2 点境界値問題の数理，コロナ社 2006

洋書

[8] O. Axelsson and V. A. Barker : Finite Element Solution of Boundary Value Problems , SIAM 2001

[9] R. A. Horn and C. R. Johnson : Matrix Analysis , Cambridge Univ. Press 1985

[10] R. A. Horn and C. R. Johnson : Topics in Matrix Analysis , Cambridge Univ. Press 1989

[11] A. S. Householder : The Theory of Matrices in Numerical Analysis , Blaisdell Publishing Company 1964

[12] M. A. Krasnosel'skii, G. Vainikko, P.P.Zabreiko, Ya.B.Rutiskii and V.Ya. Stetsenko : Approximate Solution of Operator Equations , Wolters-Noordhoff Publishing 1972

[13] C. C. Macduffee : The Theory of Matrices , Chelsea Publishing 1946

[14] M. Marcus and H. Minc : A Survey of Matrix Theory and Matrix Inequalities , Allyn and Bascon, Inc. 1964

[15] D. Serre : Matrices : Theory and Applications , Springer 2002

[16] J. C. Strikwerda : Finite Difference Schemes and Partial Differential Equations, 2nd ed. , SIAM 2004

[17] A. E. Taylor : Introduction to Functional Analysis , MacGraw-Hille 1962

[18] J. Todd (ed.) : Survey of Numerical Analysis, 2nd ed. , Mac Graw-Hill 1962

[19] R.S.Varga : Matrix Iterative Analysis, 2nd ed. , Springer 2000

[20] D. M. Young : Iterative Solution of Large Linear Systems , Academic Press 1971

[21] F. Zhang : Matrix Theory , Springer 1999

論文 その他

[22] C. Davis : On a theorem of Yamamoto , Numer. Math.14 (1970), 297-298

[23] M. Fiedler and V. Pták : On matrices with non-positive off-diagonal elements and positire principal minors , Czech Math. J. 12 (1962), 382-400

[24] M. Fiedler and V. Pták : Some generalization of positive definiteness and monotonicity , Numer. Math. 9 (1966), 163-172

[25] A. J. Hoffman and H. Wielandt : The variation of the spectrum of a normal matrix , Duke Math. J. 20 (1953), 37-39

[26] Y. Ikebe, I. Inagaki and S. Miyamoto : The monotonicity theorem, Cauchy's interlace

theorem and the Courant-Fischer theorem , Amer. Math. Monthly 94 (1987), 352-354
[27] C. R. Johnson and P. Nylen : Yamamoto's theorem for generalized singular values , Linear Algebra Appl. 128 (1990), 147-158
[28] C. R. Johnson and P. Nylen : Erratum , Linear Algebra Appl. 180 (1993), 4
[29] H. König : A formula for the eigenvalues of a compact operator , Studio Math. 65 (1979), 141-146
[30] R. Mathias : Two theorems on singular values and eigenvalues , Amer. Math. Monthly 97 (1990), 47-50
[31] L. Mirsky : Proofs of two theorems on doubly-stochastic matrices , Proc. A.M.S. 9 (1958), 371-374
[32] T. Nishi : A theorem on an Ω-Matrix which is a generalization of the P-matrix , IEICE Trans. Fundamentals E79-A (1996) 1522-1529
[33] 西岡久美子：Jordan 標準形のわかりやすい求め方 , 数学 55 (2003), 424-429
[34] D. G. Peradze : The distribution of errors in solving a system of linear algebraic equations by the iteration method(Russian) , Sokharth. SSR Mech. Akad. Moambe 50(1968), 289-294
[35] N. V. Petri and K. D. Ikramov : Extremal properties of some matrix norms , USSR Comp. Math. & Math. Physics 8 (1968), 219-230
[36] I. W. Sandberg and A. N. Willson, Jr. : Some theorems on properties of DC equations of nonlinear networks , The Bell System Technical J. 48 (1969), 1-34
[37] I. W. Sandberg and A. N. Willson, Jr. : Some network-theoretic properties of nonlinear DC transister networks , The Bell System Technical J. 48 (1969), 1273-1311
[38] T. Y. Tam and H. Huang : An extension of Yamamoto's theorem on the eigenvalues and singular values of matrices , J. Math. Soc. Japan 58 (2006), 1197-1202
[39] O. Taussky and M. Marcus : Eigenvalues of finite matrices ([18] , 279-313)
[40] H. Weyl : Inequalities between the two kinds of eigenvalues of a linear transformation , Proc. Nat. Acad. Sci. 35 (1949), 408-411
[41] H. Wielandt : Unzerlegbare, nicht negativen Matrizen , Math. Z. 52 (1950), 642-648
[42] H. K. Wimmer : Powers of a matrix and a formula for the moduli of its eigenvalues , Linear & Multilinear Alg.71 (1985), 339-340
[43] T. Yamamoto : On the extreme values of the roots of matrices , J. Math. Soc. Japan 19 (1967), 173-178
[44] T. Yamamoto : A note on the spectral mapping theorem , SIAM J. Math. Anal. 2 (1971), 49-51
[45] T. Yamamoto : On the distribution of errors in the iterative solution of a system of linear equations , Numer. Math. 24 (1975), 71-79
[46] T. Yamamoto : On the distribution of errors in the iterative solution of a system of linear equations, II , Numer. Math. 25 (1976), 461-463
[47] 山本哲朗：行列と数値解析 , 数理科学 No.407 (1997 年 5 月号), 54-61
[48] 山本哲朗：行列雑話 , 数理科学 No.540 (2008 年 6 月号), 52-53

主要 定理 (系) 一覧

主要 定理 (系)		記載の節
Autonne の定理	(オートン)	§3.4 定理 3.6
Binet-Cauchy の定理	(ビネ・コーシー)	§1.3 定理 1.4
Birkhoff の定理	(バーコフ)	§1.13 定理 1.22
Cauchy-Schwarz の不等式	(コーシー・シュワルツ)	§1.7 定理 1.14
Courant-Fischer の定理	(クーラン・フィッシャー).....	§6.3 定理 6.8
Cramer の公式	(クラメール)	§1.3
Fan-Hoffman の定理	(ファン・ホフマン)	§6.2 定理 6.7
Fiedler-Pták の定理	(フィードラー・プタック).....	§8.3 定理 8.8
		§9.1 定理 9.1
		§9.1 系 9.1.3
Frobenius の定理	(フロベニウス)	§7.1 系 7.2.1
		§7.3 定理 7.5
		§7.4 定理 7.6, 7.7
		§7.5 定理 7.8
Gerschgorin の定理	(ゲルシュゴリン)	§6.1 定理 6.1
Hahn の定理	(ハーン)	§10.3 定理 10.8
Hall-Porshing の定理	(ホール・ポーシング)	§6.5 定理 6.10
		§8.5 定理 8.11
Hamilton-Cayley の定理	(ハミルトン・ケイリー).....	§2.3 定理 2.5
Hoffman-Wielandt の定理	(ホフマン・ウィーラント)	§6.6 定理 6.13, 6.14
Hölder の不等式	(ヘルダー)	§4.3 補題 4.1
Kahan の定理	(カハン)	§12.3 定理 12.2
Kantorovich の不等式	(カントロヴィッチ)	§5.5 定理 5.8, 5.10
Krasnosel'skii-Krein の定理 ..	(クラスノセルスキー・クライン)	§12.7 定理 12.10
Liapunov の定理	(リアプノフ)	§10.3 定理 10.10
Minkowski の不等式	(ミンコフスキー)	§4.3 補題 4.2
西の定理	(ニシ)	§9.4 定理 9.7

主要 定理 (系)		記載の節
Oldenburger の定理	(オルデンブルガー)	§4.4 系 4.5.2
Ostrowski の定理	(オストロフスキー)	§6.1 定理 6.3
		§12.3 定理 12.3
		§12.4 定理 12.4
Parker の定理	(パーカー)	§6.2 系 6.4.1
Perron-Frobenius の定理	(ペロン・フロベニウス).....	§7.2 定理 7.4
Petri-Ikramov の定理	(ペトリ・イクラモフ)	§2.5 注意
Rheinboldt の不等式	(レインボルト)	§5.5 定理 5.9, 5.10
Sandberg-Willson, Jr. の定理	(サンドベルグ・ウィルソン)....	§9.2 定理 9.2
		§9.1 系 9.1.1
		§9.2 系 9.3.2
		§9.2 定理 9.4
		§9.3 定理 9.5
Schur の定理	(シュア)	§2.2 定理 2.3
		§2.5 定理 2.9
SOR 法の収束 (Young の理論)	(ヤング)	§12.4 定理 12.5
Stein の定理	(スタイン)	§10.2 系 10.7.1
Sylvester の慣性則第 1 型	(シルベスター)	§2.8 定理 2.15
Sylvester の慣性則第 2 型	(シルベスター)	§2.8 系 2.15.1
Taussky の定理	(タウスキー)	§6.1 定理 6.2
Weyl の定理	(ワイル)	§2.5 注意
		§6.5 定理 6.12
Wielandt の補題	(ウィーラント)	§7.1 定理 7.3
Young の定理	(ヤング)	§12.4 定理 12.5
収束定理		§12.5 定理 12.6
		§12.5 定理 12.9
摂動定理 (perturbation theorem)		§6.5 定理 6.10
分離定理 (separation theorem)		§6.4 定理 6.9

索　引

ア

アセント (昇数)　30
アダマール積　166
イエンゼンの不等式　82
一般逆行列　68
一般固有空間　→ 広義固有空間
ウィーラントの補題　125, 126
上 3 角行列　59
エルミート行列　1, 51
エルミート形式　58, 110
エルミート 2 次形式　→ エルミート形式
オートンの定理　66
オストロフスキーの定理　109, 188, 189
オメガ行列　161
オルデンブルガー　86

カ

開円板　89
解空間　27
階数　→ 行列のランク
解の定性的評価　92
ガウス・ザイデル法　187
　—の緩和　187
可換　96
可換 (交換可能)　2
核　25
過小緩和パラメータ　188
過大緩和パラメータ　188
カハンの定理　188
可約　34
関数ノルム　25
慣性数 $\mathrm{In}(A)$　56, 96
カントロヴィッチの不等式　104, 106
緩和パラメータ　188
　過小—　188
　過大—　188
幾何学的重複度　40
奇置換　5
基底　13
基本直交行列　65
基本不等式　110, 111
基本ベクトル　4
既約　34
既約強優対角行列　36, 109
既約行列　33
逆行列　3
　一般—　68
逆置換　3
既約非負行列　129, 132
既約優対角行列　36
行 (行列の)　1
境界値問題　171
狭義優対角行列　36
共通因子　11
共役　196
共役勾配法 (CG 法)　195, 197
共役転置行列　1
共役転置ベクトル　1
強優対角行列　36
行列　1
　実—　1
　複素—　1
行列関数　87
行列式　4
行列式の微分　93
行列の 3 角化　41
行列の演算　2
行列の積分　92
行列の対角化　41
行列の微分　92
行列のひろがり (spread)　111
行列のランク (階数)　27
行列ノルム　77, 85
行列表現　14
行列方程式　166, 168

極分解　68, 69
偶置換　5
クーラン・フィッシャーのミニマックス定理　113
クラスノセルスキー・クラインの定理　201
グラスマン積　174, 175
グラム・シュミットの直交化法　19, 20
グラム行列　20
グラム行列の性質　55
クラメル (クラーメル) の公式　8
クロネッカー積　→ テンソル積
クロネッカーの記号　19
ゲルシュゴリンの定理　107
原始的　132
交換可能　→ 可換
広義固有空間　40
広義単調増加関数　157
降数　→ ディセント
交代行列　2
恒等置換　5
コーシー・シュワルツの不等式　23
コーシー列　22
互換　5
誤差評価　194
固有空間　40
固有値　39, 85, 94
固有ベクトル　40
コレスキー分解　64

サ

最急降下法　193
　—の加速　195
最小多項式　46
最小 2 乗解　68
最大値ノルム　24
最大ノルム　77
差分近似行列　171
作用素　78
作用素ノルム　→ 従属する行列ノルム
サンドベルグ・ウィルソンの定理　150, 152, 154–156
次元等式　15
下 3 角行列　59
実対称行列　1
実内積　18
実内積空間　18
シュアの定理　41, 50
シュア補元　32, 121
終結式　10
従属する行列ノルム　78, 80
収束定理　193, 200
首座小行列　6, 28
首座小行列式　7
主小行列　6
主小行列式　7
巡回行列　132
昇数　→ アセント
ジョルダンの標準形　69, 74
ジョルダンブロック　74
ジョルダン分解　69, 74
シルベスターの慣性則　56, 58
　—第 1 型　56
　—第 2 型　58
シルベスターの行列式　10
スカラー積　2
スタインの定理　168
スチルチェス行列　141
スペクトルノルム　77
スペクトル半径　85
正規行列　3
正規直交基底　19
正規直交系　19
正規直交固有ベクトル　110
正規直交底　19
正行列　123, 126
整合順序行列　188, 189
生成元　13
正則 M 行列　144
正則行列　3
正対角行列　141
正定値行列　52
正ベクトル　123
正方行列　1
摂動定理　116
ゼロ行列　→ 零行列
ゼロ空間　→ 零空間
ゼロ元　→ 零元
線形　13
線形空間　11
　実—　12
　複素—　12

タ

対角化　43

対角行列　2
対角優位行列　36
対角要素　2
対称行列　1
代数的重複度　40, 94
タウスキーの定理　108
単位上 3 角行列　59
単位行列　2
単位下 3 角行列　59
単位置換　5
単調行列　137
値域　25
置換行列　3
逐次過大緩和法 (SOR 法)　188
重複度　40
直和　16
直交行列　2
直交補空間　18
ディセント (降数)　31
テイラー級数　89
テンソル積　163, 166
転置行列　1
転置ベクトル　1
特異 M 行列　144
特異値　50
特異値分解　66, 67
特性多項式　39, 94
特性方程式　39
凸 1 次結合　37
凸集合　37
トレース (trace)　39

ナ

内積　18
内積空間　18
西の定理　162
ノルム　21
ノルム空間　21
ノルム線形空間　21

ハ

パーカー　111
バーコフの定理　37
ハーンの定理　168
ハウスホルダー行列　65
ハミルトン・ケイリーの定理　45
半正定値 Hermite 行列　99
半正定値行列　52
反復法　195
ピカール反復　93
非対角要素　2
ビネ・コーシーの定理　8
非負行列　123, 145
非負行列に対する分離定理　145
非負対角行列　141
非負ベクトル　123
ひろがり　→ 行列のひろがり
ファン・ホフマンの定理　112
フィードラー・プタックの定理　142, 148, 150
複合行列　174, 178, 182
複素 Hermite 行列　1
複素内積空間　18
符号数　56
部分空間　14, 18
ブロック 3 重対角行列　189
ブロック行列　31
ブロック小行列　31
ブロック対角行列　31
フロベニウスの定理　125, 128, 129, 131, 132
フロベニウスノルム　77
フロベニウスノルムの拡張　80
分離定理　114, 115, 145
平方根 $\sqrt{A}$　101
ベクトルノルム　24, 78
ペトリ・イクラモフの定理　51
ヘルダーの逆向き不等式　84
ヘルダーの不等式　81
ペロン根　126, 128
ペロン・フロベニウス根　128
ペロン・フロベニウスの定理　126
ペロン・フロベニウスの理論　123
ペロンベクトル　126
ホール・ポーシングの定理　116, 146
ホフマン・ウィーラントの定理　38, 118, 120, 121

マ

前処理行列　201
ミニマックス定理　113
ミンコフスキーの逆向き不等式　84
ミンコフスキーの不等式　24, 81
無限次元空間　13
無限ノルム $\|\cdot\|_\infty$　22

ヤ

ヤコビ行列　186
ヤコビ法　186
　緩和ヤコビ法　186
ヤングの定理　190
ユークリッドノルム　24
有限差分方程式　172
有限次元線形空間　13
有限次元ノルム空間　22
有限次元ノルムの同値性　22
有限有向グラフ　35
優対角行列　36
ユニタリ行列　2
余因子　7
余因子展開　7

ラ

リアプノフの定理　170
両立する行列ノルム　80
零行列 (ゼロ行列)　1
零空間 (ゼロ空間)　25, 27
零元 (ゼロ元)　12
レインボルトの不等式　105, 106
列 (行列の)　1
連続関数　159
連立 1 階線形常微分方程式　91
連立準線形方程式　156

ワ

歪 Hermite 行列　2
歪対称行列　1
ワイルの定理　51, 118

欧字

1 次従属　12
1 次独立　12
1 ノルム　77
2 次形式　55
2 重確率行列　37
3 重対角行列　49
3 対角行列　→ 3 重対角行列

Autonne の定理　66

Binet-Cauchy の定理　8
Birkhoff の定理　37

Cauchy-Schwarz の不等式　23
Cauchy 列　22
CG 法　195, 197
Cholesky 分解　64
Courant-Fischer のミニマックス定理　113
Cramer の公式　8

Euclid ノルム　24

Fan-Hoffman の定理　112
Fiedler-Pták の定理　142, 148, 150
Frobenius の定理　125, 128, 129, 131, 132
Frobenius ノルム　77
Frobenius ノルムの拡張　80

Gauss-Seidel 法　187
　—の緩和　187
Gerschgorin の定理　107
Gram-Schmidt の直交化法　19, 20
Gram 行列　20
Gram 行列の性質　55
Grassmann 積　174, 175

H 行列　141
Hadamard 積　166
Hahn の定理　168
Hall-Porshing の定理　116, 146
Hamilton-Cayley の定理　45
Hermite 行列　1, 51
Hermite 形式　58, 110
Hermite 2 次形式　→ Hermite 形式
Hoffman-Wielandt の定理　38, 118, 120, 121
Hölder の逆向き不等式　84
Hölder の不等式　81
Householder 行列　65

Jacobi 行列　186
Jacobi 法　186
Jensen の不等式　82
Jordan の標準形　69, 74
Jordan ブロック　74
Jordan 分解　69, 74

Kahan の定理　188
Kantorovich の不等式　104, 106
Krasnosel'skii-Krein の定理　201
Kronecker 積　→ テンソル積
Kronecker の記号　19

L 行列　145
LDU 分解　62
LDV 分解　62
Liapunov の定理　170
LL^t 分解　63
LU 分解　59, 61, 136

M 行列　136, 137, 141
Minkowski の逆向き不等式　84
Minkowski の不等式　24, 81

Ω 行列　161
Oldenburger　86
Ostrowski の定理　109, 188, 189

P 行列　148
P_0 行列　148, 151
Parker　111
Perron-Frobenius 根　128
Perron-Frobenius の定理　126
Perron-Frobenius の理論　123
Perron 根　126, 128
Perron ベクトル　126
Petri-Ikramov の定理　51
Picard 反復　93
p 次複合行列　178
p 乗根 $\sqrt[p]{A}$　101
p 乗根の計算法　102

QR 分解　65

Rheinboldt の不等式　105, 106

S 行列　141
Sandberg-Willson, Jr. の定理　150, 152, 154–156
Schur の定理　41, 50
Schur 補元　32, 121
SOR 法 (逐次過大緩和法)　188
SOR 法の収束 (Young の理論)　190
Stein の定理　168
Stieltjes 行列　141
Sylvester の慣性則　56, 58
　—第 1 型　56
　—第 2 型　58
Sylvester の行列式　10

Taussky の定理　108
Taylor 級数　89

Weyl の定理　51, 118
Wielandt の補題　125, 126

Young の定理　190

Z 行列　136

著者略歴

山本 哲朗
やまもとてつろう

1961 年　広島大学大学院理学研究科修士課程（数学専攻）修了
1968 年　理学博士
1975 年　愛媛大学理学部教授
2002 年　愛媛大学名誉教授
　　早稲田大学理工学部専任客員教授
2005 年　早稲田大学退職（2007 年まで非常勤客員教授）
　専門　数値解析

主要著書

数値解析演習（共著，サイエンス社，1991）
英語で学ぶ数値解析（共著，コロナ社，2002）
数値解析入門 [増訂版]（サイエンス社，2003）
2 点境界値問題の数理（コロナ社，2006）

臨時別冊・数理科学 SGC ライブラリ - 79
『行列解析の基礎　Advanced 線形代数』
著　者　山本 哲朗
2010 年12月25日　初版発行
2014 年 4 月10日　初版第 2 刷発行

数 理 科 学 編 集 部　　発行人　木 下 敏 孝
TEL.(03)5474–8816
FAX.(03)5474–8817

ホームページ　http://www.saiensu.co.jp
ご意見・ご要望は　sk@saiensu.co.jp まで.　　表紙デザイン：長谷部貴志

発行所　© 株式会社 サ イ エ ン ス 社　　TEL.(03)5474–8500(代表)
〒 151–0051　東京都渋谷区千駄ヶ谷 1–3–25　振替 00170–7–2387　(03)5474–8600(広告部)

本誌の内容を無断で複写複製・転載することは，著作者および出版者の権利を侵害することがありますので，その場合にはあらかじめサイエンス社著作権担当者あて許諾をお求めください.

組版　ビーカム
印刷　(株) シナノ　中澤 眞
製本　(株) 関川製本所　関川安博