ベクトルと行列の演算について

前川眞一

000313,30,0403,13,0502,030721,1210,041013,050304,070805, 080206,08,09,10,11,14,090308,14,15,
20191220, 20200422,30,0604,18,0709,20230417, 20240701,02,03,20250627, 2025. 6. 27

目 次

1 ベクトル **3**
1.1 ベクトルの演算 . . . 4
1.1.1 基本演算 . . . 4
1.1.2 ベクトルの内積 . . . 4

2 行列 **5**
2.1 行列の転置 . . . 6
2.2 行列の種類 . . . 7
2.3 行列の演算 . . . 7
2.3.1 基本演算 . . . 7
2.3.2 行列とベクトルの積 . . . 7
2.3.3 行列同士の積 1 . . . 8
2.3.4 行列同士の積 2 . . . 10
2.3.5 行列の積の一般形 . . . 11
2.3.6 演算規則 . . . 14
2.4 簡易表記と添字演算子 . . . 14
2.5 ベクトルの２次形式 . . . 15
2.6 行列のトレース . . . 15
2.7 diag 演算子 . . . 15
2.8 並べ替え行列 . . . 16

3 逆行列 **17**
3.1 逆行列の求め方 . . . 17
3.1.1 ガウスの消去法 Gaussian Elimination . . . 17
3.1.2 スウィープ演算子 (Sweep Operator) の利用 . . . 20
3.1.3 その他の方法 . . . 23
3.2 偶に使う逆行列の公式 . . . 24
3.3 行列式 . . . 25

4 行列の列空間・基底・直交化 **26**
4.1 行列の列空間 . . . 26
4.2 回帰分析と射影子 . . . 27
4.2.1 射影子の幾何学的解釈 . . . 30
4.3 行列の直交基底 . . . 34
4.3.1 グラム・シュミットの直交化 The Gram-Schmidt Orthogonalization . . . 34
4.3.2 列ブロック単位での直交化 . . . 37
4.4 ベクトルの線形独立と行列の階数 . . . 38

5 一般逆行列について **40**
5.1 はじめに . . . 40
5.2 一般逆行列の性質 . . . 41
5.3 一般逆行列と射影演算子 . . . 42
5.3.1 一般逆行列の性質 2 . . . 44
5.4 一般逆行列の生成 . . . 45
5.4.1 一般逆行列から一般逆行列 . . . 47
5.5 ムーアペンローズ逆行列 . . . 47
5.5.1 ムーアペンローズ逆行列の生成 . . . 47
5.5.2 ムーアペンローズ逆行列の性質 . . . 48

6 連立一次方程式の一般解と一般逆行列 **49**
6.1 連立一次方程式の一般解 . . . 49
6.1.1 列フルランクの場合 . . . 50
6.1.2 共通の係数行列 $\boldsymbol{A}$ を持つ複数の連立方程式 . . . 51
6.2 等号成立条件 . . . 52

7 行列の利用 **53**
7.1 座標軸や図形の回転 . . . 53
7.2 度数分布表・クロス集計表（同時確率・周辺確率・条件付き確率） . . . 53
7.3 デザイン行列 . . . 54
7.4 距離行列 . . . 56
7.5 推移確率行列 . . . 56
7.6 ソシオメトリの行列 . . . 57
7.7 データ行列の要約統計量 . . . 57
7.8 確率変数と確率変数ベクトル . . . 59
7.9 因子分析 . . . 60
7.10 多変量正規分布の密度関数 . . . 61
7.11 多項分布の確率 . . . 61

8 その他の演算子 **62**
8.1 分割行列の演算 . . . 62
8.2 ssq 演算子 . . . 64
8.3 Kronecker 積 . . . 66
8.4 vec 演算子 . . . 67
8.5 逆行列の演算規則といくつかの公式 . . . 68

9 スイープ演算子と逆行列ならびに行列式 **69**
9.0.1 スイープ演算子の定義 . . . 69
9.0.2 分割行列のスイープ . . . 71
9.0.3 スイープ演算子を利用した様々な公式 . . . 72

10 行列の固有値と特異値分解 **74**
10.1 固有値分解 . . . 75
10.2 固有値分解の応用例 . . . 77
10.3 特異値分解 . . . 79
10.3.1 完全な特異値分解 full-size svd, complete svd . . . 79
10.3.2 サブコンパクトな特異値分解 . . . 80
10.3.3 コンパクトな特異値分解 compact, economy, or thin svd . . . 81
10.3.4 切断された特異値分解 compact, economy, or thin svd . . . 81

11 行列の微分 **81**

12 線形モデルと最小２乗解 **84**
12.1 線形モデルと最小２乗解 . . . 84
12.2 微分による解法 . . . 86

A 補足 **87**
A.1 実対称行列の固有値は実数である . . . 87
A.2 実対称行列の階数と 0 でない固有値の数 . . . 87
A.3 正定値行列の固有値は正である。 . . . 87
A.4 $\boldsymbol{AA}'$ が正定値であること . . . 88
A.5 $\boldsymbol{AA}'$ と $\boldsymbol{A}'\boldsymbol{A}$ の固有値分解の関係 . . . 88
A.5.1 特異値分解と行列の階数 . . . 89

1　ベクトル

数字を縦または横に一列に並べたものをまとめてベクトル (vector) と呼ぶ。たとえば、

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 6 \end{pmatrix} \qquad \boldsymbol{b} = \begin{pmatrix} 2, & 4, & 5 \end{pmatrix} \tag{1}$$

の様なものであり、$\boldsymbol{a}$ を列ベクトル（または縦ベクトル）、$\boldsymbol{b}$ を行ベクトル（または横ベクトル）と呼ぶ。行ベクトルがきれいにフォーマットされていれば、コンマを省いても良いが、通常は区切りにコンマを用いる。

一般的に、ベクトルの中に含まれる数値のことを**要素** (element) と呼び、その数が縦方向に r 個のものを $r \times 1$ の**列ベクトル** (column vector)（ r 行 1 列のベクトル）、横方向に 個のものを $1 \times c$ の**行ベクトル** (row vector)（ 1 行 c 列のベクトル）と呼ぶ。上記の $\boldsymbol{a}$ は 4×1 の列ベクトル、$\boldsymbol{b}$ は 1×3 の行ベクトルである。

一般的にベクトルの名前には太字の英小字一文字をあて、その要素を添え字付きの英小文字で

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \tag{2}$$

の様に表す。もしくは、 a_i を要素とする $r \times 1$ の列ベクトルという表現をする。いろいろなスタイルがあるが、多くの場合、単に $\boldsymbol{a}$ と書くと列ベクトルを表すことが多い。

$r \times 1$ の列ベクトルを $1 \times r$ の行ベクトルに書き直す作業、もしくはその逆の作業をベクトルを**転置**する (transpose) と言い、ベクトルの右肩にアポストロフィ ($'$) をつけて表す。（他にも右肩や左肩に T を付ける流儀もある。）たとえば、

$$\boldsymbol{a}' = \begin{pmatrix} 1, & 3, & 5, & 6 \end{pmatrix} \qquad \boldsymbol{b}' = \begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix} \tag{3}$$

である。この記法を用いれば、 (1) 式 に現れる列ベクトル $\boldsymbol{a}$ は

$$\boldsymbol{a} \quad = \quad \begin{pmatrix} 1, & 3, & 5, & 6 \end{pmatrix}' \tag{4}$$

または

$$\boldsymbol{a}' \quad = \quad \begin{pmatrix} 1, & 3, & 5, & 6 \end{pmatrix} \tag{5}$$

という形書くことが出来、紙面の節約となる。また、これを用いると $r \times 1$ のベクトルは

$$\boldsymbol{a} \quad = \quad \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{pmatrix} \qquad \begin{array}{c} \boldsymbol{a} = (a_1, a_2, \ldots, a_r)' \\ \text{または} \\ \boldsymbol{a}' = (a_1, a_2, \ldots, a_r) \end{array} \tag{6}$$

と書ける。なお、ベクトル $\boldsymbol{a}$ の第 i 番目の要素を a_i というように添え字 (subscript) を付けた形で表す。

転置演算子に関しては、明らかに

$$(\boldsymbol{a}')' = \boldsymbol{a} \tag{7}$$

である。

全ての要素が 1 である $r \times 1$ のベクトル r 次の**単位ベクトル** (unit vector) と呼び $\mathbf{1}_r$ で表す。サイズが明らかな場合には単に $\mathbf{1}$ を書く場合がある。同様に、全ての要素が 0 であるベクトルを**ゼロベクトル**と呼び、$\mathbf{0}_r$ もしくは $\mathbf{0}$ と記す。なお、単位ベクトルという名称は様々な意味で用いられており、たとえば、長さが 1 のベクトルや、一つの要素が 1 で他の要素は全て 0 のベクトルをさすこともある。

以下の示すベクトルの演算では、それが列ベクトルであるか行ベクトルであるかを、常に意識していることが大切である、この意味で、全てのベクトルは列ベクトルとし、行ベクトルは列ベクトルの転置として表すことがおおい。本稿でも、特に但し書きを付けない限り、ベクトルは列ベクトルとする。

1.0.0.1　例　ある人の身体測定の結果（身長 h、体重 w、座高 s を記録したもの）はベクトル $\boldsymbol{m} = (h, w, s)'$ で表され、これは 3 次元のベクトルとなる。また、平面上の点の座標 $\boldsymbol{x} = (x, y)'$ は 2 次元のベクトルとなる。

1.1 ベクトルの演算

1.1.1 基本演算

サイズが等しい二つの $r \times 1$ のベクトル

$$\boldsymbol{a} = \begin{pmatrix} a_1, & a_2, & \ldots, & a_r \end{pmatrix}' \qquad \boldsymbol{b} = \begin{pmatrix} b_1, & b_2, & \ldots, & b_r \end{pmatrix}' \tag{8}$$

の**和**もしくは**差**は

$$\boldsymbol{a} \pm \boldsymbol{b} = \begin{pmatrix} a_1 \pm b_1, & a_2 \pm b_2, & \ldots, & a_r \pm b_r \end{pmatrix}' \tag{9}$$

となり、 $r \times 1$ のベクトルとなる。

ベクトルとスカラー数 (scalar) λ の**積** は、各要素とそのスカラー数との積で定義される。

$$\lambda\boldsymbol{a} = \begin{pmatrix} \lambda a_1, & \lambda a_2, & \ldots, & \lambda a_r \end{pmatrix}' \tag{10}$$

なお、λ で割ること **商** は $\frac{1}{\lambda}$ を掛けることに等しい。

ベクトルの演算には以下の規則が成り立つ。

$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}, \qquad (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}), \tag{11}$$

$$\lambda\boldsymbol{a} = \boldsymbol{a}\lambda \qquad (\lambda + \mu)\boldsymbol{a} = \lambda\boldsymbol{a} + \mu\boldsymbol{a}, \qquad \lambda(\boldsymbol{a} + \boldsymbol{b}) = \lambda\boldsymbol{a} + \lambda\boldsymbol{b}, \qquad \lambda(\mu\boldsymbol{a}) = (\lambda\mu)\boldsymbol{a} \tag{12}$$

また、スカラーとベクトルの加減算は、$r \times 1$ のベクトル $\boldsymbol{a}$ の全ての要素にスカラー量 λ を足したり引いたりする演算は、単位ベクトル $\mathbf{1}_r$ を用いて

$$\boldsymbol{a} \pm \lambda\mathbf{1}_r \qquad \text{or} \qquad \boldsymbol{a} \pm +\mathbf{1}_r\lambda \tag{13}$$

と書くことが出来る。

1.1.1.1 例

$$\begin{pmatrix} 1, & 2, & 3 \end{pmatrix} + \begin{pmatrix} 4, & 5, & 6 \end{pmatrix} = \begin{pmatrix} 5, & 7, & 9, \end{pmatrix} \tag{14}$$

二人の人の身体測定値 $\boldsymbol{m}_1 = (h_1, w_1, s_1)'$ と $\boldsymbol{m}_2 = (h_2, w_2, s_2)'$ の平均は

$$m. = \frac{1}{2}(\boldsymbol{m}_1 + \boldsymbol{m}_2) = \frac{1}{2}\left((h_1, w_1, s_1)' + (h_2, w_2, s_2)'\right) \tag{15}$$

$$= \frac{1}{2}(h_1 + h_2, w_1 + w_2, s_1 + s_2)' = \left(\frac{h_1 + h_2}{2}, \frac{w_1 + w_2}{2}, \frac{s_1 + s_2}{2}\right)' \tag{16}$$

となる。一般的に、 p 個のベクトル $\boldsymbol{m}_i, i = 1, 2, \ldots, p$ の平均値は

$$m. = \frac{1}{p}\sum_{i=1}^{p}\boldsymbol{m}_i \tag{17}$$

となる。ただし、 $\boldsymbol{m}_i$ は添え字が付いているが、それ自身

$$\boldsymbol{m}_i = \begin{pmatrix} h_i, w_i, s_i \end{pmatrix}' \tag{18}$$

と表される、第 i 番目の人の測定値を表すベクトルである。

このように、ベクトルを用いることにより、スカラーのみで表すと表現が煩雑になるものを簡素に表現することが出来る。

1.1.2 ベクトルの内積

$1 \times r$ の横ベクトル $\boldsymbol{a}'$ に $r \times 1$ の縦ベクトル $\boldsymbol{b}$ を右から掛けたものをベクトルの**内積** (scalar product, inner product) と呼び、

$$\boldsymbol{a}'\boldsymbol{b} = \begin{pmatrix} a_1, & a_2, & \ldots, & a_r \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \end{pmatrix} = \sum_{i=1}^{r} a_i b_i \tag{19}$$

$$= a_1 \times b_1 + a_2 \times b_2 + \ldots + a_r \times b_r \tag{20}$$

と定義する。内積はスカラー数であるため、**スカラー積**とも呼ばれる。したがって、

$$\boldsymbol{a}'\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{a} \tag{21}$$

である。（内積を $(\boldsymbol{a}, \boldsymbol{b})$ と記す場合があるが本稿では使用しない。）
　特に、行ベクトルと単位ベクトルの積は

$$\boldsymbol{a}'\mathbf{1} \;=\; \mathbf{1}'\boldsymbol{a} \;=\; \sum_{i=1}^{r} a_i \times 1 \;=\; \sum_{i=1}^{r} a_i \tag{22}$$

となり、全ての要素の和となる。また、自分自身の転置との内積は

$$\boldsymbol{a}'\boldsymbol{a} \;=\; \sum_{i=1}^{r} a_i \times a_i \;=\; \sum_{i=1}^{r} a_i^2 \tag{23}$$

となり、全ての要素の２乗和となる。この平方根をベクトルの長さ（ノルム）(norm) と呼び、$\|\boldsymbol{a}\|$ と書く。

1.1.2.1　例

$$\begin{array}{ccc} & & \begin{pmatrix} 3 \\ 7 \end{pmatrix} \\ & \times & \\ \begin{pmatrix} 4 & 5 \end{pmatrix} & = & 4 \times 3 + 5 \times 7 \quad = \quad 47 \end{array} \tag{24}$$

　ベクトル $\boldsymbol{a}$ の要素の**平均** m と**分散** s^2 は

$$m = \frac{1}{r}\,\boldsymbol{a}'\mathbf{1}, \qquad s^2 = \frac{1}{r}\boldsymbol{a}'\boldsymbol{a} - m^2 \tag{25}$$

と書ける。
　内積に関しては以下の演算規則が成立する。

$$\boldsymbol{a}'(\boldsymbol{b}+\boldsymbol{c}) \;=\; \boldsymbol{a}'\boldsymbol{b} + \boldsymbol{a}'\boldsymbol{c} \tag{26}$$

$$(\lambda\boldsymbol{a})'\boldsymbol{b} \;=\; \lambda(\boldsymbol{a}'\boldsymbol{b}) \tag{27}$$

また、幾何学的に二つのベクトル $\boldsymbol{a}$ と $\boldsymbol{b}$ が r 次元空間において成す角を θ とすれば、

$$\cos\theta = \frac{\boldsymbol{a}'\boldsymbol{b}}{\|\boldsymbol{a}\|\;\|\boldsymbol{b}\|} \tag{28}$$

となる。

2　行列

　同じサイズ $r \times 1$ の c 個の列ベクトル

$$\boldsymbol{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{rj} \end{pmatrix},\; j = 1, 2, \ldots, c \tag{29}$$

を横に並べたもの、もしくは同じサイズ $1 \times c$ の r 個の行ベクトル

$$\boldsymbol{b}_i = \begin{pmatrix} b_{i1}, & b_{i2}, & \ldots, & b_{ic} \end{pmatrix}',\; i = 1, 2, \ldots, r \tag{30}$$

を縦に並べたもの

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1, & \boldsymbol{a}_2, & \ldots, & \boldsymbol{a}_c \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} \boldsymbol{b}_1' \\ \boldsymbol{b}_2' \\ \vdots \\ \boldsymbol{b}_r' \end{pmatrix} \tag{31}$$

を $r \times c$ の行列 (matrix)（ r 行 c 列の行列）と呼ぶ。この行列には全部で $r \times c$ 個の**要素**が含まれているから、それらの要素を $a_{ij}, i = 1, 2, \ldots, r,\ j = 1, 2, \ldots, c$ もしくは $b_{ij}, i = 1, 2, \ldots, r,\ j = 1, 2, \ldots, c$ と記す。行列の名称には、通常、太字の英大文字を用い、要素名には同じ英小文字を用いる。

逆に、$\boldsymbol{A}$ を a_{ij} を要素として持つ $r \times c$ の行列という具合に定義することもあり、

$$\boldsymbol{A} = \{a_{ij}\}, i = 1, 2, \ldots, r,\ j = 1, 2, \ldots, c \tag{32}$$

と書く。二つの添え字のうち、左側 i が行の番号を、右側 j が列の番号を表す。この場合、これを r 個の行ベクトルの集まり、もしくは c 個の列ベクトルの集まりとみることが出来る。

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1j} & \ldots & a_{1c} \\ a_{21} & a_{22} & \ldots & a_{2j} & \ldots & a_{2c} \\ & & \vdots & & & \\ a_{i1} & a_{i2} & \ldots & a_{ij} & \ldots & a_{ic} \\ & & \vdots & & & \\ a_{r1} & a_{r2} & \ldots & a_{rj} & \ldots & a_{rc} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}'_1 \\ \boldsymbol{a}'_2 \\ \vdots \\ \boldsymbol{a}'_i \\ \vdots \\ \boldsymbol{a}'_r \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_{(1)}, & \boldsymbol{a}_{(2)}, & \ldots, \boldsymbol{a}_{(j)}, & \ldots & \boldsymbol{a}_{(c)} \end{pmatrix} \tag{33}$$

最後の表現に付いている括弧付きの添え字は $\boldsymbol{a}_{(j)}$ が元の行列の第 j 列を列ベクトルとして表していることを示している。

2.0.0.1　例　たとえば、平面上の r 個の点の座標 $\boldsymbol{x}_i, i = 1, 2, \ldots, r$ を転置して縦に並べると、以下に示す $r \times 2$ の座標行列となる。また、 n 人の被験者の p 個の変数による測定値ベクトル $\boldsymbol{y}'_i, i = 1, 2, \ldots, n$ を縦に並べると、以下に示す $n \times p$ の座標行列となる。

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}'_1 \\ \boldsymbol{x}'_2 \\ \vdots \\ \boldsymbol{x}'_r \end{pmatrix} \qquad \boldsymbol{Y} = \begin{pmatrix} \boldsymbol{y}'_1 \\ \boldsymbol{y}'_2 \\ \vdots \\ \boldsymbol{y}'_r \end{pmatrix} \tag{34}$$

このほか、所謂、表計算ソフトのスプレッドシートに入っている数値やクロス集計表、成績表等は全て、行列の要素であると見なすことが出来る。

2.1　行列の転置

$r \times c$ の行列 $\boldsymbol{A}$ の行と列を交換したものを行列 $\boldsymbol{A}$ の**転置行列**と呼び $\boldsymbol{A}'$ と書く。これは $c \times r$ の行列となる。

$$\boldsymbol{A}' = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{i1} & \ldots & a_{r1} \\ a_{12} & a_{22} & \ldots & a_{i2} & \ldots & a_{r2} \\ & & \vdots & & & \\ a_{1j} & a_{2j} & \ldots & a_{ij} & \ldots & a_{rj} \\ & & \vdots & & & \\ a_{1c} & a_{2c} & \ldots & a_{ic} & \ldots & a_{rc} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}'_{(1)} \\ \boldsymbol{a}'_{(2)} \\ \vdots \\ \boldsymbol{a}'_{(j)} \\ \vdots \\ \boldsymbol{a}'_{(c)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1, & \boldsymbol{a}_2, & \ldots, \boldsymbol{a}_i, & \ldots, & \boldsymbol{a}_r \end{pmatrix} \tag{35}$$

要素ごとにみれば、元の行列の a_{ij} 要素は転置行列の a_{ji} 要素となっている。これは、要素に名前の付いていない行列（たとえば $\boldsymbol{B}$ ） の ij 要素を $(\boldsymbol{B})_{ij}$ と書くとすれば、$(\boldsymbol{A}')_{ij} = a_{ji}$ ということである。視覚的には、行列のすべての要素を、行列の左上の角から右下方向へ向かう**対角要素** $a_{11}, a_{22}, a_{33}, \ldots$ を対称軸として、180 度回転した形となっている。

定義から、

$$(\boldsymbol{A}')' = \boldsymbol{A} \tag{36}$$

である。

2.1.0.1　例

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}' = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \tag{37}$$

2.2 行列の種類

行列は、その形状や性質により様なざま形で分類されることがある。以下にその主なものをあげる。

- 正方行列 (square)　　行と列の数が等しい行列（絵に描くと正方形）
- 対称行列 (symmetric)　　正方行列で主対角線（左上から右下）を境に対応する要素が等しい行列
  $a_{ij} = a_{ij}, i = 1, 1, \ldots, r, j = 1, 2, \ldots, c, \quad \boldsymbol{A}' = \boldsymbol{A}$
- 対角行列 (diagonal)　　正方行列で主対角線上の要素（対角要素）以外の要素が 0 であるもの
  $a_{ij} = 0, i \neq j$
  なお、本稿では、ある正方行列 $\boldsymbol{A}$ が対角行列であるということを強調するために、その行列を $\mathrm{diag}\,(\boldsymbol{A})$ と書く場合がある。この $\mathrm{diag}\,()$ は、後節で述べる、非対角要素を 0 と置く演算してである。
- 単位行列 (unit)　　対角行列でその対角要素が全て 1 のもの　　$\boldsymbol{I}$ と書く
- ゼロ行列 (zero)　　正方行列でその要素が全て 0 のもの　　$\boldsymbol{O}$ と書く（$\boldsymbol{A} - \boldsymbol{A} = \boldsymbol{O}$ である。）
- 三角行列 (triangular)　　正方行列で主対角線より右上もしくは左下の要素が 0 であるもの
  $a_{ij} = 0, i \geq j$ または $i \leq j$ （等号はなくても良い。）
- ダミー行列 (dummy)　要素が 0 と 1 の値しかとらない行列（デザイン行列）
- 直交行列 (orthogonal)　自分自身の転置との積が単位行列になる行列

したがって、対称行列や対角行列は転置をしても変わらない。($\boldsymbol{A}' = \boldsymbol{A}$) また、下三角行列 (lower triangular matrix) を転置すると上三角行列 (upper triangular matrix) になる。

2.3 行列の演算

2.3.1 基本演算

同じサイズの行列が二つある場合、それらの行列の**加算**と**減算**は、以下のように定義される。

$$\boldsymbol{A} \pm \boldsymbol{B} = \{a_{ij} \pm b_{ij}\}, i = 1, 2, \ldots, r, \ j = 1, 2, \ldots, c \tag{38}$$

すなわち、行列の加減算の結果は、対応する要素間の加減算で定義される要素を持つ行列となる。

このことは、行列を列ベクトルの集まりとみれば、行列の加減算の結果は、対応する列ベクトル間の加減算で定義されるベクトルを元に形に並べたもの、また、行列を行ベクトルの集まりとみれば、対応する行ベクトル間の加減算で定義されるベクトルを元に形に並べたもの、となる。

$$\boldsymbol{A} \pm \boldsymbol{B} \quad = \quad \left(\ \boldsymbol{a}_{(1)} \pm \boldsymbol{b}_{(1)}, \quad \boldsymbol{a}_{(2)} \pm \boldsymbol{b}_{(2)}, \quad \ldots, \quad \boldsymbol{a}_{(c)} \pm \boldsymbol{b}_{(c)} \ \right) = \begin{pmatrix} \boldsymbol{a}'_1 \pm \boldsymbol{b}'_1 \\ \boldsymbol{a}'_2 \pm \boldsymbol{b}'_2 \\ \vdots \\ \boldsymbol{a}'_r \pm \boldsymbol{b}'_r \end{pmatrix} \tag{39}$$

行列とスカラーの**積**は以下のように定義される。

$$\lambda \boldsymbol{A} = \boldsymbol{A} \lambda = \{\lambda a_{ij}\}, i = 1, 2, \ldots, r, \ j = 1, 2, \ldots, c \tag{40}$$

ベクトルの演算が、その要素であるスカラー数の演算をまとめて記したものであるのと同様に、行列の演算とは、その要素であるベクトルの演算をまとめて記したものである。

2.3.2 行列とベクトルの積

$r \times c$ の行列 $\boldsymbol{A}$ と $c \times 1$ のベクトル $\boldsymbol{b}$ との間には積が定義され、以下のようになる。

$$\boldsymbol{A}\boldsymbol{b} = \left\{ \sum_{k=1}^{c} a_{ik} b_k \right\}, i = 1, 2, \ldots, r \tag{41}$$

これは、以下のように、左側の行列を a 個の行ベクトルの集まりとみた場合、それらのベクトル間の内積を、全ての r の組み合わせに関して計算し、その結果を $r \times 1$ の行列の形に並べたものというように解釈することもできる。

$$\boldsymbol{A} \quad = \quad \begin{pmatrix} \boldsymbol{a}'_1 \\ \boldsymbol{a}'_2 \\ \vdots \\ \boldsymbol{a}'_a \end{pmatrix} \tag{42}$$

ならば

$$\boldsymbol{Ab} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1c} \\ a_{21} & a_{22} & \dots & a_{2c} \\ & \vdots & & \\ a_{i1} & a_{i2} & \dots & a_{ic} \\ & \vdots & & \\ a_{a1} & a_{a2} & \dots & a_{ac} \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_c \end{pmatrix} = \left\{ \sum_{k=1}^{b} a_{ik} b_k \right\} \tag{43}$$

$$= \begin{pmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_i' \\ \vdots \\ \boldsymbol{a}_a' \end{pmatrix} \times \boldsymbol{b} = \begin{pmatrix} \boldsymbol{a}_1'\boldsymbol{b} \\ \boldsymbol{a}_2'\boldsymbol{b} \\ \vdots \\ \boldsymbol{a}_i'\boldsymbol{b} \\ \vdots \\ \boldsymbol{a}_a'\boldsymbol{b} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{b} a_{1k} b_k \\ \sum_{k=1}^{b} a_{2k} b_k \\ \vdots \\ \sum_{k=1}^{b} a_{ik} b_k \\ \vdots \\ \sum_{k=1}^{b} a_{rk} b_k \end{pmatrix} \tag{44}$$

したがって、r 元の連立方程式 (simultaneous eqaution) は、

$$\begin{array}{llll} a_{11}x_1+ & a_{12}x_2+ & \dots+ & a_{1r}x_r = c_1 \\ a_{21}x_1+ & a_{22}x_2+ & \dots+ & a_{2r}x_r = c_2 \\ & & \vdots & \\ a_{i1}x_1+ & a_{i2}x_2+ & \dots+ & a_{ir}x_r = c_i \\ & & \vdots & \\ a_{r1}x_1+ & a_{r2}x_2+ & \dots+ & a_{rr}x_r = c_r \end{array} \tag{45}$$

は、係数を要素として持つ $r \times r$ の係数行列を $\boldsymbol{A}$ とすると以下のように書ける。

$$\boldsymbol{Ax} = \boldsymbol{c} \tag{46}$$

2.3.2.1 例

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \times \begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \times 3 + 2 \times 7 \\ 4 \times 3 + 5 \times 7 \end{pmatrix} = \begin{pmatrix} 17 \\ 47 \end{pmatrix} \tag{47}$$

2.3.3 行列同士の積 1

隣り合うサイズが等しい行列、たとえば、 $a \times b$ の $\boldsymbol{A}$ 行列と $b \times c$ の $\boldsymbol{B}$ 行列との間には積が定義され、結果は以下に示すものを要素として持つ $a \times c$ の行列となる。

$$\boldsymbol{AB} = \left\{ \sum_{k=1}^{b} a_{ik} b_{kj} \right\}, i = 1, 2, \dots, a, \ j = 1, 2, \dots, c \tag{48}$$

これは、右側の行列 $\boldsymbol{B}$ を c 個の列ベクトル $\boldsymbol{b}_{(j)}, j = 1, 2, \dots, c$ の集まりと見たときに、これらの各ベクトルごとに行列とベクトルの掛け算 $\boldsymbol{Ab}_{(j)}$ を行い（その結果は $a \times 1$ のベクトル）、その結果を横にに並べたものと見ることもできる。また、行列とベクトルの掛け算は、その行列を構成する行ベクトルと列ベクトルの内積縦に並べたものであったから、以下のように、左側の行列を a 個の行ベクトルの集まり、右側の行列を c 個の列ベクトルの集まりとみた場合、そ

れらのベクトル間の内積を、全ての $a \times c$ の組み合わせに関して計算し、その結果を $a \times c$ の行列の形に並べたものというように解釈することもできる。

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_a' \end{pmatrix} \quad \text{and} \quad \boldsymbol{B} = \begin{pmatrix} \boldsymbol{b}_{(1)}, & \boldsymbol{b}_{(2)}, & \ldots, & \boldsymbol{b}_{(c)} \end{pmatrix} \tag{49}$$

ならば

$$\begin{pmatrix} b_{11} & b_{12} & & b_{1j} & & b_{1c} \\ b_{21} & b_{22} & & b_{2j} & & b_{2c} \\ \vdots & \vdots & \cdots, & \vdots & \cdots, & \vdots \\ b_{b1} & b_{b2} & & b_{bj} & & b_{bc} \end{pmatrix}$$

$$\times$$

$$\boldsymbol{AB} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1b} \\ a_{21} & a_{22} & \ldots & a_{2b} \\ & \vdots & & \\ a_{i1} & a_{i2} & \ldots & a_{ib} \\ & \vdots & & \\ a_{a1} & a_{a2} & \ldots & a_{ab} \end{pmatrix} = \left\{ \sum_{k=1}^{b} a_{ik} b_{kj} \right\} \tag{50}$$

$$= \boldsymbol{A} \times \begin{pmatrix} \boldsymbol{b}_{(1)}, & \boldsymbol{b}_{(2)}, & \ldots, & \boldsymbol{b}_{(j)}, & \ldots, & \boldsymbol{b}_{(c)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Ab}_{(1)}, & \boldsymbol{Ab}_{(2)}, & \ldots, & \boldsymbol{Ab}_{(j)}, & \ldots, & \boldsymbol{Ab}_{(c)} \end{pmatrix} \tag{51}$$

$$= \begin{pmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \vdots \\ \boldsymbol{a}_i' \\ \vdots \\ \boldsymbol{a}_a' \end{pmatrix} \times \begin{pmatrix} \boldsymbol{b}_{(1)}, & \boldsymbol{b}_{(2)}, & \ldots, & \boldsymbol{b}_{(j)}, & \ldots, & \boldsymbol{b}_{(c)} \end{pmatrix} \tag{52}$$

$$\begin{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{b1} \end{pmatrix}, & \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{b2} \end{pmatrix}, & \ldots, & \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{bj} \end{pmatrix}, & \ldots, & \begin{pmatrix} b_{1c} \\ b_{2c} \\ \vdots \\ b_{bc} \end{pmatrix} \end{pmatrix}$$

$$\times$$

$$= \begin{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1b} \\ a_{21} & a_{22} & \ldots & a_{2b} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{ib} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} a_{a1} & a_{a2} & \ldots & a_{ab} \end{pmatrix} \end{pmatrix} = \left\{ \boldsymbol{a}_i' \boldsymbol{b}_{(j)} \right\} \tag{53}$$

$$\tag{54}$$

行列の掛け算は、その要素であるベクトルの内積をとる作業をまとめて記したものであるということが出来る。

2.3.3.1 例

$$\begin{pmatrix} 1 & 4 & 7 & -1 \\ 2 & 5 & 8 & -2 \\ 3 & 6 & 9 & -3 \end{pmatrix}$$

$$\times$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & -4 & 6 \end{pmatrix} = \begin{pmatrix} 14 & 32 & 50 & -14 \\ 12 & 24 & 36 & -12 \end{pmatrix} \tag{55}$$

ただし、$(\boldsymbol{AB})_{11} = 1\times 1 + 2\times 2 + 3\times 3 = 14,\ (\boldsymbol{AB})_{23} = 2\times 7 + -4\times 8 + 6\times 9 = 36,\ \ldots$ である。右から掛ける行列が対角行列の場合は

$$\begin{array}{rcl} & & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{pmatrix} \\ & \times & \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & -4 & 6 \end{pmatrix} & = & \begin{pmatrix} 1 & 10 & 27 \\ 2 & -20 & 54 \end{pmatrix} \end{array} \tag{56}$$

となるが、これは左側の行列の各列を対応する対角要素の値で定数倍したものとなっている。同様に、左から対角行列を掛けると、右側の行列の各行が定数倍されることになる。一般的には $\boldsymbol{D}$ を対角行列とすれば、

$$\boldsymbol{AD} = \begin{pmatrix} \boldsymbol{a}_{(1)}d_{11}, & \boldsymbol{a}_{(2)}d_{22}, & \ldots, & \boldsymbol{a}_{(j)}d_{jj}, & \ldots, & \boldsymbol{a}_{(b)}d_{bb} \end{pmatrix} = \{a_{ij}d_{jj}\}, \tag{57}$$

$$\boldsymbol{DA} = \begin{pmatrix} d_{11}\boldsymbol{a}'_1 \\ d_{22}\boldsymbol{a}'_2 \\ \vdots \\ d_{ii}\boldsymbol{a}'_i \\ \vdots \\ d_{aa}\boldsymbol{a}'_a \end{pmatrix} = \{d_{ii}a_{ij}\} \tag{58}$$

である。

したがって、単位行列を右から掛けても左から掛けても元の行列の値は変わらない。

$$\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A} \tag{59}$$

自分自身の転置を左から掛けたものは

$$\boldsymbol{A}'\boldsymbol{A} = \left\{ \sum_{k=1}^{b} a_{ki}a_{kj} \right\} \tag{60}$$

となるが、これを**積和行列 SSCP matrix, sum of squares and cross products matrix** と呼ぶ。逆に自分自身の転置を右から掛けると

$$\boldsymbol{A}\boldsymbol{A}' = \left\{ \sum_{k=1}^{b} a_{ik}a_{jk} \right\} \tag{61}$$

となる。なお、積和行列の対角要素には各列の２乗和 $\sum_{k=1}^{b} a_{kj}^2$ が入り、$\boldsymbol{AA}$' の対角要素には各行の２乗和 $\sum_{k=1}^{b} a_{ik}^2$ が入る。$\boldsymbol{A}$ がどのような行列であっても、$\boldsymbol{A}$'$\boldsymbol{A}$ と $\boldsymbol{AA}$' は共に対称行列となる。

積和行列が対角行列になる場合、 $\boldsymbol{A}$ の各列は直交するという。また、

$$\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}' = \boldsymbol{I} \tag{62}$$

を満たす正方行列を**（正規）直交行列 orthonormal matrix** と呼ぶ。

2.3.4 行列同士の積 2

行列という概念はその特殊ケースとしてベクトルを含むから、縦ベクトルは列の数が 1 の行列、横ベクトルは行の数が 1 の行列ということになる。したがって、すでに見たように、ある行列に右から縦ベクトルを掛けたものは縦ベクトルに、左から行ベクトルを掛けたものは行ベクトルになる。

また、$r\times 1$ の縦ベクトル $\boldsymbol{a}$ に右から $1\times c$ の行ベクトル $\boldsymbol{b}'$ を掛けると

$$\begin{array}{rcl} & & \begin{pmatrix} b_1, & b_2, & \ldots, & b_j, & \ldots, & b_c \end{pmatrix} \\ & \times & \\ \boldsymbol{ab}' = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_i \\ \vdots \\ a_r \end{pmatrix} & = & \{\, a_ib_j \,\} \end{array} \tag{63}$$

ベクトルの内積は行ベクトルの右から縦ベクトルを掛けて得られるスカラー量であったが、上記の掛け算はそれとは異なる。また、これを外積とも呼ばない。当然であるが、 $\boldsymbol{a}\boldsymbol{b}' \neq \boldsymbol{b}\boldsymbol{a}'$ である。

この操作を利用すれば、$a \times b$ と $b \times c$ の行列の掛け算は以下のように考えることもできる。

$$\boldsymbol{AB} = \begin{pmatrix} \boldsymbol{a}_{(1)}, & \boldsymbol{a}_{(2)}, & \ldots, & \boldsymbol{a}_{(i)}, & \ldots, & \boldsymbol{a}_{(b)} \end{pmatrix} \times \begin{pmatrix} \boldsymbol{b}'_1 \\ \boldsymbol{b}'_2 \\ \vdots \\ \boldsymbol{b}'_j \\ \vdots \\ \boldsymbol{b}'_b \end{pmatrix} = \sum_{k=1}^{b} \boldsymbol{a}_{(k)} \boldsymbol{b}'_k \tag{64}$$

2.3.4.1 例

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & -4 & 6 \end{pmatrix} \times \begin{pmatrix} 1 & 4 & 7 & -1 \\ 2 & 5 & 8 & -2 \\ 3 & 6 & 9 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 7 & -1 \\ 2 & 8 & 14 & -2 \end{pmatrix} + \begin{pmatrix} 4 & 10 & 16 & -4 \\ -8 & -20 & -32 & 8 \end{pmatrix} + \begin{pmatrix} 9 & 18 & 27 & -9 \\ 18 & 36 & 54 & -18 \end{pmatrix} \tag{65}$$

この方法を利用すると、$c = 1$ の場合、すなわち $\boldsymbol{B}$ 行列が一つの列ベクトル $\boldsymbol{b}$ からなる場合

$$\boldsymbol{Ab} = \sum_{j=1}^{b} b_j \times \boldsymbol{a}_{(j)} \tag{66}$$

となるが、$\boldsymbol{A}$ が正方行列、すなわち $a = b$ の場合、この表現は行列 $\boldsymbol{A}$ によって作られる線形写像を理解する際に役立つ。すなわち、$\boldsymbol{e}_j, j = 1, 2, \ldots, a$ を a 次元空間の直行基底（単位行列の第 j 列）とすれば、$\boldsymbol{A}$ によりこの直交基底は $\boldsymbol{A}$ 行列の各列に変換されるため、$\boldsymbol{b}$ ベクトルの各要素がこの新しい基底を用いる場合の座標となっていることが分かる。

2.3.5 行列の積の一般形

$\boldsymbol{a}$ と $\boldsymbol{b}$ を $r \times 1$ の縦ベクトルとする。この二つのベクトルの和や差を $\boldsymbol{c}$ とすると、

$$\boldsymbol{c} = \boldsymbol{a} \pm \boldsymbol{b}, \quad c_i = a_i \pm b_i \tag{67}$$

となる。二つのベクトルの内積を $\boldsymbol{c}$ とすると

$$c = \boldsymbol{a}'\boldsymbol{b} = \boldsymbol{b}'\boldsymbol{a}, \quad c = \sum_{i=1}^{r} a_i b_i \tag{68}$$

である。

$\boldsymbol{A}$ と $\boldsymbol{B}$ を $r \times c$ の行列とする。この二つの行列の和や差を $\boldsymbol{C}$ とすると、

$$\boldsymbol{C} = \boldsymbol{A} \pm \boldsymbol{B}, \quad c_{ij} = a_{ij} \pm b_{ij} \tag{69}$$

となる。$\boldsymbol{C}$ 行列は元の行列と同じサイズである。$\boldsymbol{A}$ を $ra \times ca$ $\boldsymbol{B}$ を $rb \times cb$ の行列とする。いま、 $ca = rb$ の場合、積 $\boldsymbol{AB}$ が定義でき、

$$\boldsymbol{C} = \boldsymbol{AB}, \quad c_{ij} = \sum_{k=1}^{ca} a_{ik} b_{kj}, \quad i = 1, 2, \ldots, ra, \ j = 1, 2, \ldots, cb \tag{70}$$

であり、$\boldsymbol{C}$ 行列のサイズは $ra \times \ cb$ である。また、$cb = ra$ の場合、積 $\boldsymbol{BA}$ が定義でき、

$$\boldsymbol{C} = \boldsymbol{BA}, \quad c_{ij} = \sum_{k=1}^{ra} b_{ik} a_{kj}, \quad i = 1, 2, \ldots, rb, \ j = 1, 2, \ldots, ca \tag{71}$$

でり、$\boldsymbol{C}$ 行列のサイズは $rb \times \ ca$ である。この場合、スカラー同士の掛け算は順序に依存しないため

$$\boldsymbol{C} = \boldsymbol{BA}, \quad c_{ij} = \sum_{k=1}^{ra} a_{kj} b_{ik}, \quad i = 1, 2, \ldots, rb, \ j = 1, 2, \ldots, ca \tag{72}$$

と書いても良い。また、$ra = rb$ の場合、積 $\boldsymbol{A}$'$\boldsymbol{B}$ が定義でき、

$$\boldsymbol{C} = \boldsymbol{A}'\boldsymbol{B}, \quad c_{ij} = \sum_{k=1}^{ra} a_{ki}b_{kj}, \quad i = 1, 2, \ldots, ca, \ j = 1, 2, \ldots, cb \tag{73}$$

であり、$\boldsymbol{C}$ 行列のサイズは $ca \times \ cb$ である。すなわち、(70) 式 のパタンに於いて、転置となっている行列 $\boldsymbol{A}$ の添字を交換すればよい。また、$rb = ra$ の場合、積 $\boldsymbol{B}$'$\boldsymbol{A}$ が定義でき、

$$\boldsymbol{C} = \boldsymbol{B}'\boldsymbol{A}, \quad c_{ij} = \sum_{k=1}^{rb} b_{ki}a_{kj} = \sum_{k=1}^{rb} a_{kj}b_{ki}, \quad i = 1, 2, \ldots, cb, \ j = 1, 2, \ldots, ca \tag{74}$$

であり、$\boldsymbol{C}$ 行列のサイズは $cb \times \ ca$ である。これも、(71) 式 の $\boldsymbol{B}$ の添字を交換してものとなっている。

自分自身の転置との積は常に定義可能で、

$$\boldsymbol{C} = \boldsymbol{A}'\boldsymbol{A}, \quad c_{ij} = \sum_{k=1}^{ra} a_{ki}a_{kj}, \quad i = 1, 2, \ldots, ca, \ j = 1, 2, \ldots, ca \tag{75}$$

$$\boldsymbol{C} = \boldsymbol{A}\boldsymbol{A}', \quad c_{ij} = \sum_{k=1}^{ca} a_{ik}a_{jk}, \quad i = 1, 2, \ldots, ra, \ j = 1, 2, \ldots, ra \tag{76}$$

となるが、$\boldsymbol{A}$'$\boldsymbol{A}$ のサイズは $ca \times \ ca$、$\boldsymbol{A}\boldsymbol{A}$' のサイズは $ra \times \ ra$ である。

また、対角行列が入っている場合は、対角行列の掛け算に関する性質を用いて、対角行列の個数だけ総和記号を減らすことが出来る。例えば、$\boldsymbol{A}$ 行列が対角行列の場合は、 $\boldsymbol{A}\boldsymbol{B}$ は $\boldsymbol{B}$ の各行を $\boldsymbol{A}$ の対角要素倍したものであり、同様に、$\boldsymbol{B}$ 行列が対角の場合は、 $\boldsymbol{A}\boldsymbol{B}$ は $\boldsymbol{A}$ の各列 $\boldsymbol{B}$ の対角要素倍したものとなる。したがっていずれの場合にも総和記号は必要でなくなり、 (70) 式 は以下のようになる。なお、対角行列を強調するために diag () の記号を用いている。

$$\boldsymbol{C} = \mathrm{diag}\,(\boldsymbol{A})\boldsymbol{B}, \quad c_{ij} = a_{ii}b_{ij} \tag{77}$$

$$\boldsymbol{C} = \boldsymbol{A}\mathrm{diag}\,(\boldsymbol{B}), \quad c_{ij} = a_{ij}b_{jj} \tag{78}$$

$$\boldsymbol{C} = \mathrm{diag}\,(\boldsymbol{A})\mathrm{diag}\,(\boldsymbol{B}), \quad c_{ij} = \begin{cases} a_{ii}b_{ii} & i = j \\ 0 & i \neq j \end{cases} \tag{79}$$

なお、最後の式は、対角行列同士の積がまた対角行列になることを示している。

行列とベクトルとの積は、上記の公式で、ベクトルに相当する行列が、列ベクトルの場合には行の添字が常に 1 、すなわち $ca = 1$ または $cb = 1$、行ベクトルの場合には列の添字が常に 1 、すなわち $ra = 1$ または $rb = 1$、であると考えればよい。例えば、 (70) 式 に於いて $\boldsymbol{B}$ が縦ベクトルの場合、$cb = 1$ であるから、

$$\boldsymbol{c} = \boldsymbol{A}\boldsymbol{b}, \quad c_i = c_{i1} = \sum_{k=1}^{ca} a_{ik}b_{k1} = \sum_{k=1}^{ca} a_{ik}b_k \tag{80}$$

となり、$\boldsymbol{c}$ は $ra \times \ 1$ の縦ベクトルとなる。 (73) 式 に於いて $\boldsymbol{A}$ が列ベクトルの場合は $ca = 1$ であるから

$$\boldsymbol{c}' = \boldsymbol{a}'\boldsymbol{B}, \quad c_i = c_{1i} = \sum_{k=1}^{ca} a_{k1}b_{kj} = \sum_{k=1}^{ca} a_k b_{kj} \tag{81}$$

となり $\boldsymbol{c}$ は $1 \times \ cb$ の行ベクトルとなる。

この特殊形として、ベクトルが単位ベクトル $\mathbf{1}$ の場合を考えてみると、

$$\boldsymbol{c} = \boldsymbol{A}\mathbf{1}, \quad c_i = c_{i1} = \sum_{k=1}^{ca} a_{ik}1 = \sum_{k=1}^{ca} a_{ik} \tag{82}$$

$$\boldsymbol{c}' = \mathbf{1}'\boldsymbol{B}, \quad c_i = c_{1i} = \sum_{k=1}^{ca} 1b_{kj} = \sum_{k=1}^{ca} b_{kj} \tag{83}$$

となる。すなわち、行列の行や列の和がえられる。

３つの行列の積、例えば、 $\boldsymbol{D} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{C}$ は、これを $\boldsymbol{D} = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$ もしくは $\boldsymbol{D} = (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C}$ と考え、上記の公式を２度適用すればよい。前者の場合、$\boldsymbol{B}\boldsymbol{C}$ の結果として得られる行列の pq 要素を $[\boldsymbol{B}\boldsymbol{C}]_{pq}$ と書けば、

$$\boldsymbol{D} = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C}), \quad d_{ij} = \sum_{k=1}^{ca} a_{ik}[BC]_{kj} = \sum_{k=1}^{ca} a_{ik}(\sum_{\ell=1}^{cb} b_{k\ell}c_{\ell_j}) \tag{84}$$

$$= \sum_{k=1}^{ca} a_{ik} \sum_{\ell=1}^{cb} b_{k\ell}c_{\ell_j} = \sum_{k=1}^{ca}\sum_{\ell=1}^{cb} a_{ik}b_{k\ell}c_{\ell_j} \tag{85}$$

となる。また、後者の場合、$\boldsymbol{AB}$ の結果として得られる行列の pq 要素を $[\boldsymbol{AB}]_{pq}$ と書けば、

$$\boldsymbol{D} = (\boldsymbol{AB})\boldsymbol{C}, \quad d_{ij} = \sum_{k=1}^{cb}[\boldsymbol{AB}]_{ik}c_{kj} = \sum_{k=1}^{cb}(\sum_{\ell=1}^{ca}a_{i\ell}b_{\ell_k})c_{kj} \tag{86}$$

$$= \sum_{k=1}^{cb}\sum_{\ell=1}^{ca}a_{i\ell}b_{\ell_k}c_{kj} = \sum_{\ell=1}^{ca}a_{i\ell}\sum_{k=1}^{cb}b_{\ell_k}c_{kj} = \sum_{k=1}^{ca}a_{ik}\sum_{\ell=1}^{cb}b_{k\ell}c_{\ell_j} \tag{87}$$

となるが、最後の式の展開は単なる添字の付け替えである。

より一般的には、複数の行列の掛け算は、以下の 5 つの行列の掛け算

$$\boldsymbol{F} = \boldsymbol{ABCDE} \tag{88}$$

$$f_{ij} = \sum_{i1=1}^{ca}\sum_{i2=1}^{cb}\sum_{i3=1}^{cc}\sum_{i4=1}^{cd}a_{i,i1}b_{i1,i2}c_{i2,i3}d_{i3,i4}e_{i4,j} \tag{89}$$

$$= \sum_{i1=1}^{ca}a_{i,i1}\sum_{i2=1}^{cb}b_{i1,i2}\sum_{i3=1}^{cc}c_{i2,i3}\sum_{i4=1}^{cd}d_{i3,i4}e_{i4,j} \tag{90}$$

に示される様なパタンを行列の数により拡大縮小した形となっている。また、いくつかの行列が、転置の形で入っている場合、それらの行列の添字を入れ替えればよい。例えば、

$$\boldsymbol{F} = \boldsymbol{AB'CD'E} \tag{91}$$

$$f_{ij} = \sum_{i1=1}^{ca}\sum_{i2=1}^{rb}\sum_{i3=1}^{cc}\sum_{i4=1}^{rd}a_{i,i1}b_{i2,i1}c_{i2,i3}d_{i4,i3}e_{i4,j} \tag{92}$$

$$= \sum_{i1=1}^{ca}a_{i,i1}\sum_{i2=1}^{rb}b_{i2,i1}\sum_{i3=1}^{cc}c_{i2,i3}\sum_{i4=1}^{rd}d_{i4,i3}e_{i4,j} \tag{93}$$

また、対角行列が入っている場合は、先に述べた対角行列の性質を使って、対角行列の個数だけ総和記号を減らすことが出来る。例えば、$\boldsymbol{A}$ 行列と $\boldsymbol{D}$ 行列が対角行列の場合は、まず、$\mathrm{diag}\,(bA)\boldsymbol{B}$ をまとめることにより、総和記号が一つ減る。次に、$\boldsymbol{D}$ 行列であるが、これを $\boldsymbol{E}$ 行列の左からかかっていると見るか、 $\boldsymbol{C}$ 行列の右からかかっていると見るかで以下のような二つの同値な結果を得る。

$$\boldsymbol{F} = \mathrm{diag}\,(\boldsymbol{A})\boldsymbol{BC}\mathrm{diag}\,(\boldsymbol{D})\boldsymbol{E} = (\mathrm{diag}\,(\boldsymbol{A})\boldsymbol{B})\boldsymbol{C}(\mathrm{diag}\,(\boldsymbol{D})\boldsymbol{E}) \tag{94}$$

$$f_{ij} = \sum_{i1=1}^{ca}\sum_{i2=1}^{cc}(a_{i,i}b_{i,i1})c_{i1,i2}(d_{i2,i2}e_{i2,j}) \tag{95}$$

$$= \sum_{i1=1}^{ca}a_{i,i}b_{i,i1}\sum_{i2=1}^{cc}c_{i1,i2}d_{i2,i2}e_{i2,j} \tag{96}$$

$$\boldsymbol{F} = \mathrm{diag}\,(\boldsymbol{A})\boldsymbol{BC}\mathrm{diag}\,(\boldsymbol{D})\boldsymbol{E} = (\mathrm{diag}\,(\boldsymbol{A})\boldsymbol{B})(\boldsymbol{C}\mathrm{diag}\,(\boldsymbol{D}))\boldsymbol{E} \tag{97}$$

$$f_{ij} = \sum_{i1=1}^{ca}\sum_{i2=1}^{cc}(a_{i,i}b_{i,i1})(c_{i1,i2}d_{i2,i2})e_{i2,j} \tag{98}$$

$$= \sum_{i1=1}^{ca}a_{i,i}b_{i,i1}\sum_{i2=1}^{cc}c_{i1,i2}d_{i2,i2}e_{i2,j} \tag{99}$$

最後に、演算結果がスカラーとなる場合は、それを総和記号の外に出すことが出来る。例えば、以下の例で、 $\boldsymbol{C}$ 以

外が列ベクトルであれば、$\boldsymbol{B}'\boldsymbol{C}\boldsymbol{D}$ はスカラーとなるため、それを前に出すことが出来る。

$$\boldsymbol{F} = \boldsymbol{a}(\boldsymbol{b}'\boldsymbol{C}\boldsymbol{d})\boldsymbol{e}' = (\boldsymbol{b}'\boldsymbol{C}\boldsymbol{d})\boldsymbol{a}\boldsymbol{b}' \tag{100}$$

$$f_{ij} = \sum_{i1=1}^{1}\sum_{i2=1}^{rb}\sum_{i3=1}^{cc}\sum_{i4=1}^{1} a_{i,i1}b_{i2,i1}c_{i2,i3}d_{i3,i4}e_{j,i4} \tag{101}$$

$$= \sum_{i1=1}^{1} a_{i,i1}\sum_{i2=1}^{rb} b_{i2,i1}\sum_{i3=1}^{cc} c_{i2,i3}\sum_{i4=1}^{1} d_{i3,i4}e_{j,i4} \tag{102}$$

$$= a_{i,1}\sum_{i2=1}^{rb} b_{i2,1}\sum_{i3=1}^{cc} c_{i2,i3}d_{i3,1}e_{j,1} \tag{103}$$

$$= a_i e_j \tag{104}$$

$$= (\sum_{i2=1}^{rb} b_{i2}\sum_{i3=1}^{cc} c_{i2,i3}d_{i3})a_i e_j \tag{105}$$

2.3.6 演算規則

行列の演算に関しては以下の関係が成り立つ。

$$\boldsymbol{A}+\boldsymbol{B}+\boldsymbol{C} = (\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C} = \boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{B}+(\boldsymbol{A}+\boldsymbol{C}) \tag{106}$$

$$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{A}\boldsymbol{B}+\boldsymbol{A}\boldsymbol{C} \quad \text{and} \quad (\boldsymbol{B}+\boldsymbol{C})\boldsymbol{A} = \boldsymbol{B}\boldsymbol{A}+\boldsymbol{C}\boldsymbol{A} \tag{107}$$

$$\boldsymbol{A}\boldsymbol{B}\boldsymbol{C} = (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C}) \tag{108}$$

$$(\boldsymbol{A}+\boldsymbol{B})' = \boldsymbol{A}'+\boldsymbol{B}', \quad (\boldsymbol{A}\boldsymbol{B})' = \boldsymbol{B}'\boldsymbol{A}', \quad (\boldsymbol{A}')' = \boldsymbol{A} \tag{109}$$

$r \times c$ の行列 $\boldsymbol{A}$ の全ての要素にスカラー量 λ を足したり引いたりする演算は、

$$\boldsymbol{A} \pm \lambda\mathbf{1}_r\mathbf{1}_c' \quad \text{or} \quad \boldsymbol{A} \pm \mathbf{1}_r\mathbf{1}_c'\lambda \quad \text{or} \quad \boldsymbol{A} \pm \mathbf{1}_r\lambda\mathbf{1}_c' \tag{110}$$

と書ける。ただし、$\mathbf{1}_r\mathbf{1}_c'$ は、その要素が全て 1 である $r \times c$ の行列となる。

2.4 簡易表記と添字演算子

数学的には、サイズの合わない行列同士の演算は定義されない。しかし、行列演算子を持ついくつかのソフトウエアでは、以下の様な簡易表記が用いられている。$\boldsymbol{A}$ を $r \times c$ の行列、$\boldsymbol{r}$ を $r \times 1$ の列ベクトル、$\boldsymbol{c}$ を $1 \times c$ の行ベクトル、λ をスカラーとする。

$$\boldsymbol{r} \pm \lambda\mathbf{1}_r \rightarrow \boldsymbol{r} \pm \lambda, \quad \boldsymbol{c} \pm \lambda\mathbf{1}_c' \rightarrow \boldsymbol{c} \pm \lambda \tag{111}$$

$$\boldsymbol{A} \pm \lambda\mathbf{1}_r\mathbf{1}_c' \rightarrow \boldsymbol{A} \pm \lambda \tag{112}$$

また、以下の様な添字演算子も利用されることがある。

$$\boldsymbol{A}_{+j} = \sum_{i=1}^{r} a_{ij}, \quad \boldsymbol{A}_{i+} = \sum_{j=1}^{c} a_{ij} \tag{113}$$

前節でも述べたが、本稿では、行列演算の結果として得られる行列の要素に名前を付けることなく示したい場合、例えば、$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}$ の様な演算結果の第 ij 要素を

$$\text{演算式の } ij \text{ 要素} = [\text{演算式}]_{ij} \tag{114}$$

と略記する。

2.5 ベクトルの２次形式

$\boldsymbol{b}'\boldsymbol{A}\boldsymbol{b}$ の形をベクトル $\boldsymbol{b}$ の **2 次形式**と呼ぶ。この計算は以下のようになる。

$$\boldsymbol{b}'\boldsymbol{A}\boldsymbol{b} = \boldsymbol{b}'(\boldsymbol{A}\boldsymbol{b}) = \sum_{i=1}^{r} b_i \sum_{j=1}^{r} a_{ij} b_j \tag{115}$$

$$= \sum_{i=1}^{r} a_{ii} b_i^2 + 2\sum_{i<j} a_{ij} b_i b_j \qquad (\boldsymbol{A} \text{ が対称の場合}) \tag{116}$$

なお、$\boldsymbol{A}$ が対称行列の場合、どのような $\boldsymbol{b}$, ただし $\boldsymbol{b} \neq \boldsymbol{0}$、 を選んでも 2 次形式が正となる行列 $\boldsymbol{A}$ を**正定値行列**、2 次形式が負とならない行列を**半正定値行列（非負行列）**と呼ぶ。たとえば、積和行列 $\boldsymbol{A}$'$\boldsymbol{A}$ は少なくとも非負行列である。

2.6 行列のトレース

$r \times r$ の正方行列の対角要素の和

$$\mathrm{tr}\,(\boldsymbol{A}) = \sum_{j=1}^{r} a_{ii} \tag{117}$$

を**トレース** (trace) と呼び $\mathrm{tr}\,(\boldsymbol{A})$ と書く。
トレース演算子は以下の様な性質を持っている。

$$\mathrm{tr}\,(\lambda) = \lambda. \quad \mathrm{tr}\,(\lambda\boldsymbol{A}) = \lambda \mathrm{tr}\,(\boldsymbol{A}), \quad \mathrm{tr}\,(\boldsymbol{A}) = \mathrm{tr}\,(\boldsymbol{A}') \tag{118}$$

$$\mathrm{tr}\,(\boldsymbol{A} + \boldsymbol{B}) = \mathrm{tr}\,(\boldsymbol{A}) + \mathrm{tr}\,(\boldsymbol{B}) \tag{119}$$

$$\mathrm{tr}\,(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}\,(\boldsymbol{B}\boldsymbol{A}) = \mathrm{tr}\,(\boldsymbol{B}'\boldsymbol{A}') = \mathrm{tr}\,(\boldsymbol{A}'\boldsymbol{B}'), \tag{120}$$

$$\mathrm{tr}\,(\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}) = \mathrm{tr}\,(\boldsymbol{B}\boldsymbol{C}\boldsymbol{A}) = \mathrm{tr}\,(\boldsymbol{C}\boldsymbol{A}\boldsymbol{B}) \tag{121}$$

$$= \mathrm{tr}\,(\boldsymbol{C}'\boldsymbol{B}'\boldsymbol{A}') = \mathrm{tr}\,(\boldsymbol{B}'\boldsymbol{A}'\boldsymbol{C}') = \mathrm{tr}\,(\boldsymbol{A}'\boldsymbol{C}'\boldsymbol{B}') \tag{122}$$

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = \mathrm{tr}\,(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}) = \mathrm{tr}\,(\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}') \tag{123}$$

どのような行列 $\boldsymbol{A}$ でも、自分自身の転置との積のトレースは

$$\mathrm{tr}\,(\boldsymbol{A}'\boldsymbol{A}) = \mathrm{tr}\,(\boldsymbol{A}\boldsymbol{A}') = \sum_{i=1}^{r}\sum_{j=1}^{c} a_{ij}^2 \tag{124}$$

と、要素の２乗和となる。したがって、残差の２乗和 (RSS, Residual Sum of Squares) は

$$\mathrm{RSS} = \sum_{i=1}^{r} (y_i - \hat{y}_i)^2 = (\boldsymbol{y} - \widehat{\boldsymbol{y}})'(\boldsymbol{y} - \widehat{\boldsymbol{y}}) \tag{125}$$

$$= \mathrm{tr}\,((\boldsymbol{y} - \widehat{\boldsymbol{y}})'(\boldsymbol{y} - \widehat{\boldsymbol{y}})) = \mathrm{tr}\,((\boldsymbol{y} - \widehat{\boldsymbol{y}})(\boldsymbol{y} - \widehat{\boldsymbol{y}})') \tag{126}$$

$$\mathrm{RSS} = \sum_{i=1}^{r}\sum_{j=1}^{c} (y_{ij} - \hat{y}_{ij})^2 = \mathrm{tr}\left((\boldsymbol{Y} - \widehat{\boldsymbol{Y}})'(\boldsymbol{Y} - \widehat{\boldsymbol{Y}})\right) \tag{127}$$

$$= \mathrm{tr}\left((\boldsymbol{Y} - \widehat{\boldsymbol{Y}})(\boldsymbol{Y} - \widehat{\boldsymbol{Y}})'\right) \tag{128}$$

と書くことが出来る。

2.7 diag 演算子

$\mathrm{diag}\,(\boldsymbol{A})$ で、正方行列 $\boldsymbol{A}$ の非対角要素を全て 0 と置いた対角行列を表し、$\mathrm{Diag}\,(\boldsymbol{a})$ で、ベクトル $\boldsymbol{a}$ の各要素を対角要素として持つ対角行列を表す。

なお、$\mathrm{Diag}\,()$ を使わずに $\mathrm{diag}\,()$ のみを用い、それに行列を与えた場合には対角要素以外を 0 とした行列を返し、ベクトルを与えたときには上記の $\mathrm{Diag}\,()$ と同じ働きをするものと定義しても良い。

対角行列をある行列の右から掛けることは、その行列の各列を対応する対角行列の対角要素倍することになる。すなわち $\boldsymbol{A} = [\boldsymbol{a}_{(1)}, \boldsymbol{a}_{(2)}, \ldots, \boldsymbol{a}_{(c)}]$ の場合、$\boldsymbol{D}$ を d_i を対角要素として持つ対角行列とすると、

$$\boldsymbol{A}\boldsymbol{D} = [\boldsymbol{a}_{(1)} d_1, \boldsymbol{a}_{(2)} d_2, \ldots, \boldsymbol{a}_{(c)} d_c] \tag{129}$$

となる。また、対角行列をある行列の左から掛けることは、その行列の各行を対応する対角行列の対角要素倍することになる。すなわち $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_c]'$ の場合、

$$\boldsymbol{DA} = [d_1\boldsymbol{a}_1, d_2\boldsymbol{a}_2, \ldots, d_r\boldsymbol{a}_c]' \tag{130}$$

となる。

2.8 並べ替え行列

$r \times c$ の行列 $\boldsymbol{X}$ の第 i 行と第 j 行を入れ替えたものを $\boldsymbol{Y}$ とする。この行を入れ替える作業は $r \times r$ の単位行列の第 (i,i) 要素と (j,j) 要素を 0 に、また、(i,j) 要素と (j,i) 要素を 1 に置き換えた行列 $\boldsymbol{P}_{ij}$ を用いて $\boldsymbol{Y} = \boldsymbol{P}_{ij}\boldsymbol{X}$ と書くことが出来る。以下に $r = 4, i = 1, j = 4$ と $r = 4, i = 2, j = 3$ の場合の例を示す。

$$\boldsymbol{P}_{34} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \boldsymbol{P}_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{131}$$

さらに、この $\boldsymbol{Y}$ 行列の第 i' 行と第 j' 行を入れ替えたものを $\boldsymbol{Z}$ とすると、

$$\boldsymbol{Z} = \boldsymbol{P}_{i'j'}\boldsymbol{Y} = \boldsymbol{P}_{i'j'}\boldsymbol{P}_{ij}\boldsymbol{X} = \boldsymbol{PX} \quad (\boldsymbol{P} = \boldsymbol{P}_{i'j'}\boldsymbol{P}_{ij}) \tag{132}$$

である。例えば、上記の二つの行列の積は

$$\boldsymbol{P} = \boldsymbol{P}_{34} \times \boldsymbol{P}_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{133}$$

となる。このような行列は**並べ替え行列** (permutation matrix) と呼ばれるが、その特徴は、各列と行に一カ所だけ 1 が有りそれ以外の要素はゼロであるということである。また、並べ替え行列は直交行列である。したがって、行を元に戻すには $\boldsymbol{X} = \boldsymbol{P}'\boldsymbol{Z}$ とすればよい。

逆に列を入れ替えたい場合には同様に $c \times c$ の行列 $\boldsymbol{P}$ を定義して、 $\boldsymbol{XP}$ とすればよい。

3 逆行列

$r \times r$ の正方行列 $\boldsymbol{A}$ の逆行列 (inverse matrix) とは

$$\boldsymbol{B}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{B} = \boldsymbol{I} \tag{134}$$

を満たす同じサイズの正方行列 $\boldsymbol{B}$ のことで、それを $\boldsymbol{A}^{-1}$ と書く。すなわち、

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I} \tag{135}$$

である。特に、$\boldsymbol{A}$ がスカラー数の場合は $a^{-1} = 1/a$ となり見慣れた逆数の定義となる。(1×1 の単位行列は 1 である。)

2×2 の行列の逆行列は以下のようにして計算する事が出来る。

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \tag{136}$$

ただし、$a_{11}a_{22} - a_{12}a_{21}$ がゼロとなるような行列は逆行列を持たない。このような行列は**特異 singular** であるという。なお、特異でない行列、すなわち逆行列を持つ行列を**正則 regular** な行列という。正則な行列の逆行列は一意に定まる。3 x 3 以上の行列の逆行列の計算法に関しては次節に述べる。

対角行列の逆行列はまた対角行列となり、その対角要素は元の対角要素の逆数となる。

$$\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ d & d & d_3 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & 0 \\ 0 & \frac{1}{d_2} & 0 \\ 0 & 0 & \frac{1}{d_3} \end{bmatrix} \tag{137}$$

(正規) 直交行列の逆行列はその転置行列である。

$$\boldsymbol{A}^{-1} = \boldsymbol{A}' \quad \text{ただし} \quad \boldsymbol{A}\boldsymbol{A}' = \boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I} \tag{138}$$

また、三角行列の逆行列は三角行列となる。

3.0.0.1 例

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/3 \end{pmatrix} \tag{139}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 3/2 & 1/2 & -5/2 \\ -1/2 & -1/2 & 3/2 \end{pmatrix} \tag{140}$$

$$\begin{pmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{pmatrix}^{-1} = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix} \tag{141}$$

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^{-1} = \begin{bmatrix} 1/a_{11} & 0 & 0 \\ -a_{21}/(a_{11}a_{22}) & 1/a_{22} & 0 \\ -a_{31}/(a_{11}a_{33}) + a_{32}a_{21}/(a_{11}a_{22}a_{33}) & -a_{32}/(a_{22}a_{33}) & 1/a_{33} \end{bmatrix} \tag{142}$$

逆行列は以下の様な演算規則を持つ。

$$\lambda^{-1} = \frac{1}{\lambda}, \quad (\lambda\boldsymbol{A})^{-1} = \frac{1}{\lambda}\boldsymbol{A}^{-1}, \quad \boldsymbol{I}^{-1} = \boldsymbol{I} \tag{143}$$

$$(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}, \quad (\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})', \quad (\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \tag{144}$$

3.1 逆行列の求め方

3.1.1 ガウスの消去法 Gaussian Elimination

逆行列を求める方法には様々なのものがあるが、ここではガウスの消去法と呼ばれる方法に関して説明する。ガウスの消去法は $\boldsymbol{T}_1, \boldsymbol{T}_2, \ldots, \boldsymbol{T}_r$ を行に関する基本演算を行う行列とした時に、逆行列を

$$\boldsymbol{A}^{-1} = \boldsymbol{T}_r \times \boldsymbol{T}_{r-1} \times \cdots \times \boldsymbol{T}_2 \times \boldsymbol{T}_1 \tag{145}$$

として求める方法であるが、 $\boldsymbol{T}_k,\ k=1,2,\ldots,r$ は、$\boldsymbol{B}_0=\boldsymbol{A}$ として

$$\boldsymbol{B}_k=\boldsymbol{T}_k\times\boldsymbol{B}_{k-1} \tag{146}$$

の第 k 列が第 k 番目の要素が 1 でそれ以外が 0 となる様に定められる。すなわち、$\boldsymbol{T}_k$ 行列は単位行列の第 k 列を以下に示すように書き換えたものとなっている。

このようにして $\boldsymbol{B}_k$ 行列を順次求めていくと、最終的に得られる $\boldsymbol{B}_r$ 行列は単位行列となっており、

$$\boldsymbol{B}_r \quad = \quad \boldsymbol{T}_r\times\boldsymbol{T}_{r-1}\times\cdots\times\boldsymbol{T}_2\times\boldsymbol{T}_1\times\boldsymbol{A}\ =\ \boldsymbol{I} \tag{147}$$

であるから (145) 式を得る。

例えば

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1j} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2j} & \ldots & a_{2r} \\ & & \vdots & & \vdots & \\ a_{i1} & a_{i2} & \ldots & a_{ij} & \ldots & a_{ir} \\ & & \vdots & & \vdots & \\ a_{r1} & a_{r2} & \ldots & a_{rj} & \ldots & a_{rr} \end{bmatrix} \tag{148}$$

の場合、$\boldsymbol{T}_k$ 行列は以下の通りである。

- $\boldsymbol{T}_1$

  $\boldsymbol{T}_1$ は それを $\boldsymbol{A}$ 行列の左から掛けたときに得られる行列 $\boldsymbol{B}_1=\boldsymbol{T}_1\boldsymbol{A}$ の第 1 列の第 1 行目の要素を 1 とし、それ以外の要素を 0 とする様な行の基本変形行列で、

$$\boldsymbol{T}_1=\begin{bmatrix} 1/a_{11} & 0 & \ldots & 0 & \ldots & 0 \\ -a_{21}/a_{11} & 1 & \ldots & 0 & \ldots & 0 \\ & & \vdots & & \vdots & \\ -a_{i1}/a_{11} & 0 & \ldots & 0 & \ldots & 0 \\ & & \vdots & & \vdots & \\ -a_{r1}/a_{11} & 0 & \ldots & 0 & \ldots & 1 \end{bmatrix} \tag{149}$$

  という形をしている。これを $\boldsymbol{A}$ 行列の左から掛け $\boldsymbol{B}_1$ 行列を作ると、その第 1 列は

$$\begin{aligned} \boldsymbol{T}_1\boldsymbol{a}_{(1)} &= \boldsymbol{T}_1\ (a_{11},a_{21},\ldots,a_{i1},\ldots,a_{r1})' && (150)\\ &= ((1/a_{11})\times a_{11}, -a_{21}/a_{11}\times a_{21}+a_{21},\ldots, && (151)\\ &\quad -a_{i1}/a_{11}\times a_{11}+a_{i1},\ldots,-a_{r1}/a_{11}\times a_{11}+a_{r1})' && (152)\\ &= (1,0,\ldots,0,\ldots,0)' && (153) \end{aligned}$$

  であるから、

$$\boldsymbol{B}_1 \quad = \quad \boldsymbol{T}_1\times\boldsymbol{A}\ =\ \begin{bmatrix} 1 & a_{12}/a_{11} & \ldots & a_{1j}/a_{11} & \ldots & a_{1r}/a_{11} \\ 0 & a_{22}-a_{21}a_{12}/a_{11} & \ldots & a_{2j}-a_{21}a_{1j}/a_{11} & \ldots & a_{2r}-a_{21}a_{1r}/a_{11} \\ & & \vdots & & \vdots & \\ 0 & a_{i2}-a_{i1}a_{1i}/a_{11} & \ldots & a_{ij}-a_{i1}a_{1j}/a_{11} & \ldots & a_{ir}-a_{i1}a_{1r}/a_{11} \\ & & \vdots & & \vdots & \\ 0 & a_{r2}-a_{r1}a_{1r}/a_{11} & \ldots & a_{rj}-a_{r1}a_{1r}/a_{11} & \ldots & a_{rr}-a_{r1}a_{1r}/a_{11} \end{bmatrix} \tag{154}$$

  となり、その第 1 列は第 1 行目に 1 が、それ以外に 0 が入った形となる。

- $\boldsymbol{T}_2$

  $\boldsymbol{T}_2$ は $\boldsymbol{B}_1$ 行列の第 1 列の第 2 行目の要素を 1 とし、それ以外の要素を 0 とする様な行の基本変形行列で、$\boldsymbol{B}_k$ 行列の第 i,j 要素を $b_{ij}^{(k)}$ と書くと、

$$\boldsymbol{T}_2=\begin{bmatrix} 1 & -b_{12}^{(1)}/b_{22}^{(1)} & 0 & \ldots & 0 & \ldots & 0 \\ 0 & 1/b_{22}^{(1)} & 0 & \ldots & 0 & \ldots & 0 \\ & & & \vdots & & \vdots & \\ 0 & -b_{i2}^{(1)}/b_{22}^{(1)} & 0 & \ldots & 0 & \ldots & 0 \\ & & & \vdots & & \vdots & \\ 0 & -b_{r2}^{(1)}/b_{22}^{(1)} & 0 & \ldots & 0 & \ldots & 1 \end{bmatrix} \tag{155}$$

という形をしている。これを $\boldsymbol{B}_1$ 行列の左から掛け $\boldsymbol{B}_2$ 行列を作ると

$$\boldsymbol{B}_2 \;=\; \boldsymbol{T}_2\boldsymbol{B}_1 \;=\; \begin{bmatrix} 1 & 0 & \dots & b_{1j}^{(1)} - b_{12}^{(1)}b_{2j}^{(1)}/b_{22} & \dots & b_{1r}^{(1)} - b_{12}^{(1)}b_{2r}^{(1)}/b_{22}^{(1)} \\ 0 & 1 & \dots & b_{2j}^{(1)}/b_{22}^{(1)} & \dots & b_{2r}^{(1)}/b_{22}^{(1)} \\ & & \vdots & & \vdots & \\ 0 & 0 & \dots & b_{ij}^{(1)} - b_{i2}^{(1)}b_{2j}^{(1)}/b_{22}^{(1)} & \dots & b_{ir}^{(1)} - b_{i2}^{(1)}b_{2r}^{(1)}/b_{22}^{(1)} \\ & & \vdots & & \vdots & \\ 0 & 0 & \dots & b_{rj}^{(1)} - b_{r2}^{(1)}b_{2j}^{(1)}/b_{22}^{(1)} & \dots & b_{rr}^{(1)} - b_{r2}^{(1)}b_{2r}^{(1)}/b_{22}^{(1)} \end{bmatrix} \tag{156}$$

となり、その第 2 列は第 2 行目に 1 が、それ以外に 0 が入っている。

- $\boldsymbol{T}_k$

一般的に $\boldsymbol{T}_k$ は $\boldsymbol{B}_{k-1}$ 行列の第 k 列の第 k 行目の要素を 1 とし、それ以外の要素を 0 とする様な行の基本変形行列で

$$\boldsymbol{T}_k = \begin{bmatrix} 1 & 0 & \dots & -b_{1k}^{(k-1)}/b_{kk}^{(k-1)} & \dots & 0 \\ 0 & 1 & \dots & -b_{2k}^{(k-1)}/b_{kk}^{(k-1)} & \dots & 0 \\ & & \vdots & & \vdots & \\ 0 & 0 & \dots & 1/b_{kk}^{(k-1)} & \dots & 0 \\ & & \vdots & & \vdots & \\ 0 & 0 & \dots & -b_{rk}^{(k-1)}/b_{kk}^{(k-1)} & \dots & 1 \end{bmatrix} \tag{157}$$

という形をしている。これを $\boldsymbol{B}_{k-1}$ 行列の左から掛け $\boldsymbol{B}_k$ 行列を作ると

$$\boldsymbol{B}_k = \boldsymbol{T}_k\boldsymbol{B}_{k-1} = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & b_{1r}^{(k-1)} - b_{1k}^{(k-1)}b_{kr}^{(k-1)}/b_{kk}^{(k-1)} \\ 0 & 1 & \dots & 0 & \dots & b_{2r}^{(k-1)} - b_{2k}^{(k-1)}b_{kr}^{(k-1)}/b_{kk}^{(k-1)} \\ & & \vdots & & \vdots & \\ 0 & 0 & \dots & 1 & \dots & b_{kr}^{(k-1)}/b_{kk}^{(k-1)} \\ & & \vdots & & \vdots & \\ 0 & 0 & \dots & 0 & \dots & b_{rr}^{(k-1)} - b_{rk}^{(k-1)}b_{kr}^{(k-1)}/b_{kk}^{(k-1)} \end{bmatrix} \tag{158}$$

となり、その第 k 列は第 k 行目に 1 が、それ以外に 0 が入っている。

なお、$\boldsymbol{B}_{k-1}$ 行列から $\boldsymbol{B}_k$ 行列を作る際に用いられる $\boldsymbol{B}_{k-1}$ 行列の第 (k,k) 要素、すなわち $b_{kk}^{(k-1)}$ のことをピボット (pivot) と呼ぶ。

ガウスの消去法を行う際に第 k 番目のピボットが 0 となり $\boldsymbol{T}_k$ 行列を定義する際に逆数が取れない場合には、$k+1$ 行以降の第 k 列の要素、$b_{jk}^{(k-1)},\ j = k+1, k+2, \dots, k$ を調べて、もしも $b_{qk}^{(k-1)}$ がその中で絶対値が一番大きなものであれば、$\boldsymbol{B}_{k-1}$ の第 k 行と第 q 行を入れ替えたものを新たに $\boldsymbol{B}_{k-1}$ 行列として $\boldsymbol{T}_k$ 行列を求めればよい。この操作は、partial pivoting と呼ばれるが、並べ替え行列 $\boldsymbol{Q}$ を用いて

$$\boldsymbol{B}_{k-1} \leftarrow \boldsymbol{Q}\boldsymbol{B}_{k-1} \tag{159}$$

と書けるため、最後に

$$\boldsymbol{A}^{-1} = \boldsymbol{T}_r \times \boldsymbol{T}_{r-1} \times \cdots \boldsymbol{T}_k \times \boldsymbol{Q} \times \boldsymbol{T}_{k-1} \times \cdots \times \boldsymbol{T}_2 \times \boldsymbol{T}_1 \tag{160}$$

とすることにより、逆行列を求めることが出来る。もしも、そのような行が見つからない場合には、行列 $\boldsymbol{A}$ は特異 (singular) な行列であり、逆行列を持たない。

なお、実際にガウスの消去法を行う際には、全ての $\boldsymbol{T}_k$ 行列や $\boldsymbol{B}_k$ を保存する必要は無く、$\boldsymbol{A}$ 行列の右に単位行列を並べた拡大行列

$$\boldsymbol{B}_0^{(*)} = [\boldsymbol{A},\quad \boldsymbol{I}_r] \tag{161}$$

を定義し、その各行を $\boldsymbol{T}_k$ 行列を用いて

$$\boldsymbol{B}_1^{(*)} \;=\; \boldsymbol{T}_1\,\boldsymbol{B}_0^{(*)} \;=\; [\boldsymbol{T}_1 \times \boldsymbol{A},\quad \boldsymbol{T}_1 \times \boldsymbol{I}_r] \tag{162}$$

$$\boldsymbol{B}_2^{(*)} \;=\; \boldsymbol{T}_2\,\boldsymbol{B}_1^{(*)} \;=\; [\boldsymbol{T}_2 \times \boldsymbol{B}_1,\quad \boldsymbol{T}_2 \times \boldsymbol{T}_1] \tag{163}$$

$$\boldsymbol{B}_k^{(*)} \;=\; \boldsymbol{T}_k\,\boldsymbol{B}_{k-1}^{(*)} \;=\; [\boldsymbol{T}_k \times \boldsymbol{B}_{k-1},\quad \boldsymbol{T}_k \times \boldsymbol{T}_{k-1} \times \cdots \times \boldsymbol{T}_2 \times \boldsymbol{T}_1] \tag{164}$$

という形で基本変形していけばよく、最終的には

$$\boldsymbol{T}_r\boldsymbol{T}_{r-1}\cdots\boldsymbol{T}_1\,\boldsymbol{A}^{(*)} \;=\; [\boldsymbol{T}_r\boldsymbol{T}_{r-1}\cdots\boldsymbol{T}_1\,\boldsymbol{A},\quad \boldsymbol{T}_r\boldsymbol{T}_{r-1}\cdots\boldsymbol{T}_1\,\boldsymbol{I}_r] \;=\; \left[\boldsymbol{I}_r,\quad \boldsymbol{A}^{-1}\right] \tag{165}$$

を得る。なお partial pivoting を行う際も、$\boldsymbol{B}_{k-1}$ のみならず拡大行列の行を入れ替えることになる。

3.1.2 スウィープ演算子 (Sweep Operator) の利用

逆行列は、以下に示すスウィープ演算子を用いて求めることも出来る。**スウィープ演算子** (sweep operator) は Beaton(1964) や Goodnight (1979) により提案された演算子であり以下のように定義される。$r \times r$ の行列 $\boldsymbol{A}$ が

$$\boldsymbol{A} \quad = \quad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2i} & \cdots & a_{2r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} & \cdots & a_{ir} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{ri} & \cdots & a_{rr} \end{bmatrix} \begin{matrix} \\ \\ \\ i\,\text{行} \\ \\ \\ \end{matrix} \tag{166}$$
$$\qquad\qquad\qquad i\,\text{列}$$

として与えられている場合、行列 $\boldsymbol{A}$ の第 i, i 要素によるスイープを $\mathrm{Swp}\,[\boldsymbol{A} \mid i]$ と記し、それを以下のように定義する。

$$\mathrm{Swp}\,[\boldsymbol{A} \mid i] \;=$$
$$\begin{bmatrix} a_{11}-a_{1i}a_{i1}/a_{ii} & a_{12}-a_{1i}a_{i2}/a_{ii} & \cdots & -a_{1i}/a_{ii} & \cdots & a_{1k}-a_{1i}a_{ik}/a_{ii} & \cdots & a_{1r}-a_{1i}a_{ir}/a_{ii} \\ a_{21}-a_{2i}a_{i1}/a_{ii} & a_{22}-a_{2i}a_{i2}/a_{ii} & \cdots & -a_{2i}/a_{ii} & \cdots & a_{2k}-a_{2i}a_{ik}/a_{ii} & \cdots & a_{2r}-a_{2i}a_{ir}/a_{ii} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \cdots & \vdots \\ a_{i1}/a_{ii} & a_{i2}/a_{ii} & \cdots & 1/a_{ii} & \cdots & a_{ik}/a_{ii} & \cdots & a_{ir}/a_{ii} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \cdots & \vdots \\ a_{j1}-a_{ji}a_{i1}/a_{ii} & a_{j2}-a_{ji}a_{i2}/a_{ii} & \cdots & -a_{ji}/a_{ii} & \cdots & a_{jk}-a_{ji}a_{ik}/a_{ii} & \cdots & a_{jr}-a_{ji}a_{ir}/a_{ii} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \cdots & \vdots \\ a_{r1}-a_{ri}a_{i1}/a_{ii} & a_{r2}-a_{ri}a_{i2}/a_{ii} & \cdots & -a_{ri}/a_{ii} & \cdots & a_{rk}-a_{ri}a_{ik}/a_{ii} & \cdots & a_{rr}-a_{ri}a_{ir}/a_{ii} \end{bmatrix} \begin{matrix} \\ \\ \\ i\,\text{行} \\ \\ j\,\text{行} \\ \\ \\ \end{matrix}$$
$$\qquad\qquad i\,\text{列} \qquad\qquad\qquad k\,\text{列} \tag{167}$$

すなわち行列 $\boldsymbol{A}$ の第 i, i 要素を自身の逆数で、対角を除く第 i 行を $a_{ij}/a_{ii},\ j \neq i$ で、第 i 列を $-a_{ij}/a_{ii},\ j \neq i$ で、それ以外の要素を $a_{jk} - a_{ji}a_{ik}/a_{ii},\ j \neq i, k \neq i$ で置き換えている。

このスイープ演算子は $a_{ii} \neq 0$ の場合、可逆的

$$\mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid i] \mid i] = \boldsymbol{A} \tag{168}$$

であり、また、各スイープにおいて分母となる対角要素（ピボット）が 0 でない場合は可換的、すなわち、順序に依存しないこと

$$\mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid i] \mid j] = \mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid j] \mid i] \tag{169}$$

が知られている。また、一つの要素ではなく、複数の要素に関するスイープを

$$\mathrm{Swp}\,[\boldsymbol{A} \mid i, j, \ldots, k] = \mathrm{Swp}\,[\cdots\ \mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid i] \mid j] \cdots \mid k] \tag{170}$$

と定義すると、ピボットが 0 でない限り、$\mathrm{Swp}\,[\boldsymbol{A} \mid i, j, \ldots, k]$ はその順序に依存せずに、加算的である。

スイープ演算子を用いると、行列の逆行列を求めることが出来る。すなわち、$r \times r$ の行列の全ての対角要素に関してスイープを行った結果は、ピボットが一度も 0 にならなければその行列の逆行列となる。

$$\mathrm{Swp}\,[\boldsymbol{A} \mid 1, 2, \ldots, r] = \mathrm{Swp}\,[\ldots \mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid 1] \mid 2] \ldots \mid r] = \boldsymbol{A}^{-1} \tag{171}$$

もしも、$a_{1,1} = 0$ であったり、$\mathrm{Swp}\,[\boldsymbol{A} \mid 1, 2, \ldots, i-1]$ の i 番目の要素、すなわち第 i 番目のピボットが 0 となった場合には、$i < j$ として、i 行の第 j 番目の要素が 0 でない場合は i 列と j 列を、また、i 列の第 j 番目の要素が 0 でない場合には i 行と j 行を交換してスイープを続けていけば良い。このように、可能な限り行や列を入れ替えることによりスイープを続けていき、最後まで 0 のピボットを避けられた場合、最終的に得られたスイープの結果を並べ替えれば、それが $\boldsymbol{A}$ の逆行列となっている。

また、毎回、$\mathrm{Swp}\,[\boldsymbol{A} \mid 1, 2, \ldots, i-1]$ の第 i 行の $(i, j),\ j = i+1, i+2, \ldots, r$ 要素の絶対値の最大値を与える列 k を探した後に i 列と k 列を入れ替えてスイープを行う方法を partial pivoting と呼び、数値的により安定した方法であることが知られている。

しかし、このような行や列の入れ替えを行っても第 i 番目のピボットが 0 となる場合は、$\boldsymbol{A}$ 行列は特異 (singular) であり、逆行列は存在しない。しかし、第 i 番目のピボットが 0 になった場合、第 i 行と第 i 列を全て 0 として次のピボット $(i+1, i+1)$ から計算を続けていくと、最終的に一般逆行列 (g12-inverse, 反射型一般逆行列) が得られる。

以下に 2×2 の正則な行列の例を示す。

$$\mathrm{Swp}\left[\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mid 1\right] = \begin{bmatrix} 1/a_{11} & a_{12}/a_{11} \\ -a_{21}/a_{11} & a_{22} - a_{21} * a_{12}/a_{11} \end{bmatrix} \tag{172}$$

$$\mathrm{Swp}\left[\begin{bmatrix} 1/a_{11} & a_{12}/a_{11} \\ -a_{21}/a_{11} & a_{22} - a_{21} * a_{12}/a_{11} \end{bmatrix} \mid 2\right] = \tag{173}$$

$$\begin{bmatrix} 1/a_{11} + a_{12}/a_{11} * a_{21}/a_{11}/(a_{22} - a_{21} * a_{12}/a_{11}) & -a_{12}/a_{11}/(a_{22} - a_{21} * a_{12}/a_{11}) \\ -a_{21}/a_{11}/(a_{22} - a_{21} * a_{12}/a_{11}) & 1/(a_{22} - a_{21} * a_{12}/a_{11}) \end{bmatrix} = \tag{174}$$

$$\frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ a_{21} & a_{11} \end{bmatrix} \tag{175}$$

3.1.2.1 分割行列のスイープ

$r \times r$ の行列 $\boldsymbol{A}$ の分割を、

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1i} & \cdots & \boldsymbol{A}_{1k} & \cdots & \boldsymbol{A}_{1G} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2i} & \cdots & \boldsymbol{A}_{2k} & \cdots & \boldsymbol{A}_{2G} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & & \vdots \\ \boldsymbol{A}_{i1} & \boldsymbol{A}_{i2} & \cdots & \boldsymbol{A}_{ii} & \cdots & \boldsymbol{A}_{ik} & \cdots & \boldsymbol{A}_{iG} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & & \vdots \\ \boldsymbol{A}_{j1} & \boldsymbol{A}_{j2} & \cdots & \boldsymbol{A}_{ji} & \cdots & \boldsymbol{A}_{jk} & \cdots & \boldsymbol{A}_{jG} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & & \vdots \\ \boldsymbol{A}_{G1} & \boldsymbol{A}_{G2} & \cdots & \boldsymbol{A}_{Gi} & \cdots & \boldsymbol{A}_{Gk} & \cdots & \boldsymbol{A}_{GG} \end{bmatrix} \tag{176}$$

とする。ただし、行列 $\boldsymbol{A}_{ij}, i = 1, 2, \ldots, G, j = 1, 2, \ldots, G$ は $r_i \times r_j$ の行列で $\sum_{i=1}^{G} r_i = r$ である。この場合、第 i 番目の対角ブロック $\boldsymbol{A}_{ii}$ によるスイープ $\mathrm{Swp}\left[\boldsymbol{A} \mid (i)\right]$ を

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (i)\right] = \begin{bmatrix} \boldsymbol{A}_{11} - \boldsymbol{A}_{1i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i1} & \boldsymbol{A}_{12} - \boldsymbol{A}_{1i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i2} & \cdots & -\boldsymbol{A}_{1i}\boldsymbol{A}_{ii}^{-1} & \cdots & \boldsymbol{A}_{1k} - \boldsymbol{A}_{1i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{ik} & \cdots & \boldsymbol{A}_{1G} - \boldsymbol{A}_{1i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{iG} \\ \boldsymbol{A}_{21} - \boldsymbol{A}_{2i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i1} & \boldsymbol{A}_{22} - \boldsymbol{A}_{2i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i2} & \cdots & -\boldsymbol{A}_{2i}\boldsymbol{A}_{ii}^{-1} & \cdots & \boldsymbol{A}_{2k} - \boldsymbol{A}_{2i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{ik} & \cdots & \boldsymbol{A}_{2G} - \boldsymbol{A}_{2i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{iG} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \cdots & \vdots \\ \boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i1} & \boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i2} & \cdots & \boldsymbol{A}_{ii}^{-1} & \cdots & \boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{ik} & \cdots & \boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{iG} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \cdots & \vdots \\ \boldsymbol{A}_{j1} - \boldsymbol{A}_{ji}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i1} & \boldsymbol{A}_{j2} - \boldsymbol{A}_{ji}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i2} & \cdots & -\boldsymbol{A}_{ji}\boldsymbol{A}_{ii}^{-1} & \cdots & \boldsymbol{A}_{jk} - \boldsymbol{A}_{ji}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{ik} & \cdots & \boldsymbol{A}_{jG} - \boldsymbol{A}_{ji}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{iG} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \cdots & \vdots \\ \boldsymbol{A}_{G1} - \boldsymbol{A}_{Gi}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i1} & \boldsymbol{A}_{G2} - \boldsymbol{A}_{Gi}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i2} & \cdots & -\boldsymbol{A}_{Gi}\boldsymbol{A}_{ii}^{-1} & \cdots & \boldsymbol{A}_{Gk} - \boldsymbol{A}_{Gi}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{ik} & \cdots & \boldsymbol{A}_{GG} - \boldsymbol{A}_{Gi}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{iG} \end{bmatrix} \tag{177}$$

と定義するが、これは、第 i 番目の行列の全ての対角要素に関するスイープに等しい。

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (i)\right] = \mathrm{Swp}\left[\boldsymbol{A} \mid \textstyle\sum_{j=1}^{i-1} r_j + 1, \sum_{j=1}^{i-1} r_j + 2, \ldots, \sum_{j=1}^{i} r_j\right] \tag{178}$$

たとえば、行列が 2 分割されている場合には

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (1)\right] = \begin{bmatrix} \boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} \\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} \end{bmatrix} \tag{179}$$

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (2)\right] = \begin{bmatrix} \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21} & -\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1} \\ \boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21} & \boldsymbol{A}_{22}{}^{-1} \end{bmatrix} \tag{180}$$

であり、行列が 3 分割されている場合には

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (1)\right] = \begin{bmatrix} \boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} & \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{13} \\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} & \boldsymbol{A}_{23} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{13} \\ -\boldsymbol{A}_{31}\boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{32} - \boldsymbol{A}_{31}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} & \boldsymbol{A}_{33} - \boldsymbol{A}_{31}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{13} \end{bmatrix} \tag{181}$$

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (2)\right] = \begin{bmatrix} \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21} & -\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1} & \boldsymbol{A}_{13} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{23} \\ \boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21} & \boldsymbol{A}_{22}{}^{-1} & \boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{23} \\ \boldsymbol{A}_{31} - \boldsymbol{A}_{32}\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21} & -\boldsymbol{A}_{32}\boldsymbol{A}_{22}{}^{-1} & \boldsymbol{A}_{33} - \boldsymbol{A}_{32}\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{23} \end{bmatrix} \tag{182}$$

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (3)\right] = \begin{bmatrix} \boldsymbol{A}_{11} - \boldsymbol{A}_{13}\boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{31} & \boldsymbol{A}_{12} - \boldsymbol{A}_{13}\boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{32} & -\boldsymbol{A}_{13}\boldsymbol{A}_{33}{}^{-1} \\ \boldsymbol{A}_{21} - \boldsymbol{A}_{23}\boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{31} & \boldsymbol{A}_{22} - \boldsymbol{A}_{23}\boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{32} & -\boldsymbol{A}_{23}\boldsymbol{A}_{33}{}^{-1} \\ \boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{31} & \boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{32} & \boldsymbol{A}_{33}{}^{-1} \end{bmatrix} \tag{183}$$

となる。

可逆性、可換性や加算性等の性質は分割行列の場合も同様であり、

$$\mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (i)\right] \mid (i)\right] = \boldsymbol{A} \tag{184}$$

$$\mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (i)\right] \mid (j)\right] = \mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (j)\right] \mid (i)\right] \tag{185}$$

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (i,j,\ldots,k)\right] = \mathrm{Swp}\left[\cdots\ \mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (i)\right] \mid (j)\right]\cdots \mid (k)\right] \tag{186}$$

である。また、全ての対角ブロックに関してスウィープを行った結果は基の行列の逆行列となる。すなわち

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (1,2,\ldots,G)\right] = \boldsymbol{A}^{-1} \tag{187}$$

である。

分割行列の逆行列

$\boldsymbol{A}_{11}$ と $\boldsymbol{A}_{22}$ を正方行列として行列 $\boldsymbol{A}$ が

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \tag{188}$$

と分割されている場合、スウィープ演算子の可換性から、$\boldsymbol{A}$ の逆行列は、以下のふた通りに書くことが出来る。

$$\boldsymbol{A}^{-1} = \mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (1)\right] \mid (2)\right] \tag{189}$$

$$= \begin{bmatrix} \boldsymbol{A}_{11}{}^{-1} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} & -\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1} \\ -\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{22.1}{}^{-1} \end{bmatrix} \tag{190}$$

$$= \mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (2)\right] \mid (1)\right] \tag{191}$$

$$= \begin{bmatrix} \boldsymbol{A}_{11.2}{}^{-1} & -\boldsymbol{A}_{11.2}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1} \\ -\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11.2}{}^{-1} & \boldsymbol{A}_{22}{}^{-1} + \boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11.2}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1} \end{bmatrix} \tag{192}$$

ただし、残差積和行列 (residual sum of squares and cros-product matrix) を

$$\boldsymbol{A}_{ii.j} = \boldsymbol{A}_{ii} - \boldsymbol{A}_{ij}\boldsymbol{A}_{jj}^{-1}\boldsymbol{A}_{ji} \tag{193}$$

と置いている。

残差積和行列の逆行列

上式の (2,2) 要素を比べると

$$\boldsymbol{A}_{ii.j}^{-1} = \boldsymbol{A}_{ii}^{-1} + \boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{ij}\boldsymbol{A}_{jj.i}^{-1}\boldsymbol{A}_{ji}\boldsymbol{A}_{ii}^{-1} \tag{194}$$

を得る。

Woodburry-Lindley-Smith の公式

残差積和行列の逆行列公式において、

$$\boldsymbol{A}_{ii} \to \boldsymbol{A},\ \boldsymbol{A}_{ij} \to \boldsymbol{B},\ -\boldsymbol{A}_{jj}^{-1} \to \boldsymbol{C},\ \boldsymbol{A}_{ji} \to \boldsymbol{D} \tag{195}$$

とおけば、Woodburry-Lindley-Smith の公式

$$(\boldsymbol{A} + \boldsymbol{BCD})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}\Big(\boldsymbol{C}^{-1} + \boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}\Big)^{-1}\boldsymbol{D}\boldsymbol{A}^{-1} \tag{196}$$

を得る。

ブロック下三角行列の逆行列

また、$\boldsymbol{A}_{12} = \boldsymbol{O}$ とおくと ブロック下三角行列の逆行列

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{O} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{A}_{11}^{-1} & \boldsymbol{O} \\ -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{A}_{22}^{-1} \end{bmatrix} \tag{197}$$

を得る。

***** Sweep8.tex に詳細あり ******

3.1.3 その他の方法

逆行列は、様々な行列の分解を用いて計算することも可能であり、数値的にはこちらの方が安定していると言われている。例えば、行列の QR 分解が $\boldsymbol{A} = \boldsymbol{QR}$ の場合には

$$\boldsymbol{A}^{-1} = \boldsymbol{R}^{-1}\boldsymbol{Q}' \tag{198}$$

であり、行列の特異値分解が $\boldsymbol{A} = \boldsymbol{UDV}'$ の場合には

$$\boldsymbol{A}^{-1} = \boldsymbol{V}\boldsymbol{D}^{-1}\boldsymbol{U}' \tag{199}$$

となる。また、行列が対称行列の場合は、固有値分解が $\boldsymbol{A} = \boldsymbol{P\Lambda P}'$ の場合には

$$\boldsymbol{A}^{-1} = \boldsymbol{P\Lambda}^{-1}\boldsymbol{P}' \tag{200}$$

となる。

3.2 偶に使う逆行列の公式

以下の掛け算が可能で、かつ全ての逆行列が取れる場合の公式。

1. $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}\boldsymbol{B}^{-1}$
2. $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}\boldsymbol{A}^{-1}$
3. $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}\boldsymbol{A}^{-1}$
4. $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1} - \boldsymbol{B}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}\boldsymbol{B}^{-1}$
5. $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1}$
6. $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} - (\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}$
7. $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1} - \boldsymbol{B}^{-1}\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}$
8. $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1} - (\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}$
9. $(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}$
   (Woodburry-Lindley-Smith の公式)
10. $(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1}\boldsymbol{B} = \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}^{-1}$
11. $\boldsymbol{D}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1} = \boldsymbol{C}^{-1}(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}$
12. $\boldsymbol{D}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1}\boldsymbol{B} = \boldsymbol{C}^{-1} - \boldsymbol{C}^{-1}(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}^{-1}$

（註）
公式 1-4 は $\boldsymbol{A}$ と $\boldsymbol{B}$ を $\boldsymbol{A}^{-1}$ と $\boldsymbol{B}^{-1}$ で置き換える。
公式 5-8 は $(\boldsymbol{A}+\boldsymbol{B})^{-1}$ を残す。
公式 11-14 は $\boldsymbol{A}$ と $\boldsymbol{C}$ を $\boldsymbol{A}^{-1}$ と $\boldsymbol{C}^{-1}$ で置き換える。
公式 1-4 は、たとえば、

$$(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1} = \boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B} \tag{201}$$

という形で用いる場合が多い。
公式 5-8 は、 $\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}$ の様な形の式を単純化するときに利用されることが多い。

$$\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{I} - \boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1} \quad \text{formula5} \tag{202}$$
$$(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A} = \boldsymbol{I} - (\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B} \quad \text{formula6} \tag{203}$$
$$\boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1} \quad \text{formula7} \tag{204}$$
$$(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B} = \boldsymbol{I} - (\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A} \quad \text{formula8} \tag{205}$$

である。

3.3 行列式

行列式の直感的な説明は以下の通り。

r 次元空間の標準基底 $\boldsymbol{e}$（座標軸を表す長さが 1 の r 本のベクトル）の頂点で形成される r 次元の超立方体（hypercube、r=2 の時は正方形、r=3 の時は立方体） の体積は 1 である。この直交基底で作られる超立方体を r 次の正方行列 $\boldsymbol{A}$ による線形変換で

$$\boldsymbol{v} = \boldsymbol{A}\boldsymbol{e} \tag{206}$$

と変換した超立方体 $\boldsymbol{v}$ の体積の大きさを表すのが $\boldsymbol{A}$ の行列式 $|\boldsymbol{A}| = det(\boldsymbol{A})$ の絶対値であり、その符号は、変換に際して何度座標軸の向きが反転されたかを示す。したがって、この変換により得られる図形の次元が r より小さいものになれば、体積は 0 となるので行列式も 0 となる。これは、正方形を変換して直線が得られる場合、また、立方体が平面に潰されるような場合である。

スウィープ演算子を用いて partial pivoting を利用して逆行列を求める場合に得られる全てのピボットを掛け合わせたものはその行列の行列式となることが知られている。たとえば、スカラーの行列式はそのスカラーそのもの、また、2×2 の行列の行列式は

$$\left| \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right| \;=\; a_{11} \times \; (a_{11}a_{22} - a_{12}a_{21})/a_{11} \;=\; a_{11}a_{22} - a_{21}a_{12} \tag{207}$$

である。partial pviting を行ってもピボットが 0 となる場合には行列式の値は 0 となる。すなわち、特異行列の行列式がの値は 0 である。また、三角行列の行列式は、その対角要素を掛け合わせたものとなる。

行列式に関しては以下の演算規則が成り立つ。ただし、全ての行列はそのサイズが r の正方行列とする。

$$|\boldsymbol{A}\boldsymbol{B}| \;=\; |\boldsymbol{A}| \times |\boldsymbol{B}| \;=\; |\boldsymbol{B}\boldsymbol{A}| \tag{208}$$

$$\left|\boldsymbol{A}^{-1}\right| \;=\; \frac{1}{|\boldsymbol{A}|}, \qquad |\boldsymbol{A}| \;=\; \left|\boldsymbol{A}'\right|, \qquad |\lambda\boldsymbol{A}| \;=\; \lambda^r|\boldsymbol{A}| \tag{209}$$

また、直交行列の行列式の絶対値は 1 である。

分割行列の場合には、ピボットとなる行列の行列式を掛け合わせたものが、元の行列の行列式となる。たとえば、2×2 の分割行列の場合、 (513) 式 の $\mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid (1)] \mid (2)]$ を用いて逆行列を計算する際には、最初のピボットが $\boldsymbol{A}_{11}$, その結果として得られる行列が (613) 式 であるから、２番目のピボットは $\boldsymbol{A}_{22} - \boldsymbol{A}_{21}{\boldsymbol{A}_{11}}^{-1}\boldsymbol{A}_{12}$ となる。したがって、この行列の行列式は、

$$|\boldsymbol{A}| \;=\; |\boldsymbol{A}_{11}| \times \left|\boldsymbol{A}_{22} - \boldsymbol{A}_{21}{\boldsymbol{A}_{11}}^{-1}\boldsymbol{A}_{12}\right| \tag{210}$$

となる。同様に $\mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid (2)] \mid (1)]$ を用いた場合は

$$|\boldsymbol{A}| \;=\; |\boldsymbol{A}_{22}| \times \left|\boldsymbol{A}_{11} - \boldsymbol{A}_{12}{\boldsymbol{A}_{22}}^{-1}\boldsymbol{A}_{21}\right| \tag{211}$$

となり、この二つは同値である。したがって、上式において $\boldsymbol{A}_{11} = \boldsymbol{I}_r,\ \boldsymbol{A}_{22} = \boldsymbol{I}_c$ とすれば行列式に関して以下の公式を得る。

$$|\boldsymbol{I}_r - \boldsymbol{A}_{21}\boldsymbol{A}_{12}| \;=\; |\boldsymbol{I}_c - \boldsymbol{A}_{12}\boldsymbol{A}_{21}| \tag{212}$$

4　行列の列空間・基底・直交化

4.1　行列の列空間

$n \times q$ の行列 $\boldsymbol{X}$ の各列を

$$\boldsymbol{X} \quad = \quad [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_q] \tag{213}$$

と記すとき、行列 $\boldsymbol{X}$ の列空間とは、n 個の列ベクトルの線形結合

$$\boldsymbol{b} \ = \ \sum_{j=1}^{q} (\beta_j \times \boldsymbol{x}_j) \ = \ \boldsymbol{X}\boldsymbol{\beta} \tag{214}$$

として作られる全てのベクトルの集合のことである。ただし、β_j は $\boldsymbol{X}$ 行列の第 j 列、$\boldsymbol{x}_j$、への任意の重みであり、$\boldsymbol{\beta}$ はそれを並べた $q \times 1$ のベクトルである。

あるベクトル $\boldsymbol{b}$ が行列 $\boldsymbol{X}$ の列空間に含まれることを

$$\boldsymbol{b} \in \ \mathcal{C}(\boldsymbol{X}) \tag{215}$$

と書く。このとき、(214) 式 を満たすベクトル $\boldsymbol{\beta}$ が存在する。

$\boldsymbol{b}$ が $\boldsymbol{X}$ の列空間に含まれる場合、正則な行列 $\boldsymbol{A}$ を用いて作られる行列

$$\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{A} \tag{216}$$

を定義すると、

$$\boldsymbol{b} = \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}\boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{\beta} = (\boldsymbol{X}\boldsymbol{A})\left(\boldsymbol{A}^{-1}\boldsymbol{\beta}\right) = \boldsymbol{Z}\boldsymbol{\beta}^{*} \tag{217}$$

となり、$\boldsymbol{b}$ は行列 $\boldsymbol{Z}$=$\boldsymbol{X}\boldsymbol{A}$ の列空間に含まれることが分かる。また、このことは、$\boldsymbol{b} \in \mathcal{C}(\boldsymbol{X})$ を満たす全てのベクトルに関して成り立つため

$$\boldsymbol{b} \in \mathcal{C}(\boldsymbol{X}) \longleftrightarrow \ \boldsymbol{b} \in \ \mathcal{C}(\boldsymbol{X}\boldsymbol{A}) \tag{218}$$

である。すなわち、

1：行列の列空間

$\boldsymbol{X}$ の列空間と $\boldsymbol{Z}$=$\boldsymbol{X}\boldsymbol{A}$ の列空間は、$\boldsymbol{A}$ が正則の場合等しくなる。

このことは、$\boldsymbol{X}$ の列空間に含まれるベクトルは $\boldsymbol{X}$ 行列の各列 $\boldsymbol{x}_j, j = 1, 2, \ldots, q$ の線形結合だけではなく正則な行列 $\boldsymbol{A}$ を用いて作られる新しい行列 $\boldsymbol{Z}$ の各列 $\boldsymbol{z}_j, j = 1, 2, \ldots, q$ の線形結合としても表されることを示しているが、このような、列空間に含まれるベクトルを表現するベクトルのセットのことを列空間の基底ベクトルと呼ぶ。

基底ベクトルは、これまで見てきたように、一意には定まらず、$\boldsymbol{A}$ としてどのような行列を用いるかに依存している。ここでは、$n \geq \ q$ の場合に、回帰分析を利用した、$\boldsymbol{Z}$ の各列が直交するような基底、すなわち直交基底を選ぶ方法について説明する。

4.2 回帰分析と射影子

$$\boldsymbol{y} = \widehat{\boldsymbol{y}} + \boldsymbol{\varepsilon} \tag{219}$$
$$\widehat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\beta} \tag{220}$$

という回帰分析モデルを考える。このモデルは $\boldsymbol{y}$ のモデルによる予測値 $\widehat{\boldsymbol{y}}$ が、説明変数行列 $\boldsymbol{X}$ の列空間に含まれるということを意味している。残差の２乗和

$$\mathrm{RSS} = \sum_{i=1}^{n}(y_i - \widehat{y_i})^2 = \mathrm{tr}\left[(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})\right] \tag{221}$$

を最小とするような回帰係数 $\widehat{\boldsymbol{\beta}}$ を求めると

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{222}$$

となる。そして、この回帰係数の推定値を用いて計算されるモデルの予測値は以下のように書き表すことが出来る。

$$\widehat{\boldsymbol{y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} \tag{223}$$
$$= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{224}$$
$$= \boldsymbol{P}(\boldsymbol{X})\boldsymbol{y} \tag{225}$$

ただし、

$$\boldsymbol{P}(\boldsymbol{X}) = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \tag{226}$$

は、$\boldsymbol{X}$ の列空間への射影子（射影作用素 projection operator）と呼ばれる、対称かつ冪等

$$\boldsymbol{P}(\boldsymbol{X}) = \boldsymbol{P}(\boldsymbol{X})',\ \boldsymbol{P}(\boldsymbol{X})^2 = \boldsymbol{P}(\boldsymbol{X}) \tag{227}$$

な正方行列であり、その階数は、$\boldsymbol{X}$ の階数に等しい。（階数に関しては後述）また、

$$\boldsymbol{P}(\boldsymbol{X})\boldsymbol{X} = \boldsymbol{X} \tag{228}$$

である。

特に、$q = 1$ の場合、

$$\boldsymbol{P}(\boldsymbol{x}) = \boldsymbol{x}(\boldsymbol{x}'\boldsymbol{x})^{-1}\boldsymbol{x}' = \frac{1}{\boldsymbol{x}'\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}' \tag{229}$$

であるから、

$$\boldsymbol{P}(\boldsymbol{x})\boldsymbol{x} = \frac{1}{\boldsymbol{x}'\boldsymbol{x}}\boldsymbol{x}\boldsymbol{x}'\ \ \boldsymbol{x} = \frac{\boldsymbol{x}'\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}}\boldsymbol{x} = \boldsymbol{x} \tag{230}$$

である。

任意のベクトル $\boldsymbol{b}$ に $\boldsymbol{P}(\boldsymbol{X})$ を適用したものは、

$$\boldsymbol{P}(\boldsymbol{X})\boldsymbol{b} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{b} \tag{231}$$
$$= \boldsymbol{X}\boldsymbol{\beta} \tag{232}$$

ただし

$$\boldsymbol{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{b} \tag{233}$$

と書けるため、$\boldsymbol{P}(\boldsymbol{X})\boldsymbol{b}$ は $\boldsymbol{X}$ の列空間に入っていることが分かるが、これが、$\boldsymbol{X}$ の列空間への射影子という名称の由来である。

なお、行列の列空間への射影子は、その空間の基底の選び方に依存しない。すなわち、

$$\begin{aligned}
\boldsymbol{P}(\boldsymbol{X}\boldsymbol{A}) &= \boldsymbol{X}\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{X}' && (234)\\
&= \boldsymbol{X}\boldsymbol{A}\boldsymbol{A}^{-1}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{A}'^{-1}\boldsymbol{A}'\boldsymbol{X}' && (235)\\
&= \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}' && (236)
\end{aligned}$$

である。

また、

$$\boldsymbol{Q}(\boldsymbol{X}) = \boldsymbol{I} - \boldsymbol{P}(\boldsymbol{X}) = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \tag{237}$$

で定義される行列を $\boldsymbol{X}$ の列空間の直交補空間への射影子と呼ぶが、これもまた、対称かつ冪等な正方行列となり、しかも $\boldsymbol{P}(\boldsymbol{X})$ と $\boldsymbol{Q}(\boldsymbol{X})$ は互いに直交する。

$$\boldsymbol{P}(\boldsymbol{X})\boldsymbol{Q}(\boldsymbol{X}) = \boldsymbol{Q}(\boldsymbol{X})\boldsymbol{P}(\boldsymbol{X}) = \boldsymbol{O} \tag{238}$$

したがって、

$$\boldsymbol{Q}(\boldsymbol{X})\boldsymbol{X} = \boldsymbol{Q}(\boldsymbol{X})\boldsymbol{P}(\boldsymbol{X})\boldsymbol{X} = \boldsymbol{O} \tag{239}$$

である。なお、回帰分析に於ける残差ベクトルは (225) 式 より

$$\boldsymbol{y} - \widehat{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{P}(\boldsymbol{X})\boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{X}))\boldsymbol{y} = \boldsymbol{Q}(\boldsymbol{X})\boldsymbol{y} \tag{240}$$

となっているが、これは、線形の回帰分析においては、**最小２乗法を用いて推定された回帰係数 $\widehat{\beta}$ を用いて計算された残差ベクトルは、説明変数行列 X ならびに基準変数 y の予測値 $\widehat{y}$ と直交する**ことを意味している。

いま、$\boldsymbol{Q}(\boldsymbol{X})$ の定義から

$$\boldsymbol{P}(\boldsymbol{X}) + \boldsymbol{Q}(\boldsymbol{X}) = \boldsymbol{I} \tag{241}$$

であるから、このふたつの射影子を用いれば、ベクトル $\boldsymbol{y}$ を

$$\begin{aligned}
\boldsymbol{y} &= \left(\boldsymbol{P}(\boldsymbol{X}) + \boldsymbol{Q}(\boldsymbol{X})\right)\boldsymbol{y} \;=\; \boldsymbol{P}(\boldsymbol{X})\boldsymbol{y} + \boldsymbol{Q}(\boldsymbol{X})\boldsymbol{y} && (242)\\
&= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} + (\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{y} && (243)
\end{aligned}$$

とふたつの互いに直交するベクトルの和として書くことが出来る。すなわち、このことは、**任意のベクトル y が、任意の行列 X の列空間に含まれる部分 $P(X)y$ とそれに直交する部分 $Q(X)y$ のふたつの直交するベクトルの和として書き表される**ことを示している。なお、このようなベクトルが複数ある場合、それを行列 $\boldsymbol{Y}$ の形に並べると、

$$\begin{aligned}
\boldsymbol{Y} &= \boldsymbol{P}(\boldsymbol{X})\boldsymbol{Y} + \boldsymbol{Q}(\boldsymbol{X})\boldsymbol{Y} && (244)\\
&= \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Y} + \left(\boldsymbol{I} - \boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\right)\boldsymbol{Y} && (245)
\end{aligned}$$

となるが、これは、$\boldsymbol{Y}$ の列として含まれる全てのベクトルが、$\boldsymbol{X}$ の列空間に含まれる部分（多変量重回帰のモデルの予測値）と、それの直交する部分（残差）とに分けられることを示している。

最後に、行列 $\boldsymbol{X}$ が $\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2]$ と二つに分割されていて、しかも、そのふたつの行列 $\boldsymbol{X}_1$ と $\boldsymbol{X}_2$ が直交する場合、すなわち、

$$\boldsymbol{X}_1'\boldsymbol{X}_2 = \boldsymbol{O} \tag{246}$$

の場合には、$\boldsymbol{X}$ の列空間への射影子は

$$\begin{aligned}
\boldsymbol{P}(\boldsymbol{X}) &= \boldsymbol{P}([\boldsymbol{X}_1, \boldsymbol{X}_2]) && (247)\\
&= \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{X}_2 \\ \boldsymbol{X}_2'\boldsymbol{X}_1 & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \end{bmatrix} && (248)\\
&= \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{bmatrix} \begin{bmatrix} (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} & \boldsymbol{O} \\ \boldsymbol{O} & (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \end{bmatrix} && (249)\\
&= \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' + \boldsymbol{X}_2(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2' && (250)\\
&= \boldsymbol{P}(\boldsymbol{X}_1) + \boldsymbol{P}(\boldsymbol{X}_2) && (251)
\end{aligned}$$

と、$\boldsymbol{X}_1$ の列空間への射影子と $\boldsymbol{X}_2$ への射影子行列の和として表現できる。同様に、回帰係数行列は

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &= \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{X}_2 \\ \boldsymbol{X}_2'\boldsymbol{X}_1 & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \end{bmatrix} \boldsymbol{y} && (252)\\
&= \begin{bmatrix} (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} & \boldsymbol{O} \\ \boldsymbol{O} & (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{y} \\ \boldsymbol{X}_2'\boldsymbol{y} \end{bmatrix} && (253)\\
&= \begin{bmatrix} (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{y} \\ (\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{y} \end{bmatrix} && (254)
\end{aligned}$$

となり、これは、 $\boldsymbol{y}$ を別々にそれぞれ $\boldsymbol{X}_1$ と $\boldsymbol{X}_2$ へ回帰した時に得られる回帰係数行列を縦に並べたものと等しくなる。

2：**回帰係数の比較**

以下の回帰モデル

$$\begin{aligned}
\boldsymbol{y} &= \boldsymbol{X}_1\boldsymbol{b}_1 + \boldsymbol{\varepsilon}_1, \quad \boldsymbol{y} = \boldsymbol{X}_2\boldsymbol{b}_2 + \boldsymbol{\varepsilon}_2 && (255)\\
\boldsymbol{y} &= \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon} && (256)\\
\boldsymbol{y} &= \boldsymbol{X}_1\boldsymbol{\gamma}_1 + \boldsymbol{Q}(\boldsymbol{X}_1)\boldsymbol{X}_2\boldsymbol{\gamma}_2 + \boldsymbol{\varepsilon} && (257)\\
\boldsymbol{y} &= \boldsymbol{Q}(\boldsymbol{X}_2)\boldsymbol{X}_1\boldsymbol{\lambda}_1 + \boldsymbol{X}_2\boldsymbol{\lambda}_2 + \boldsymbol{\varepsilon} && (258)\\
\boldsymbol{y} &= \boldsymbol{Q}(\boldsymbol{X}_2)\boldsymbol{X}_1\boldsymbol{\delta}_1 + \boldsymbol{\varepsilon}_3, \quad \boldsymbol{y} = \boldsymbol{Q}(\boldsymbol{X}_1)\boldsymbol{X}_2\boldsymbol{\delta}_2 + \boldsymbol{\varepsilon}_4 && (259)\\
\boldsymbol{y} &= \boldsymbol{Q}(\boldsymbol{X}_2)\boldsymbol{X}_1\boldsymbol{\kappa}_1 + \boldsymbol{Q}(\boldsymbol{X}_1)\boldsymbol{X}_2\boldsymbol{\kappa}_2 + \boldsymbol{\varepsilon} && (260)
\end{aligned}$$

におけるパラメタ間には

$$\begin{aligned}
\boldsymbol{\beta}_1 &= \boldsymbol{\delta}_1 = \boldsymbol{\lambda}_1 && (261)\\
\boldsymbol{\beta}_2 &= \boldsymbol{\delta}_2 = \boldsymbol{\gamma}_2 && (262)\\
\boldsymbol{b}_1 &= \boldsymbol{\gamma}_1, \quad \boldsymbol{b}_2 = \boldsymbol{\lambda}_2 && (263)\\
& && (264)
\end{aligned}$$

という関係がある。

4.2.1 射影子の幾何学的解釈

n 次元空間の中の任意のベクトル $\boldsymbol{u}$ を、$n \times q$ の行列 $\boldsymbol{X}$ の列空間へ射影することを考える。$\boldsymbol{X}$ は n 次元のベクトルが q 個集まったものであると考えられ、行列 $\boldsymbol{X}$ の列空間への射影子は (229) 式 で定義される。

いま、 $\boldsymbol{X}$ が $n = 2, q = 2$ でフルランクの場合を考えてみると、これは 2 次元空間の基底となっている。したがって、$\boldsymbol{u}$ を2次元空間の中の任意のベクトルとすると、それは

$$\boldsymbol{u} = \boldsymbol{X}\boldsymbol{\beta} \tag{265}$$

と書けるため、

$$\boldsymbol{P}(\boldsymbol{X})\boldsymbol{u} = \boldsymbol{P}(\boldsymbol{X})\ \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{u} \tag{266}$$

となる。これは $n = q$ である限り $n > 2$ の場合にも同様である。

一方、$\boldsymbol{X}$ として $n = 2, q = 1$ の場合、すなわち $\boldsymbol{X} = \boldsymbol{x}_1$ の場合を考えると、$\boldsymbol{X}$ の列空間、すなわち、2次元空間の中で原点から $\boldsymbol{x}_1$ で表されるベクトルの頂点方向へ向かう直線、への射影子は

$$\boldsymbol{P}(\boldsymbol{x}_1) \quad = \quad \frac{1}{\boldsymbol{x}_1'\boldsymbol{x}_1}\ \boldsymbol{x}_1\boldsymbol{x}_1' \tag{267}$$

となるが、

$$\boldsymbol{v} \quad = \quad P(\boldsymbol{X})\ \boldsymbol{u} \ = \ \frac{1}{\boldsymbol{x}_1'\boldsymbol{x}_1}\ \boldsymbol{x}_1\boldsymbol{x}_1'\ \boldsymbol{u} \tag{268}$$

は、ベクトル $\boldsymbol{u}$ の頂点からこの $\boldsymbol{x}_1$ に向かって下ろした垂線の足の元の２次元空間における座標、すなわち、$\boldsymbol{u}$ の $\boldsymbol{x}_1$ への射影 (ベクトル射影）を表すことになる。また、原点から垂線の足までの距離の２乗は

$$||\boldsymbol{v}||^2 = \boldsymbol{v}'\boldsymbol{v} = \boldsymbol{u}'\boldsymbol{P}(\boldsymbol{x}_1)\boldsymbol{u} = \frac{1}{\boldsymbol{x}_1'\boldsymbol{x}_1}\ \boldsymbol{u}'\ \boldsymbol{x}_1\boldsymbol{x}_1'\ \boldsymbol{u} = \frac{(\boldsymbol{u}'\boldsymbol{x}_1)^2}{\boldsymbol{x}_1'\boldsymbol{x}_1} \tag{269}$$

であるから、原点から垂線の足までの距離に基づくこの $\boldsymbol{x}_1$ を座標軸とした座標は、所謂スカラー射影 (scalar projection)

$$\frac{\boldsymbol{u}'\boldsymbol{x}_1}{||\boldsymbol{x}_1||} = \boldsymbol{u}'\frac{\boldsymbol{x}_1}{||\boldsymbol{x}_1||} = ||\boldsymbol{u}||\cos(\theta) \tag{270}$$

となる。ただし、 θ は $\boldsymbol{x}_1$ と $\boldsymbol{u}$ の間の角度である。スカラー射影は、長さを 1 に基準化した $\boldsymbol{x}_1$ 方向のベクトル (基底)、$\frac{1}{||\boldsymbol{x}_1||}\ \boldsymbol{x}_1$ 、と $\boldsymbol{u}$ との内積であり、直線 $\boldsymbol{x}_1$ を座標軸としてみたときの垂線の足の座標である。基底ベクトルと $\boldsymbol{u}$ の間の角が $0 \leq \theta \leq \pi/2$ の場合はこの座標は正となり、ベクトルの向き逆の場合には負となる。なお、このスカラー射影を長さが 1 に基準化された $\boldsymbol{x}_1$ 方向のベクトル (基底）に掛けたものがベクトル $\boldsymbol{u}$ のベクトル $\boldsymbol{x}_1$ への射影 (ベクトル射影）

$$\boldsymbol{u}'\frac{\boldsymbol{x}_1}{||\boldsymbol{x}_1||} \quad \frac{1}{||\boldsymbol{x}_1||}\ \boldsymbol{x}_1 \quad = \quad \frac{1}{||\boldsymbol{x}_1||^2}\ \boldsymbol{x}_1\boldsymbol{x}_1'\boldsymbol{u} \tag{271}$$

となる。

$n = 3,\ q = 1$ の場合も同様で、$\boldsymbol{X}$ として $\boldsymbol{x}_1$ を用いれば、$\boldsymbol{P}(\boldsymbol{x}_1)\boldsymbol{u}$ はベクトル $\boldsymbol{u}$ から $\boldsymbol{x}_1$ へ下ろした垂線の足の元の３次元空間における座標となる。

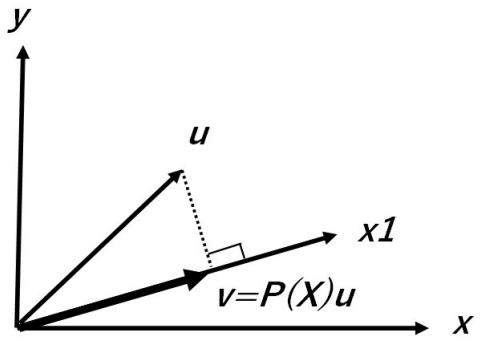

図 1: 直線への射影

次に $n = 3, q = 2$ の場合を考えると、$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2]$ としてその階数が 2 の場合には、$\boldsymbol{X}$ の列空間はベクトル $\boldsymbol{x}_1$ と $\boldsymbol{x}_2$ を基底として持つ平面を形成するため、$\boldsymbol{v} = \boldsymbol{P}(\boldsymbol{X})\boldsymbol{u}$ はベクトル $\boldsymbol{u}$ からその平面に下ろした垂線の足の３次元空間における座標であり、その長さの２乗は $||\boldsymbol{v}||^2 = \boldsymbol{u}'\boldsymbol{P}(\boldsymbol{X})\boldsymbol{u}$ となる。

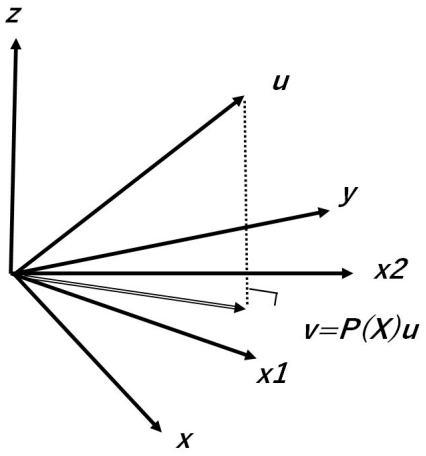

図 2: ２次元平面への射影

また、$\boldsymbol{u}$ の平面の基底ベクトルへのスカラー射影は、θ_1 と θ_2 をそれぞれ $\boldsymbol{u}$ と $\boldsymbol{x}_1$ ならびに $\boldsymbol{x}_2$ の間の角度とすれば、

$$s_1 = \boldsymbol{u}' \frac{\boldsymbol{x}_1}{||\boldsymbol{x}_1||} = ||\boldsymbol{u}|| \cos(\theta_1), \qquad s_2 = \boldsymbol{u}' \frac{\boldsymbol{x}_2}{||\boldsymbol{x}_2||} = ||\boldsymbol{u}|| \cos(\theta_2) \tag{272}$$

であるが、x_1 と x_2 が直交している場合には、(s_1, s_2) が新しい $x_1 - x_2$ 平面での座標となる。そうで無い場合は、s_1 は $\boldsymbol{v}$ から x_1 へ下ろした垂線の足の原点からの x_1 方向への距離、s_2 は $\boldsymbol{v}$ から x_2 へ下ろした垂線の足の原点からの x_2 方向への距離となる。、

一般的には、n 次元空間の中の q 次元の部分空間 (超平面： q=1 の時は直線、 q=2 の時は平面、等）の基底を $n \times q$ の行列 $\boldsymbol{X}$ で表すと、n 次元ベクトル $\boldsymbol{u}$ をこの部分空間へ射影した垂線の足は $\boldsymbol{v} = \boldsymbol{P}(\boldsymbol{X})\boldsymbol{u}$ で表される。また、この超平面上での $\boldsymbol{v}$ の座標は $\boldsymbol{u}$ の $\boldsymbol{X}$ 行列の各列へのスカラー射影となる。

回帰分析の場合は $\boldsymbol{y}$ を n 次元空間のベクトル、$\boldsymbol{X}$ の各列を同じ n 次元空間の中の q 個のベクトルと考えることになるが、$\widehat{\boldsymbol{y}} = \boldsymbol{P}(\boldsymbol{X})\boldsymbol{y}$ は $\boldsymbol{y}$ を $\boldsymbol{X}$ の q 個のベクトルで作られる列空間（q 次元の超平面）へ射影したものとなっている。

4.2.1.1 回転行列と射影

横軸に x 縦軸に y を取った平面上の点 (x, y) を考える。原点からこの点までのベクトルを $\boldsymbol{u}$ とする。２次元空間（平面）の標準基底 standard basis （平面の直交する座標軸を表す二つのベクトル) を

$$\boldsymbol{e}_1 = (1, 0)' \quad \text{and} \quad \boldsymbol{e}_2 = (0, 1)' \tag{273}$$

とすれば、これらは長さが 1 のベクトルであるため単位ベクトル unit vector と呼ばれることもある。また、これらは 2×2 の単位行列の各列に対応している。すなわち

$$\boldsymbol{I}_2 = [\boldsymbol{e}_1, \boldsymbol{e}_2] \tag{274}$$

である。２次元空間内の全てのベクトルは、この標準基底ベクトルを用いて表現することが出来るが、その表現の仕方は、頂点の座標が (u_1, u_2) であるベクトルを $\boldsymbol{u}$ とすれば

$$\boldsymbol{u} \quad = \quad u_1\, \boldsymbol{e}_1 + u_2\, \boldsymbol{e}_2 \;=\; (u_1, u_2)' \tag{275}$$

である。すなわち、ベクトルの頂点の座標を重みとしてそれぞれの標準基底ベクトルを足し合わせれば良い。(逆に、頂点の座標とは、ベクトル $\boldsymbol{u}$ から各標準基底ベクトルへのスカラー射影である。）

次に x 軸との反時計回りの角度が θ で長さが $||\boldsymbol{\varepsilon}|| = d$ のベクトル $\boldsymbol{\varepsilon}$ を考えてみると、$\boldsymbol{\varepsilon}$ と y 軸の間の角度は $\pi/2 - \theta$ であるから、このベクトルの頂点の $x - y$ 平面上の座標はスカラー射影より

$$\varepsilon_1 = d\cos(\theta), \quad \varepsilon_2 = d\cos(\pi/2 - \theta) = d\sin(\theta) \tag{276}$$

となるが、このベクトルの長さを $d = 1$ としたものを $\boldsymbol{e}_1^{(*)}$ とする。すなわち

$$\boldsymbol{e}_1^{(*)} = (\cos(\theta), \sin(\theta))' \tag{277}$$

である。同様に、 y 軸との反時計回りの角度が θ で長さが 1 のベクトル $\boldsymbol{e}_2^{(*)}$ を考えてみると、このベクトルと x 軸との成す角は $\pi/2 + \theta$ であるから、その頂点の $x - y$ 平面上の座標は

$$\boldsymbol{e}_2^{(*)} = (\cos(\pi/2 + \theta), \cos(\theta))' = (-\sin(\theta), \cos(\theta))' \tag{278}$$

となる。この二つのベクトルは直交しているため、この二つのベクトルを横に並べたもの

$$\boldsymbol{E} = \left[\boldsymbol{e}_1^{(*)}, \boldsymbol{e}_2^{(*)}\right] = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \tag{279}$$

をこの平面の新しい標準基底（新しい座標軸）として用いることが出来る。

ここで $\boldsymbol{E}$ の列空間への射影子を考えると

$$\boldsymbol{P}(\boldsymbol{E}) = \boldsymbol{I}_2 \tag{280}$$

となるが、これはものと標準基底 $\boldsymbol{I}_2$ と $\boldsymbol{E}$ の列空間が同じであることを示している。したがって、どのような２次元ベクトル $\boldsymbol{u}$ を $\boldsymbol{E}$ の列空間へ射影してもその垂線の足の座標は変わらない。（平面上のベクトルを平面に射影した時の垂線の足は頂点そのもの。）

しかし、$\boldsymbol{E}$ の各列を新しい座標軸として用いた場合の $\boldsymbol{u}$ の座標は以下のようになる。すなわち、$\boldsymbol{u}$ の $\boldsymbol{e}_1^{(*)}$ と $\boldsymbol{e}_2^{(*)}$ とへのスカラー射影を計算すればよいが、それは、$\boldsymbol{u}$ と $\boldsymbol{e}_1^{(*)}$ ならびに $\boldsymbol{e}_2^{(*)}$ との成す角度を γ_1, γ_2 とすれば、

$$u_1^{(*)} \quad = \quad ||\boldsymbol{u}||\cos(\gamma_1) = ||\boldsymbol{u}|| \times \frac{\boldsymbol{u}'\boldsymbol{e}_1^{(*)}}{||\boldsymbol{u}|| \times ||\boldsymbol{e}_1^{(*)}||} = \boldsymbol{u}'\boldsymbol{e}_1^{(*)} \tag{281}$$

$$u_1^{(*)} \quad = \quad ||\boldsymbol{u}||\cos(\gamma_2) = ||\boldsymbol{u}|| \times \frac{\boldsymbol{u}'\boldsymbol{e}_2^{(*)}}{||\boldsymbol{u}|| \times ||\boldsymbol{e}_2^{(*)}||} = \boldsymbol{u}'\boldsymbol{e}_2^{(*)} \tag{282}$$

となるが、これをまとめれば

$$\boldsymbol{u}^{(*)\prime} \;=\; \boldsymbol{u}'\boldsymbol{E} \qquad \text{or} \qquad \boldsymbol{u}^{(*)} \;=\; \boldsymbol{E}'\boldsymbol{u} \tag{283}$$

となる。この $\boldsymbol{E}$ 行列の転置行列

$$\boldsymbol{E}' \;=\; \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \equiv \boldsymbol{T}_A \tag{284}$$

を座標軸の回転行列と呼んでいる。

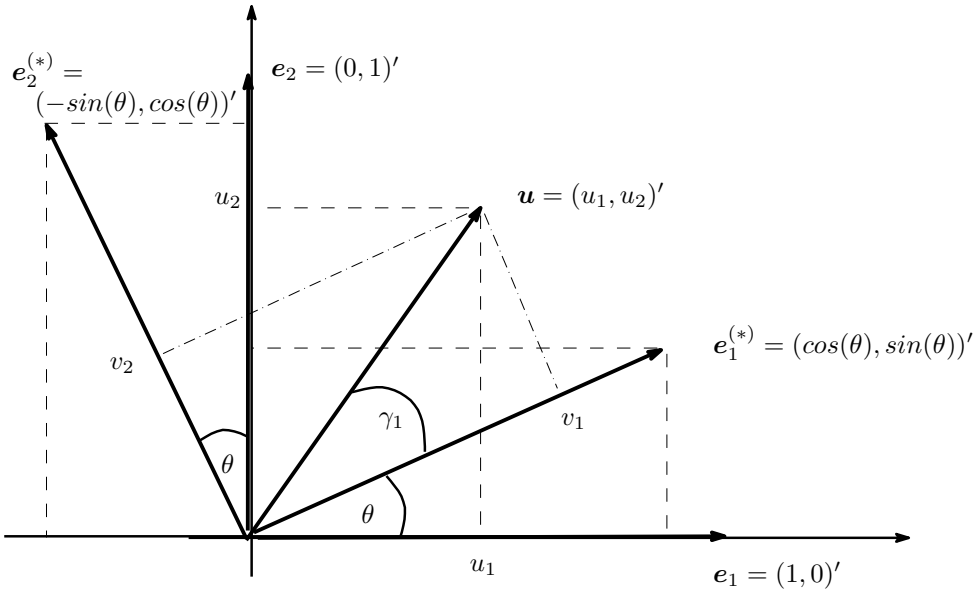

図 3: 標準基底の変換

4.3 行列の直交基底

4.3.1 グラム・シュミットの直交化 The Gram-Schmidt Orthogonalization

回帰分析の残差は説明変数に直交することとや、任意のベクトルが説明変数の列空間に含まれる部分とそれに直交する部分の和に分解できること等を用いれば、以下のような操作をすることにより、$n \times q$ の行列 $\boldsymbol{X}$ の直交基底を求めることが出来る。なお、本節では行列 $\boldsymbol{X}$ を

$$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_q] \tag{285}$$

という形で列ベクトルを横に並べたもの考えるが、列ベクトルの添え字に括弧 () を付けない。

$\boldsymbol{X}$ の直交基底 $\boldsymbol{Q}$ の元となる行列を $\boldsymbol{U}$ とすれば、それは以下のようにして求めることが出来る。

- 第 1 列はそのまま
  $\boldsymbol{u}_1 = \boldsymbol{x}_1$
- 第 2 列は $\boldsymbol{x}_2$ を $\boldsymbol{x}_1$ へ回帰した残差とする。
  $\boldsymbol{u}_2 = \boldsymbol{Q}(\boldsymbol{x}_1)\, \boldsymbol{x}_2$
- 第 3 列は $\boldsymbol{x}_3$ を $\boldsymbol{x}_1$ と $\boldsymbol{x}_2$ へ回帰した残差とする。
  $\boldsymbol{u}_3 = \boldsymbol{Q}(\boldsymbol{x}_1, \boldsymbol{x}_2)\, \boldsymbol{x}_3$
  $\vdots$ $\vdots$
- 第 j 列は $\boldsymbol{x}_j$ を $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{j-1}$ へ回帰した残差とする。
  $\boldsymbol{u}_j = \boldsymbol{Q}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{j-1})\, \boldsymbol{x}_j$
  $\vdots$ $\vdots$
- 第 q 列は $\boldsymbol{x}_q$ を $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{q-1}$ へ回帰した残差とする。
  $\boldsymbol{u}_q = \boldsymbol{Q}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{q-1})\, \boldsymbol{x}_q$

上記のプロセスを経て $n \times q$ の $\boldsymbol{U}$ 行列が求まったなら、それの各列の長さを 1 としたもの

$$\boldsymbol{Q} \quad = \quad \boldsymbol{U}\, \mathrm{diag}\left(\boldsymbol{U}'\boldsymbol{U}\right)^{-\frac{1}{2}} \tag{286}$$

が $\boldsymbol{X}$ 行列の直交基底行列となる。（この $\boldsymbol{Q}$ は射影子を表す $\boldsymbol{Q}$ ではないことに注意。単なる行列で、伝統的に $\boldsymbol{Q}$ と呼ばれるものである。また、グラム・シュミットの直交化で得られる $\boldsymbol{U}$ 行列は後述する特異値分解で得られる $\boldsymbol{U}$ とは異なるものである。）

そして、この直交基底は

$$\boldsymbol{U} \quad = \quad \boldsymbol{X}\, \boldsymbol{A} \tag{287}$$

ただし、$\boldsymbol{A}$ は $q \times q$ の上三角行列、という形で書き表される。なぜなら、$\boldsymbol{x}_j$ を $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{j-1}$ へ回帰したときの回帰係数ベクトルを $\boldsymbol{\beta}_j = [\beta_{1j}, \beta_{2j}, \ldots, \beta_{j-1,j}]$ とすれば、

$$\boldsymbol{u}_j = \boldsymbol{x}_j - \sum_{k=1}^{j-1} \beta_{kj}\, \boldsymbol{x}_k \tag{288}$$

であるから、まとめて

$$[\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4, \ldots, \boldsymbol{u}_q] \;=\; [\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4, \ldots, \boldsymbol{x}_q] \times \begin{bmatrix} 1 & -\beta_{12} & -\beta_{13} & -\beta_{14} & \cdots & -\beta_{1q} \\ 0 & 1 & -\beta_{23} & -\beta_{24} & \cdots & -\beta_{2q} \\ 0 & 0 & 1 & -\beta_{34} & \cdots & -\beta_{3q} \\ 0 & 0 & 0 & 1 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{289}$$

と書けるが、この上三角行列が (287) 式 の $\boldsymbol{A}$ 行列となる。したがって、

$$\boldsymbol{Q} = \boldsymbol{X}\boldsymbol{A}\mathrm{diag}\left(\boldsymbol{A}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}\right)^{-\frac{1}{2}} \;=\; \boldsymbol{X}\boldsymbol{A}^{*} \tag{290}$$

であるため、 $\boldsymbol{X}$ の列空間と $\boldsymbol{U}$ そして $\boldsymbol{Q}$ の列空間は等しくなることがわかる。

また、上式を $\boldsymbol{X}$ に関して解けば

$$\boldsymbol{X} = \boldsymbol{Q}\,{\boldsymbol{A}^{*}}^{-1} \;=\; \boldsymbol{Q}\,\boldsymbol{R} \tag{291}$$

ただし、

$$\boldsymbol{R} = {\boldsymbol{A}^{*}}^{-1} \;=\; \mathrm{diag}\left(\boldsymbol{A}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A}\right)^{\frac{1}{2}}\boldsymbol{A}^{-1} \tag{292}$$

を得るが、この $\boldsymbol{R}$ も $q \times q$ の上三角行列である。

このプロセスはグラム・シュミットの直交化と呼ばれるが、その結果得られた $\boldsymbol{X}$ 行列の直交行列 $\boldsymbol{Q}$ と上三角行列 $\boldsymbol{R}$ の積への分解を $\boldsymbol{X}$ 行列の **QR 分解 (QR decomposition)** と呼ぶ。

表 1 に 8×5 の行列 $\boldsymbol{X}$ のグラム・シュミットの直交化の結果を示す。ただし、この結果は $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{R}$ かつ $\boldsymbol{U} = \boldsymbol{X}\boldsymbol{A}$ を与えるもので、$\boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}$ となる $\boldsymbol{Q}$ 行列はこの $\boldsymbol{U}$ 行列の各列を (290) 式 で正規化したもの、それに対応する $\boldsymbol{R}$ は (292) 式 として与えられる。

表 1: グラム・シュミットの直交化：列フルランクの場合

$\boldsymbol{X}$

	X1	X2	X3	X4	X5
1)	7	3	1	6	9
2)	6	9	5	9	2
3)	6	9	7	9	0
4)	3	6	5	7	8
5)	5	9	8	0	2
6)	7	1	7	1	8
7)	7	8	4	5	4
8)	7	4	1	9	0

$\boldsymbol{U}$

	U1	U2	U3	U4	U5
1)	7	-3.560	-2.280	-0.683	4.687
2)	6	3.377	-1.014	1.460	-0.561
3)	6	3.377	0.986	2.926	-3.710
4)	3	3.189	1.246	3.931	5.900
5)	5	4.315	2.241	-4.605	-0.447
6)	7	-5.560	4.717	0.220	-0.371
7)	7	1.440	-1.771	-3.255	1.207
8)	7	-2.560	-2.778	1.563	-4.072

$\boldsymbol{R}$

	U1	U2	U3	U4	U5
U1	1	0.937	0.722	0.914	0.639
U2	0	1	0.498	0.389	-0.288
U3	0	0	1	-0.733	0.481
U4	0	0	0	1	0.127
U5	0	0	0	0	1

$\boldsymbol{A}$

	U1	U2	U3	U4	U5
X1	1	-0.937	-0.255	-0.736	-0.692
X2	0	1	-0.498	-0.754	0.623
X3	0	0	1	0.733	-0.575
X4	0	0	0	1	-0.127
X5	0	0	0	0	1

なお、上記の手続きを行う際に、ある時点で回帰の残差 $\boldsymbol{u}_j$ が $\boldsymbol{0}$ となった場合、それは $\boldsymbol{x}_j$ が $[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{j-1}]$ の列空間に含まれることを示しているが、この場合には、以降の回帰係数の計算には、$\boldsymbol{x}_j$ を省いて行い、

$\boldsymbol{\beta}_{i,j}, i = 1, 2, \ldots$ と $\boldsymbol{\beta}_{j,i}, i = j+1, j+2, \ldots$ を全て 0 として先に進んで行く。すると、最終的にいくつかの全ての要素が 0 である列を持つ $\boldsymbol{U}$ 行列と全ての要素が 0 である行を持つ $\boldsymbol{A}$ 行列が生成されるが、この、$\boldsymbol{U}$ 行列の中の全てが 0 でない列の数を、行列の**階数 (rank)** と呼び $\mathrm{rank}\,(\boldsymbol{X})$ と記す。また、行列の階数がその列数に等しい場合を**列フルランク**、行数に等しい場合を**行フルランク**と呼ぶ。行フルランクな行列の各列は互いに**線型独立**である。そして、$\boldsymbol{u}_j = \boldsymbol{0}$ に対応した $\boldsymbol{x}_j$ は、$[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{j-1}, \boldsymbol{x}_{j+1}, \ldots, \boldsymbol{x}_p]$ に**線形従属**であるという。

表 2: グラム・シュミットの直交化：列フルランクではない場合

X

	X1	X2	X3	X4	X5
1)	7	14	1	8	9
2)	6	12	5	11	2
3)	6	12	7	13	0
4)	3	6	5	8	8
5)	5	10	8	13	2
6)	7	14	7	14	8
7)	7	14	4	11	4
8)	7	14	1	8	0

U

	U1	U2	U3	U4	U5
1)	7	0	-4.053	0	4.955
2)	6	0	0.669	0	-1.905
3)	6	0	2.669	0	-4.117
4)	3	0	2.834	0	5.783
5)	5	0	4.391	0	-1.660
6)	7	0	1.947	0	3.321
7)	7	0	-1.053	0	-0.362
8)	7	0	-4.053	0	-4.045

R

	X1	X2	X3	X4	X5
U1	1	2	0.722	1.722	0.639
U2	0	0	0	0	0
U3	0	0	1	1	0.106
U4	0	0	0	0	0
U5	0	0	0	0	1

A

	U1	U2	U3	U4	U5
X1	1	0	-0.722	0	-0.563
X2	0	0	0	0	0
X3	0	0	1	0	-0.106
X4	0	0	0	0	0
X5	0	0	0	0	1

表 3: 縮小化されたグラム・シュミットの直交化

U

	U1	U3	U5
1)	7	-4.053	4.955
2)	6	0.669	-1.905
3)	6	2.669	-4.117
4)	3	2.834	5.783
5)	5	4.391	-1.660
6)	7	1.947	3.321
7)	7	-1.053	-0.362
8)	7	-4.053	-4.045

R

	X1	X2	X3	X4	X5
U1	1	2	0.722	1.722	0.639
U3	0	0	1	1	0.106
U5	0	0	0	0	1

A

	U1	U3	U5
X1	1	-0.722	-0.563
X2	0	0	0
X3	0	1	-0.106
X4	0	0	0
X5	0	0	1

表 2 に $\boldsymbol{X}$ 行列の第 2 列と 4 列がそれ以前の列の線形結合で表される場合、すなわち、行列 $\boldsymbol{X}$ の階数が $\mathrm{rank}\,(\boldsymbol{X}) = 5 - 2 = 3$ である場合の例を示した。この例からも明らかなように、行列が列フルランクでない場合、すなわち $\mathrm{rank}\,(\boldsymbol{X}) \leq q$ の場合には $\boldsymbol{U}'\boldsymbol{U}$ は対角行列ではあるが、対角要素に $q - \mathrm{rank}\,(\boldsymbol{X})$ 個の 0 を持ち、それを正規化した $\boldsymbol{Q}$ 行列も同様である。そこで、$\boldsymbol{U}$ 行列の全てが 0 である列を取り除いた $n \times \mathrm{rank}\,(\boldsymbol{X})$ の行列 $\boldsymbol{U}^\dagger$ とそれを (290) 式 で正規化して得られる $\boldsymbol{Q}^\dagger$ ならびに $\boldsymbol{A}$ 行列の全てが 0 の行を取り除いた $q \times \mathrm{rank}\,(\boldsymbol{X})$ の $\boldsymbol{A}^\dagger$ それらから (292) 式 を用いて計算される $\boldsymbol{R}^\dagger$ 行列を定義すると、

$$\boldsymbol{X} = \boldsymbol{Q}^\dagger \boldsymbol{R}^\dagger, \quad \boldsymbol{U}^\dagger = \boldsymbol{X}\boldsymbol{A}^\dagger, \quad \boldsymbol{Q}^{\dagger'}\boldsymbol{Q}^\dagger = \boldsymbol{I} \tag{293}$$

となる。表 2 に対応するこれらの行列を 表 3 に示した。

$\boldsymbol{X}$ 行列が列フルランクである場合には、QR 分解はその各列の符号を除いて一意に定まる。

最後に、グラム・シュミットの直交化の結果として得られた $\mathbf{0}$ でない $\boldsymbol{u}_j$ に対応する $\boldsymbol{x}_j$ を $n \times \mathrm{rank}\,(\boldsymbol{X})$ の行列の形に並べたものを $\boldsymbol{Q}^*$ とすると、この行列には $\boldsymbol{X}$ の列の中で他の列で説明できない部分が入っていることになる。そして、これを用いて

$$\boldsymbol{X} \quad = \quad \boldsymbol{Q}^*\boldsymbol{R}^*, \quad \text{ただし} \quad \boldsymbol{R}^* = \left(\boldsymbol{Q}^{*'}\boldsymbol{Q}^*\right)^{-1}\boldsymbol{Q}^{*'}\boldsymbol{X} \tag{294}$$

という分解が考えられるが、これを $\boldsymbol{X}$ の**階数分解 (rank decomposition, rank factorization)** と呼ぶ。なお、行列の階数分解の方法は様々なものが有り、必ずしも $\boldsymbol{R}^*$ が上三角行列であるという制約を置くわけでは無い。すなわち、**行列の階数分解は一意では無い。**

4.3.2 列ブロック単位での直交化

前節では行列の列を単位として行列の直交化を行ったが、ブロック、すなわち列の集合を単位として行列の直交化を行うことも出来る。

いま、$n \times q$ の行列 $\boldsymbol{X}$ が以下のように G 個のブロックに分割されているとし、

$$\boldsymbol{X} \quad = \quad [\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_G] \tag{295}$$

$q_j,\ j = 1, 2, \ldots, q$ をそれぞれのブロックの列数とする。このとき、第 j ブロック $\boldsymbol{X}_j$ を第 1 ブロックから第 $j-1$ ブロックまで、すなわち $\boldsymbol{X}_{1:(j-1)} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_{j-1}]$ に回帰した時の $\sum_{k=1}^{j-1} q_k \times n_j$ の回帰係数行列を $\boldsymbol{B}_j$ とし、その行列を $\boldsymbol{X}$ 行列のブロックに対応した形で分割したものを

$$\begin{aligned} \boldsymbol{B}_j \quad = \quad & \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{X}_2 & \ldots & \boldsymbol{X}_1'\boldsymbol{X}_{j-1} \\ \boldsymbol{X}_2'\boldsymbol{X}_1 & \boldsymbol{X}_2'\boldsymbol{X}_2 & \ldots & \boldsymbol{X}_2'\boldsymbol{X}_{j-1} \\ & \vdots & \vdots & \\ \boldsymbol{X}_{j-1}'\boldsymbol{X}_1 & \boldsymbol{X}_{j-1}'\boldsymbol{X}_2 & \ldots & \boldsymbol{X}_{j-1}'\boldsymbol{X}_{j-1} \end{bmatrix}^{-1} \times \begin{bmatrix} \boldsymbol{X}_{1j}'\boldsymbol{X}_j \\ \boldsymbol{X}_{2j}'\boldsymbol{X}_j \\ \cdots \\ \boldsymbol{X}_{j-1,j}'\boldsymbol{X}_j \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_{1j} \\ \boldsymbol{B}_{2j} \\ \vdots \\ \boldsymbol{B}_{j-1,j} \end{bmatrix}, \\ & j = 2, 3, \ldots, G \end{aligned} \tag{296}$$

とすれば、$\boldsymbol{X}_j$ を $\boldsymbol{X}_{1\to(j-1)}$ へ回帰した時の $n \times q_j$ の残差行列 $\boldsymbol{U}_j$ は

$$\boldsymbol{U}_j \quad = \quad \boldsymbol{X}_j - \boldsymbol{X}_{1:(j-1)}\boldsymbol{B}_j \;=\; \boldsymbol{X}_j - \sum_{k=1}^{j-1} \boldsymbol{X}_k\boldsymbol{B}_{kj} \tag{297}$$

と書くことが出来る。そして、これらはまとめて、以下のように $\boldsymbol{A}$ 行列を定義すれば

$$\boldsymbol{U} \quad = \quad [\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_G] \;=\; \boldsymbol{X}\boldsymbol{A} \tag{298}$$

という形で $\boldsymbol{U}$ 行列を作成することが出来る。

$$\begin{aligned} \boldsymbol{A} \quad = \quad & \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{B}_{12} & \cdots & -\boldsymbol{B}_{1,j-1} & -\boldsymbol{B}_{1,j} & \cdots & -\boldsymbol{B}_{1,G} \\ \boldsymbol{O} & \boldsymbol{I} & \cdots & -\boldsymbol{B}_{2,i-1} & -\boldsymbol{B}_{2,j} & \cdots & -\boldsymbol{B}_{2,G} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{I} & -\boldsymbol{B}_{j-1,j} & \cdots & -\boldsymbol{B}_{j-1,G} \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{O} & \boldsymbol{I} & \cdots & -\boldsymbol{B}_{j,G} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{I} \end{bmatrix} \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j \text{ 列} \end{aligned} \tag{299}$$

このとき、$\boldsymbol{U}'\boldsymbol{U}$ 行列はブロック対角行列となる。この操作は Gram-Schumidt の直交化を列ブロックを単位として行ったことに相当するが、列ベクトルを単位とした通常と Gram-Schumidt の直交化との違いは、この方法では各ブロックの中での直交性は保証されていないことである。

表 4 に先ほどの $\boldsymbol{X}$ 行列を $\boldsymbol{X}_1 = [\boldsymbol{x}_1, \boldsymbol{x}_2], \quad \boldsymbol{X}_2 = [\boldsymbol{x}_3, \boldsymbol{x}_4], \quad \boldsymbol{X}_3 = \boldsymbol{x}_5]$ としたときの例を示す。

表 4: 列ブロック単位の直交化

$\boldsymbol{U}$

	Q1	Q2	Q3	Q4	Q5
1)	7	3	-2.280	0.987	4.687
2)	6	9	-1.014	2.203	-0.561
3)	6	9	0.986	2.203	-3.710
4)	3	6	1.246	3.018	5.900
5)	5	9	2.241	-6.247	-0.447
6)	7	1	4.717	-3.236	-0.371
7)	7	8	-1.771	-1.957	1.207
8)	7	4	-2.778	3.598	-4.072

$\boldsymbol{A}$

	Q1	Q2	Q3	Q4	Q5
X1	1	0	-0.255	-0.550	-0.692
X2	0	1	-0.498	-0.389	0.623
X3	0	0	1	0	-0.575
X4	0	0	0	1	-0.127
X5	0	0	0	0	1

$\boldsymbol{U}'\boldsymbol{U}$

	Q1	Q2	Q3	Q4	Q5
Q1	302	283	0	0	0
Q2	283	369	0	0	0
Q3	0	0	47	-34	0
Q4	0	0	-34	86	0
Q5	0	0	0	0	89

***** (299) 式 の別の表現は OrthLinEqETC5.tex で詳説 ******

4.4 ベクトルの線形独立と行列の階数

$a \times b$ の行列 $\boldsymbol{A} = [\boldsymbol{a}_{(1)}, \boldsymbol{a}_{(2)}, \ldots, \boldsymbol{a}_{(b)}]$ のグラム・シュミットの直交化を行った場合、途中で残差が 0 となる列が出現する場合がある。例えば第 j 番目で残差が 0 になった場合は、$\boldsymbol{a}_{(j)}$ が 行列 $[\boldsymbol{a}_{(1)}, \boldsymbol{a}_{(2)}, \ldots, \boldsymbol{a}_{(j-1)}]$ の各列の線形結合で書き表されることを意味しているが、これを $\boldsymbol{a}_{(j)}$ は行列 $[\boldsymbol{a}_{(1)} \quad \boldsymbol{a}_{(2)} \quad \ldots \quad \boldsymbol{a}_{(j-1)}]$ に従属しているという。そのような場合は $\boldsymbol{u}$ を $\boldsymbol{0}$ としてグラム・シュミットの直交化を進めると、最終的に完成する $\boldsymbol{Q}$ 行列のいくつかの列が $\boldsymbol{0}$ となっている場合がある。このような場合、行列 $\boldsymbol{A}$ の列は線形独立では無いという。

また、$\boldsymbol{A}$ の列数 b から $\boldsymbol{Q}$ 行列の全てが 0 である列の数を引いたものは行列の階数 (rank) と呼ばれる。したがって、行列の階数とは、その行列の列の中で互いに線形独立である列の最大数になる。

$\boldsymbol{Q}$ 行列が $\boldsymbol{0}$ を一つも含まない場合、すなわち、階数が列数に等しい場合、$\boldsymbol{A}$ の b 個の列ベクトルは線形独立であるという。また、階数が b より小さい場合は、いくつかのベクトルが他のベクトルの線形結合で書き表されるという意味で、線形従属であると呼ぶ。

このようにして列ベクトルを用いて定義される行列の階数は行ベクトルを用いて定義される行列の階数と等しい。(証明略) したがって、階数 $\leq a, b$ である。また、$a > b$ の場合、行列 $\boldsymbol{A}$ の階数がその列数 b に等しい場合、その行列は列フルランク (full column rank) であるといい、$a < b$ の場合、行列 $\boldsymbol{A}$ の階数がその行数 a に等しい場合、その行列は行フルランク (full row rank) であるという。また、正方行列の場合、階数が行数、したがって列数に等しい行列を正則行列 (nonsingular matrix) 、フルランクでない正方行列を特異行列 (singular matrix) と呼ぶ。

ある行列に正則な行列を掛けた行列の階数は元の行列の階数に等しい。また、$\boldsymbol{A}'\boldsymbol{A}$ と $\boldsymbol{A}\boldsymbol{A}'$ の階数は $\boldsymbol{A}$ の階数に等しい。正方行列が正則の場合、すなわちその各列ベクトルが線形独立である場合、その逆行列

$\boldsymbol{A}^{-1}$ が存在する。また、$\boldsymbol{A}$ が列フルランクの場合 $(\boldsymbol{A}'\boldsymbol{A})^{-1}$ が、 $\boldsymbol{A}$ が行フルランクの場合、 $(\boldsymbol{A}\boldsymbol{A}')^{-1}$ が存在する。

線形従属な行列の分解 $a \times b$ の行列 $\boldsymbol{A}$ の階数が r である場合、それを、$a \times r$ の列フルランクの行列 $\boldsymbol{P}$ と $b \times r$ の列フルランクの行列 $\boldsymbol{Q}$ の積に分解することが出来る。

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{Q}' \tag{300}$$

ただし、このような $\boldsymbol{P}, \boldsymbol{Q}$ は一位には定まらない。このような分解法の一つが QR 分解であり階数分解である。また固有値と固有ベクトルを利用した分解方法も存在する。

3：行列の階数

- $\boldsymbol{A}$ のサイズが $a \times b$ の場合、$\mathrm{rank}\,(\boldsymbol{A}) \leq \min(a, b)$
- $\mathrm{rank}\,(\boldsymbol{A}) = \mathrm{rank}\left(\boldsymbol{A}'\right) = \mathrm{rank}\left(\boldsymbol{A}'\boldsymbol{A}\right) = \mathrm{rank}\left(\boldsymbol{A}\boldsymbol{A}'\right) = \mathrm{rank}\,(c\boldsymbol{A})$
- $\mathrm{rank}\,(\boldsymbol{A}\boldsymbol{B}) \leq \min(\mathrm{rank}\,(\boldsymbol{A}), \mathrm{rank}\,(\boldsymbol{B}))$
- $\boldsymbol{B}$ のサイズが $b \times r$, $r \leq b$ で $\mathrm{rank}\,(\boldsymbol{B}) = r \leq \mathrm{rank}\,(\boldsymbol{A})$ の場合 $\mathrm{rank}\,(\boldsymbol{A}\boldsymbol{B}) = r$
- $\boldsymbol{P}$ と $\boldsymbol{Q}$ が正則の場合、$\mathrm{rank}\,(\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}) = \mathrm{rank}\,(\boldsymbol{A})$
- $\boldsymbol{A}$ が特異の場合、行列式の値は 0 となる。
- $\boldsymbol{A}$ が実対称の場合 $\mathrm{rank}\,(\boldsymbol{A}) = \boldsymbol{A}$ の 0 でない固有値の数
- $\boldsymbol{A}$ が半正定値の場合 $\mathrm{rank}\,(\boldsymbol{A}) = \boldsymbol{A}$ の 0 でない固有値の数= 正の固有値の数
- $\mathrm{rank}\,([\boldsymbol{A}, \boldsymbol{B}]) \leq \mathrm{rank}\,(\boldsymbol{A}) + \mathrm{rank}\,(\boldsymbol{B})$ (Harville(5.1))
- $\mathrm{rank}\,(\boldsymbol{A} + \boldsymbol{B}) \leq \mathrm{rank}\,(\boldsymbol{A}) + \mathrm{rank}\,(\boldsymbol{B})$ (Harville(5.7))

証明: $\mathrm{rank}\,(\boldsymbol{A}\boldsymbol{B}) \leq \min(\mathrm{rank}\,(\boldsymbol{A}), \mathrm{rank}\,(\boldsymbol{B}))$

$\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$ とすると $\boldsymbol{C}$ 行列の列は全て $\boldsymbol{A}$ 行列の列空間に含まれる。したがって、 $\mathrm{rank}\,(\boldsymbol{C}) \leq \mathrm{rank}\,(\boldsymbol{A})$ である。また、$\boldsymbol{C}' = \boldsymbol{B}'\boldsymbol{A}'$ とすると $\boldsymbol{C}'$ 行列の列は全て $\boldsymbol{B}'$ 行列の列空間に含まれる。したがって、$\mathrm{rank}\,(\boldsymbol{C}) = \mathrm{rank}\left(\boldsymbol{C}'\right) \leq \mathrm{rank}\left(\boldsymbol{B}'\right) = \mathrm{rank}\,(\boldsymbol{B})$ である。 **証明終わり**

証明: $\mathrm{rank}\,(\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}) = \mathrm{rank}\,(\boldsymbol{A})$

$\mathrm{rank}\,(\boldsymbol{P}\boldsymbol{A}) \leq \mathrm{rank}\,(\boldsymbol{A})$ である。また、 $\boldsymbol{I} = \boldsymbol{P}^{-1}\boldsymbol{P}$ を用いて $\mathrm{rank}\,(\boldsymbol{A}) = \mathrm{rank}\left(\boldsymbol{P}^{-1}\boldsymbol{P}\boldsymbol{A}\right) \leq \mathrm{rank}\,(\boldsymbol{P}\boldsymbol{A})$ である。 **証明終わり**

5　一般逆行列について

5.1　はじめに

階数が r の $a \times b$ の行列 $\boldsymbol{A}$ に関して、$b \times a$ の行列 $\boldsymbol{A}^{-}$ が以下の条件のうち少なくとも最初のものを満たすならば、それは $\boldsymbol{A}$ の一般逆行列 (g-inverse, g1-inverse, conditional inverse, generalized inverse) であるという。

4：一般逆行列

$$g1) \quad \boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} \tag{301}$$

$$g2) \quad \boldsymbol{A}^{-} = \boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{A}^{-} \tag{302}$$

$$g3) \quad \boldsymbol{A}\boldsymbol{A}^{-} = \left(\boldsymbol{A}\boldsymbol{A}^{-}\right)' \tag{303}$$

$$g4) \quad \boldsymbol{A}^{-}\boldsymbol{A} = \left(\boldsymbol{A}^{-}\boldsymbol{A}\right)' \tag{304}$$

ただし、このような行列は常に存在するが、一般的には一意には定まるとはかぎらない。

また、条件 g1 に加えて、条件 g2 を満たす行列を反射型一般逆行列（g12-inverse, reflexive generalized inverse）と呼ぶ。また、反射型一般逆行列のなかで、$\boldsymbol{A}\boldsymbol{A}^{-}$ と $\boldsymbol{A}^{-}\boldsymbol{A}$ が対称という条件（ g3 と g4 ）をも満たすものをムーアペンローズ逆行列（g1234-inverse, Moor-Penrose inverse, the pseud inverse, 疑似逆行列）と呼ぶが、この行列は一意に定まり、特に $\boldsymbol{A}^{+}$ と記される。その他の条件の組み合わせとしては、g14 (minimum norm inverse), g124 (niminum norm reflactive inverse), g13 (least squares inverse), g123 (lease squares reflective inverse) が存在する。 Seaber (2008)

なお $\boldsymbol{A}$ が正則な場合は $\boldsymbol{A}^{+} = \boldsymbol{A}^{-} = \boldsymbol{A}^{-1}$ である。

数値例として、以下の対称な正方行列 $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{bmatrix} 8 & 8 & 0 \\ 8 & 9 & 1 \\ 0 & 1 & 1 \end{bmatrix} \tag{305}$$

のいくつかの一般逆行列とムーア・ペンローズ逆行列を示すが、$\boldsymbol{B}_1$ は条件 g1 を、$\boldsymbol{B}_2$ は条件 g1 と g2 を、$\boldsymbol{B}_3$ は条件 g1, g2, と g3 を、$\boldsymbol{B}_4$ は条件 g1, g2, と g4 を満たす。

$$\boldsymbol{B}_1 = \begin{bmatrix} 65/8 & -8 & 7 \\ 6 & -6 & 7 \\ 3 & -3 & 3 \end{bmatrix}, \quad \boldsymbol{B}_2 = \begin{bmatrix} 9/8 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{B}_3 = \begin{bmatrix} 5/12 & -7/24 & -17/24 \\ -1/3 & 1/3 & 2/3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{B}_4 = \begin{bmatrix} 5/12 & -1/3 & 0 \\ -7/24 & 1/3 & 0 \\ -17/24 & 2/3 & 0 \end{bmatrix}, \qquad \boldsymbol{A}^{+} = \begin{bmatrix} 1/6 & -1/12 & -1/4 \\ -1/12 & 1/8 & 5/24 \\ -1/4 & 5/24 & 11/24 \end{bmatrix} \tag{306}$$

また、矩形行列

$$\boldsymbol{A} = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \tag{307}$$

のいくつかの一般逆行列とムーア・ペンローズ逆行列を示すが、$\boldsymbol{B}_1$ は条件 g1 を、$\boldsymbol{B}_2$ は条件 g1 と g2 を、

$\boldsymbol{B}_3$ は条件 g1, g2, と g3 を満たす。

$$\boldsymbol{B}_1 = \begin{bmatrix} -5/2 & 3 & 7 & -1 \\ -5 & 5 & 7 & 1 \\ -7 & 7 & 3 & 0 \end{bmatrix}, \quad \boldsymbol{B}_2 = \begin{bmatrix} 1/2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{B}_3 = \begin{bmatrix} 1/4 & 1/4 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{A}^{+} = \begin{bmatrix} 1/6 & 1/6 & 0 & -1/3 \\ 1/12 & 1/12 & 0 & 1/3 \\ -1/12 & -1/12 & 0 & 2/3 \end{bmatrix} \tag{308}$$

5.2 一般逆行列の性質

一般逆行列に関しては以下の性質が知られている。

1. ある行列 $\boldsymbol{A}$ の一般逆行列を転置したもの $\left(\boldsymbol{A}^{-}\right)'$ は $\boldsymbol{A}'$ の一般逆行列となっている。$\boldsymbol{A}$ が対称行列の場合、$\boldsymbol{A}^{-}$ は必ずしも対称では無いが、 $\left(\boldsymbol{A}^{-}\right)'$ は $\boldsymbol{A}$ の一般逆行列である。**証明:**
   $\boldsymbol{A}'$ の一般逆行列は

   $$\boldsymbol{A}'(\boldsymbol{A}')^{-}\boldsymbol{A}' = \boldsymbol{A}' \tag{309}$$

   を満たさなければならないが、これに $\left(\boldsymbol{A}^{-}\right)'$ を代入すると

   $$\boldsymbol{A}'(\boldsymbol{A}')^{-}\boldsymbol{A}' = \boldsymbol{A}'\left(\boldsymbol{A}^{-}\right)'\boldsymbol{A}' = \left(\boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A}\right)' = \boldsymbol{A}' \tag{310}$$

   を得る。 **証明終わり**

2. 以下に示す行列は冪等である。

   $$\boldsymbol{A}\boldsymbol{A}^{-}, \quad \boldsymbol{A}^{-}\boldsymbol{A} \tag{311}$$

   証明:

   $$\boldsymbol{A}\boldsymbol{A}^{-}\ \ \boldsymbol{A}\boldsymbol{A}^{-} = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A}\ \ \boldsymbol{A}^{-} = \boldsymbol{A}\boldsymbol{A} \tag{312}$$

   $$\boldsymbol{A}^{-}\boldsymbol{A}\ \ \boldsymbol{A}^{-}\boldsymbol{A} = \boldsymbol{A}^{-}\ \ \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} = \boldsymbol{A}^{-}\boldsymbol{A} \tag{313}$$

   証明終わり

3. $\boldsymbol{B}$ と $\boldsymbol{C}$ を正則行列とすれば

   $$(\boldsymbol{B}\boldsymbol{A}\boldsymbol{C})^{-} = \boldsymbol{C}^{-1}\boldsymbol{A}^{-}\boldsymbol{B}^{-1} \tag{314}$$

   である。
   証明:
   $\boldsymbol{B}\boldsymbol{A}\boldsymbol{C}$ の一般逆行列は条件

   $$\boldsymbol{B}\boldsymbol{A}\boldsymbol{C}\ \ (\boldsymbol{B}\boldsymbol{A}\boldsymbol{C})^{-}\ \ \boldsymbol{B}\boldsymbol{A}\boldsymbol{C} \tag{315}$$

   を満たさなければならないが、これに (314) 式 を代入すると

   $$\boldsymbol{B}\boldsymbol{A}\boldsymbol{C}(\boldsymbol{B}\boldsymbol{A}\boldsymbol{C})^{-}\boldsymbol{B}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{B}\boldsymbol{A}\boldsymbol{C}\ \ \boldsymbol{C}^{-1}\boldsymbol{A}^{-}\boldsymbol{B}^{-1}\ \ \boldsymbol{B}\boldsymbol{A}\boldsymbol{C} \tag{316}$$

   $$= \boldsymbol{B}\ \ \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A}\ \ \boldsymbol{C} = \boldsymbol{B}\boldsymbol{A}\boldsymbol{C} \tag{317}$$

   を得る。 **証明終わり**

4. 行列 $\boldsymbol{B}$ が列フルランク（階数が列数に等しい）であり、行列 $\boldsymbol{C}$ が行フルランク（階数が行数に等しい）の場合、

$$\boldsymbol{B}^{-}\boldsymbol{B} = \boldsymbol{I} \quad \text{and} \quad \boldsymbol{C}\boldsymbol{C}^{-} = \boldsymbol{I} \tag{318}$$

であり、この場合には、

$$(\boldsymbol{BAC})^{-} \;=\; \boldsymbol{C}^{-}\boldsymbol{A}^{-}\boldsymbol{B}^{-} \tag{319}$$

である。

証明:

$\boldsymbol{B}$ の一般逆行列が満たすべき条件で $\boldsymbol{B}^{-}\boldsymbol{B}$ を $\boldsymbol{K}$ と置けば、

$$\boldsymbol{BK} \;=\; \boldsymbol{B} \tag{320}$$

を得る。また、この等式の両辺の左から $\boldsymbol{B}'$ を掛けると

$$\boldsymbol{B}'\boldsymbol{BK} \;=\; \boldsymbol{B}'\boldsymbol{B} \tag{321}$$

を得るが $\boldsymbol{B}'\boldsymbol{B}$ は逆行列を持つため $\boldsymbol{K} = \boldsymbol{I}$ を得る。また、

$$\boldsymbol{BAC}\;\;\boldsymbol{C}^{-}\boldsymbol{A}^{-}\boldsymbol{B}^{-}\;\;\boldsymbol{BAC} \;=\; \boldsymbol{BA}\;\;\boldsymbol{CC}^{-}\;\;\boldsymbol{A}^{-}\;\;\boldsymbol{B}^{-}\boldsymbol{B}\;\;\boldsymbol{AC} \;=\; \boldsymbol{BAC} \tag{322}$$

であるが、これは $\boldsymbol{C}^{-}\boldsymbol{A}^{-}\boldsymbol{B}^{-}$ が $\boldsymbol{BAC}$ の一般逆行列であることを示している。

証明終わり

5.3 一般逆行列と射影演算子

どのような $a \times 1$ のベクトル $\boldsymbol{x}$ も、ある $a \times b$ の行列 $\boldsymbol{A}$ の列の列の線形結合として書ける部分 $\boldsymbol{x}_1$ とそれに直交する部分 $\boldsymbol{x}_2$ に分解することが出来る。すなわち、$\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2, \quad , \boldsymbol{x}_1{}'\boldsymbol{x}_2 = 0$ と書ける。このようなふたつの直交するベクトルは

$$\boldsymbol{x}_1 = \boldsymbol{P}(\boldsymbol{A})\boldsymbol{x}, \qquad \text{and} \qquad \boldsymbol{x}_2 = \boldsymbol{Q}(\boldsymbol{A})\boldsymbol{x}, \tag{323}$$

ただし、

$$\boldsymbol{P}(\boldsymbol{A}) \;=\; \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}' \tag{324}$$

$$\boldsymbol{Q}(\boldsymbol{A}) \;=\; \boldsymbol{I} - \boldsymbol{P}(\boldsymbol{A}) \;=\; \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}' \tag{325}$$

として求めることが出来るが、この行列 $\boldsymbol{P}(\boldsymbol{A})$ を行列 $\boldsymbol{A}$ の列空間への **射影子行列（射影演算子行列）** (projection matrix)、 $\boldsymbol{Q}_A$ を行列 $\boldsymbol{A}$ の直交補空間への射影子行列と呼ぶ。

なお、以下に示すように、射影子行列は一般逆行列の種類によらず一意に定まる。

射影子行列は以下の性質を持つ。

5：射影子行列の性質

$$\boldsymbol{P}(\boldsymbol{A}) = \boldsymbol{P}(\boldsymbol{A})' \qquad \text{（対称）} \tag{326}$$

$$\boldsymbol{P}(\boldsymbol{A})\;\boldsymbol{P}(\boldsymbol{A}) = \boldsymbol{P}(\boldsymbol{A}) \qquad \text{（冪等）} \tag{327}$$

$$\boldsymbol{P}(\boldsymbol{A})\boldsymbol{A} = \boldsymbol{A} \qquad \text{（自身への射影）} \tag{328}$$

$$\operatorname{rank}(\boldsymbol{P}(\boldsymbol{A})) = \operatorname{tr}(\boldsymbol{P}(\boldsymbol{A})) = \operatorname{rank}(\boldsymbol{A}) \qquad \text{（階数）} \tag{329}$$

$$\tag{330}$$

$$\boldsymbol{A}'\;\boldsymbol{P}(\boldsymbol{A}) = \boldsymbol{A}' \qquad \text{（左から）} \tag{331}$$

証明:

(328) 式 の自分自身の射影は自分自身になることを示すには、$\boldsymbol{A}'\boldsymbol{A}$ の一般逆行列の性質

$$\boldsymbol{A}'\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{A} \quad = \quad \boldsymbol{A}'\boldsymbol{A} \tag{332}$$

を

$$\boldsymbol{A} \leftarrow \boldsymbol{A}, \quad \boldsymbol{B} \leftarrow (\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{A}, \quad \boldsymbol{C} \leftarrow \boldsymbol{I} \tag{333}$$

と置いた等号成立条件 (7) 式 と解釈すれば

$$\boldsymbol{A} \ (\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{A} \quad = \quad \boldsymbol{A} \ \boldsymbol{I} \ = \ \boldsymbol{A} \tag{334}$$

を得る。

(327) 式 の冪等性は、射影子行列の性質 (328) 式 から

$$\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'\boldsymbol{A} \ (\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}' \quad = \quad \boldsymbol{A} \ (\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}' \tag{335}$$

となることで示せる。なお、上式は、$(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'$ が $\boldsymbol{A}$ の一般逆行列となっていることを示している。同様に、$\boldsymbol{P}(\boldsymbol{A})$ の対称性から $\left((\boldsymbol{A}'\boldsymbol{A})^{-}\right)'\boldsymbol{A}'$ も $\boldsymbol{A}$ の一般逆行列である。

(326) 式 の対称性を示すには、まず、射影子行列が一般逆行列の種類によらずに一意に定まることを以下の様にして示す。すなわち、今、$\boldsymbol{A}'\boldsymbol{A}$ の２種類の異なる一般逆行列を $\boldsymbol{G}$、 $\boldsymbol{F}$ とすれば、 (328) 式 から $\boldsymbol{A}\boldsymbol{G}\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{A}\boldsymbol{F}\boldsymbol{A}'\boldsymbol{A}$ であるが、等号成立条件 (8) 式 より、このことは $\boldsymbol{A}\boldsymbol{G}\boldsymbol{A}' = \boldsymbol{A}\boldsymbol{F}\boldsymbol{A}'$ を示している。すなわち、どのような一般逆行列を用いても射影子行列は一意に定まる。

したがって、$\boldsymbol{G}$ を $\boldsymbol{A}'\boldsymbol{A}$ の対称な一般逆行列とすれば、

$$(\boldsymbol{A}\boldsymbol{G}\boldsymbol{A}')' \quad = \quad \boldsymbol{A}\boldsymbol{G}'\boldsymbol{A} \ = \ \boldsymbol{A}\boldsymbol{G}\boldsymbol{A}' \tag{336}$$

となり射影子行列が対称行列となることが示せるが、射影子行列は一般逆行列の種類にはよらすーに定まるので、全ての射影子行列は対称である。

(331) 式 は、射影演算子の対称性より、

$$\boldsymbol{A}'\boldsymbol{P}(\boldsymbol{A}) = \boldsymbol{A}'\boldsymbol{P}(\boldsymbol{A})' = (\boldsymbol{P}(\boldsymbol{A})\boldsymbol{A})' = \boldsymbol{A}' \tag{337}$$

である。

(329) 式 の証明は以下の通り。$\boldsymbol{P}(\boldsymbol{A})$ は冪等であるため、λ と $\boldsymbol{p}$ をその固有値と固有ベクトルとすると

$$\boldsymbol{P}(\boldsymbol{A})\boldsymbol{p} = \lambda\boldsymbol{p} \tag{338}$$

であるが、この両辺の左から $\boldsymbol{P}(\boldsymbol{A})$ を掛けると左辺は

$$\boldsymbol{P}(\boldsymbol{A})^2\boldsymbol{p} = \boldsymbol{P}(\boldsymbol{A})\boldsymbol{p} = \lambda\boldsymbol{p} \tag{339}$$

であり、右辺は

$$\boldsymbol{P}(\boldsymbol{A})\lambda\boldsymbol{p} = \lambda\boldsymbol{P}(\boldsymbol{A})\boldsymbol{p} = \lambda^2\boldsymbol{p} \tag{340}$$

となるため

$$\lambda\boldsymbol{p} = \lambda^2\boldsymbol{p} \tag{341}$$

を得るが、これは $\lambda = \lambda^2$ を意味しており、λ は 0 か 1 である。また、行列の階数は行列の 0 でない固有値の数である。また、実対称行列は直交行列を用いて対角化が可能であり、トレースの性質を用いると固有値の和がトレースとなる。 **証明終わり**

また、定義から、$\boldsymbol{P}(\boldsymbol{A})\boldsymbol{Q}_A = \boldsymbol{O}$ となる。射影子行列 $\boldsymbol{P}(\boldsymbol{A})$ の階数は $\boldsymbol{A}$ の階数に等しく、固有値は全て非負であり正の固有値は全て 1 である。したがって、射影子行列のトレースがその行列の階数を表す。

$\boldsymbol{A}$ の行空間への射影子は

$$\boldsymbol{P}(\boldsymbol{A}') \quad = \quad \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-}\boldsymbol{A} \tag{342}$$

であるが、

$$\boldsymbol{P}(\boldsymbol{A}')\boldsymbol{A}' = \boldsymbol{A}' \tag{343}$$

を満たす。

また、$\boldsymbol{A}\boldsymbol{A}^{-}$ と $\boldsymbol{A}^{-}\boldsymbol{A}$ は冪等であるが、これらは対称とは限らないので、直交射影演算子ではない。ただ、$\boldsymbol{A}^{-}$ として $(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}$ や $\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-}$ を用いた場合には直交射影演算子となる。

最後に射影子の例を挙げると、N 次元空間のなかのベクトル $\mathbf{1}_N$ への射影子行列は

$$\boldsymbol{P}(\mathbf{1}) = \mathbf{1}_N(\mathbf{1}_N{}'\mathbf{1}_N)^{-1}\mathbf{1}_N{}' = \frac{1}{N}\mathbf{1}_N\mathbf{1}_N{}' \tag{344}$$

であり、これに直交する空間への射影子行列は

$$\boldsymbol{J} = \boldsymbol{I}_N - \boldsymbol{P}(\mathbf{1}) = \boldsymbol{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N{}' \tag{345}$$

となるが、これは先に見た中心化演算子である。

5.3.1 一般逆行列の性質 2

以下の 3 つの表現はそれぞれ $\boldsymbol{A}$ の一般逆行列であるが、(Graville, Harville)

$$\boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{A}^{-}, \quad (\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}', \quad \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-}, \quad \left((\boldsymbol{A}'\boldsymbol{A})^{-}\right)'\boldsymbol{A}', \quad \boldsymbol{A}'\left((\boldsymbol{A}\boldsymbol{A}')^{-}\right)' \tag{346}$$

最後の四つは反射型一般逆行列 (g12-inverse) となる。

証明:

$\boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{A}^{-}$ の場合

$$\begin{aligned} \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} &= \boldsymbol{A}\ \boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{A}^{-}\ \boldsymbol{A} && (347)\\ &= \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A}\ \ \boldsymbol{A}^{-}\boldsymbol{A} = \boldsymbol{A}\ \ \boldsymbol{A}^{-}\boldsymbol{A} = \boldsymbol{A} && (348) \end{aligned}$$

$(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'$ の場合

$$\begin{aligned} \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} &= \boldsymbol{A}\ \ (\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'\ \ \boldsymbol{A} && (349)\\ &= \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'\ \ \boldsymbol{A} = \boldsymbol{A} && (350)\\ \boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{A}^{-} &= (\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'\ \ \boldsymbol{A}\ \ (\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}' && (351)\\ &= (\boldsymbol{A}'\boldsymbol{A})^{-}\ \ \boldsymbol{A}'\ \ \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}' = (\boldsymbol{A}'\boldsymbol{A})^{-}\ \ \boldsymbol{A}' && (352) \end{aligned}$$

$\left((\boldsymbol{A}'\boldsymbol{A})^{-}\right)'\boldsymbol{A}'$ の場合

$\left((\boldsymbol{A}'\boldsymbol{A})^{-}\right)'$ は $\boldsymbol{A}'\boldsymbol{A}$ の一般逆行列であり、射影演算子は一般逆行列の形に依存しないため

$$\boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} \quad = \quad \boldsymbol{A}\ \ \left((\boldsymbol{A}'\boldsymbol{A})^{-}\right)'\boldsymbol{A}'\ \ \boldsymbol{A} = \boldsymbol{P}(\boldsymbol{A})\,\boldsymbol{A} = \boldsymbol{A} \tag{353}$$

となる。また、

$$\boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{A}^{-} = \left((\boldsymbol{A}'\boldsymbol{A})^{-}\right)'\boldsymbol{A}' \quad \boldsymbol{A} \quad \left((\boldsymbol{A}'\boldsymbol{A})^{-}\right)'\boldsymbol{A}' = \left(\boldsymbol{A} \quad (\boldsymbol{A}'\boldsymbol{A})^{-}(\boldsymbol{A}'\boldsymbol{A})'(\boldsymbol{A}'\boldsymbol{A})^{-}\right)' \tag{354}$$

$$= \left(\boldsymbol{A} \quad (\boldsymbol{A}'\boldsymbol{A})^{-}\right)' = \left((\boldsymbol{A}'\boldsymbol{A})^{-}\right)'\boldsymbol{A}' \tag{355}$$

である。 **証明終わり**

(346) 式 の第 2 式と第 3 式の特殊例として $\boldsymbol{A}$ が行フルランクもしくは列フルランクの場合には、右逆行列または左逆行列が一般逆行列となる。

$$\boldsymbol{A}^{-} = \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1} \qquad \text{or} \qquad \boldsymbol{A}^{-} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}' \tag{356}$$

しかも、これらの行列は g3) と g4) の条件も満たすため、ムーアペンローズ逆行列となっている。

5.4 一般逆行列の生成

(301) 式 を満たす階数が r の $\boldsymbol{A}$ の一般逆行列は、$\boldsymbol{P}$ と $\boldsymbol{Q}$ をそれぞれ行と列を並べ替える行列として以下のような分割行列 $\boldsymbol{B}$ を作り、

$$\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{bmatrix} = \boldsymbol{B} \tag{357}$$

$r \times r$ の行列 $\boldsymbol{B}_{11}$ が正則行列（ただし、r は行列 $\boldsymbol{A}$ の階数）であるとすれば、その一般逆行列は

$$\boldsymbol{B}^{-} = \begin{bmatrix} \boldsymbol{B}_{11}{}^{-1} & \boldsymbol{O}_{21} \\ \boldsymbol{O}_{12} & \boldsymbol{O}_{22} \end{bmatrix} \tag{358}$$

ただし、$\boldsymbol{O}_{ij}$ は対応する $\boldsymbol{B}_{ij}$ の転置行列の要素を全て 0 と置いた行列、として得られることから、

$$\boldsymbol{A}^{-} = \boldsymbol{Q}\boldsymbol{B}^{-}\boldsymbol{P} \tag{359}$$

として得られる。なお、$\boldsymbol{B}$ 行列が矩形で列フルランクの場合には $\boldsymbol{B}_{12}$ と $\boldsymbol{B}_{22}$ 行列が、行フルランクの場合には $\boldsymbol{B}_{21}$ と $\boldsymbol{B}_{22}$ 行列が存在しないが、その場合は対応する $\boldsymbol{O}_{ij}$ 行列も存在しないものとする。

このようにして得られる $\boldsymbol{B}^{-}$ が $\boldsymbol{B}$ の一般逆行列となっていることは以下のようにして示される。

$$\boldsymbol{B}\boldsymbol{B}^{-}\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{21}\boldsymbol{B}_{11}{}^{-1}\boldsymbol{B}_{12} \end{bmatrix} \tag{360}$$

であるが、$\boldsymbol{B}$ 行列の第 2 列ブロックは、第 1 列ブロックに従属であるため、

$$\begin{pmatrix} \boldsymbol{B}_{12} \\ \boldsymbol{B}_{22} \end{pmatrix} = \begin{pmatrix} \boldsymbol{B}_{11} \\ \boldsymbol{B}_{21} \end{pmatrix} \boldsymbol{C} \tag{361}$$

と書けるが、これは

$$\boldsymbol{B}_{12} = \boldsymbol{B}_{11}\boldsymbol{C}, \qquad \text{and} \qquad \boldsymbol{B}_{22} = \boldsymbol{B}_{21}\boldsymbol{C} \tag{362}$$

を意味し、$\boldsymbol{C} = \boldsymbol{B}_{11}{}^{-1}\boldsymbol{B}_{12}$ を得るため、 (360) 式 の第 2-2 要素は $\boldsymbol{B}_{22}$ となり、一般逆行列が満たすべき条件 (301) 式 が満たされることを示している。また、並べ替え行列は自分自身が自分自身の逆行列となっているため (357) 式 より

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{B}\boldsymbol{Q} \tag{363}$$

となるが、これと (359) 式 より、

$$\boldsymbol{A}\ \ \boldsymbol{A}^{-}\ \ \boldsymbol{A} \ = \ \boldsymbol{PBQ}\ \ \boldsymbol{QB}^{-}\boldsymbol{P}\ \ \boldsymbol{PBQ} \tag{364}$$

$$= \ \boldsymbol{P}\ \ \boldsymbol{BB}^{-}\boldsymbol{B}\ \ \boldsymbol{Q} \ = \ \boldsymbol{PBQ} \ = \ \boldsymbol{A} \tag{365}$$

を得る。

(301) 式 と (302) 式 を共に満たす一般逆行列 g12-inerse はスイープ演算子を用いて簡単に計算することが出来る。基本的にスイープ演算子は、正方行列の中の正則部分の逆行列を計算し、それ以外の部分を 0 と置く操作をするため、行列が矩形の場合は、正方となるように全てが 0 の列や行を付け加えてからスイープ演算子を用いたのちに、元の矩形行列に戻せば良い。

より一般的には、$\boldsymbol{U}$, $\boldsymbol{V}$, そして $\boldsymbol{W}$ を任意の行列として

$$\begin{bmatrix} \boldsymbol{B}^{-} = \boldsymbol{B}_{11}{}^{-1} - \boldsymbol{U}\boldsymbol{B}_{21}\boldsymbol{B}_{11}{}^{-1} - \boldsymbol{B}_{11}{}^{-1}\boldsymbol{B}_{12}\boldsymbol{V} - \boldsymbol{B}_{11}{}^{-1}\boldsymbol{B}_{12}\boldsymbol{W}\boldsymbol{B}_{21}\boldsymbol{B}_{11}{}^{-1} & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{bmatrix} \tag{366}$$

が $\boldsymbol{B}$ の一般逆行列となっていることが知られている。 (Harvill. p219, eq21)

5.4.0.1 行列の QR 分解を用いる方法

また、階数が r の $a \times b$ の行列 $\boldsymbol{A}$ の QR 分解を

$$\boldsymbol{A} = \boldsymbol{QR} \tag{367}$$

ただし、 $\boldsymbol{Q}$ は $a \times r$ の直交行列、$\boldsymbol{R}$ は $r \times b$ の上三角行列、とすれば、

$$\boldsymbol{A}^{-} = \boldsymbol{R}^{-}\boldsymbol{Q}' \tag{368}$$

である。(証明略)

階数 r の行列 $\boldsymbol{A}$ の階数分解 $\boldsymbol{A} = \boldsymbol{Q}^*\boldsymbol{R}^*$ が与えられた場合は

$$\boldsymbol{A}^{-} = \boldsymbol{R}^{*-}\boldsymbol{Q}^{*-} \tag{369}$$

である。

証明:

階数分解と QR 分解の間には、$\boldsymbol{T}$ を正則行列として、

$$\boldsymbol{Q}^* = \boldsymbol{QT} \qquad \text{and} \qquad \boldsymbol{R}^* = \boldsymbol{T}^{-1}\boldsymbol{R} \tag{370}$$

という関係があるため、

$$\boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} \ = \ \boldsymbol{A}\,\boldsymbol{R}^{*-}\boldsymbol{Q}^{*-}\,\boldsymbol{A} \ = \ \boldsymbol{A}\,\boldsymbol{R}^{-}\boldsymbol{T}\,\boldsymbol{T}^{-1}\boldsymbol{Q}^{-}\,\boldsymbol{A} \tag{371}$$

$$= \ \boldsymbol{QR}\,\boldsymbol{R}^{-}\boldsymbol{Q}^{-}\,\boldsymbol{QR} \ = \ \boldsymbol{QR}\,\boldsymbol{R}^{-}\,\boldsymbol{R} \ = \ \boldsymbol{QR} \tag{372}$$

を得る。 **証明終わり**

5.4.0.2 行列の特異値分解を用いる方法

$\boldsymbol{A}$ 行列のフルサイズの特異値分解が

$$\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix} \boldsymbol{D}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}\boldsymbol{V}' \tag{373}$$

であるとき、任意の行列 $\boldsymbol{D}_{21}, \boldsymbol{D}_{12}, \boldsymbol{D}_{22}$ を用いて

$$\boldsymbol{A}^{-} = \boldsymbol{V} \begin{bmatrix} \boldsymbol{D}_{11}{}^{-1} & \boldsymbol{D}_{12} \\ \boldsymbol{D}_{21} & \boldsymbol{D}_{22} \end{bmatrix} \boldsymbol{U}' \tag{374}$$

である。なぜなら、$\boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}, \quad \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}$ であるから、

$$\begin{aligned} \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} &= \boldsymbol{U} \begin{bmatrix} \boldsymbol{D}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix} \begin{bmatrix} \boldsymbol{D}_{11}{}^{-1} & \boldsymbol{D}_{12} \\ \boldsymbol{D}_{21} & \boldsymbol{D}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{D}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix} \boldsymbol{V}' && (375) \\ &= \boldsymbol{U} \begin{bmatrix} \boldsymbol{D}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix} \boldsymbol{V}' && (376) \end{aligned}$$

となるからである。

5.4.1 一般逆行列から一般逆行列

$\boldsymbol{Y}$ を任意の $b \times a$ の行列とすると、以下の四つの行列はまた $\boldsymbol{A}$ の一般逆行列となっている。

$$\boldsymbol{B}_1 = \boldsymbol{A}^{-} + \boldsymbol{Y}\,(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^{-}), \qquad \boldsymbol{B}_2 = \boldsymbol{A}^{-} + (\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A})\,\boldsymbol{Y} \tag{377}$$

$$\boldsymbol{B}_3 = \boldsymbol{A}^{-} + \boldsymbol{Y}\,\boldsymbol{Q}(\boldsymbol{A}'), \qquad \boldsymbol{B}_4 = \boldsymbol{A}^{-} + \boldsymbol{Q}(\boldsymbol{A}')\,\boldsymbol{Y} \tag{378}$$

なぜなら、以下のものが全て $\boldsymbol{O}$ となるからである。

$$(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^{-})\boldsymbol{A}, \qquad \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A}), \qquad \boldsymbol{Q}(\boldsymbol{A})\boldsymbol{A}, \qquad \boldsymbol{A}\boldsymbol{Q}(\boldsymbol{A}') \tag{379}$$

また、$\boldsymbol{S}$ と $\boldsymbol{T}$ を任意の行列として、

$$\boldsymbol{B}_5 = \boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{A}^{-} + (\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A})\boldsymbol{T} + \boldsymbol{S}(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^{-}) \tag{380}$$

も $\boldsymbol{A}$ の一般逆行列となる。 (Harville, p220, eq23)

5.5 ムーアペンローズ逆行列

5.5.1 ムーアペンローズ逆行列の生成

階数が r の行列 $\boldsymbol{A}$ のムーアペンローズ逆行列は、行列 $\boldsymbol{A}$ の特異値分解を

$$\boldsymbol{A} = \boldsymbol{U}_r \boldsymbol{D}_r \boldsymbol{V}_r', \quad \boldsymbol{U}_r'\boldsymbol{U}_r = \boldsymbol{V}_r'\boldsymbol{V} = \boldsymbol{I}_r, \quad \boldsymbol{D} \text{ は対角} \tag{381}$$

とすると

$$\boldsymbol{A}^{+} = \boldsymbol{V}_r \boldsymbol{D}_r{}^{-1} \boldsymbol{U}_r' \tag{382}$$

で与えられる。（証明略）

また、$\boldsymbol{A}$ の階数分解を (294) 式 とすると

$$\boldsymbol{A}^{+} = \boldsymbol{R}^{*'}(\boldsymbol{R}^{*}\boldsymbol{R}^{*'})^{-1}(\boldsymbol{Q}^{*'}\boldsymbol{Q}^{*})^{-1}\boldsymbol{Q}^{*'} \tag{383}$$

である。なお、$\boldsymbol{L}$ が列フルランク、$\boldsymbol{R}$ が行フルランクの場合

$$\boldsymbol{L}_L{}^{-1} = (\boldsymbol{L}'\boldsymbol{L})^{-1}\boldsymbol{L}' \quad \text{and} \quad \boldsymbol{R}_R{}^{-1} = \boldsymbol{R}'(\boldsymbol{R}\boldsymbol{R}')^{-1} \tag{384}$$

を $\boldsymbol{L}$ の左逆行列、$\boldsymbol{R}$ の右逆行列と呼び、それぞれ、元の行列のムーアペンローズ逆行列となっている。

また、ムーアペンローズ逆行列は任意の一般逆行列を用いて

$$\boldsymbol{A}^{+} = \boldsymbol{A}'\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A}\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}' \tag{385}$$

$$\boldsymbol{A}^{+} = \boldsymbol{A}\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}'\boldsymbol{A}\boldsymbol{A}')^{-}\boldsymbol{A} \tag{386}$$

$$\boldsymbol{A}^{+} = \boldsymbol{A}'\big(\boldsymbol{A}\boldsymbol{A}'\big)^{-}\boldsymbol{A}\big(\boldsymbol{A}'\boldsymbol{A}\big)^{-}\boldsymbol{A}' \tag{387}$$

と計算することも出来る。（証明略）この計算式は、たとえ $\boldsymbol{A}$ が矩形行列であっても、正方で対称な行列の一般逆行列が計算できればムーアペンローズ逆行列が計算できることを示している。

なお、

$$\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{A}^{+}, \quad \boldsymbol{A}'\big(\boldsymbol{A}\boldsymbol{A}'\big)^{+}, \quad \big(\boldsymbol{A}'\boldsymbol{A}\big)^{+}\boldsymbol{A}' \tag{388}$$

は $\boldsymbol{A}$ 行列のムーアペンローズ逆行列となる。（証明略）

5.5.2 ムーアペンローズ逆行列の性質

ムーアペンローズ一般逆行列のムーアペンローズ逆行列は元の行列となる。

$$\Big(\boldsymbol{A}^{+}\Big)^{+} = \boldsymbol{A} \tag{389}$$

また、以下のことが言える。

$$(\boldsymbol{A}'\boldsymbol{A})^{+} = \boldsymbol{A}^{+}(\boldsymbol{A}')^{+} \tag{390}$$

$$\Big(\boldsymbol{A}\boldsymbol{A}^{+}\Big)^{+} = \boldsymbol{A}\boldsymbol{A}^{+} \tag{391}$$

$$\Big(\boldsymbol{A}^{+}\boldsymbol{A}\Big)^{+} = \boldsymbol{A}^{+}\boldsymbol{A} \tag{392}$$

$$\boldsymbol{A}\big(\boldsymbol{A}'\boldsymbol{A}\big)^{-}\boldsymbol{A}' = \boldsymbol{A}\boldsymbol{A}^{+} \tag{393}$$

$$(\boldsymbol{A}\otimes\boldsymbol{B})^{+} = \boldsymbol{A}^{+}\otimes\boldsymbol{B}^{+} \tag{394}$$

6　連立一次方程式の一般解と一般逆行列

6.1　連立一次方程式の一般解

$\boldsymbol{A}$ を $a \times b$ の係数行列、 $\boldsymbol{c}$ を $b \times 1$ のベクトルとした場合、連立 1 次方程式

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c} \tag{395}$$

が解を持つためには、$\boldsymbol{c}$ が $\boldsymbol{A}$ 行列の列空間に含まれること、すなわち、$\boldsymbol{c}$ が $\boldsymbol{A}$ の各列に線形従属であること、すなわち、$\boldsymbol{A}$ の階数が、 $\boldsymbol{A}$ の横に $\boldsymbol{c}$ を付け加えたものの階数に等しい

$$\mathrm{rank}\,(\boldsymbol{A}) = \mathrm{rank}\,([\boldsymbol{A}, \boldsymbol{c}]) \tag{396}$$

ことが必要である。

係数行列 $\boldsymbol{A}$ が正則である（特異でない）場合、すなわち $a = b = \mathrm{rank}\,(A)$ の場合には $\boldsymbol{A}$ の逆行列が存在するため

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{c} \tag{397}$$

が唯一の解となる。また、$b \times b$ の $\boldsymbol{A}$ 行列が正則な場合はどのような $b \times 1$ のベクトルも $\boldsymbol{A}$ の列空間に含まれるため、 (396) 式 が満たされる。

係数行列 $\boldsymbol{A}$ の逆行列が存在しない場合（$\boldsymbol{A}$ 行列が特異行列の場合）には、 (396) 式 は $\boldsymbol{c}$ が $\boldsymbol{A}$ の列空間に含まれることを意味し、その場合には、解は一意には定まらないが方程式を満たす $\boldsymbol{x}$ は存在する、すなわち、解は不定 (indeterminate) ということになる。そして、 (396) 式 が満たされていなければこの方程式は不能 (inconsistent) 、すなわち、方程式を満たす $\boldsymbol{x}$ は存在しない場合となる。

解が存在する場合、それらの解の全体、すなわち一般解は、任意のベクトル $\boldsymbol{y}$ を用いて

$$\boldsymbol{x}^* = \boldsymbol{A}^{-}\boldsymbol{c} + (\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A})\boldsymbol{y} \tag{398}$$

と書くことが出来る。だたし、 $\boldsymbol{A}^{-}$ は

$$\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} \tag{399}$$

を満たす行列で $\boldsymbol{A}$ の **一般逆行列** (generalized inverse) と呼ばれる。一般逆行列は、通常、一意に定まるものではない。なお、上式の $(\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A})$ の部分は、行列 $\boldsymbol{A}$ の零空間 (null space) とか核空間 (Kernel) と呼ばれるもので、

$$\boldsymbol{A}\,(\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A}) \;\; = \;\; \boldsymbol{A} - \boldsymbol{A}\,\boldsymbol{A}^{-}\boldsymbol{A} \;=\; \boldsymbol{A} - \boldsymbol{A} \;=\; \boldsymbol{O} \tag{400}$$

となるものである。

なお、$\boldsymbol{A}$ がフルランクの場合 $\boldsymbol{A}^{-} = \boldsymbol{A}^{-1}$ と書けるから、この一般解は (397) 式 と一致する。また、$\boldsymbol{A}$ が列フルランクの場合、階数 b と未知数 $\boldsymbol{x}$ の数は等しく

$$\boldsymbol{A}^{-} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}' \tag{401}$$

であるから、

$$(\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A}) = (\boldsymbol{I} - (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{A}) = \boldsymbol{O} \tag{402}$$

となり、解は

$$\boldsymbol{x} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{c} \tag{403}$$

と一意に定まる。

この一般解が確かに解となっていることを示すためにもとの方程式に代入すると、

$$\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{A}\left(\boldsymbol{A}^{-}\boldsymbol{c}+(\boldsymbol{I}-\boldsymbol{A}^{-}\boldsymbol{A})\boldsymbol{y}\right) = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{c}+\boldsymbol{A}(\boldsymbol{I}-\boldsymbol{A}^{-}\boldsymbol{A})\boldsymbol{y} \tag{404}$$

$$= \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{c}+\boldsymbol{A}\boldsymbol{y}-\boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{y} = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{c} \tag{405}$$

となるが、この方程式は解を持つと仮定すると、係数ベクトル $\boldsymbol{c}$ は、任意のベクトル $\boldsymbol{z}$ を用いて $\boldsymbol{c}=\boldsymbol{A}\boldsymbol{z}$ と書けなければならない。($\boldsymbol{c}$ が $\boldsymbol{A}$ の列空間の中にある。) したがって、

$$\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{c} = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\boldsymbol{z}=\boldsymbol{c} \tag{406}$$

となる。

なお、 (396) 式 が満たされていない場合はこの方程式は解を持たず、 (421) 式 は計算できるもののそれは**解では無い**ことに注意すべきである。

定数ベクトル $\boldsymbol{c}$ の要素が全て 0 の場合の方程式 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ を**斉次連立一次方程式** (homogeneous system of linear equations) もしくは**同次連立一次方程式**と呼ぶが、それが $\boldsymbol{x}=\boldsymbol{0}$ という自明な解以外の解を持つためには $\boldsymbol{A}$ が特異でなければならない。その場合はの上記の公式において $\boldsymbol{c}=\boldsymbol{0}$ として解を求めればよいが、その解は、任意の $\boldsymbol{y}$ を用いて

$$\boldsymbol{x}=(\boldsymbol{I}-\boldsymbol{A}^{-}\boldsymbol{A})\boldsymbol{y} \tag{407}$$

となる。したがって斉次連立一次方程式の自明でない解は不定である。

$\boldsymbol{x}$ が斉次連立一次方程式の一つの解であれば、a をスカラーとして $a\boldsymbol{x}$ もまた解となる。また、$\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots$ を斉次連立一次方程式 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ の解であるとすれば、

$$\sum_i \boldsymbol{A}\boldsymbol{x}_i = \boldsymbol{A}\sum_i \boldsymbol{x}_i = \boldsymbol{0} \tag{408}$$

であるから、$\sum_i \boldsymbol{x}_i$ もまたこの方程式の解となっていることが分かる。また、$\boldsymbol{x}$ が斉次連立一次方程式の一つの解であれば $-\boldsymbol{x}$ も解となっているため、上式で和を取る際に符号を変えて和をとっても良い。これを斉次連立一次方程式の解の重ね合わせと呼ぶ。

6.1.1 列フルランクの場合

係数行列 $\boldsymbol{A}$ が列フルランクの場合は、その左逆行列 $(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'$ が存在するため、それを方程式の両辺の左から掛けると

$$(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}' \quad \boldsymbol{A}\boldsymbol{x} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}' \quad \boldsymbol{c} \tag{409}$$

となり、 $\boldsymbol{x}=(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{c}$ が一意に定まる解となる。

6.1.1.1 回帰分析との関係

重回帰分析モデルは、データの数を n、 説明変数の数を q とすれば

$$\boldsymbol{y}=\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{e} \tag{410}$$

ただし、$\boldsymbol{y}$ は $n\times 1$、$\boldsymbol{X}$ は $n\times q$、$\boldsymbol{\beta}$ は $q\times 1$ 、と書ける。したがって、 $\boldsymbol{\beta}$ を求めるためには、$\boldsymbol{e}$ を $\boldsymbol{0}$ と置き

$$\boldsymbol{y}=\boldsymbol{X}\boldsymbol{\beta} \tag{411}$$

を $\boldsymbol{\beta}$ に関する連立方程式と見做して

$$\boldsymbol{\beta} = \boldsymbol{X}^{-}\boldsymbol{y} + (\boldsymbol{I} - \boldsymbol{X}^{-}\boldsymbol{X})\boldsymbol{c} \tag{412}$$

として解けば解が得られると思われるかもしれないが、これは間違いである。なぜなら、一般に、$q < n$ であるので、この方程式の解は不能になる可能性が高く、解が存在する場合は $\boldsymbol{y}$ が $\boldsymbol{X}$ の列空間に含まれる場合だけであるからである。（これはまれである。）上式の $\boldsymbol{c}$ を $\boldsymbol{0}$ と置き、左逆行列

$$\boldsymbol{X}^{-} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \tag{413}$$

を用いれば、この方程式の見かけ上の解は

$$\boldsymbol{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \quad \boldsymbol{y} \tag{414}$$

となるが、回帰係数の推定値 $\widehat{\boldsymbol{\beta}}$ はこのようにして導かれたものでは無い。

$\widehat{\boldsymbol{\beta}}$ は

$$\mathrm{RSS} = \mathrm{tr}\left[(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})\right] \tag{415}$$

を最小とする $\boldsymbol{\beta}$ の推定値として

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{416}$$

として与えられたが、これは、RSS を $\boldsymbol{\beta}$ の各要素で微分した結果を $\boldsymbol{0}$ と置いて得られる正規方程式

$$(\boldsymbol{X}'\boldsymbol{X}) \quad \boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{y} \tag{417}$$

の解として得られたものであった。この方程式は、係数行列 $(\boldsymbol{X}'\boldsymbol{X})$ は正方行列であり、方程式の数は未知数 $\boldsymbol{\beta}$ の長さと等しい。また、右辺 $\boldsymbol{X}'\boldsymbol{y}$ は必ず係数行列の列空間に含まれているため、解は存在する。ただし、係数行列の階数が q より小さい場合には、解は不定となる。（例えば、$n < q$ の場合や $\boldsymbol{X}$ の列が線形独立でない場合等。）

6.1.2 共通の係数行列 $\boldsymbol{A}$ を持つ複数の連立方程式

同じ係数行列 $\boldsymbol{A}$ を用いた連立方程式が 個存在する場合、それらは

$$\boldsymbol{A}\ \boldsymbol{X} \quad = \quad \boldsymbol{C} \tag{418}$$

ただし、$\boldsymbol{X}$ は $b \times c$ の未知数行列、$\boldsymbol{C}$ は $a \times c$ の係数行列である。

この場合、もしも $\boldsymbol{A}$ が正則であれば、解は

$$\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{C} \tag{419}$$

として一意に定まる。また、$\boldsymbol{A}$ が列フルランクの場合には

$$\boldsymbol{X} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{C} \tag{420}$$

が一意な解となる。

$\boldsymbol{A}$ 行列がフルランクで無い場合の一般解は、$\boldsymbol{Y}$ を任意の $a \times c$ の行列として以下の通りである。

$$\boldsymbol{X}^{*} = \boldsymbol{A}^{-}\boldsymbol{C} + (\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A})\boldsymbol{Y} \tag{421}$$

ただし、 方程式が不能か否かは、$\boldsymbol{C}$ 行列の各列と $\boldsymbol{A}$ 行列との関係を (396) 式 を用いて個別に判断する必要がある。

6.2 等号成立条件

スカラーの場合 $ab = 0$ は $a = 0$ または $b = 0$ を意味するし、$a(b - c) = 0$ は $a = 0$ もしくは $b = c$ を意味するが、行列の場合は必ずしもそうはならない。ここでは、いくつかの等号成立条件を挙げる。

$\boldsymbol{A} = \boldsymbol{B}$ と $\boldsymbol{A} - \boldsymbol{B} = 0$ は同値である。

6：等号成立条件 1

$r \times c$ の行列 $\boldsymbol{A}$ に関して $\boldsymbol{A} = \boldsymbol{O}$ となるための必要十分条件は $\mathrm{tr}\left(\boldsymbol{A}'\boldsymbol{A}\right)$=0 したがって、$\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{O}$ である。

なぜなら、$\mathrm{tr}\left(\boldsymbol{A}'\boldsymbol{A}\right)$ は $\boldsymbol{A}$ 行列の要素の２乗和となるからである。(Harville: Lemma 5.3.1, Corollary 5.3.2)
また、

7：等号成立条件 2

$\boldsymbol{AB} = \boldsymbol{AC}$ となるための必要十分条件は $\boldsymbol{A}'\boldsymbol{AB} = \boldsymbol{A}'\boldsymbol{AC}$ である。

なぜなら、

$$(\boldsymbol{AB} - \boldsymbol{AC})'(\boldsymbol{AB} - \boldsymbol{AC}) = (\boldsymbol{B}' - \boldsymbol{C}')(\boldsymbol{A}'\boldsymbol{AB} - \boldsymbol{A}'\boldsymbol{AC}) \tag{422}$$

であるが、$\boldsymbol{A}$'$\boldsymbol{AB} = \boldsymbol{A}$'$\boldsymbol{AC}$ の場合、上式の右辺、したがって左辺は $\boldsymbol{O}$ となる。したがって、先の等号成立条件により、$\boldsymbol{AB} = \boldsymbol{AC}$ となる。(Harville: Corollary 5.3.3, Searl: Lemma 3)

なお、この条件は、以下の条件と同値である。

8：等号成立条件 2

$$\boldsymbol{A}'\boldsymbol{B} = \boldsymbol{A}'\boldsymbol{C} \quad \longleftrightarrow \quad \boldsymbol{A}\boldsymbol{A}'\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}'\boldsymbol{C}$$
$$\boldsymbol{B}\boldsymbol{A} = \boldsymbol{C}\boldsymbol{A} \quad \longleftrightarrow \quad \boldsymbol{B}\boldsymbol{A}\boldsymbol{A}' = \boldsymbol{C}\boldsymbol{A}\boldsymbol{A}'$$
$$\boldsymbol{B}\boldsymbol{A}' = \boldsymbol{C}\boldsymbol{A}' \quad \longleftrightarrow \quad \boldsymbol{B}\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{C}\boldsymbol{A}'\boldsymbol{A}$$

****
固有値分解・特異値分解に関しては Eigen.tex を参照のこと。20250510
****

7　行列の利用

7.1　座標軸や図形の回転

平面上の一点の座標を $\boldsymbol{x} = (x_1, x_2)'$ とする。この座標軸を原点に関して角度 θ だけ回転して得られる新しい座標系において、元の点の座標は

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \times \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \boldsymbol{T}\boldsymbol{x} \tag{423}$$

で与えられる。この行列を回転行列（変換行列）と呼ぶ。回転行列は直交行列である。したがって、元の座標系に戻すには $\boldsymbol{x} = \boldsymbol{T}^{-1}\boldsymbol{y} = \boldsymbol{T}'\boldsymbol{y}$ とすればよい。また、座標系を角度 θ' だけ回転する行列を $\boldsymbol{S}$ とすれば、$\boldsymbol{y}$ をさらに θ' だけ回転したものは

$$\boldsymbol{z} = \boldsymbol{S}\boldsymbol{y} = \boldsymbol{S}(\boldsymbol{T}\boldsymbol{x}) = (\boldsymbol{S}\boldsymbol{T})\boldsymbol{x} \tag{424}$$

となるが、これは、行列 $\boldsymbol{S}\boldsymbol{T}$,

$$\boldsymbol{S}\,\boldsymbol{T} = \begin{bmatrix} \cos\theta'\cos\theta - \sin\theta'\sin\theta & -\cos\theta'\sin\theta - \sin\theta'\cos\theta \\ \sin\theta'\cos\theta + \cos\theta'\sin\theta & -\sin\theta'\sin\theta + \cos\theta'\cos\theta \end{bmatrix} \tag{425}$$

が、元の座標系を角度 $\theta + \theta'$ だけ回転する行列となっていることを示している。したがって、この行列の $(1,1)$ 要素には $\cos(\theta + \theta')$ がまた、$(2,1)$ 要素には $\sin(\theta + \theta')$ が入っているはずである。すなわち、三角関数の加法定理を表している。

p 個の点を同時に変換する場合には、点の座標を行列の形に並べたものを $\boldsymbol{X}$ $(p \times 2)$ とすると、

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{T}' \tag{426}$$

となる。

逆に、元の座標系で点 $\boldsymbol{x}$ を原点の周りに角度 θ だけ回転して得られる点 $\boldsymbol{z}$ の座標は、座標軸を $-\theta$ だけ回転することと同じだから、

$$\boldsymbol{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \times \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \boldsymbol{T}'\boldsymbol{x} \tag{427}$$

となる。

7.2　度数分布表・クロス集計表（同時確率・周辺確率・条件付き確率）

r 個と c 個のカテゴリを持つ二つの名義尺度による 2 元のクロス集計表を $r \times c$ の $\boldsymbol{F}$ で表す。ただし、f_{ij} には、第 1 変数の第 i 番目のカテゴリと、第 2 変数の第 j 番目のカテゴリとを同時に選んだ人の度数が入っている。この場合、周辺度数並びに総反応数、そして周辺反応確率は

$$f_{r_i} = \sum_{j=1}^{c} f_{ij}, \quad \text{and} \quad f_{c_j} = \sum_{i=1}^{r} f_{ij}, \quad \text{and} \quad n = \sum_{i=1}^{r}\sum_{j=1}^{c} f_{ij} \tag{428}$$

$$p_{ij} = \frac{1}{n} f_{ij}, \quad \text{and} \quad p_{r_i} = \frac{1}{n}\sum_{j=1}^{c} f_{ij}, \quad \text{and} \quad p_{c_j} = \frac{1}{n}\sum_{i=1}^{r} f_{ij} \tag{429}$$

となるが、これを行列を用いて表すと

$$\boldsymbol{f}_r = \boldsymbol{F}\boldsymbol{1}_c \quad \text{and} \quad \boldsymbol{f}_c = (\boldsymbol{1}_r'\boldsymbol{F})' \quad \text{and} \quad n = \boldsymbol{1}_r'\boldsymbol{F}\boldsymbol{1} \tag{430}$$

$$\boldsymbol{P} = \frac{1}{n}\boldsymbol{F} \quad \text{and} \quad \boldsymbol{p}_r = \frac{1}{n}\boldsymbol{f}_r \quad \text{and} \quad \boldsymbol{p}_c = \frac{1}{n}\boldsymbol{f}_c \tag{431}$$

となる。二つの変数が独立であれば、

$$\boldsymbol{P} = \boldsymbol{p}_r \boldsymbol{p}_c' \tag{432}$$

となる。また、これらを用いると、列に配置された変数で条件付けられた反応確率行列 $\boldsymbol{C}$, ただし、$\boldsymbol{C}$ は $\sum_{i=1}^{r} c_{ij} = 1\ , j = 1, 2, \ldots, c$ を満たす、と、行に配置された変数で条件付けられた反応確率行列 $\boldsymbol{R}$ ただし、$\boldsymbol{R}$ は $\sum_{j=1}^{c} r_{ij} = 1,\ i = 1, 2, \ldots, r$ を満たす、は

$$\boldsymbol{C} = \boldsymbol{P}\ \mathrm{Diag}\left(\boldsymbol{f}_c\right)^{-1} \quad \text{and} \quad \boldsymbol{R} = \mathrm{Diag}\left(\boldsymbol{f}_r\right)^{-1} \boldsymbol{P} \tag{433}$$

となる。ただし、 $\mathrm{Diag}\left(\boldsymbol{a}\right)$ はベクトル $\boldsymbol{a}$ の要素を対角要素として持つ対角行列である。

逆に、列に配置された変数で条件付けられた反応確率行列 $\boldsymbol{C}$ と列に配置されたカテゴリの周辺確率 $\boldsymbol{f}_c$ が与えられた場合、行に配置された変数のカテゴリの周辺分布は

$$\boldsymbol{f}_r = \boldsymbol{C}\boldsymbol{f}_c \quad \left(= \boldsymbol{P}\ \mathrm{Diag}\left(\boldsymbol{f}_c\right)^{-1} \boldsymbol{f}_c = \boldsymbol{P}\boldsymbol{1}\right) \tag{434}$$

となる。

7.3 デザイン行列

水準数が A の１元配置の分散分析モデルは、各水準での観測値数を $n_j, j = 1, 2, \ldots, A$ とすると、

$$y_{ij} = \widehat{y}_{ij} + \varepsilon_{ij}, \quad \widehat{y}_{ij} = \mu + \alpha_j, \quad i = 1, 2, \ldots, n_j,\ j = 1, 2, \ldots, A \tag{435}$$

である。これを線形モデル $\widehat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{\beta}$ の形で書くと

$$\begin{pmatrix} \widehat{y}_{11} \\ \widehat{y}_{21} \\ \vdots \\ \widehat{y}_{n_1 1} \\ \widehat{y}_{12} \\ \widehat{y}_{22} \\ \vdots \\ \widehat{y}_{1A} \\ \vdots \\ \widehat{y}_{n_A A} \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 1 & \ldots & 0 \\ 1 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & 0 & 0 & \ldots & 1 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & 0 & 0 & \ldots & 1 \end{bmatrix} \times \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_A \end{pmatrix} \tag{436}$$

となる。なお、このデザイン行列は、その第 2 列目から最後までを加えると第 1 列目の単位ベクトルとなるが、このように、行列のある列が、残りの列の線形結合として表されるような場合にその行列の各列は１次独立ではない（１次従属である）と言う。その場合、$\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}$ は存在しない。従って、パラメタ $\boldsymbol{\beta}$ の推定値は一意には定まらない。このデザイン行列のように、各行に一つだけ 1 があり他の要素は全て 0 と言う行列をダミー行列と呼ぶことがある。

通常の制約条件 $\sum_{j=1}^{A}\alpha_j = 0$ を入れると、$\alpha_A = -\sum_{j=1}^{A-1}\alpha_j$ であるから、同じモデルを

$$\begin{pmatrix} \widehat{y}_{11} \\ \widehat{y}_{21} \\ \vdots \\ \widehat{y}_{n_1 1} \\ \widehat{y}_{12} \\ \widehat{y}_{22} \\ \vdots \\ \widehat{y}_{1A} \\ \vdots \\ \widehat{y}_{n_A A} \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & -1 & -1 & \dots & -1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & -1 & -1 & \dots & -1 \end{bmatrix} \times \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{A-1} \end{pmatrix} \tag{437}$$

と書くこともできる。このデザイン行列の各列は１次独立である。

２元配置の場合のモデルは

$$y_{ijk} = \widehat{y}_{ijk} + \varepsilon_{ijk}, \quad \widehat{y}_{ijk} = \mu + \alpha_j + \beta_k, \quad i = 1, 2, \dots, n_{ij},\ j = 1, 2, \dots, A\ k = 1, 2, \dots, B \tag{438}$$

であるが、全ての n_{ij} を 1 とした場合のデザイン行列は

$$\begin{pmatrix} \widehat{y}_{111} \\ \widehat{y}_{112} \\ \vdots \\ \widehat{y}_{11B} \\ \widehat{y}_{121} \\ \widehat{y}_{122} \\ \vdots \\ \widehat{y}_{1A1} \\ \vdots \\ \widehat{y}_{1AB} \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 1 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 0 & 0 & \dots & 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 1 \end{bmatrix} \times \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_A \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_B \end{pmatrix} \tag{439}$$

となる。このデザイン行列の場合、要因 A に対応する列の和、並びに要因 B に対応する列の和がそれぞれ単位ベクトルになり、二つの線形従属関係がある。従って、パラメタの推定値は一意には定まらない。なお、このデザイン行列を $\boldsymbol{X} = \{\boldsymbol{1},\ \boldsymbol{A},\ \boldsymbol{B}\}$ と分割し、すべてのデータの分だけ生成すると、二つの要因の水準間の同時度数行列 $\boldsymbol{F}$ は

$$\boldsymbol{F} = \boldsymbol{A}\boldsymbol{B}' \tag{440}$$

で与えられる。

通常の制約条件 $\sum_{j=1}^{A}\alpha_j=0,\ \sum_{k=1}^{B}\beta_k=0$ を入れると、$\beta_B=-\sum_{k=1}^{B-1}\beta_k$ であるから、同じモデルを

$$\begin{pmatrix} \widehat{y}_{111} \\ \widehat{y}_{112} \\ \vdots \\ \widehat{y}_{11B} \\ \widehat{y}_{121} \\ \widehat{y}_{122} \\ \vdots \\ \widehat{y}_{1A1} \\ \vdots \\ \widehat{y}_{1AB} \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 0 & \dots & 0 & -1 & -1 & \dots & -1 \\ 1 & 0 & 1 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & -1 & -1 & \dots & -1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & -1 & -1 & \dots & -1 & -1 & -1 & \dots & -1 \end{bmatrix} \times \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{A-1} \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{B-1} \end{pmatrix} \tag{441}$$

と書くこともできる。このデザイン行列の各列は１次独立である。、

なお、モデルに交互作用が組み込まれている場合には、交互作用のパラメタ $\alpha\beta_{jk}$ に対応するデザイン行列の列を、 α_j と β_k に対応する $\boldsymbol{X}$ の２つの列 $\boldsymbol{x}_{(1+j)},\boldsymbol{x}_{(1+A+k)}$ の要素毎の積を要素として持つベクトルとして定義していけばよい。

7.4 距離行列

二つの点 i,j の間のユークリッド距離（の２乗）は、各点の座標を $\boldsymbol{x}_j,\ ,j=1,2,\dots,N$ とすれば、

$$d_{ij}^2 \ = \ \sum_{a=1}^{q}(x_{ia}-x_{ja})^2 \ = \ (\boldsymbol{x_i}-\boldsymbol{x}_j)'(\boldsymbol{x_i}-\boldsymbol{x}_j) \tag{442}$$

と書ける。これをまとめて、d_{ij}^2 を要素として持つ $N\times N$ の行列を $\boldsymbol{D}^2$ とすれば,

$$\boldsymbol{D}^2=\mathrm{diag}\,(\boldsymbol{XX}')\mathbf{1}\mathbf{1}'-2\boldsymbol{XX}'+\mathbf{1}\mathbf{1}'\mathrm{diag}\,(\boldsymbol{XX}') \tag{443}$$

と書くことが出来る。

7.5 推移確率行列

たとえば、 第 1 - 3 次産業に従事する人たちの産業間の移動確率が

$$\boldsymbol{P}=\begin{bmatrix} .5 & .1 & .1 \\ .3 & .6 & .1 \\ .2 & .3 & .8 \end{bmatrix} \tag{444}$$

で与えられているとする。この推移確率行列 $\boldsymbol{P}$ に要素 p_{ij} は、一定期間後に第 j 次産業から第 i 次産業へ移動する人の割合を示している。従って、 $\mathbf{1}'\boldsymbol{P}=\mathbf{1}'$ である。すなわち、第 1 次産業の人で一定期間後にそのまま第 1 次産業に従事している人が 50 %, 第 2 次産業に移動する人が 30 %, 第 3 次産業に移動する人が 20 % いることを示している。最初の各産業に従事する人の割合が $\boldsymbol{p}_0,\ \mathbf{1}'\boldsymbol{p}_0=1$ であったとすると、時点を重ねるに連れその割合は

$$\boldsymbol{p}_1=\boldsymbol{P}\boldsymbol{p}_0,\quad \boldsymbol{p}_2=\boldsymbol{P}\boldsymbol{p}_1=\boldsymbol{P}^2\boldsymbol{p}_0,\quad \dots \tag{445}$$

と変化していき、最終的には

$$\boldsymbol{P}^{\infty} = \begin{bmatrix} .16667 & .16667 & .16667 \\ .26667 & .26667 & .26667 \\ .56667 & .56667 & .56667 \end{bmatrix} \tag{446}$$

を得る。

7.6 ソシオメトリの行列

n 人の人のうち、 i さんが j さんに話しかける場合に $s_{ij} = 1$ そうでない場合には $s_{ij} = 0$ と定義した $n \times n$ の行列を考える。ただし $s_{ii} = 1$ とする。この場合、$\boldsymbol{S}^{\ell}$ の (i, j) 要素が 0 であるということは、間に ℓ-1 人を介しても i さんから j さんに情報が伝わらないことを示している。また、i さんが発信した情報が間に ℓ-1 人の人を挟んで j さんに伝わる場合、その異なる経路の数を $d_{ij}^{(\ell)}$ で表すと以下のようになる。

$$\boldsymbol{D}^1 = \boldsymbol{S}, \quad d_{ij}^{(\ell)} = \sum_{k \neq i,j} s_{ik} d_{kj}^{(\ell-1)}, \tag{447}$$

$$\boldsymbol{D} = \boldsymbol{S}\boldsymbol{D}^{(\ell-1)} - \mathrm{diag}\,(\boldsymbol{S})\boldsymbol{D}^{(\ell-1)} - \boldsymbol{S}\mathrm{diag}\left(\boldsymbol{D}^{(\ell-1)}\right) + \mathrm{diag}\,(\boldsymbol{S})\mathrm{diag}\left(\boldsymbol{D}^{(\ell-1)}\right) \tag{448}$$

7.7 データ行列の要約統計量

個体の変数による測定値をベクトルや行列で表した場合の標本平均や標本分散等の標本統計量の計算については以下の通りである。

$N \times 1$ のベクトル $\boldsymbol{x}$ の要素の平均（スカラー量）は

$$\overline{\boldsymbol{x}} = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{1}{N}\mathbf{1}'\boldsymbol{x} = \frac{1}{N}\boldsymbol{x}'\mathbf{1} \tag{449}$$

である。したがって、元のベクトルから平均を引いて得られる $N \times 1$ の平均偏差ベクトル $\{x_i - \bar{\boldsymbol{x}}\}$ は

$$\boldsymbol{x} - \mathbf{1}\,(\frac{1}{N}\mathbf{1}'\boldsymbol{x}) = \boldsymbol{x} - \frac{1}{N}\,\mathbf{1}\mathbf{1}'\,\boldsymbol{x} = \left(\boldsymbol{I} - \frac{1}{N}\,\mathbf{1}\mathbf{1}'\right)\boldsymbol{x} = \boldsymbol{J}\boldsymbol{x} \tag{450}$$

と書ける。この $N \times N$ の行列 $\boldsymbol{J} = \boldsymbol{I} - \frac{1}{N}\mathbf{1}\mathbf{1}'$ は **中心化演算子** (column centoring operattor) と呼ばれ $\boldsymbol{J}\boldsymbol{J} = \boldsymbol{J},\ \boldsymbol{J}\mathbf{1} = \mathbf{0}$ となる。平均偏差ベクトルの自分自身との内積を作り N で除すれば分散および標準偏差を得る。

$$var\,(\boldsymbol{x}) = \frac{\sum_{i=1}^{N}(x_i - \overline{\boldsymbol{x}})^2}{N} = \frac{1}{N}(\boldsymbol{J}\boldsymbol{x})'\boldsymbol{J}\boldsymbol{x} = \frac{1}{N}\boldsymbol{x}'\boldsymbol{J}\boldsymbol{x} = \frac{1}{N}\|\boldsymbol{J}\boldsymbol{x}\|^2 = [std\,(\boldsymbol{x})]^2 \tag{451}$$

と定義する。また、$\boldsymbol{y}$ を $\boldsymbol{x}$ と同じサイズのベクトルとすれば、$\boldsymbol{x}$ と $\boldsymbol{y}$ の共分散を

$$cov\,(\boldsymbol{x}, \boldsymbol{y}) = \frac{\sum_{i=1}^{N}(x_i - \overline{\boldsymbol{x}})(y_i - \overline{\boldsymbol{y}})}{N} = \frac{1}{N}\boldsymbol{x}'\boldsymbol{J}\boldsymbol{y} \tag{452}$$

相関係数を

$$corr\,(\boldsymbol{x}, \boldsymbol{y}) = \frac{cov\,(\boldsymbol{x}, \boldsymbol{y})}{std\,(\boldsymbol{x})std\,(\boldsymbol{y})} \tag{453}$$

と定義する。 $\bar{}$, $var\,(\cdot)$, $cov\,(\cdot,\cdot)$ および $corr\,(\cdot,\cdot)$ を演算子と見た場合、これらは引数の線形変換に関して以下のような性質を持つ。

$$\begin{aligned}
\overline{a+b\boldsymbol{x}} &= a+b\overline{\boldsymbol{x}} && (454)\\
var\,(a+b\boldsymbol{x}) &= b^2 var\,(\boldsymbol{x}) && (455)\\
cov\,(a+b\boldsymbol{x}, c+d\boldsymbol{y}) &= bd\ cov\,(\boldsymbol{x},\boldsymbol{y}) && (456)\\
cov\,(\boldsymbol{x},\boldsymbol{x}) &= var\,(\boldsymbol{x}) && (457)\\
corr\,(a+b\boldsymbol{x}, c+d\boldsymbol{y}) &= \mathrm{sign}(bd)\ corr\,(\boldsymbol{x},\boldsymbol{y}) && (458)
\end{aligned}$$

変数が複数ある場合には以下のようになる。$N\times\ p$ の行列 $\boldsymbol{X}$ の列平均を $p\times\ 1$ の形に並べたものは以下のようになる。

$$\overline{\boldsymbol{X}} = \left\{\frac{1}{N}\sum_{i=1}^{N} x_{ij}\right\} = \left\{\overline{\boldsymbol{x}_{(j)}}\right\} = \frac{1}{N}(\mathbf{1}'\boldsymbol{X})' = \frac{1}{N}\boldsymbol{X}'\mathbf{1} \tag{459}$$

従って、各列ごとの平均偏差ベクトルからなる行列は $\boldsymbol{JX}$ となる。ゆえに、分散共分散行列（対角要素に分散、非対角要素に共分散が入った $p\times\ p$ の対称行列）は

$$\boldsymbol{D}(\boldsymbol{X}) = \left\{\frac{1}{N}\sum_{i=1}^{N}(x_{ij}-\overline{\boldsymbol{x}_{(j)}})(x_{ik}-\overline{\boldsymbol{x}_{(k)}})\right\} = \frac{1}{N}(\boldsymbol{JX})'\boldsymbol{JX} = \frac{1}{N}\boldsymbol{X}'\boldsymbol{JX} \tag{460}$$

と書ける。次に、$\mathrm{diag}\,(\boldsymbol{A})$ で、正方行列 $\boldsymbol{A}$ の非対角部分をゼロとおいた対角行列を、また、$\boldsymbol{A}$ が非負の要素を持つ対角行列の場合、$(\boldsymbol{A})^{\frac{1}{2}}$ で $\boldsymbol{A} = (\boldsymbol{A})^{\frac{1}{2}}(\boldsymbol{A})^{\frac{1}{2}}$ となるような行列（もとの行列の対角要素の平方根からなる行列）、を表せば、

$$\left[(\mathrm{diag}\,(\boldsymbol{D}(\boldsymbol{X})))^{\frac{1}{2}}\right]^{-1} = (\mathrm{diag}\,(\boldsymbol{D}(\boldsymbol{X})))^{-\frac{1}{2}} \tag{461}$$

は各変数の標準偏差の逆数を対角要素として持つ対角行列となる。従って、各要素を列毎に標準化したものを要素として持つ行列 $\boldsymbol{Z} = \left\{(x_{ij}-\overline{\boldsymbol{x}_{(j)}})/std\left(\boldsymbol{x}_{(j)}\right)\right\}$ は

$$\boldsymbol{Z} = \left(\boldsymbol{X} - \frac{1}{N}\mathbf{1}\overline{\boldsymbol{X}}'\right)(\mathrm{diag}\,(\boldsymbol{D}(\boldsymbol{X})))^{-\frac{1}{2}} \tag{462}$$

と書ける。

また、$\boldsymbol{Y}$ を $N\times\ q$ の行列とすれば、$\boldsymbol{X}$ と $\boldsymbol{Y}$ の共分散行列 $(p\times\ q)$ は

$$\boldsymbol{C}(\boldsymbol{X},\boldsymbol{Y}) = \left\{\frac{1}{N}\sum_{i=1}^{N}(x_{ij}-\overline{\boldsymbol{x}_{(j)}})(y_{ik}-\overline{\boldsymbol{y}_{(k)}})\right\} = \frac{1}{N}(\boldsymbol{JX})'\boldsymbol{JY} = \frac{1}{N}\boldsymbol{X}'\boldsymbol{JY} \tag{463}$$

相関係数行列は

$$\begin{aligned}
\boldsymbol{R}(\boldsymbol{X},\boldsymbol{Y}) &= \left\{\frac{1}{N}\frac{\sum_{i=1}^{N}(x_{ij}-\overline{\boldsymbol{x}_{(j)}})(y_{ik}-\overline{\boldsymbol{y}_{(k)}})}{std\left(\boldsymbol{x}_{(j)}\right)std\left(\boldsymbol{x}_{(k)}\right)}\right\} && (464)\\
&= \frac{1}{N}(\boldsymbol{JX}\left[(\mathrm{diag}\,(\boldsymbol{D}(\boldsymbol{X})))^{\frac{1}{2}}\right]^{-1})'\boldsymbol{JX}\left[(\mathrm{diag}\,(\boldsymbol{D}(\boldsymbol{Y})))^{\frac{1}{2}}\right]^{-1} && (465)\\
&= \left[(\mathrm{diag}\,(\boldsymbol{D}(\boldsymbol{X})))^{\frac{1}{2}}\right]^{-1}\boldsymbol{C}(\boldsymbol{X},\boldsymbol{Y})\left[(\mathrm{diag}\,(\boldsymbol{D}(\boldsymbol{Y})))^{\frac{1}{2}}\right]^{-1} && (466)
\end{aligned}$$

と書くことが出来る。なお、

$$\boldsymbol{C}(\boldsymbol{X},\boldsymbol{X}) = \boldsymbol{D}(\boldsymbol{X}) \tag{467}$$

$\boldsymbol{D}(\cdot)$, $\boldsymbol{C}(\cdot,\cdot)$ および $\boldsymbol{R}(\cdot,\cdot)$ は以下のような性質を持つ。

$$\begin{aligned}
\overline{\mathbf{1}\boldsymbol{a}' + \boldsymbol{X}\boldsymbol{B}'} &= \boldsymbol{a} + \boldsymbol{B}\overline{\boldsymbol{X}} && (468)\\
\boldsymbol{D}(\mathbf{1}\boldsymbol{a}' + \boldsymbol{X}\boldsymbol{B}') &= \boldsymbol{B}\boldsymbol{D}(\boldsymbol{X})\boldsymbol{B}' && (469)\\
\boldsymbol{C}(\mathbf{1}\boldsymbol{a}' + \boldsymbol{X}\boldsymbol{B}', \mathbf{1}\boldsymbol{c}' + \boldsymbol{Y}\boldsymbol{D}') &= \boldsymbol{B}\boldsymbol{C}(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{D}' && (470)\\
\overline{\boldsymbol{X}+\boldsymbol{Y}} &= \overline{\boldsymbol{X}} + \overline{\boldsymbol{Y}} && (471)\\
\boldsymbol{C}(\boldsymbol{X}, \boldsymbol{Y}+\boldsymbol{Z}) &= \boldsymbol{C}(\boldsymbol{X},\boldsymbol{Y}) + \boldsymbol{C}(\boldsymbol{X},\boldsymbol{Z}) && (472)\\
\boldsymbol{D}(\boldsymbol{Y}+\boldsymbol{X}) &= \boldsymbol{D}(\boldsymbol{Y}) + \boldsymbol{D}(\boldsymbol{X}) + \boldsymbol{C}(\boldsymbol{Y},\boldsymbol{X}) + \boldsymbol{C}(\boldsymbol{X},\boldsymbol{Y}) && (473)
\end{aligned}$$

7.8 確率変数と確率変数ベクトル

X をスカラーの離散的確率変数とし、その値域を $\{x_1, x_2, \ldots, x_m\}$ とする。ただし、$x_k,\ k = 1, 2, \ldots, m$ は数値である。また、X がそれぞれの値を取る確率を $\Pr(X = x_k) = \pi_k,\ k = 1, 2, \ldots, m$ とする。ただし、$\sum_{k=1}^{m} \pi_k = 1$ である。この場合、X の平均と分散は

$$\begin{aligned}
\boldsymbol{\mathcal{E}}(X) &= \sum_{k=1}^{m} x_i \pi_k = \boldsymbol{x}'\boldsymbol{\pi} && (474)\\
var(X) &= \sum_{k=1}^{m} (x_k - \boldsymbol{\mathcal{E}}(X))^2 \pi_k \ = \ \sum_{k=1}^{m} x_k^2 \pi_k + \left(\sum_{k=1}^{m} x_k \pi_k\right)^2 && (475)\\
&= \boldsymbol{x}'\mathrm{Diag}(\boldsymbol{\pi})\boldsymbol{x} - (\boldsymbol{x}'\boldsymbol{\pi})^2 \ = \ (\boldsymbol{x}\odot\boldsymbol{x})'\boldsymbol{\pi} - (\boldsymbol{x}'\boldsymbol{\pi})^2 && (476)
\end{aligned}$$

と書ける。ただし、$\boldsymbol{x}$ や $\boldsymbol{\pi}$ は x_k や π_k を要素として持つ $m \times 1$ のベクトルである。X が連続的確率変数の場合、その密度関数を $f(x)$ とすると、

$$\begin{aligned}
\boldsymbol{\mathcal{E}}(X) &= \int_{-\infty}^{+\infty} x f(x) \dot{\mathrm{x}} && (477)\\
var(X) &= \int_{-\infty}^{+\infty} (x - \boldsymbol{\mathcal{E}}(()X))^2 f(x) dx && (478)\\
&= \int_{-\infty}^{+\infty} x^2 f(x) dx - \boldsymbol{\mathcal{E}}(X)^2 && (479)
\end{aligned}$$

確率変数が二つある場合、以下の様に共分散を定義することが出来る。ただし、X の値域を $\{x_1, x_2, \ldots, x_m x\}$ Y の値域を $\{y_1, y_2, \ldots, y_m y\}$ とし、$\Pr(X = x_k, Y = y_j) = \pi_{kj},\ k = 1, 2, \ldots, mx,\ j = 1, 2, \ldots, my$, $\sum_{k=1}^{mx}\sum_{j=1}^{my} \pi_{kj} = 1$ とする。

$$\begin{aligned}
cov(X,Y) &= \sum_{k=1}^{mx}\sum_{j=1}^{my} (x_k - \boldsymbol{\mathcal{E}}(X)))\,(y_j - \boldsymbol{\mathcal{E}}(Y)))\,\pi_{kj} && (480)\\
&= \sum_{k=1}^{mx}\sum_{j=1}^{my} x_k y_j \pi_{kj} - \boldsymbol{\mathcal{E}}(X)\boldsymbol{\mathcal{E}}(Y) && (481)\\
&= \boldsymbol{x}'\boldsymbol{\Pi}\boldsymbol{y} - \boldsymbol{x}'\boldsymbol{\pi}\boldsymbol{y}'\boldsymbol{\pi} && (482)
\end{aligned}$$

ただし、$\boldsymbol{\Pi}$ は π_{jk} を要素として持つ $mx \times my$ の行列である。なお、X と Y の周辺確率は、

$$\Pr(X = x_i = \boldsymbol{\Pi}[i+]), \qquad \text{and} \qquad \Pr(Y = y_j = \boldsymbol{\Pi}[+j]) \tag{483}$$

で与えられる。また、連続的確率変数の場合は、同時密度関数を $f(x,y)$ として、

$$cov\,(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\boldsymbol{\mathcal{E}}(X))\,(y-\boldsymbol{\mathcal{E}}(Y))\,f(x,y)dydx \tag{484}$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)dydx-\boldsymbol{\mathcal{E}}(X)\boldsymbol{\mathcal{E}}(Y) \tag{485}$$

である。

なお、通常の数理統計学のテキストでは、スカラーの確率変数は大文字のアルファベットで、その取る値もしくは実現値を小文字のアルファベットで表すが、以下に示す確率変数ベクトルの場合は、この記法が使いにくくなる。なぜなら、確率変数ベクトルを大文字のボールドのアルファベットで表すと、それが行列に見えるからである。したがって、以下の記法では、確率変数とその実現値を区別せず、ボールドの小文字はベクトル、ボールドの大文字は行列を表すものとする。

r 個の確率変数を要素として持つ $r_x\times 1$ のベクトル $\boldsymbol{x}$ を確率変数ベクトルと呼び、その平均（期待値）ベクトルと分散共分散行列を

$$\boldsymbol{\mathcal{E}}(\boldsymbol{x})\quad(r_x\times 1)\qquad\text{and}\qquad\boldsymbol{\mathcal{D}}(\boldsymbol{x})=\boldsymbol{\mathcal{E}}(\boldsymbol{x}\boldsymbol{x}')-\boldsymbol{\mathcal{E}}(\boldsymbol{x})\boldsymbol{\mathcal{E}}(\boldsymbol{x})'\quad(r_x\times r_x) \tag{486}$$

と記す。ただし、 $\boldsymbol{\mathcal{E}}(\boldsymbol{x})$ はその第 i 番目の要素として $\boldsymbol{\mathcal{E}}(x_i)$ を持つ $r_x\times 1$ のベクトル、$\boldsymbol{\mathcal{D}}(\boldsymbol{x})$ はその ij 要素として x_i と x_j の共分散 $cov\,(x_i,x_j)$ を持つ $r_x\times r_x$ の対称行列である。

この確率変数を $r_y\times 1$ のベクトル $\boldsymbol{a}$ と $r_y\times r_x$ の行列 $\boldsymbol{B}$ を用いて

$$\boldsymbol{y}=\boldsymbol{a}+\boldsymbol{B}\boldsymbol{x} \tag{487}$$

と変換したもの $\boldsymbol{y}$ の期待値と分散共分散行列は以下のようになる。

$$\boldsymbol{\mathcal{E}}(\boldsymbol{y})=\boldsymbol{a}+\boldsymbol{B}\boldsymbol{\mathcal{E}}(\boldsymbol{x})\qquad\text{and}\qquad\boldsymbol{\mathcal{D}}(\boldsymbol{y})=\boldsymbol{B}\boldsymbol{\mathcal{D}}(\boldsymbol{x})\boldsymbol{B}' \tag{488}$$

また、別の確率変数ベクトル $\boldsymbol{y}$, $\boldsymbol{z}$ との間の共分散行列を $\boldsymbol{\mathcal{C}}ov\,(\boldsymbol{x},\boldsymbol{y})$, $\boldsymbol{\mathcal{C}}ov\,(\boldsymbol{x},\boldsymbol{z})$ とすれば、

$$\boldsymbol{\mathcal{E}}(\boldsymbol{B}\boldsymbol{x}+\boldsymbol{A}\boldsymbol{y}) = \boldsymbol{B}\boldsymbol{\mathcal{E}}(\boldsymbol{x})+\boldsymbol{A}\boldsymbol{\mathcal{E}}(\boldsymbol{y}),\quad \boldsymbol{\mathcal{C}}ov\,(\boldsymbol{B}\boldsymbol{x},\boldsymbol{A}\boldsymbol{y}) = \boldsymbol{B}\boldsymbol{\mathcal{C}}ov\,(\boldsymbol{x},\boldsymbol{y})\boldsymbol{A}' \tag{489}$$

$$\boldsymbol{\mathcal{E}}(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{\mathcal{E}}(\boldsymbol{x})+\boldsymbol{\mathcal{E}}(\boldsymbol{y}),\quad \boldsymbol{\mathcal{C}}ov\,(\boldsymbol{x},\boldsymbol{y}+\boldsymbol{z}) = \boldsymbol{\mathcal{C}}ov\,(\boldsymbol{x},\boldsymbol{y})+\boldsymbol{\mathcal{C}}ov\,(\boldsymbol{x},\boldsymbol{z}) \tag{490}$$

$$\boldsymbol{\mathcal{D}}(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{\mathcal{D}}(\boldsymbol{x})+\boldsymbol{\mathcal{D}}(\boldsymbol{y})+\boldsymbol{\mathcal{C}}ov\,(\boldsymbol{x},\boldsymbol{y})+\boldsymbol{\mathcal{C}}ov\,(\boldsymbol{y},\boldsymbol{x}) \tag{491}$$

なお、期待値の場合は確率変数を要素として持つ行列 $\boldsymbol{X}$ に関してベクトルの場合と同様に

$$\boldsymbol{\mathcal{E}}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{X})=\boldsymbol{A}+\boldsymbol{B}\boldsymbol{\mathcal{E}}(\boldsymbol{X}) \tag{492}$$

となる。すなわち、期待値演算子は線形の演算子である。

また、２次形式 $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ の期待値は、トレースも線形演算子であることから、

$$\boldsymbol{\mathcal{E}}(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{\mathcal{E}}(\mathrm{tr}\,(\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}')) = \mathrm{tr}\,(\boldsymbol{\mathcal{E}}(\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}')) = \mathrm{tr}\,(\boldsymbol{A}\boldsymbol{\mathcal{E}}(\boldsymbol{x}\boldsymbol{x}')) = \mathrm{tr}\left(\boldsymbol{A}\left(\boldsymbol{\mathcal{D}}(\boldsymbol{x})+\boldsymbol{\mathcal{E}}(\boldsymbol{x})\boldsymbol{\mathcal{E}}(\boldsymbol{x})'\right)\right) \tag{493}$$

$$= \mathrm{tr}\,(\boldsymbol{A}\boldsymbol{\mathcal{D}}(\boldsymbol{x}))+\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{\mathcal{E}}(\boldsymbol{x})\boldsymbol{\mathcal{E}}(\boldsymbol{x})'\right) = \boldsymbol{\mathcal{E}}(\boldsymbol{x})'\boldsymbol{A}\boldsymbol{\mathcal{E}}(\boldsymbol{x})+\mathrm{tr}\,(\boldsymbol{A}\boldsymbol{\mathcal{D}}(\boldsymbol{x})) \tag{494}$$

となる。

7.9 因子分析

因子分析モデルは

$$\boldsymbol{y} = \boldsymbol{\mu}+\boldsymbol{\Lambda}\boldsymbol{f}+\boldsymbol{\varepsilon} \tag{495}$$

$$\boldsymbol{\mathcal{E}}(\boldsymbol{f}) = \boldsymbol{0},\quad \boldsymbol{\mathcal{D}}(\boldsymbol{f})=\boldsymbol{\Phi},\quad \boldsymbol{\mathcal{E}}(\boldsymbol{\varepsilon})=\boldsymbol{0},\quad \boldsymbol{\mathcal{D}}(\boldsymbol{\varepsilon})=\boldsymbol{\Psi},\quad \boldsymbol{\mathcal{C}}ov\,(\boldsymbol{f},\boldsymbol{\varepsilon})=\boldsymbol{O} \tag{496}$$

と書くことが出来るが、これから $\boldsymbol{y}$ の分散共分散行列

$$\boldsymbol{\mathcal{D}}(\boldsymbol{y}) = \boldsymbol{\Lambda}\Phi\boldsymbol{\Lambda}' + \boldsymbol{\Psi} \tag{497}$$

が導ける。

7.10 多変量正規分布の密度関数

$\boldsymbol{y}$ が母平均 $\boldsymbol{\mu}$ 母分散共分散行列 $\boldsymbol{\Sigma}$ を持つ p 変数の多変量正規確率変数の場合、その密度関数は

$$f(\boldsymbol{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\left(\sqrt{2\pi}\right)^p \sqrt{|\boldsymbol{\Sigma}|}} \exp\left(-\tfrac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right) \tag{498}$$

と書ける。ただし、$|\boldsymbol{\Sigma}|$ は母分散共分散行列の行列式、$\boldsymbol{\Sigma}^{-1}$ はその逆行列である。従って、独立なデータ $\boldsymbol{y}_i, i = 1, 2, \ldots, n$ が与えられたときの尤度は残差の２乗和の分解公式等を用いて

$$\begin{aligned}
L &= \prod_{i=1}^{n} f(\boldsymbol{y}_i \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) && (499)\\
&= \left(\frac{1}{\sqrt{|\boldsymbol{\Sigma}|}}\right)^n \exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}(\boldsymbol{y}_i-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu})\right) && (500)\\
&\propto \exp\left(-\tfrac{1}{2}\sum_{i=1}^{n}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu})(\boldsymbol{y}_i-\boldsymbol{\mu})'\right)\right) && (501)\\
&= \exp\left(-\tfrac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\textstyle\sum_{i=1}^{n}(\boldsymbol{y}_i-\boldsymbol{\mu})(\boldsymbol{y}_i-\boldsymbol{\mu})'\right)\right) && (502)\\
&= \exp\left(-\tfrac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\mathbf{1}\boldsymbol{\mu}')'(\boldsymbol{Y}-\mathbf{1}\boldsymbol{\mu}')\right)\right) && (503)\\
&= \exp\left(-\tfrac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\left(n(\boldsymbol{\mu}-\overline{\boldsymbol{Y}})(\boldsymbol{\mu}-\overline{\boldsymbol{Y}})' + (\boldsymbol{Y}-\mathbf{1}\overline{\boldsymbol{Y}}')'(\boldsymbol{Y}-\mathbf{1}\overline{\boldsymbol{Y}}')\right)\right)\right) && (504)\\
&= \exp\left(-\frac{n}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}-\overline{\boldsymbol{Y}})(\boldsymbol{\mu}-\overline{\boldsymbol{Y}})'\right)\right) \times \exp\left(-\frac{n}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{C}(\boldsymbol{Y})\right)\right) && (505)
\end{aligned}$$

標本平均と標本分散共分散行列を用いて書けることになる。

7.11 多項分布の確率

$U_k,\ k = 0, 1, 2, \ldots, m$ をたがいに背反な $m+1$ 個の事象 $A_k,\ k = 0, 1, 2, \ldots, m$ の生起を 1 または 0 で記録する確率変数とする。事象は互いに背反であるから、 縦ベクトル $\boldsymbol{U}$ の $m+1$ 個の要素のうち、どれか一つが 1 となり、残りは 0 である。また、それぞれの事象の生起確率を $m+1$ 個の要素を持つ縦ベクトル $\boldsymbol{\pi} = \{\pi_k\}$ で表すと、$\sum_{k=0}^{m}\pi_k = 1$ である。このと

$$\boldsymbol{\mathcal{E}}(U_k) = \pi_k,\ var\,(U_k) = \pi_k^2 - \pi_k,\ cov\,(U_k, U_j) = -\pi_k\pi_j \tag{506}$$

であるが、これを行列で表すと、平均ベクトルと分散共分散行列は

$$\boldsymbol{\mathcal{E}}(\boldsymbol{U}) = \boldsymbol{\pi},\ \boldsymbol{\mathcal{D}}(\boldsymbol{U}) = \mathrm{Diag}\,(\boldsymbol{\pi}) - \boldsymbol{\pi}\boldsymbol{\pi}' \tag{507}$$

と書ける。

8 その他の演算子

8.1 分割行列の演算

行列 $\boldsymbol{A}$ を分割し小行列 $\boldsymbol{A}_{ik}$ $r_{A_{ik}} \times c_{A_{ik}}, i = 1, 2, \ldots, r_A, k = 1, 2, \ldots, c_A$ を作る。同様に、行列 $\boldsymbol{B}$ を分割し小行列 $\boldsymbol{B}_{kj}$ $r_{B_{kj}} \times c_{B_{kj}}, k = 1, 2, \ldots, r_B, j = 1, 2, \ldots, c_B$ を作る。したがって、これらの小行列を要素と考えれば、行列 $\boldsymbol{A}$ は $r_A \times c_A$ の行列、行列 $\boldsymbol{B}$ は $r_B \times c_B$ の行列であると考えることが出来る。（ただしその要素がまた行列になっている。）このような行列のことを **分割行列** (partitioned matrix) と呼ぶ。この場合、もしも、$r_A = r_B, c_A = c_B$ でかつ全ての対応する小行列が同じサイズであれば、これらの二つの行列の間に和と差が定義され、それは、対応する小行列の和（または差）を要素として持つ分割行列となる。すなわち、

$$(\boldsymbol{A} + \boldsymbol{B})_{ik} = \boldsymbol{A}_{ik} + \boldsymbol{B}_{ik} \tag{508}$$

である。

また、$c_A = r_B$ であり、かつ、全ての組み合わせに関して $c_{A_{ik}} = r_{B_{kj}}$ であれば、行列 $\boldsymbol{A}$ と $\boldsymbol{B}$ の積は、通常の行列の積と同様に、以下のように定義する事ができる。

$$\boldsymbol{AB} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \ldots & \boldsymbol{A}_{1c_A} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \ldots & \boldsymbol{A}_{2c_A} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{i1} & \boldsymbol{A}_{i2} & \ldots & \boldsymbol{A}_{ic_A} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{r_A1} & \boldsymbol{A}_{r_A2} & \ldots & \boldsymbol{A}_{r_Ac_A} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \ldots & \boldsymbol{B}_{1j} & \ldots & \boldsymbol{B}_{1c_B} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \ldots & \boldsymbol{B}_{2j} & \ldots & \boldsymbol{B}_{2c_B} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{B}_{r_B1} & \boldsymbol{B}_{r_B2} & \ldots & \boldsymbol{B}_{r_Bj} & \ldots & \boldsymbol{B}_{r_Bc_B} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} & \ldots & \boldsymbol{C}_{1j} & \ldots & \boldsymbol{C}_{1c_C} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} & \ldots & \boldsymbol{C}_{2j} & \ldots & \boldsymbol{C}_{2c_C} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{C}_{i1} & \boldsymbol{C}_{i2} & \ldots & \boldsymbol{C}_{ij} & \ldots & \boldsymbol{C}_{ic_C} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{C}_{r_C1} & \boldsymbol{C}_{r_C2} & \ldots & \boldsymbol{C}_{r_Cj} & \ldots & \boldsymbol{C}_{r_Cc_C} \end{bmatrix} \tag{509}$$

ただし

$$(\boldsymbol{AB})_{ij} = \boldsymbol{C}_{ij} = \sum_{k=1}^{c_{A_{ik}}} \boldsymbol{A}_{ik}\boldsymbol{B}_{kj} \tag{510}$$

分割行列の演算規則で通常の場合と異なるものは以下の通りである。

分割行列の転置

$$\boldsymbol{A}' = \begin{bmatrix} \boldsymbol{A}'_{11} & \boldsymbol{A}'_{21} & \ldots & \boldsymbol{A}'_{r_A1} \\ \boldsymbol{A}'_{12} & \boldsymbol{A}'_{22} & \ldots & \boldsymbol{A}'_{r_A2} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}'_{1j} & \boldsymbol{A}'_{2j} & \ldots & \boldsymbol{A}'_{r_Aj} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}'_{1c_A} & \boldsymbol{A}'_{2c_A} & \ldots & \boldsymbol{A}'_{c_Ar_A} \end{bmatrix} \tag{511}$$

分割行列のトレース

$$\mathrm{tr}\,(\boldsymbol{A}) = \sum_{i=1}^{r} \mathrm{tr}\,(\boldsymbol{A}_{ii}) \tag{512}$$

分割行列の逆行列

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{A}_{11}{}^{-1} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} & -\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1} \\ -\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{22.1}{}^{-1} \end{bmatrix} \tag{513}$$

$$= \begin{bmatrix} \boldsymbol{A}_{11.2}{}^{-1} & -\boldsymbol{A}_{11.2}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1} \\ -\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11.2}{}^{-1} & \boldsymbol{A}_{22}{}^{-1} + \boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11.2}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1} \end{bmatrix} \tag{514}$$

ただし、残差積和行列 (resudual sum of squares and cros-product matrix) を

$$\boldsymbol{A}_{ii.j} = \boldsymbol{A}_{ii} - \boldsymbol{A}_{ij}\boldsymbol{A}_{jj}{}^{-1}\boldsymbol{A}_{ji} \tag{515}$$

と置いている。

分割行列の逆行列の公式の証明 (513) 式 を用いて $\boldsymbol{A}^{-1}\boldsymbol{A}$ と $\boldsymbol{A}\boldsymbol{A}^{-1}$ を計算し、それらが単位行列になることを示す。

$\boldsymbol{A}^{-1}\boldsymbol{A}$ の各要素：

$$\left[\boldsymbol{A}^{-1}\boldsymbol{A}\right]_{11} = \left(\boldsymbol{A}_{11}{}^{-1} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\right)\boldsymbol{A}_{11} - \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21} \tag{516}$$

$$= \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{11} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{11} - \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21} \tag{517}$$

$$= \boldsymbol{I} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21} - \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21} = \boldsymbol{I} \tag{518}$$

$$\left[\boldsymbol{A}^{-1}\boldsymbol{A}\right]_{12} = \left(\boldsymbol{A}_{11}{}^{-1} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\right)\boldsymbol{A}_{12} - \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{22} \tag{519}$$

$$= \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} - \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{22} \tag{520}$$

$$= \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\left(\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} - \boldsymbol{A}_{22}\right) \tag{521}$$

$$= \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} - \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{22.1} \tag{522}$$

$$= \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} - \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} = \boldsymbol{O} \tag{523}$$

$$\left[\boldsymbol{A}^{-1}\boldsymbol{A}\right]_{21} = \left(-\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\right)\boldsymbol{A}_{11} + \boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21} \tag{524}$$

$$= -\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{11} + \boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21} \tag{525}$$

$$= -\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21} + \boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21} = \boldsymbol{O} \tag{526}$$

$$\left[\boldsymbol{A}^{-1}\boldsymbol{A}\right]_{22} = \left(-\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\right)\boldsymbol{A}_{12} + \boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{22} \tag{527}$$

$$= -\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} + \boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{22} \tag{528}$$

$$= \boldsymbol{A}_{22.1}{}^{-1}\left(-\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} + \boldsymbol{A}_{22}\right) \tag{529}$$

$$= \boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{22.1} = \boldsymbol{I} \tag{530}$$

$\boldsymbol{A}\boldsymbol{A}^{-1}$ の各要素：

$$\left[\boldsymbol{A}\boldsymbol{A}^{-1}\right]_{11} = \boldsymbol{A}_{11}\left(\boldsymbol{A}_{11}{}^{-1} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\right) - \boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} \tag{531}$$

$$= \boldsymbol{A}_{11}\boldsymbol{A}_{11}{}^{-1} + \boldsymbol{A}_{11}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} - \boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} \tag{532}$$

$$= \boldsymbol{I} + \boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} - \boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} = \boldsymbol{I} \tag{533}$$

$$\left[\boldsymbol{A}\boldsymbol{A}^{-1}\right]_{12} = -\boldsymbol{A}_{11}\left(\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\right) + \boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1} \tag{534}$$

$$= -\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1} + \boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1} = \boldsymbol{O} \tag{535}$$

$$\left[\boldsymbol{A}\boldsymbol{A}^{-1}\right]_{21} = \boldsymbol{A}_{21}\left(\boldsymbol{A}_{11}{}^{-1} + \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\right) - \boldsymbol{A}_{22}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} \tag{536}$$

$$= \boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} + \boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} - \boldsymbol{A}_{22}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} \tag{537}$$

$$= \boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} + \left(\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} - \boldsymbol{A}_{22}\right)\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} \tag{538}$$

$$= \boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} + (-\boldsymbol{A}_{22.1})\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} \tag{539}$$

$$= \boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} = \boldsymbol{O} \tag{540}$$

$$\left[\boldsymbol{A}\boldsymbol{A}^{-1}\right]_{22} = -\boldsymbol{A}_{21}\left(\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\right) + \boldsymbol{A}_{22}\boldsymbol{A}_{22.1}{}^{-1} \tag{541}$$

$$= \left(-\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} + \boldsymbol{A}_{22}\right)\boldsymbol{A}_{22.1}{}^{-1} \tag{542}$$

$$= \boldsymbol{A}_{22.1}\boldsymbol{A}_{22.1}{}^{-1} = \boldsymbol{I} \tag{543}$$

証明終わり

(513) 式 に関しても同様に、 (515) 式 に相当するものを作る形で式を変形していけば証明することが出来る。

8.2 ssq 演算子

$ssq(\boldsymbol{A})$ で行列もしくはベクトル $\boldsymbol{A}$ の要素の２乗和を表す。

$$ssq(\boldsymbol{A}) = \mathrm{tr}\left[\boldsymbol{A}'\boldsymbol{A}\right] = \|\boldsymbol{A}\|^2 \quad \text{and} \quad ssq(\boldsymbol{a}) = \boldsymbol{a}'\boldsymbol{a} = \|\boldsymbol{a}\|^2 \tag{544}$$

例えば、回帰分析の残差の２乗和は

$$ssq(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}) = \mathrm{tr}\left[(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B})'(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B})\right] \tag{545}$$

また、重み付きの残差の２乗和は

$$ssq(\boldsymbol{W}^{\frac{1}{2}}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B})) = \mathrm{tr}\left((\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B})'\boldsymbol{W}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B})\right) \tag{546}$$

と書くことが出来る。また、この残差の２乗和のは以下の様に分解される（残差の２乗和の分解公式)。

$$(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B})'\boldsymbol{W}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{B}) = (\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{B}})'\boldsymbol{W}(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{B}}) + (\boldsymbol{B} - \widehat{\boldsymbol{B}})'\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X}(\boldsymbol{B} - \widehat{\boldsymbol{B}}) \tag{547}$$

ただし、$\widehat{\boldsymbol{B}} = (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y}$ である。

また、この演算子との関連でスカラーの場合の平方完成の公式

$$ay^2 + by + c = a\left(y + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a} \tag{548}$$

$$A(x-a)^2 + B(x-b)^2 = (A+B)\left(x - \frac{aA + bB}{A+B}\right)^2 + \frac{AB}{A+B}(a-b)^2 \tag{549}$$

のベクトル・行列版は、

$$\boldsymbol{x}'\boldsymbol{S}\boldsymbol{x} - 2\boldsymbol{c}'\boldsymbol{x} = (\boldsymbol{x}-\boldsymbol{d})'\boldsymbol{S}(\boldsymbol{x}-\boldsymbol{d}) - \boldsymbol{c}'\boldsymbol{S}\boldsymbol{c} \tag{550}$$

ただし、

$$\boldsymbol{d} = \boldsymbol{S}^{-1}\boldsymbol{c} \tag{551}$$

または

$$\mathrm{tr}\left(\boldsymbol{X}'\boldsymbol{S}\boldsymbol{X}\right) - 2\mathrm{tr}\left(\boldsymbol{C}'\boldsymbol{X}\right) = \mathrm{tr}\left((\boldsymbol{X}-\boldsymbol{D})'\boldsymbol{S}(\boldsymbol{X}-\boldsymbol{D})\right) - \mathrm{tr}\left(\boldsymbol{C}'\boldsymbol{S}\boldsymbol{C}\right) \tag{552}$$

ただし、

$$\boldsymbol{D} = \boldsymbol{S}^{-1}\boldsymbol{C} \tag{553}$$

となる。また、この平方完成の公式を使うと二つの $\boldsymbol{x}$ に関する２次形式を以下の様にまとめることが可能である。

$$\begin{aligned}(\boldsymbol{x}-\boldsymbol{a})'\boldsymbol{C}(\boldsymbol{x}-\boldsymbol{a}) + (\boldsymbol{x}-\boldsymbol{b})'\boldsymbol{D}(\boldsymbol{x}-\boldsymbol{b}) &= \left(\boldsymbol{x} - (\boldsymbol{C}+\boldsymbol{D})^{-1}(\boldsymbol{C}\boldsymbol{a}+\boldsymbol{D}\boldsymbol{b})\right)'(\boldsymbol{C}+\boldsymbol{D})\left(\boldsymbol{x} - (\boldsymbol{C}+\boldsymbol{D})^{-1}(\boldsymbol{C}\boldsymbol{a}+\boldsymbol{D}\boldsymbol{b})\right) \\ &\quad +(\boldsymbol{a}-\boldsymbol{b})'\boldsymbol{C}(\boldsymbol{C}+\boldsymbol{D})^{-1}\boldsymbol{D}(\boldsymbol{a}-\boldsymbol{b}) \qquad (554) \\ &= \left(\boldsymbol{x} - (\boldsymbol{C}+\boldsymbol{D})^{-1}(\boldsymbol{C}\boldsymbol{a}+\boldsymbol{D}\boldsymbol{b})\right)'(\boldsymbol{C}+\boldsymbol{D})\left(\boldsymbol{x} - (\boldsymbol{C}+\boldsymbol{D})^{-1}(\boldsymbol{C}\boldsymbol{a}+\boldsymbol{D}\boldsymbol{b})\right) \\ &\quad +(\boldsymbol{a}-\boldsymbol{b})'\left(\boldsymbol{C}^{-1}+\boldsymbol{D}^{-1}\right)^{-1}(\boldsymbol{a}-\boldsymbol{b}) + \qquad (555)\end{aligned}$$

なお、第２の形は $\boldsymbol{C}$ 行列と $\boldsymbol{D}$ 行列が正則であることを仮定し、後に示す、逆行列の公式を用いたものである。

同様に、残差の２乗和の場合は以下の様になる。

$$\begin{aligned}&(\boldsymbol{Y}_1 - \boldsymbol{X}_1\boldsymbol{B})'\boldsymbol{W}_1(\boldsymbol{Y}_1 - \boldsymbol{X}_1\boldsymbol{B}) + (\boldsymbol{Y}_2 - \boldsymbol{X}_2\boldsymbol{B})'\boldsymbol{W}_2(\boldsymbol{Y}_2 - \boldsymbol{X}_2\boldsymbol{B}) = \\ &(\boldsymbol{B} - \boldsymbol{A}^{-1}(\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{Y}1 + \boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{Y}_2))'\boldsymbol{A}(\boldsymbol{B} - \boldsymbol{A}^{-1}(\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{Y}1 + \boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{Y}_2)) \\ &\quad +(\boldsymbol{Y}_1 - \boldsymbol{X}_1\boldsymbol{B}_1)'\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{X}_1(\boldsymbol{Y}_1 - \boldsymbol{X}_1\boldsymbol{B}_1) \\ &\quad +(\boldsymbol{Y}_2 - \boldsymbol{X}_2\boldsymbol{B}_2)'\boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{X}_2(\boldsymbol{Y}_2 - \boldsymbol{X}_2\boldsymbol{B}_2) \\ &\quad +(\boldsymbol{B}_1 - \boldsymbol{B}_2)'\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{X}_1\boldsymbol{A}^{-1}\boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{X}_2(\boldsymbol{B}_1 - \boldsymbol{B}_2) \qquad (556) \\ &\quad = \\ &(\boldsymbol{B} - \boldsymbol{A}^{-1}(\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{Y}1 + \boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{Y}_2))'\boldsymbol{A}(\boldsymbol{B} - \boldsymbol{A}^{-1}(\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{Y}1 + \boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{Y}_2)) \\ &\quad +(\boldsymbol{Y}_1 - \boldsymbol{X}_1\boldsymbol{B}_1)'\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{X}_1(\boldsymbol{Y}_1 - \boldsymbol{X}_1\boldsymbol{B}_1) \\ &\quad +(\boldsymbol{Y}_2 - \boldsymbol{X}_2\boldsymbol{B}_2)'\boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{X}_2(\boldsymbol{Y}_2 - \boldsymbol{X}_2\boldsymbol{B}_2) \\ &\quad +(\boldsymbol{B}_1 - \boldsymbol{B}_2)'\boldsymbol{C}^{-1}(\boldsymbol{B}_1 - \boldsymbol{B}_2) \qquad (557)\end{aligned}$$

ただし、

$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{X}_1 + \boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{X}_2 \qquad (558) \\ \boldsymbol{B}_1 &= \left(\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{X}_1\right)^{-1}\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{Y}_1 \qquad (559) \\ \boldsymbol{B}_2 &= \left(\boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{X}_2\right)^{-1}\boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{Y}_2 \qquad (560) \\ \boldsymbol{C} &= \left(\boldsymbol{X}_1'\boldsymbol{W}_1\boldsymbol{X}_1\right)^{-1} + \left(\boldsymbol{X}_2'\boldsymbol{W}_2\boldsymbol{X}_2\right)^{-1} \qquad (561)\end{aligned}$$

である。

8.3 Kronecker 積

$\boldsymbol{A}$ を $ra \times ca$ の行列 $\boldsymbol{B}$ を $rb \times cb$ の行列とする。例えば、 $ra = 2, ca = 2, rb = 3, rc = 2$ の場合、

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \text{and} \quad \boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \tag{562}$$

であるが、$\boldsymbol{A}$ と $\boldsymbol{B}$ のクロネッカー積を以下の様な行列としてに定義する。

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} \end{bmatrix} \tag{563}$$

すなわち、$\odot$ 記号の左側にある（第 1 引数） $\boldsymbol{A}$ 行列の各要素に、右側にある（第 2 引数）$\boldsymbol{B}$ 行列を掛け合わせたものを、$\boldsymbol{A}$ 行列の形に配置したものである。この形は、 $ra \times ca$ の分割行列の形であるが、全体としてみると、以下の様な $(ra \times rb) \times (ca \times cb)$ の行列となる。

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} & a_{12}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \\ a_{21}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} & a_{22}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{31} & a_{12}b_{22} \\ a_{11}b_{31} & a_{11}b_{32} & a_{12}b_{31} & a_{12}b_{32} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \\ a_{21}b_{31} & a_{21}b_{32} & a_{22}b_{31} & a_{22}b_{32} \end{bmatrix} \tag{564}$$

クロネッカー積は (563) 式 の形（スカラー掛ける行列）から、通常の行列の積とは異なり、どのような行列やベクトルの間にも定義されるものである。なお、$\boldsymbol{A} \otimes \boldsymbol{B} \neq \boldsymbol{B} \otimes \boldsymbol{A}$ であり、

$$\boldsymbol{B} \otimes \boldsymbol{A} = \begin{bmatrix} b_{11}\boldsymbol{A} & b_{12}\boldsymbol{A} \\ b_{21}\boldsymbol{A} & b_{22}\boldsymbol{A} \\ b_{31}\boldsymbol{A} & b_{32}\boldsymbol{A} \end{bmatrix} \tag{565}$$

となる。

クロネッカー積は単位行列や単位ベクトルと組み合わせて利用されることが多く、$\boldsymbol{\phi} = [\phi_1, \phi_2, \phi_3]$ の時

$$\boldsymbol{I}_3 \odot \boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\phi} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\phi} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\phi} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\phi} \odot \boldsymbol{I} = \begin{bmatrix} \phi_1\boldsymbol{I}_3 \\ \phi_2\boldsymbol{I}_3 \\ \phi_3\boldsymbol{I}_3 \end{bmatrix} \tag{566}$$

$$\mathbf{1}_3 \otimes \boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\phi} \\ \boldsymbol{\phi} \\ \boldsymbol{\phi} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\phi} \otimes \mathbf{1}_3 = \begin{bmatrix} \phi_1\mathbf{1}_3 \\ \phi_2\mathbf{1}_3 \\ \phi_3\mathbf{1}_3 \end{bmatrix} \tag{567}$$

また、この拡張として、$\boldsymbol{A}$ 行列が対角行列の場合、$\boldsymbol{A} \otimes \boldsymbol{B}$ は $\boldsymbol{B}$ を対角要素に持つブロック対角行列となる。

クロネッカー積は以下の性質を持つ。

$$\boldsymbol{a} \otimes \boldsymbol{b}' = \boldsymbol{b}' \otimes \boldsymbol{a} = \boldsymbol{a}\boldsymbol{b}' \tag{568}$$

$$(\lambda \boldsymbol{A}) \otimes \boldsymbol{B} = \boldsymbol{A} \otimes (\lambda \boldsymbol{B}) = \lambda(\boldsymbol{A} \otimes \boldsymbol{B}) \tag{569}$$

$$(\boldsymbol{A} + \boldsymbol{B}) \otimes \boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{C} + \boldsymbol{B} \otimes \boldsymbol{C} \tag{570}$$

$$\boldsymbol{A} \otimes (\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A} \otimes \boldsymbol{B} + \boldsymbol{A} \otimes \boldsymbol{C} \tag{571}$$

$$\boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C}) = (\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C} \tag{572}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{A}\boldsymbol{C}) \otimes (\boldsymbol{B}\boldsymbol{C}) \tag{573}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B})' = \boldsymbol{A}' \otimes \boldsymbol{B}' \tag{574}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1} \tag{575}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B})^{-} = \boldsymbol{A}^{-} \otimes \boldsymbol{B}^{-} \tag{576}$$

$$\mathrm{tr}\,(\boldsymbol{A} \otimes \boldsymbol{B}) = \mathrm{tr}\,(\boldsymbol{A})\mathrm{tr}\,(\boldsymbol{B}) \tag{577}$$

$$\mathrm{rank}\,(\boldsymbol{A} \otimes \boldsymbol{B}) = \mathrm{rank}\,(\boldsymbol{A})\mathrm{tr}\,(\boldsymbol{B}) \tag{578}$$

$\boldsymbol{A}$ 行列が分割されている場合のクロネッカー積は通常のクロネッカー積の定義を以下の様に変更したものとなる。

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} \boldsymbol{A}_{11} \otimes \boldsymbol{B} & \boldsymbol{A}_{12} \otimes \boldsymbol{B} & \dots & \boldsymbol{A}_{1c_A} \otimes \boldsymbol{B} \\ \boldsymbol{A}_{21} \otimes \boldsymbol{B} & \boldsymbol{A}_{22} \otimes \boldsymbol{B} & \dots & \boldsymbol{A}_{2c_A} \otimes \boldsymbol{B} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{i1} \otimes \boldsymbol{B} & \boldsymbol{A}_{i2} \otimes \boldsymbol{B} & \dots & \boldsymbol{A}_{ic_A} \otimes \boldsymbol{B} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{r_A1} \otimes \boldsymbol{B} & \boldsymbol{A}_{r_A2} \otimes \boldsymbol{B} & \dots & \boldsymbol{A}_{r_Ac_A} \otimes \boldsymbol{B} \end{bmatrix} \tag{579}$$

8.4 vec 演算子

vec () 演算子は引数として与えられた行列の各列を縦に並べるものである。たとえば、

$$\boldsymbol{A} = [\boldsymbol{a}_{(1)}, \boldsymbol{a}_{(2)}, \boldsymbol{a}_{(3)}] = \begin{pmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \end{pmatrix} \tag{580}$$

の場合、

$$\mathrm{vec}\,(\boldsymbol{A}) = \begin{pmatrix} \boldsymbol{a}_{(1)} \\ \boldsymbol{a}_{(2)} \\ \boldsymbol{a}_{(3)} \end{pmatrix} \tag{581}$$

なる。 逆に、行方向に並べて横ベクトルを作りたい場合には

$$\mathrm{vec}\,\left(\boldsymbol{A}'\right)' = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{pmatrix}' \tag{582}$$

とすればよい。ただし、$\boldsymbol{a}$ が列ベクトルの場合は

$$\mathrm{vec}\,(\boldsymbol{a}) = \mathrm{vec}\,(\boldsymbol{a}') = \boldsymbol{a} \tag{583}$$

であり、またもう一つの列ベクトル $\boldsymbol{b}$ を用いて

$$\mathrm{vec}\,(\boldsymbol{b}\boldsymbol{a}') = \boldsymbol{a} \otimes \boldsymbol{b} \tag{584}$$

という関係がある。

$r \times c$ の行列 $\boldsymbol{A}$ から作られる二つのベクトル $\mathrm{vec}\,(\boldsymbol{A})$ と $\mathrm{vec}\left(\boldsymbol{A}'\right)$ との間には転換行列 (commutation matrix) $\boldsymbol{K}(r,c)$ を用いて

$$\begin{aligned} \mathrm{vec}\,(\boldsymbol{A}) &= \boldsymbol{K}(c,r)\mathrm{vec}\left(\boldsymbol{A}'\right) && (585)\\ \mathrm{vec}\left(\boldsymbol{A}'\right) &= \boldsymbol{K}(r,c)\mathrm{vec}\,(\boldsymbol{A}) && (586)\\ & && (587) \end{aligned}$$

という関係がある。たとえば、

$$\boldsymbol{K}(2,3) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{K}(3,2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{588}$$

であるが、見て分かる様に $\boldsymbol{K}(r,c) = \boldsymbol{K}(c,r)' = \boldsymbol{K}(c,r)^{-1}$ である。

vec ( ) 演算子は以下の性質を持つ。

$$\mathrm{vec}\,(\lambda\boldsymbol{A}) = \lambda\mathrm{vec}\,(\boldsymbol{A}) \quad \text{and} \quad \mathrm{vec}\,(\boldsymbol{A}+\boldsymbol{B}) = \mathrm{vec}\,(\boldsymbol{A}) + \mathrm{vec}\,(\boldsymbol{B}) \tag{589}$$

行列の積に vec () 演算子を作用させると以下の様になる。

$$\begin{aligned} \mathrm{vec}\,(\boldsymbol{ABC}) &= (\boldsymbol{C}'\otimes\boldsymbol{A})\mathrm{vec}\,(\boldsymbol{B}) && (590)\\ \mathrm{vec}\,(\boldsymbol{AB}) &= (\boldsymbol{I}\otimes\boldsymbol{A})\mathrm{vec}\,(\boldsymbol{B}) = (\boldsymbol{B}'\otimes\boldsymbol{A})\mathrm{vec}\,(\boldsymbol{I}) = (\boldsymbol{B}'\otimes\boldsymbol{I})\mathrm{vec}\,(\boldsymbol{A}) && (591) \end{aligned}$$

vec () 演算子を用いると行列の積のトレースを様々な形で書き表すことが出来る。

$$\begin{aligned} \mathrm{tr}\,(\boldsymbol{ABCD}) &= \mathrm{vec}\left(\boldsymbol{D}'\right)'(\boldsymbol{C}'\otimes\boldsymbol{A})\mathrm{vec}\,(\boldsymbol{B}) && (592)\\ &= \mathrm{vec}\,(\boldsymbol{D})'(\boldsymbol{A}\otimes\boldsymbol{C}')\mathrm{vec}\left(\boldsymbol{B}'\right) && (593) \end{aligned}$$

また、行列のトレースの性質を用いると、上記以外にも様々な形を得ることが出来る。行列が 3 つの場合には、上記のどれかを単位行列と於けば良いが、その一つの表現は以下のものとなる。

$$\mathrm{tr}\,(\boldsymbol{ABC}) = \mathrm{tr}\,(\boldsymbol{BCIA}) = \mathrm{vec}\left(\boldsymbol{A}'\right)'(\boldsymbol{I}\otimes\boldsymbol{B})\mathrm{vec}\,(\boldsymbol{C}) \tag{594}$$

二つの行列の積のトレース（トレースが取れるためには $\boldsymbol{A}$' と $\boldsymbol{B}$ のサイズが同じでなければならない）は対応する要素の２乗和であるから、

$$\mathrm{tr}\,(\boldsymbol{AB}) = \mathrm{vec}\left(\boldsymbol{A}'\right)'\mathrm{vec}\,(\boldsymbol{B}) \tag{595}$$

を得る。

8.5 逆行列の演算規則といくつかの公式

以下の掛け算が可能で、かつ全ての逆行列が取れる場合の公式。

1. $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}\boldsymbol{B}^{-1}$

2. $(\boldsymbol{A}+\boldsymbol{B})^{-1}=\boldsymbol{B}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}\boldsymbol{A}^{-1}$
3. $(\boldsymbol{A}+\boldsymbol{B})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}\boldsymbol{A}^{-1}$
4. $(\boldsymbol{A}+\boldsymbol{B})^{-1}=\boldsymbol{B}^{-1}-\boldsymbol{B}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}\boldsymbol{B}^{-1}$
5. $(\boldsymbol{A}+\boldsymbol{B})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1}$
6. $(\boldsymbol{A}+\boldsymbol{B})^{-1}=\boldsymbol{A}^{-1}-(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}$
7. $(\boldsymbol{A}+\boldsymbol{B})^{-1}=\boldsymbol{B}^{-1}-\boldsymbol{B}^{-1}\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}$
8. $(\boldsymbol{A}+\boldsymbol{B})^{-1}=\boldsymbol{B}^{-1}-(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A}\boldsymbol{B}^{-1}$
9. $(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}$
10. $(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1}\boldsymbol{B}=\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}^{-1}$
11. $\boldsymbol{D}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1}=\boldsymbol{C}^{-1}(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}$
12. $\boldsymbol{D}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1}\boldsymbol{B}=\boldsymbol{C}^{-1}-\boldsymbol{C}^{-1}(\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}^{-1}$

（註）
公式 1-4 は $\boldsymbol{A}$ と $\boldsymbol{B}$ を $\boldsymbol{A}^{-1}$ と $\boldsymbol{B}^{-1}$ で置き換える。
公式 5-8 は $(\boldsymbol{A}+\boldsymbol{B})^{-1}$ を残す。
公式 11-14 は $\boldsymbol{A}$ と $\boldsymbol{C}$ を $\boldsymbol{A}^{-1}$ と $\boldsymbol{C}^{-1}$ で置き換える。
公式 1-4 は、たとえば、

$$(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}=\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B} \tag{596}$$

という形で用いる場合が多い。
公式 5-8 は、 $\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}$ の様な形の式を単純化するときに利用されることが多い。

$$\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{I}-\boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1} \quad \text{formula5} \tag{597}$$
$$(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A} = \boldsymbol{I}-(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B} \quad \text{formula6} \tag{598}$$
$$\boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{I}-\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1} \quad \text{formula7} \tag{599}$$
$$(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B} = \boldsymbol{I}-(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A} \quad \text{formula8} \tag{600}$$

である。

9　スイープ演算子と逆行列ならびに行列式

9.0.1　スイープ演算子の定義

スウィープ演算子 (sweep operator) は Beaton(1964) により提案された演算子であり以下のように定義

される。$n \times n$ の行列 $\boldsymbol{A}$ が

$$\boldsymbol{A} \quad = \quad \left[\begin{array}{cccccc} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{ni} & \cdots & a_{nn} \end{array} \right] \begin{array}{l} \\ \\ \\ i\text{ 行} \\ \\ \\ \end{array} \qquad (601)$$

$$\begin{array}{c} \\ \\ \\ \\ \\ \\ i\text{ 列} \end{array}$$

として与えられている場合、行列 $\boldsymbol{A}$ の第 i,i 要素によるスイープを $\mathrm{Swp}\,[\boldsymbol{A} \mid i]$ と記し、それを以下のように定義する。

$$\mathrm{Swp}\,[\boldsymbol{A} \mid i] \quad = \quad \left[\begin{array}{cccccc} a_{11} - a_{1i}a_{i1}/a_{ii} & a_{12} - a_{1i}a_{i2}/a_{ii} & \cdots & -a_{1i}/a_{ii} & \cdots & a_{1n} - a_{1i}a_{in}/a_{ii} \\ a_{21} - a_{2i}a_{i1}/a_{ii} & a_{22} - a_{2i}a_{i2}/a_{ii} & \cdots & -a_{2i}/a_{ii} & \cdots & a_{2n} - a_{2i}a_{in}/a_{ii} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1}/a_{ii} & a_{i2}/a_{ii} & \cdots & 1/a_{ii} & \cdots & a_{in}/a_{ii} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} - a_{ni}a_{i1}/a_{ii} & a_{n2} - a_{ni}a_{i2}/a_{ii} & \cdots & -a_{ni}/a_{ii} & \cdots & a_{nn} - a_{ni}a_{in}/a_{ii} \end{array} \right] \begin{array}{l} \\ \\ \\ i\text{ 行} \\ \\ \\ \end{array}$$

$$\begin{array}{c} \\ \\ \\ \\ \\ \\ i\text{ 列} \end{array} \qquad (602)$$

すなわち行列 $\boldsymbol{A}$ の第 i 要素を自身の逆数で、対角を除く第 i 行を $a_{ij}/a_{ii},\ j \neq i$ で、第 i 列を $-a_{ij}/a_{ii},\ j \neq i$ で、それ以外の要素を $a_{jk} - a_{ji}a_{ik}/a_{ii},\ j \neq i, k \neq i$ で置き換えている。

このスイープ演算子は $a_{ii} \neq 0$ の場合、可逆的

$$\mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid i] \mid i] = \boldsymbol{A} \qquad (603)$$

であり、また、各スイープにおいて分母となる対角要素（ピボット）が 0 でない場合は対称的、すなわち、順序に依存しないこと

$$\mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid i] \mid j] = \mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid j] \mid i] \qquad (604)$$

が知られている。また、一つの要素ではなく、複数の要素に関するスイープを

$$\mathrm{Swp}\,[\boldsymbol{A} \mid \{i, j, \ldots, k\}] = \mathrm{Swp}\,[\cdots\ \mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid i] \mid j] \cdots \mid k] \qquad (605)$$

と定義すると、ピボットが 0 でない限り、$\mathrm{Swp}\,[\boldsymbol{A} \mid \{i, j, \ldots, k\}]$ はその順序に依存せずに、加算的である。

スイープ演算子を用いると、行列の逆行列を求めることが出来る。すなわち、$n \times n$ の行列の全ての対角要素に関してスイープを行った結果

$$\mathrm{Swp}\,[\boldsymbol{A} \mid \{1, 2, \ldots, n\}] = \mathrm{Swp}\,[\ldots \mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A} \mid 1] \mid 2] \ldots \mid n] = \boldsymbol{A}^{-1} \qquad (606)$$

は、ピボットが一度も 0 にならなければその行列の逆行列となる。また、全てのピボットを掛け合わせたものはその行列の行列式となることが知られている。たとえば、スカラーの行列式はそのスカラーそのもの、また、2×2 の行列の行列式は

$$\left| \left[\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right] \right| \quad = \quad a_{11} \times\ (a_{11}a_{22} - a_{12}a_{21})/a_{11} \ = \ a_{11}a_{22} - a_{21}a_{12} \qquad (607)$$

である。行列式が 0 となるような行列を特異行列と呼ぶ。

任意の正則な行列を $\boldsymbol{Q}$, $\boldsymbol{R}$ とすれば、

$$(\boldsymbol{QAR})^{-1} = \boldsymbol{R}^{-1}\boldsymbol{A}^{-1}\boldsymbol{Q}^{-1} \tag{608}$$

であるが、 $\boldsymbol{Q} = \boldsymbol{P}$, $\boldsymbol{R} = \boldsymbol{P}'$ として並べ替え行列 $\boldsymbol{P}$ を用いると、第 1-1 要素に 0 を持つ行列の逆行列も求めることが出来る。このような並べ替えを行っても、第 i 番目のピボットが 0 となる場合は、0 となるピボットに対応する列と行の要素をすべて 0 とおいて、計算を続けていくと、最終的に一般逆行列 (反射型一般逆行列) が得られる。なお、このような 0 のピボットを持つ行列のことを特異行列 (singular matrix) と呼び、その行列式は 0 となる、。

9.0.2 分割行列のスイープ

$n \times n$ の行列 $\boldsymbol{A}$ の分割を、

$$\boldsymbol{A} = \begin{array}{c} \left[\begin{array}{cccccc} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1i} & \cdots & \boldsymbol{A}_{1G} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2i} & \cdots & \boldsymbol{A}_{2G} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{i1} & \boldsymbol{A}_{i2} & \cdots & \boldsymbol{A}_{ii} & \cdots & \boldsymbol{A}_{iG} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{G1} & \boldsymbol{A}_{G2} & \cdots & \boldsymbol{A}_{Gi} & \cdots & \boldsymbol{A}_{GG} \end{array}\right] \begin{array}{l} \\ \\ \\ i \text{ 行} \\ \\ \\ \end{array} \\ \begin{array}{c} \\ \\ \\ \\ \\ \\ i \text{ 列} \end{array} \end{array} \tag{609}$$

とする。ただし、行列 $\boldsymbol{A}_{ij}, i = 1, 2, \ldots, G, j = 1, 2, \ldots, G$ は $n_i \times n_j$ の行列で $\sum_{i=1}^{G} n_i = n$ である。この場合、第 i 番目の対角ブロック $\boldsymbol{A}_{ii}$ によるスイープ $\mathrm{Swp}\,[\boldsymbol{A} \mid (i)]$ を

$$\mathrm{Swp}\,[\boldsymbol{A} \mid (i)] = \begin{array}{c} \left[\begin{array}{cccccc} \boldsymbol{A}_{11} - \boldsymbol{A}_{1i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i1} & \boldsymbol{A}_{12} - \boldsymbol{A}_{1i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i2} & \cdots & -\boldsymbol{A}_{1i}\boldsymbol{A}_{ii}^{-1} & \cdots & \boldsymbol{A}_{1G} - \boldsymbol{A}_{1i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{iG} \\ \boldsymbol{A}_{21} - \boldsymbol{A}_{2i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i1} & \boldsymbol{A}_{22} - \boldsymbol{A}_{2i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i2} & \cdots & -\boldsymbol{A}_{2i}\boldsymbol{A}_{ii}^{-1} & \cdots & \boldsymbol{A}_{2G} - \boldsymbol{A}_{2i}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{iG} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i1} & \boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i2} & \cdots & \boldsymbol{A}_{ii}^{-1} & \cdots & \boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{iG} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{A}_{G1} - \boldsymbol{A}_{Gi}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i1} & \boldsymbol{A}_{G2} - \boldsymbol{A}_{Gi}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{i2} & \cdots & -\boldsymbol{A}_{Gi}\boldsymbol{A}_{ii}^{-1} & \cdots & \boldsymbol{A}_{GG} - \boldsymbol{A}_{Gi}\boldsymbol{A}_{ii}^{-1}\boldsymbol{A}_{iG} \end{array}\right] \begin{array}{l} \\ \\ \\ i \text{ 行} \\ \\ \\ \end{array} \\ \begin{array}{c} \\ \\ \\ \\ \\ \\ i \text{ 列} \end{array} \end{array} \tag{610}$$

と定義するが、これは、第 i 番目の行列の全ての対角要素に関するスイープに等しい。

$$\mathrm{Swp}\,[\boldsymbol{A} \mid (i)] = \mathrm{Swp}\left[\boldsymbol{A} \mid \{\textstyle\sum_{j=1}^{i-1} n_j + 1, \sum_{j=1}^{i-1} n_j + 2, \ldots, \sum_{j=1}^{i} n_j\}\right] \tag{611}$$

たとえば、行列が

$$\boldsymbol{A} = \left[\begin{array}{cc} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{array}\right] \tag{612}$$

の様に 2 分割されている場合には

$$\mathrm{Swp}\,[\boldsymbol{A} \mid (1)] = \left[\begin{array}{cc} \boldsymbol{A}_{11}^{-1} & \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \end{array}\right] \tag{613}$$

$$\mathrm{Swp}\,[\boldsymbol{A} \mid (2)] = \left[\begin{array}{cc} \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} & -\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} \\ \boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} & \boldsymbol{A}_{22}^{-1} \end{array}\right] \tag{614}$$

であり、行列が 3 分割されている場合には

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (1)\right] = \begin{bmatrix} \boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} & \boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{13} \\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12} & \boldsymbol{A}_{23}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{13} \\ -\boldsymbol{A}_{31}\boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{32}-\boldsymbol{A}_{31}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{13} & \boldsymbol{A}_{33}-\boldsymbol{A}_{31}\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{13} \end{bmatrix} \tag{615}$$

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (2)\right] = \begin{bmatrix} \boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21} & -\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1} & \boldsymbol{A}_{13}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{23} \\ \boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21} & \boldsymbol{A}_{22}{}^{-1} & \boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{23} \\ \boldsymbol{A}_{31}-\boldsymbol{A}_{32}\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21} & -\boldsymbol{A}_{32}\boldsymbol{A}_{22}{}^{-1} & \boldsymbol{A}_{33}-\boldsymbol{A}_{32}\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{23} \end{bmatrix} \tag{616}$$

$$\mathrm{Swp}\left[\boldsymbol{A} \mid (3)\right] = \begin{bmatrix} \boldsymbol{A}_{11}-\boldsymbol{A}_{13}\boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{31} & \boldsymbol{A}_{12}-\boldsymbol{A}_{13}\boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{32} & -\boldsymbol{A}_{13}\boldsymbol{A}_{33}{}^{-1} \\ \boldsymbol{A}_{21}-\boldsymbol{A}_{23}\boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{31} & \boldsymbol{A}_{22}-\boldsymbol{A}_{23}\boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{32} & -\boldsymbol{A}_{23}\boldsymbol{A}_{33}{}^{-1} \\ \boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{31} & \boldsymbol{A}_{33}{}^{-1}\boldsymbol{A}_{32} & \boldsymbol{A}_{33}{}^{-1} \end{bmatrix} \tag{617}$$

となる。可逆性、対象性、加算性等の性質は分割行列の場合も同様であり、

$$\mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (i)\right] \mid (i)\right] = \boldsymbol{A} \tag{618}$$

$$\mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (i)\right] \mid (j)\right] = \mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (j)\right] \mid (i)\right] \tag{619}$$

$$\mathrm{Swp}\left[\boldsymbol{A} \mid \{(i),(j),\ldots,(k)\}\right] = \mathrm{Swp}\left[\cdots\ \mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (i)\right] \mid (j)\right]\cdots \mid (k)\right] \tag{620}$$

である。

9.0.3 スイープ演算子を利用した様々な公式

スイープ演算子の性質を用いて様々な行列演算の公式を導出することが出来る。例えば、2 分割の行列の逆行列を求めるためには

$$\boldsymbol{A}^{-1} = \mathrm{Swp}\left[\boldsymbol{A}|\{(1),(2)\}\ |=\right]\mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (1)\right] \mid (2)\right] = \mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (2)\right] \mid (1)\right] \tag{621}$$

という 3 通りの方法が考えられるが、この最後の二つのスイープの結果 $\mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (1)\right] \mid (2)\right]$ と $\mathrm{Swp}\left[\mathrm{Swp}\left[\boldsymbol{A} \mid (2)\right] \mid (1)\right]$ を書き下せば

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{A}_{11}{}^{-1}+\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} & -\boldsymbol{A}_{11}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22.1}{}^{-1} \\ -\boldsymbol{A}_{22.1}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}{}^{-1} & \boldsymbol{A}_{22.1}{}^{-1} \end{bmatrix} \tag{622}$$

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{A}_{11.2}{}^{-1} & -\boldsymbol{A}_{11.2}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1} \\ -\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11.2}{}^{-1} & \boldsymbol{A}_{22}{}^{-1}+\boldsymbol{A}_{22}{}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11.2}{}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}{}^{-1} \end{bmatrix} \tag{623}$$

となるが、これは、**分割行列の逆行列の公式**として知られているものである。ただし、残差積和行列 (resudual sum of squares and cros-product matrix) を

$$\boldsymbol{A}_{ii.j} = \boldsymbol{A}_{ii} - \boldsymbol{A}_{ij}\boldsymbol{A}_{jj}{}^{-1}\boldsymbol{A}_{ji} \tag{624}$$

とおいている。

(622) 式 と (623) 式 を比べることにより残差積和行列の逆行列

$$\boldsymbol{A}_{ii.j}{}^{-1} = \boldsymbol{A}_{ii}{}^{-1} + \boldsymbol{A}_{ii}{}^{-1}\boldsymbol{A}_{ij}\boldsymbol{A}_{jj.i}{}^{-1}\boldsymbol{A}_{ji}\boldsymbol{A}_{ii}{}^{-1} \tag{625}$$

を得る。また、残差積和行列の逆行列公式において、

$$\boldsymbol{A}_{ii} \to \boldsymbol{A},\ \boldsymbol{A}_{ij} \to \boldsymbol{B},\ -\boldsymbol{A}_{jj}{}^{-1} \to \boldsymbol{C},\ \boldsymbol{A}_{ji} \to \boldsymbol{D} \tag{626}$$

とおけば、Woodburry-Lindley-Smith の公式

$$(\boldsymbol{A}+\boldsymbol{BCD})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}\Big(\boldsymbol{C}^{-1}+\boldsymbol{DA}^{-1}\boldsymbol{B}\Big)^{-1}\boldsymbol{DA}^{-1} \tag{627}$$

を得る。また、$\boldsymbol{A}_{12}=\boldsymbol{O}$ とおくと **下三角行列の逆行列**

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{O} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} {\boldsymbol{A}_{11}}^{-1} & \boldsymbol{O} \\ -{\boldsymbol{A}_{22}}^{-1}\boldsymbol{A}_{21}{\boldsymbol{A}_{11}}^{-1} & {\boldsymbol{A}_{22}}^{-1} \end{bmatrix} \tag{628}$$

を得る。

また、スイープ演算子を用いて逆行列を求める際に現れるピボットとなる行列の行列式を掛け合わせたものが元の **行列式**となることが知られている。たとえば、(612) 式 で定義された分割行列の逆行列を (621) 式 の $\mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A}\mid(1)]\mid(2)]$ を用いて計算する際には、最初のピボットが $\boldsymbol{A}_{11}$, その結果として得られる行列が (613) 式 であるから、２番目のピボットは $\boldsymbol{A}_{22}-\boldsymbol{A}_{21}{\boldsymbol{A}_{11}}^{-1}\boldsymbol{A}_{12}$ となる。したがって、この行列の行列式は、

$$|\boldsymbol{A}| \;=\; |\boldsymbol{A}_{11}|\times\left|\boldsymbol{A}_{22}-\boldsymbol{A}_{21}{\boldsymbol{A}_{11}}^{-1}\boldsymbol{A}_{12}\right| \tag{629}$$

となる。同様に $\mathrm{Swp}\,[\mathrm{Swp}\,[\boldsymbol{A}\mid(2)]\mid(1)]$ を用いた場合は

$$|\boldsymbol{A}| \;=\; |\boldsymbol{A}_{22}|\times\left|\boldsymbol{A}_{11}-\boldsymbol{A}_{12}{\boldsymbol{A}_{22}}^{-1}\boldsymbol{A}_{21}\right| \tag{630}$$

となり、この二つは同値である。したがって、$r\times c$ の行列 $\boldsymbol{P}$ と $c\times r$ の行列 $\boldsymbol{Q}$ に関して以下の公式を得る。

$$|\boldsymbol{I}_r-\boldsymbol{PQ}| \;=\; |\boldsymbol{I}_c-\boldsymbol{QP}| \tag{631}$$

行列式に関しては以下の演算規則が成り立つ。ただし、全ての行列はそのサイズが r の正方行列とする。

$$|\boldsymbol{AB}| \;=\; |\boldsymbol{A}|\times|\boldsymbol{B}| \;=\; |\boldsymbol{BA}| \tag{632}$$

$$\left|\boldsymbol{A}^{-1}\right| \;=\; \frac{1}{|\boldsymbol{A}|} \tag{633}$$

$$|\lambda\boldsymbol{A}| \;=\; \lambda^r|\boldsymbol{A}| \qquad \text{and} \qquad |\boldsymbol{A}| \;=\; \left|\boldsymbol{A}'\right| \tag{634}$$

10　行列の固有値と特異値分解

n 次の正方行列 $\boldsymbol{A}$ に関して

$$\boldsymbol{A}\,\boldsymbol{p} = \lambda\,\boldsymbol{p} \tag{635}$$

を満たすスカラー λ を $\boldsymbol{A}$ の**固有値** または固有根 (eigen value, characteristic root) 、ベクトル $\boldsymbol{p}$ を λ に対応する**固有ベクトル** (eigen vector) と呼ぶ。

上式を変形すれば、

$$(\boldsymbol{A} - \lambda\boldsymbol{I})\,\boldsymbol{p} = \boldsymbol{0} \tag{636}$$

となるが、これは $\boldsymbol{p}$ に関する斉次方程式であり、$\boldsymbol{p} = 0$ 以外の解が存在するためには係数行列 $(\boldsymbol{A} - \lambda\boldsymbol{I})$ の行列式が 0

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = 0 \tag{637}$$

でなければならいが、これを λ の方程式とみたものを固有方程式と呼ぶ。例えば、$n = 2$ の場合

$$|\boldsymbol{A} - \lambda\boldsymbol{I}| = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12} * a_{21} \;=\; 0 \tag{638}$$

という λ に関する２次方程式となるが、一般的に、固有方程式は固有値の多項式として表される。

上記の定義から明らかなように固有ベクトルの長さは一意に定まらず、通常はそのノルムを１ $(\|\boldsymbol{p}\|^2 = 1)$ に規準化したものを用いる。一般に、n 次の正方行列の固有値は重根も含めれば n 個存在するが、それらのうち q 個を（大きさの順に並べ）対角行列の形に書いたものを $\boldsymbol{\Lambda}_q$, それに対応する固有ベクトルを横に並べた行列を $\boldsymbol{P}_q$ とすれば、上記の関係はまとめて

$$\boldsymbol{A}\boldsymbol{P}_q = \boldsymbol{P}_q\boldsymbol{\Lambda}_q \tag{639}$$

と書くことが出来る。

通常、多変量解析では、実数の要素を持つ対称行列の固有値問題のみを利用するため、以下の議論では $\boldsymbol{A}$ は対称行列とする。対称行列の固有値は全て実数である。対称行列の規準化された固有ベクトルのうち、ゼロでない異なる固有値に対応するものは互いに直交する。すなわち、 $\boldsymbol{\Lambda}_q$ が 0 を含まないならば、

$$\boldsymbol{P}_q{}'\boldsymbol{P}_q = \boldsymbol{I} \tag{640}$$

であるため、

$$\boldsymbol{P}'_q\boldsymbol{A}\boldsymbol{P}_q = \boldsymbol{\Lambda}_q \tag{641}$$

である。

証明

$$\boldsymbol{A}\boldsymbol{p}_1 = \lambda_1\boldsymbol{p}_1 \qquad \text{and} \qquad \boldsymbol{A}\boldsymbol{p}_2 = \lambda_2\boldsymbol{p}_2 \tag{642}$$

であるが、２番目の式の転置の右から $\boldsymbol{p}_1$ を掛けたものは以下のふた通りに変形出来る。

$$\boldsymbol{p}'_2\boldsymbol{A}\boldsymbol{p}_1 = \lambda_1\,\boldsymbol{p}'_2\boldsymbol{p}_1 \;= \lambda_2\,\boldsymbol{p}'_2\boldsymbol{p}_1 \tag{643}$$

したがって、

$$(\lambda_1 - \lambda_2)\boldsymbol{p}'_2\boldsymbol{p}_1 = 0 \tag{644}$$

を得るが、これは $\lambda_1 \neq \lambda_2$ の場合には $\boldsymbol{p}'_2\boldsymbol{p}_1 = 0$ を示している。　　**証明終わり**

10.1 固有値分解

対称行列 $\boldsymbol{A}$ が正則の場合、（$\boldsymbol{A}$ の固有値に 0 が含まれない場合） n 個の全ての固有ベクトルを並べたものを $\boldsymbol{P}$ とすると、

$$\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{I} \quad \text{and} \quad \boldsymbol{P}\boldsymbol{P}' = \boldsymbol{I} \tag{645}$$

である。したがって、

$$\boldsymbol{P}'\ \boldsymbol{A}\ \boldsymbol{P} = \boldsymbol{\Lambda} \tag{646}$$

であり、

$$\boldsymbol{A} \quad = \quad \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}' \tag{647}$$

と書けるが、これを $\boldsymbol{A}$ の固有値分解（スペクトル分解） (complete eigen decomposition, spectral decomposition) と呼ぶ。

すなわち、対称行列 $\boldsymbol{A}$ は、その固有ベクトルを左右から掛けることにより対角化することが可能となり、その階数はその行列のゼロでない固有値の個数に等しい。

これより、

$$\sum_{i=1}^{n} a_{ii} = \mathrm{tr}\,(\boldsymbol{A}) = \mathrm{tr}\left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'\right) = \mathrm{tr}\left(\boldsymbol{\Lambda}\boldsymbol{P}'\boldsymbol{P}\right) = \mathrm{tr}\,(\boldsymbol{\Lambda}) = \sum_{i=1}^{n} \lambda_i \tag{648}$$

すなわち、行列の対角要素の和は固有値の和に等しいことが分かる。また、行列の積の行列式は行列式の積であり、直交行列の行列式は 1 であるため、

$$|\boldsymbol{A}| = \left|\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'\right| = |\boldsymbol{P}||\boldsymbol{\Lambda}|\left|\boldsymbol{P}'\right| = |\boldsymbol{\Lambda}| = \prod_{i=1}^{n} \lambda_i \tag{649}$$

である。

行列の固有値分解を用いると、$\boldsymbol{A}$ の逆行列は

$$\boldsymbol{A}^{-1} \quad = \quad \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}' \tag{650}$$

と書ける。また、行列の冪乗は

$$\boldsymbol{A}^n = \boldsymbol{P}\ \boldsymbol{\Lambda}^n\ \boldsymbol{P}' \tag{651}$$

となる。すなわち、行列をべき乗したものの固有値は元の固有値をべき乗したものとなり、固有ベクトルは変わらない。

$\boldsymbol{A}$ の階数が $r, r < n$ の場合には、$n - r$ 個の固有値は 0 となる。いま、$\boldsymbol{\Lambda}$ の対角要素には r 個の固有値が大きい順に入っているとし、残りは 0 であるとする。またそれに対応して固有ベクトル行列 $\boldsymbol{P}$ を $\boldsymbol{P} = [\boldsymbol{P}_1, \boldsymbol{P}_2]$ と分割する。ただし、$\boldsymbol{P}_1$ は $\boldsymbol{P}_1'\boldsymbol{P}_1 = \boldsymbol{I}_r$ を満たす直交する固有ベクトル行列である。このとき、$\boldsymbol{P}_2$ は、$\boldsymbol{P}_2'\boldsymbol{P}_2 = \boldsymbol{I}_{n-r}$ ならびに $\boldsymbol{P}_2'\boldsymbol{P}_1 = \boldsymbol{O}$ を満たせば自由に選ぶことが出来る。

$\boldsymbol{A}$ の階数が r、すなわち、$\boldsymbol{A}$ のゼロでない固有値が全部で r 個しかない場合、それらを対角要素に持つ対角行列を $\boldsymbol{\Lambda}_r$ とし、それらに対応する固有ベクトルを列に持つ $n \times r$ の行列を $\boldsymbol{P}_r$ とすれば

$$\boldsymbol{A} = \boldsymbol{P}_r\boldsymbol{\Lambda}_r\boldsymbol{P}_r{}', \quad \boldsymbol{P}_r'\boldsymbol{P}_r = \boldsymbol{I}, \quad \boldsymbol{P}_r\boldsymbol{P}_r' \neq \boldsymbol{I} \tag{652}$$

と書くことが出来るが、これを対称行列 $\boldsymbol{A}$ のコンパクトな、もしくは、縮小された (reduced, compact) 固有値分解と呼ぶ。

階数 r の非負定値対称行列 $\boldsymbol{A}$ はそれを階数 r の列フルランク行列 $\boldsymbol{B}$ の積として

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{B}' \tag{653}$$

の形に分解することが出来る。ただし、

$$\boldsymbol{B} = \boldsymbol{P}_r \boldsymbol{\Lambda}_r^{\frac{1}{2}} \boldsymbol{T} \tag{654}$$

であり、$\boldsymbol{\Lambda}_r^{\frac{1}{2}}$ は $\boldsymbol{A}$ の正の固有値の平方根を要素として持つ行列、$\boldsymbol{P}_r$ はそれに対応する固有ベクトル、また、$\boldsymbol{T}$ は任意の直交行列である。なお、この分解は任意に $\boldsymbol{T}$ を選べるため一意ではない。また、

$$\boldsymbol{B} = \boldsymbol{P}_r \boldsymbol{\Lambda}_r^{\frac{1}{2}} \boldsymbol{P}_r' \tag{655}$$

で定義される行列は

$$\boldsymbol{B}\boldsymbol{B} \;=\; \boldsymbol{P}_r \boldsymbol{\Lambda}_r^{\frac{1}{2}} \boldsymbol{P}_r' \; \boldsymbol{P}_r \boldsymbol{\Lambda}_r^{\frac{1}{2}} \boldsymbol{P}_r' \;=\; \boldsymbol{P}_r \boldsymbol{\Lambda}_r^{\frac{1}{2}} \quad \boldsymbol{P}_r' \boldsymbol{P}_r \quad \boldsymbol{\Lambda}_r^{\frac{1}{2}} \boldsymbol{P}_r' \;=\; \boldsymbol{P}_r \boldsymbol{\Lambda}_r^{\frac{1}{2}} \boldsymbol{\Lambda}_r^{\frac{1}{2}} \boldsymbol{P}_r' \;=\; \boldsymbol{A} \tag{656}$$

となるため、行列 $\boldsymbol{A}$ の対称な平方根行列 (symmetric square root matrix) と呼ばれることがある。

$\boldsymbol{A}$ 行列の階数が $r < n$ の場合、 $\boldsymbol{A}$ の逆行列は存在しないが、

$$\boldsymbol{A}^{+} = \boldsymbol{P}_r {\boldsymbol{\Lambda}_r}^{-1} \boldsymbol{P}_r' \tag{657}$$

は

$$\boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{A} \;=\; \boldsymbol{P}_r \boldsymbol{\Lambda}_r \boldsymbol{P}_r' \quad \boldsymbol{P}_r {\boldsymbol{\Lambda}_r}^{-1} \boldsymbol{P}_r' \quad \boldsymbol{P}_r \boldsymbol{\Lambda}_r \boldsymbol{P}_r' \;=\; \boldsymbol{P}_r \boldsymbol{\Lambda}_r \boldsymbol{P}_r' \;=\; \boldsymbol{A} \tag{658}$$

$$\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{A}^{+} \;=\; \boldsymbol{P}_r {\boldsymbol{\Lambda}_r}^{-1} \boldsymbol{P}_r' \quad \boldsymbol{P}_r \boldsymbol{\Lambda}_r \boldsymbol{P}_r' \quad \boldsymbol{P}_r {\boldsymbol{\Lambda}_r}^{-1} \boldsymbol{P}_r' \;=\; \boldsymbol{P}_r {\boldsymbol{\Lambda}_r}^{-1} \boldsymbol{P}_r' \;=\; \boldsymbol{A}^{+} \tag{659}$$

等を示せるため、$\boldsymbol{A}$ のムーア・ペンローズ逆行列 (Moor-Penrose Generalized Inverse) になる。

なお、一般に、$\boldsymbol{P}_q, \boldsymbol{\Lambda}_q, q < r \leq n$ を用いた場合には

$$\boldsymbol{A} \neq \boldsymbol{P}_q \boldsymbol{\Lambda}_q \boldsymbol{P}_q' \tag{660}$$

であるが、これを切断された固有値分解 (truncated eigen decomposition) と呼ぶことがある。

また、この場合、右辺を $\widehat{\boldsymbol{A}}$ と置けば、それは $\boldsymbol{A}$ 行列の階数 q の最小２乗近似、すなわち

$$\mathrm{RSS} = \mathrm{tr}\left[(\boldsymbol{A} - \widehat{\boldsymbol{A}})'(\boldsymbol{A} - \widehat{\boldsymbol{A}})\right], \quad \mathrm{rank}\left(\widehat{\boldsymbol{A}}\right) = q \tag{661}$$

を最小とするものとなっている。

$\boldsymbol{0}$ では無いどのようなベクトル $\boldsymbol{x}$ を用いても２次形式が正

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} > 0 \tag{662}$$

であるような対称行列を **正定値行列** (または正値行列 positive definete matrix)（等号が付く場合は半正定値行列または非負定値行列, positive semidefinite matrix, nonnegative matrix）と呼ぶが、正定値行列の固有値は全て正、半正定値行列の固有値は全て非負である。特に、$\boldsymbol{A}'\boldsymbol{A}$ 行列や $\boldsymbol{A}\boldsymbol{A}'$ 行列は非負定値行列でありこれらの行列の正の固有値は等しい。

10.2 固有値分解の応用例

9 :

$\boldsymbol{A}$ が $n \times r,\ n \geq r$ の行列、 $\boldsymbol{D}$ が $r \times r$ の非負の対称行列の時

$$\boldsymbol{B} \quad = \quad \boldsymbol{AT} \tag{663}$$

で作られる行列 $\boldsymbol{B}$ が

$$\boldsymbol{B}'\boldsymbol{B} \quad = \quad \boldsymbol{D} \tag{664}$$

となるような $r \times r$ の $\boldsymbol{T}$ 行列を求めよ。

解

$\boldsymbol{A}'\boldsymbol{A}$ は半正定値の対称行列であるのでその固有値分解を

$$\boldsymbol{A}'\boldsymbol{A} \quad = \quad \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}', \quad \boldsymbol{P}'\boldsymbol{P} = \boldsymbol{I}_r \tag{665}$$

また、$\boldsymbol{D}$ の固有値分解を

$$\boldsymbol{D} \quad = \quad \boldsymbol{Q}\boldsymbol{\Delta}\boldsymbol{Q}', \quad \boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}_r \tag{666}$$

とすると、$\boldsymbol{R}$ を任意の $r \times r$ の直交行列として

$$\boldsymbol{T} \quad = \quad \boldsymbol{P}\,\boldsymbol{\Lambda}^{-\frac{1}{2}}\,\boldsymbol{R}\,\boldsymbol{\Delta}^{\frac{1}{2}}\,\boldsymbol{Q}' \tag{667}$$

が求める解である。（一意には定まらない。）

なぜなら、

$$\begin{aligned}
\boldsymbol{B}'\boldsymbol{B} \quad &= \quad \boldsymbol{T}'\boldsymbol{A}'\boldsymbol{A}\boldsymbol{T} \;=\; \boldsymbol{T}'\,\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'\,\boldsymbol{T} && (668)\\
&= \quad \boldsymbol{Q}\,\boldsymbol{\Delta}^{\frac{1}{2}}\boldsymbol{R}'\,\boldsymbol{\Lambda}^{-\frac{1}{2}}\,\boldsymbol{P}' \quad \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}' \quad \boldsymbol{P}\,\boldsymbol{\Lambda}^{-\frac{1}{2}}\,\boldsymbol{R}\,\boldsymbol{\Delta}^{\frac{1}{2}}\,\boldsymbol{Q}' && (669)\\
&= \quad \boldsymbol{Q}\,\boldsymbol{\Delta}^{\frac{1}{2}}\boldsymbol{R}' \quad \boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{-\frac{1}{2}} \quad \boldsymbol{R}\,\boldsymbol{\Delta}^{\frac{1}{2}}\,\boldsymbol{Q}' && (670)\\
&= \quad \boldsymbol{Q}\,\boldsymbol{\Delta}^{\frac{1}{2}}\boldsymbol{R}' \quad \boldsymbol{R}\,\boldsymbol{\Delta}^{\frac{1}{2}}\,\boldsymbol{Q}' && (671)\\
&= \quad \boldsymbol{Q}\,\boldsymbol{\Delta}^{\frac{1}{2}}\boldsymbol{\Delta}^{\frac{1}{2}}\,\boldsymbol{Q}' \;=\; \boldsymbol{Q}\,\boldsymbol{\Delta}\,\boldsymbol{Q}' \;=\; \boldsymbol{D} && (672)
\end{aligned}$$

である。

直交行列の自由な要素

$\boldsymbol{R}$ を $r \times r$ として

$$\boldsymbol{R}'\boldsymbol{R} \quad = \quad \boldsymbol{I} \tag{673}$$

は対角要素が全て 1、非対角要素が全て 0 という制約であるが、$\boldsymbol{R}'\boldsymbol{R}$ は対称であるため、実質的に 0 の制約は $r(r-1)/2$ 個である。したがって制約の数は $r + r(r-1)/2 = r(r+1)/2$ であり、自由に決められる要素の数は $r^2 - r(r+1)/2 = r(r-1)/2$ である。

$\boldsymbol{B}'\boldsymbol{B} = \boldsymbol{D}$ という制約条件は対称であるため実質的に $r(r+1)/2$ 個の制約である。しかし、求めるべき行列 $\boldsymbol{T}$ は $r \times r$ の行列であるため $r^2 \;-\; r(r+1)/2 = r(r-1)/2$ 個の自由な要素があると考えられるが、これが $\boldsymbol{R}$ の自由な要素に対応している。

$\boldsymbol{D} = \boldsymbol{B}'\boldsymbol{B} =$ 対角行列 という制約

この場合の制約条件の数は、対称行列の非対角要素が全て 0 というものであるから $r(r-1)/2$ であり、対角要素は自由に決められるという意味で、$\boldsymbol{B}'\boldsymbol{B} = \boldsymbol{D}$ よりも弱い。上記の解で $\boldsymbol{D}$ を対角行列とし、$\boldsymbol{Q} = \boldsymbol{I}, \boldsymbol{\Delta} = \boldsymbol{D}$ と置き、$\boldsymbol{D}$ の対角要素は未知とすればよいが、対角要素が未知数であるため、結局

$$\boldsymbol{T} = \boldsymbol{P}\,\boldsymbol{\Lambda}^{-\frac{1}{2}}\,\boldsymbol{R} \tag{674}$$

であれば良い。何故なら、$\boldsymbol{B}'\boldsymbol{B} = \boldsymbol{\Delta}^{\frac{1}{2}}\boldsymbol{R}'\quad \boldsymbol{R}\,\boldsymbol{\Delta}^{\frac{1}{2}}$ であり、$\boldsymbol{R}$ が $\boldsymbol{R}'\boldsymbol{R}$ = 対角行列 という $r(r-1)/2$ 個の制約を満たせばこの制約は満たされることになるからである。この条件は $\boldsymbol{R}$ が列直交であれば良いことを示している。

$\boldsymbol{T}$ が直交行列の場合

直交行列 $\boldsymbol{T}$ の自由な要素の数は $r(r-1)/2$ である。制約条件 $\boldsymbol{B}'\boldsymbol{B}$ = 対角行列 が持つ制約の数と一致するため、この条件で直交行列 $\boldsymbol{T}$ が一意に定まる。すなわち、 $\boldsymbol{B}'\boldsymbol{B} = \boldsymbol{D}$ の制約は強すぎることになる。

制約条件 $\boldsymbol{B}'\boldsymbol{B}$ = 対角行列 を満たす直交行列 $\boldsymbol{T}$ は

$$\boldsymbol{T} \;=\; \boldsymbol{P} \tag{675}$$

である。なぜなら、

$$\boldsymbol{B}'\boldsymbol{B} \;=\; \boldsymbol{P}'\boldsymbol{A}'\boldsymbol{A}\boldsymbol{P} \;=\; \boldsymbol{P}'\quad \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'\quad \boldsymbol{P} \;=\; \boldsymbol{\Lambda} \;=\; \text{対角行列} \tag{676}$$

となるからである。（厳密には、対角要素の順序の入れ替えという不定性が残っている。）

10.3 特異値分解

10.3.1 完全な特異値分解 full-size svd, complete svd

$p \leq n$ として、階数が $r \leq p$ の $n \times p$ の行列 $\boldsymbol{X}$ の特異値分解 (full-sie or complete SVD, Singular Value Decomposition) を

$$\boldsymbol{X} \quad = \quad \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}' \tag{677}$$

と表す。ただし、$\boldsymbol{U}$ は $n \times n$ の直交行列、$\boldsymbol{V}$ は $p \times p$ の直交行列、そして $\boldsymbol{D}$ は $n \times p$ の行列であり、その左上の $r \times r$ の正方部分の対角要素以外は 0 である。なお、$\boldsymbol{U}$ は左特異行列、$\boldsymbol{D}$ は特異値行列、$\boldsymbol{V}$ は右特異行列と呼ばれる。

このような行列 $\boldsymbol{U}, \boldsymbol{D}, \boldsymbol{V}$ は以下の様にして与えられる。すなわち、$\boldsymbol{U}$ は $\boldsymbol{X}\boldsymbol{X}'$ の直交する固有ベクトルからなる行列、$\boldsymbol{V}$ は $\boldsymbol{X}'\boldsymbol{X}$ の直交する固有ベクトルからなる行列、そして、 $\boldsymbol{D}$ はその左上の $r \times r$ の対角要素に $\boldsymbol{X}'\boldsymbol{X}$ (したがって $\boldsymbol{X}\boldsymbol{X}$'）の r 個の固有値の平方根を持つ行列である。

証明:

いま、 (677) 式 が成り立っているとすると、この式の両辺の左から $\boldsymbol{U}'$ を、また、右から $\boldsymbol{V}$ をかけることにより

$$\boldsymbol{U}'\boldsymbol{X} = \boldsymbol{D}\boldsymbol{V}' \tag{678}$$

$$\boldsymbol{X}\boldsymbol{V} = \boldsymbol{U}\boldsymbol{D} \tag{679}$$

を得る。そして、 (679) 式 の両辺の左から $\boldsymbol{X}'$ を掛け、 (678) 式 を使うと

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{V} = \boldsymbol{X}'\boldsymbol{U}\boldsymbol{D} \ = \boldsymbol{V}\boldsymbol{D}'\boldsymbol{D} \tag{680}$$

を得るが、$\boldsymbol{D}'\boldsymbol{D}$ はその左上の対角行列の対角部分に $d_j^2,\ j = 1, 2, \ldots, r$ を持つ $p \times p$ の対角行列となるため、これは $\boldsymbol{V}$ が $\boldsymbol{X}'\boldsymbol{X}$ の固有ベクトル、$\boldsymbol{D}'\boldsymbol{D}$ が固有値を対角要素に持つ対角行列であることを示している。($r = \mathrm{rank}\,(\boldsymbol{X}) = \mathrm{rank}\left(\boldsymbol{X}'\boldsymbol{X}\right) = \mathrm{rank}\left(\boldsymbol{X}\boldsymbol{X}'\right)$ であり、$\boldsymbol{X}'\boldsymbol{X}$ は半正定値行列であるから $\boldsymbol{X}'\boldsymbol{X}$ は r 個の正の固有値を持つ。)

同様に、 (678) 式 の両辺の右から $\boldsymbol{X}'$ を掛け、 (679) 式 を使うと

$$\boldsymbol{U}'\boldsymbol{X}\boldsymbol{X}' = \boldsymbol{D}\boldsymbol{V}'\boldsymbol{X}' \ = \boldsymbol{D}\boldsymbol{D}'\boldsymbol{U}' \tag{681}$$

を得るが、$\boldsymbol{D}\boldsymbol{D}'$ はその左上の対角行列の対角部分に $d_j^2,\ j = 1, 2, \ldots, r$ を持つ $n \times n$ の対角行列となるためこれは（両辺を転置することにより） $\boldsymbol{U}$ が $\boldsymbol{X}\boldsymbol{X}'$ の固有ベクトル、$\boldsymbol{D}\boldsymbol{D}'$ が固有値を対角要素に持つ対角行列であることを示している。

ここで、$\boldsymbol{X}'\boldsymbol{X}$ と $\boldsymbol{X}\boldsymbol{X}'$ は共に半正定値行列であり、同一の r 個の正の固有値 $\lambda_j,\ j = 1, 2, \ldots, r$ を持つため、$\boldsymbol{D}$ 行列の左上の $r \times r$ の部分の対角要素は $d_j = \sqrt{\lambda},\ j = 1, 2, \ldots, r$ となり、それを $\boldsymbol{X}$ 行列の特異値と呼ぶ。

逆に、$\boldsymbol{\Lambda}$ と $\boldsymbol{U}$ を $\boldsymbol{X}\boldsymbol{X}'$ の $r \times r$ の固有値行列、$n \times r$ の固有ベクトル行列、$\boldsymbol{V}$ を $p \times r$ の $\boldsymbol{X}'\boldsymbol{X}$ の固有ベクトル行列すなわち、

$$\boldsymbol{X}\boldsymbol{X}'\boldsymbol{U} = \boldsymbol{U}\boldsymbol{\Lambda}, \quad \boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}_r \tag{682}$$

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda}, \quad \boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_r \tag{683}$$

$$\tag{684}$$

とすれば (A.17) 式 より

$$\boldsymbol{X}'\boldsymbol{U} = \boldsymbol{V}\boldsymbol{\Lambda}^{\frac{1}{2}} \tag{685}$$

であるから (682) 式 は

$$\boldsymbol{X}\boldsymbol{X}'\boldsymbol{U} = \boldsymbol{X}\boldsymbol{V}\boldsymbol{\Lambda}^{\frac{1}{2}} = \boldsymbol{U}\boldsymbol{\Lambda} \tag{686}$$

となり、この右の等式から

$$\boldsymbol{X}\boldsymbol{V} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{-\frac{1}{2}} = \boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}} \tag{687}$$

を得る。

ここで、$r = p$ の場合はそのままで良いが、$r < p$ の場合には、0 の特異値に対応する特異ベクトルは直交性を保つ限り自由に設定できるので $\boldsymbol{V}$ を $p \times p$ の直交行列に、また、$\boldsymbol{U}$ を $n \times n$ の直交行列に拡張する。そして、$\boldsymbol{D}$ 行列を、その $r \times r$ の主対角部分に $\boldsymbol{\Lambda}$ の平方根を持つ $n \times p$ の行列とすれば、(687) 式 は

$$\boldsymbol{X}\boldsymbol{V} = \boldsymbol{U}\boldsymbol{D} \tag{688}$$

となり、この両辺の右から $\boldsymbol{V}'$ を掛ければ (677) 式 を得る。

別解として、以下の様に考えることも出来る。

ここで、(687) 式 の左から $\boldsymbol{X}'$ を右から $\boldsymbol{V}'$ を掛け、

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda} \quad \text{したがって} \quad \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}' = \boldsymbol{X}'\boldsymbol{X} \tag{689}$$

の関係を使うと

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{V}\boldsymbol{V}' = \boldsymbol{X}'\boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{V}' \tag{690}$$

$$\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}' = \boldsymbol{X}'\boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{V}' \tag{691}$$

$$\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}'\boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{V}' \tag{692}$$

という関係を得る。そして、この両辺の左から、$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}$ を掛け、$\boldsymbol{U}$ は $\boldsymbol{X}$ の列空間に存在することに注意すると

$$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\ \boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\ \boldsymbol{X}'\boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{V}' \tag{693}$$

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{V}' \tag{694}$$

を得る。この表現は次節で言うコンパクトな特異値分解であるから、それぞれの行列を引き延ばせば (677) 式 を得る。 **証明終わり**

10.3.2 サブコンパクトな特異値分解

階数 r が $r < \min(n, p)$ の場合、$\boldsymbol{D}$ 行列の左上の $r \times r$ の対角行列以外の部分は全て 0 となる。今ここで $\boldsymbol{D}$ 行列をその $\min(n, p) \times \min(n, p)$ の部分だけ残し、$\boldsymbol{U}$ と $\boldsymbol{V}$ も同様に $n \times \min(n, p)$ ならびに $p \times \min(n, p)$ の行列としても (677) 式 は成立する。

この場合、$p < n$ の場合には $\boldsymbol{U}\boldsymbol{U}' \neq \boldsymbol{I}$ となり $n < p$ の場合には $\boldsymbol{V}\boldsymbol{V}' \neq \boldsymbol{I}$ となる。

これは R の svd 関数のデフォールトの仕様で、ここではサブコンパクトな特異値分解と呼んでおく。

(MATLAB では economy-sized svd と呼ばれているらしい。)

10.3.3 コンパクトな特異値分解 compact, economy, or thin svd

階数 r が $r < \min(n,p)$ の場合、$\boldsymbol{D}$ 行列の左上の $r \times r$ の対角行列以外の部分は全て 0 となる。したがって、 $\boldsymbol{D}$ 行列をこの $r \times r$ の部分だけで再定義し、$\boldsymbol{U}$ 行列の列数と $\boldsymbol{V}$ 行列の行数を r としても (677) 式 は成立する。

今、これらの行列を、添字 r を付けた形で書くと (677) 式 は

$$\boldsymbol{X} \quad = \boldsymbol{U}_r \boldsymbol{D}_r \boldsymbol{V}_r' \tag{695}$$

となる。ただし、$\boldsymbol{V}_r$ は $p \times r$ の行列、$\boldsymbol{U}_r$ は $n \times r$ の行列、$\boldsymbol{D}_r$ は $r \times r$ の対角行列である。これを行列 $\boldsymbol{X}$ のコンパクトな特異値分解もしくは「縮小された特異値分解 (compact SVD, Singular Value Decomposition) と呼ぶ。

完全な特異値分解との違いは $\boldsymbol{U}_r$ と $\boldsymbol{V}_r$ 行列が直交行列ではなく、

$$\boldsymbol{U}_r' \boldsymbol{U}_r = \boldsymbol{V}_r' \boldsymbol{V}_r = \boldsymbol{I}_r \tag{696}$$

のみを満たす列直交行列となるところである。

10.3.4 切断された特異値分解 compact, economy, or thin svd

q 等として、$q, q \leq r$ のものを用いた場合、

$$\hat{\boldsymbol{X}} = \boldsymbol{U}_q \boldsymbol{D}_q \boldsymbol{V}_q' \tag{697}$$

は

$$\mathrm{RSS} = \sum_{i=1}^{n} \sum_{j=1}^{p} (x_{ij} - \hat{x}_{ij})^2 \tag{698}$$

を最小とする階数 q の行列となることが知られている。これを行列 $\boldsymbol{X}$ の切断された特異値分解 (truncated svd) もしくは Eckart-Young Decomposition と呼ぶ。

また、これを

$$\widehat{\boldsymbol{X}} \quad = \quad \boldsymbol{F}\, \boldsymbol{A}' \tag{699}$$

ただし、

$$\boldsymbol{F} = \sqrt{n}\, \boldsymbol{U}_q \qquad \text{and} \qquad \boldsymbol{A} = \frac{1}{\sqrt{n}}\, \boldsymbol{V}_q \boldsymbol{D}_q \tag{700}$$

と書くことも可能であるが、これは主成分分析に他ならない。

11 行列の微分

いま、$\boldsymbol{X} = [x_{ij}]$ を、$a \times b$ の行列とし、$\boldsymbol{U}$ や $\boldsymbol{V}$ の要素が $u_{ij} = g_{ij}(x_{ij}), v_{ij} = h_{ij}(x_{ij})$ という形で $\boldsymbol{X}$ の要素の関数となっているとする。

そして、f を行列からスカラーへの関数とした場合に、f を行列 $\boldsymbol{X}$ の各要素で偏微分し、その行列の形に並べた $a \times b$ の行列を、

$$\frac{\partial\, f}{\partial\, \boldsymbol{X}} = \left[\frac{\partial\, f}{\partial\, x_{ij}} \right] \tag{701}$$

と書くことにする。

スカラー関数の行列に関する微分の基本公式の主なものは以下の４つである。

１）転置行列による微分

$$\frac{\partial f}{\partial \boldsymbol{X}'} = \left(\frac{\partial f}{\partial \boldsymbol{X}}\right)' \tag{702}$$

転置行列で微分したものはもとの行列で微分したものを転置したものである。

２）対角行列による微分

$$\frac{\partial f}{\partial \boldsymbol{X}} = \mathrm{diag}\left(\frac{\partial f}{\partial \boldsymbol{X}}\right) \quad \text{ただし } \boldsymbol{X} \text{ は対角行列} \tag{703}$$

これは対角行列であることを気にせずに微分し、その結果を対角化すればよいことを示している。

３）トレースの微分に関する乗法公式

$$\frac{\partial \,\mathrm{tr}\,(\boldsymbol{UV})}{\partial \boldsymbol{X}} = \frac{\partial \,\mathrm{tr}\,(\boldsymbol{UV}_c)}{\partial \boldsymbol{X}} + \frac{\partial \,\mathrm{tr}\,(\boldsymbol{U}_c\boldsymbol{V})}{\partial \boldsymbol{X}} \tag{704}$$

ただし、$\frac{\partial \,\mathrm{tr}\,(\boldsymbol{UV}_c)}{\partial \boldsymbol{X}}$ は $\boldsymbol{V}$ を $\boldsymbol{X}$ の関数ではなくあたかも定数のようにして微分することを表している。

４）Chain Rule

$$\frac{\partial f}{\partial \boldsymbol{X}} = \frac{\partial \,\mathrm{tr}\left(\left[\frac{\partial f}{\partial \boldsymbol{U}}\right]_c\right)'\boldsymbol{U}}{\partial \boldsymbol{X}} \tag{705}$$

ただし、$\left[\frac{\partial f}{\partial \boldsymbol{U}}\right]_c$ は、f を $\boldsymbol{U}$ の要素で微分したものを、$\boldsymbol{X}$ を含まない定数行列のように取り扱うことを示している。

以上の基本公式を用いれば以下のような公式を導くことができる。いずれの場合にも、$\boldsymbol{X}$ 行列が対称行列の場合には、対角要素と非対角要素を分けて考えなければならない。なお、$\boldsymbol{A}$ および $\boldsymbol{B}$ は $\boldsymbol{X}$ を含まない定数行列であり、そのサイズは、以下に示す乗算が可能なものとする。

$$\begin{aligned}
\frac{\partial \,\mathrm{tr}\,(\boldsymbol{AX})}{\partial \boldsymbol{X}} &= \boldsymbol{A}', \quad \boldsymbol{X} \text{ が非対称の場合} && (706)\\
&= \boldsymbol{A} + \boldsymbol{A}' - \mathrm{diag}\,(\boldsymbol{A}), \quad \boldsymbol{X} \text{ が対称の場合} && (707)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \,\mathrm{tr}\,(\boldsymbol{AX}')}{\partial \boldsymbol{X}} &= \boldsymbol{A}, \quad \boldsymbol{X} \text{ が非対称の場合} && (708)\\
&= \boldsymbol{A} + \boldsymbol{A}' - \mathrm{diag}\,(\boldsymbol{A}), \quad \boldsymbol{X} \text{ が対称の場合} && (709)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \,\mathrm{tr}\,(\boldsymbol{X}'\boldsymbol{AX})}{\partial \boldsymbol{X}} &= (\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{X}, \quad \boldsymbol{X} \text{ が非対称の場合} && (710)\\
&= (\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{X} + \boldsymbol{X}(\boldsymbol{A} + \boldsymbol{A}') - \mathrm{diag}\,((\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{X}), \quad \boldsymbol{X} \text{ が対称の場合} && (711)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \,\mathrm{tr}\,(\boldsymbol{XAX}')}{\partial \boldsymbol{X}} &= \boldsymbol{X}(\boldsymbol{A} + \boldsymbol{A}'), \quad \boldsymbol{X} \text{ が非対称の場合} && (712)\\
&= (\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{X} + \boldsymbol{X}(\boldsymbol{A} + \boldsymbol{A}') - \mathrm{diag}\,((\boldsymbol{A} + \boldsymbol{A}')\boldsymbol{X}), \quad \boldsymbol{X} \text{ が対称の場合} && (713)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \,\mathrm{tr}\,(\boldsymbol{X}'\boldsymbol{AXB})}{\partial \boldsymbol{X}} &= \boldsymbol{AXB} + \boldsymbol{A}'\boldsymbol{XB}', \quad \boldsymbol{X} \text{ が非対称の場合} && (714)\\
&= \boldsymbol{AXB} + \boldsymbol{BXA} + \boldsymbol{A}'\boldsymbol{XB}' + \boldsymbol{B}'\boldsymbol{XA}' - \mathrm{diag}\,(\boldsymbol{AXB}) - \mathrm{diag}\,(\boldsymbol{BXA}), && (715)\\
&\quad \boldsymbol{X} \text{ が対称の場合} && (716)
\end{aligned}$$

$$\frac{\partial \,\mathrm{tr}\,(\boldsymbol{XAXB})}{\partial\,\boldsymbol{X}} = \boldsymbol{B'X'A'} + \boldsymbol{A'X'B'}, \quad \boldsymbol{X} \text{ が非対称の場合} \tag{717}$$

$$= \boldsymbol{B'X'A'} + \boldsymbol{A'X'B'} + \boldsymbol{BXA} + \boldsymbol{AXB} - \mathrm{diag}\,(\boldsymbol{BXA}) - \mathrm{diag}\,(\boldsymbol{AXB}), \tag{718}$$

$$\boldsymbol{X} \text{ が対称の場合} \tag{719}$$

$$\frac{\partial \,\mathrm{tr}\left(\boldsymbol{X}^{-1}\boldsymbol{A}\right)}{\partial\,\boldsymbol{X}} = -(\boldsymbol{X}^{-1}\boldsymbol{AX}^{-1})', \quad \boldsymbol{X} \text{ が非対称の場合} \tag{720}$$

$$= -(\boldsymbol{X}^{-1}\boldsymbol{AX}^{-1})' - \boldsymbol{X}^{-1}\boldsymbol{AX}^{-1} + \mathrm{diag}\left(\boldsymbol{X}^{-1}\boldsymbol{A'X}^{-1}\right), \quad \boldsymbol{X} \text{ が対称の場合} \tag{721}$$

$$\frac{\partial\,|\boldsymbol{X}|^a}{\partial\,\boldsymbol{X}} = a|\boldsymbol{X}|^a(\boldsymbol{X}^{-1})', \quad \boldsymbol{X} \text{ が非対称の場合} \tag{722}$$

$$= a|\boldsymbol{X}|^a\left(2\boldsymbol{X}^{-1} - \mathrm{diag}\left(\boldsymbol{X}^{-1}\right)\right), \quad \boldsymbol{X} \text{ が対称の場合} \tag{723}$$

$$\frac{\partial\,|\boldsymbol{XX'}|}{\partial\,\boldsymbol{X}} = 2|\boldsymbol{XX'}|(\boldsymbol{XX'})^{-1}\boldsymbol{X}, \quad \boldsymbol{X} \text{ が非対称の場合} \tag{724}$$

$$= 2|\boldsymbol{XX'}|\left((\boldsymbol{XX'})^{-1}\boldsymbol{X} + ((\boldsymbol{XX'})^{-1}\boldsymbol{X})' - \mathrm{diag}\left((\boldsymbol{XX'})^{-1}\boldsymbol{X}\right)\right), \quad \boldsymbol{X} \text{ が対称の場合} \tag{725}$$

$$\frac{\partial\,|\boldsymbol{X'X}|}{\partial\,\boldsymbol{X}} = 2|\boldsymbol{X'X}|\boldsymbol{X}(\boldsymbol{X'X})^{-1}, \quad \boldsymbol{X} \text{ が非対称の場合} \tag{726}$$

$$= 2|\boldsymbol{X'X}|\left(\boldsymbol{X}(\boldsymbol{X'X})^{-1} + (\boldsymbol{X}(\boldsymbol{X'X})^{-1})' - \mathrm{diag}\left(\boldsymbol{X}(\boldsymbol{X'X})^{-1}\right)\right), \quad \boldsymbol{X} \text{ が対称の場合} \tag{727}$$

$$\frac{\partial\,|\boldsymbol{AXB}|}{\partial\,\boldsymbol{X}} = |\boldsymbol{AXB}|\boldsymbol{A'}(\boldsymbol{B'X'A'})^{-1}\boldsymbol{B'}, \quad \boldsymbol{X} \text{ が非対称の場合} \tag{728}$$

$$= |\boldsymbol{AXB}|\left(\boldsymbol{A'}(\boldsymbol{B'X'A'})^{-1}\boldsymbol{B'} + \boldsymbol{B}(\boldsymbol{AXB})^{-1}\boldsymbol{A} - \mathrm{diag}\left(\boldsymbol{B}(\boldsymbol{AXB})^{-1}\boldsymbol{A}\right)\right), \quad \boldsymbol{X} \text{ が対称の場合} \tag{729}$$

$$\frac{\partial\,\ln|\boldsymbol{X}|}{\partial\,\boldsymbol{X}} = (\boldsymbol{X}^{-1})', \quad \boldsymbol{X} \text{ が非対称の場合} \tag{730}$$

$$= 2\boldsymbol{X}^{-1} - \mathrm{diag}\left(\boldsymbol{X}^{-1}\right), \quad \boldsymbol{X} \text{ が対称の場合} \tag{731}$$

なお、$\frac{\partial\,\boldsymbol{X'A}}{\partial\,\boldsymbol{X}}$ や $\frac{\partial\,\boldsymbol{AXX'}}{\partial\,\boldsymbol{X}}$ のようなものは、さきに示したトレースの性質を用いれば、上記の公式の変形として取り扱うことができる。また、chain rule を使えば、たとえば、$\boldsymbol{Y}$ が $\boldsymbol{X}$ の関数である場合、

$$\frac{\partial \,\mathrm{tr}\left(\boldsymbol{Y}^{-1}\boldsymbol{A}\right)}{\partial\,\boldsymbol{X}} = -\frac{\partial \,\mathrm{tr}\left(\left[\boldsymbol{Y}^{-1}\boldsymbol{A'Y}^{-1}\right]_c\boldsymbol{Y}\right)}{\partial\,\boldsymbol{X}} \tag{732}$$

また、

$$\frac{\partial\,\ln|\boldsymbol{Y}|}{\partial\,\boldsymbol{X}} = \frac{\partial \,\mathrm{tr}\left(\left(\boldsymbol{Y}^{-1}\right)_c\boldsymbol{Y}\right)}{\partial\,\boldsymbol{X}} \tag{733}$$

である。

12　線形モデルと最小２乗解

12.1　線形モデルと最小２乗解

$n \times p$ の基準変数行列 Y, $n \times q$ の説明変数行列 $\boldsymbol{X}$, $q \times p$ の係数行列 $\boldsymbol{B}$, $n \times p$ の誤差行列 $\boldsymbol{E}$ を用いて作られる線形モデル

$$\boldsymbol{Y} = \widehat{\boldsymbol{Y}} + \boldsymbol{E} \qquad \text{ただし} \qquad \widehat{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{\beta}' \tag{734}$$

$$\boldsymbol{y}_{(j)} = \widehat{\boldsymbol{y}}_{(j)} + \boldsymbol{\varepsilon}_{(j)} \qquad \text{ただし} \qquad \widehat{\boldsymbol{y}}_{(j)} = \boldsymbol{X}\boldsymbol{\beta}_j \tag{735}$$

$$\boldsymbol{y}_i = \widehat{\boldsymbol{y}}_i + \boldsymbol{\varepsilon}_i \qquad \text{ただし} \qquad \widehat{\boldsymbol{y}}_i = \boldsymbol{\beta}\boldsymbol{x}_i \tag{736}$$

$$y_{ij} = \widehat{y}_{ij} + \varepsilon_{ij} \qquad \text{ただし} \qquad \widehat{y}_{ij} = \sum_{k=1}^{q} x_{ik}\beta_{jk} \tag{737}$$

を考える。

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1q} \\ x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2q} \\ & & \vdots & & & \\ x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{iq} \\ & & \vdots & & & \\ x_{n1} & x_{n2} & \dots & x_{nj} & \dots & x_{nq} \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}'_1 \\ \boldsymbol{x}'_2 \\ \vdots \\ \boldsymbol{x}'_i \\ \vdots \\ \boldsymbol{x}'_q \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_{(1)}, & \boldsymbol{x}_{(2)}, & \dots, & \boldsymbol{x}_{(n)} \end{pmatrix} \tag{738}$$

$$\boldsymbol{Y} = \begin{pmatrix} y_{11} & y_{12} & \dots & y_{1j} & \dots & y_{1p} \\ y_{21} & y_{22} & \dots & y_{2j} & \dots & y_{2p} \\ & & \vdots & & & \\ y_{i1} & y_{i2} & \dots & y_{ij} & \dots & y_{ip} \\ & & \vdots & & & \\ y_{n1} & y_{n2} & \dots & y_{nj} & \dots & y_{np} \end{pmatrix} = \begin{pmatrix} \boldsymbol{y}'_1 \\ \boldsymbol{y}'_2 \\ \vdots \\ \boldsymbol{y}'_i \\ \vdots \\ \boldsymbol{y}'_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{y}_{(1)}, & \boldsymbol{y}_{(2)}, & \dots, & \boldsymbol{y}_{(n)} \end{pmatrix} \tag{739}$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1k} & \dots & \beta_{1q} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2k} & \dots & \beta_{2q} \\ & & \vdots & & & \\ \beta_{j1} & \beta_{j2} & \dots & \beta_{jk} & \dots & \beta_{jq} \\ & & \vdots & & & \\ \beta_{p1} & \beta_{p2} & \dots & \beta_{pk} & \dots & \beta_{pq} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}'_1 \\ \boldsymbol{\beta}'_2 \\ \vdots \\ \boldsymbol{\beta}'_j \\ \vdots \\ \boldsymbol{\beta}'_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}_{(1)}, & \boldsymbol{\beta}_{(2)}, & \dots, & \boldsymbol{\beta}_{(q)} \end{pmatrix} \tag{740}$$

$$\boldsymbol{\beta}' = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1j} & \dots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2j} & \dots & \beta_{2p} \\ & & \vdots & & & \\ \beta_{k1} & \beta_{k2} & \dots & \beta_{kj} & \dots & \beta_{kp} \\ & & \vdots & & & \\ \beta_{q1} & \beta_{q2} & \dots & \beta_{qj} & \dots & \beta_{qp} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}'_{(1)} \\ \boldsymbol{\beta}'_{(2)} \\ \vdots \\ \boldsymbol{\beta}'_{(k)} \\ \vdots \\ \boldsymbol{\beta}'_{(q)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}_1, & \boldsymbol{\beta}_2, & \dots, & \boldsymbol{\beta}_p \end{pmatrix} \tag{741}$$

データ行列 $\boldsymbol{Y}$ の期待値 $\widehat{\boldsymbol{Y}}$、すなわち、行列 $\boldsymbol{X}$ に行列 $\boldsymbol{\beta}'$ を掛けたものは

$$\begin{pmatrix} \beta_{11} & \beta_{12} & & \beta_{1j} & & \beta_{1p} \\ \beta_{21} & \beta_{22} & & \beta_{2j} & & \beta_{2p} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \beta_{q1} & \beta_{q2} & & \beta_{qj} & & \beta_{qp} \end{pmatrix}$$

$$\times$$

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{\beta}' = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1q} \\ x_{21} & x_{22} & \dots & x_{2q} \\ & \vdots & & \\ x_{i1} & x_{i2} & \dots & x_{iq} \\ & \vdots & & \\ x_{n1} & x_{n2} & \dots & x_{nq} \end{pmatrix} = \left\{ \sum_{k=1}^{b} x_{ik}\beta_{kj} \right\} \tag{742}$$

$$= \boldsymbol{X} \times \begin{pmatrix} \boldsymbol{\beta}_1, & \boldsymbol{\beta}_2, & \dots, & \boldsymbol{\beta}_j, & \dots, & \boldsymbol{\beta}_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}\boldsymbol{\beta}_1, & \boldsymbol{X}\boldsymbol{\beta}_2, & \dots, & \boldsymbol{X}\boldsymbol{\beta}_j, & \dots, & \boldsymbol{X}\boldsymbol{\beta}_p \end{pmatrix} \tag{743}$$

$$= \begin{pmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_i' \\ \vdots \\ \boldsymbol{x}_n' \end{pmatrix} \times \begin{pmatrix} \boldsymbol{\beta}_1, & \boldsymbol{\beta}_2, & \dots, & \boldsymbol{\beta}_j, & \dots, & \boldsymbol{\beta}_p \end{pmatrix} \tag{744}$$

$$\begin{pmatrix} \begin{pmatrix} \beta_{11} \\ \beta_{21} \\ \vdots \\ \beta_{q1} \end{pmatrix}, & \begin{pmatrix} \beta_{12} \\ \beta_{22} \\ \vdots \\ \beta_{q2} \end{pmatrix}, & \dots, & \begin{pmatrix} \beta_{1j} \\ \beta_{2j} \\ \vdots \\ \beta_{qj} \end{pmatrix}, & \dots, & \begin{pmatrix} \beta_{1p} \\ \beta_{2p} \\ \vdots \\ \beta_{qp} \end{pmatrix} \end{pmatrix}$$

$$\times$$

$$= \begin{pmatrix} \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1q} \end{pmatrix} \\ \begin{pmatrix} x_{21} & x_{22} & \dots & x_{2q} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} x_{i1} & x_{i2} & \dots & x_{iq} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} x_{n1} & x_{n2} & \dots & x_{nq} \end{pmatrix} \end{pmatrix} = \left\{ \boldsymbol{x}_i'\boldsymbol{\beta}_j \right\} \tag{745}$$

となる。

このモデルのパラメタ $\boldsymbol{\beta}$ を求める一つの方法は、以下に示す残差の２乗和

$$\mathrm{RSS} = \sum_{j=1}^{p}\sum_{i=1}^{n}(y_{ij} - \widehat{y}_{ij})^2 = \sum_{j=1}^{p}(\boldsymbol{y}_{(j)} - \boldsymbol{X}\boldsymbol{\beta}_j)'(\boldsymbol{y}_{(j)} - \boldsymbol{X}\boldsymbol{\beta}_j) = \sum_{i=1}^{n}(\boldsymbol{y}_i - \boldsymbol{\beta}\boldsymbol{x}_i)'(\boldsymbol{y}_i - \boldsymbol{\beta}\boldsymbol{x}_i) \tag{746}$$

$$= \mathrm{tr}\left[(\boldsymbol{Y} - \widehat{\boldsymbol{Y}})'(\boldsymbol{Y} - \widehat{\boldsymbol{Y}})\right] = \mathrm{tr}\left[(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}')'(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}')\right] \tag{747}$$

を最小とする最小２乗解で、それは

$$\widehat{\boldsymbol{\beta}'} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} \tag{748}$$

で与えられる。(→ 行列の微分) これは RSS を残差の２乗和の分解公式を用いて

$$\begin{aligned}
\mathrm{tr}\left[(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta}')'(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta}')\right] &= \mathrm{tr}\left((\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{\beta}'-\widehat{\boldsymbol{\beta}}')\right)+\mathrm{tr}\left[(\boldsymbol{Y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}')'(\boldsymbol{Y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}}')\right] &(749)\\
&= \mathrm{tr}\left((\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{\beta}'-\widehat{\boldsymbol{\beta}}')\right)-\mathrm{tr}\left(\widehat{\boldsymbol{\beta}}\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}'\right)+\mathrm{tr}\,(\boldsymbol{Y}'\boldsymbol{Y}) &(750)\\
&= \mathrm{tr}\left((\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}})\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{\beta}'-\widehat{\boldsymbol{\beta}}')\right)-\mathrm{tr}\left(\boldsymbol{Y}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}'\right)+\mathrm{tr}\,(\boldsymbol{Y}'\boldsymbol{Y}) &(751)
\end{aligned}$$

と展開すれば、 $\boldsymbol{X}$'$\boldsymbol{X}$ 行列は正定値行列であるから 2 次形式の形をしている第 1 項は必ず非負であり、その最小が $\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}$ に与えられることによる。

なお、$\boldsymbol{X}$'$\boldsymbol{X}$ が特異な場合は

$$\widehat{\boldsymbol{\beta}'}=(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{Y} \tag{752}$$

で与えられるが、この場合、$\widehat{\boldsymbol{\beta}'}$ は一意に定まらない。ただし、その場合でも、$\widehat{\boldsymbol{Y}}$ の部分の最小２乗推定値 $\boldsymbol{X}\widehat{\boldsymbol{\beta}'}$ は一意に定まる。なぜなら、

$$\boldsymbol{X}\widehat{\boldsymbol{\beta}'}=\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{Y} \tag{753}$$

となるが、この $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}$' 行列は射影子（行列）と呼ばれ、一意に定まることが知られているからである。

なお、$p=1$ の単純な回帰分析の場合は

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \tag{754}$$

となる。

12.2 微分による解法

$$\begin{aligned}
\mathrm{RSS} &= \sum_{j=1}^{p}\sum_{i=1}^{n}(y_{ij}-\widehat{y}_{ij})^2=\sum_{j=1}^{p}(\boldsymbol{y}_{(j)}-\boldsymbol{X}\boldsymbol{\beta}_j)'(\boldsymbol{y}_{(j)}-\boldsymbol{X}\boldsymbol{\beta}_j)=\sum_{i=1}^{n}(\boldsymbol{y}_i-\boldsymbol{\beta}\boldsymbol{x}_i)'(\boldsymbol{y}_i-\boldsymbol{\beta}\boldsymbol{x}_i) &(755)\\
&= \mathrm{tr}\left[(\boldsymbol{Y}-\widehat{\boldsymbol{Y}})'(\boldsymbol{Y}-\widehat{\boldsymbol{Y}})\right]=\mathrm{tr}\left[(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta}')'(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\beta}')\right] &(756)\\
&= \mathrm{tr}\,(\boldsymbol{Y}'\boldsymbol{Y})-\mathrm{tr}\,(\boldsymbol{\beta}\boldsymbol{X}'\boldsymbol{Y})+\mathrm{tr}\left(\boldsymbol{Y}'\boldsymbol{X}\boldsymbol{\beta}'\right)+\mathrm{tr}\left(\boldsymbol{\beta}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}'\right) &(757)\\
&= \mathrm{tr}\,(\boldsymbol{Y}'\boldsymbol{Y})-2\mathrm{tr}\left(\boldsymbol{Y}'\boldsymbol{X}\boldsymbol{\beta}'\right)+\mathrm{tr}\left(\boldsymbol{\beta}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}'\right) &(758)\\
& &(759)
\end{aligned}$$

$$\frac{\partial\,\mathrm{RSS}}{\partial\,\boldsymbol{\beta}}=-2\boldsymbol{Y}'\boldsymbol{X}+2\boldsymbol{\beta}\boldsymbol{X}'\boldsymbol{X} \tag{760}$$

$$\boldsymbol{Y}'\boldsymbol{X}=\boldsymbol{\beta}\boldsymbol{X}'\boldsymbol{X} \tag{761}$$

$$\boldsymbol{\beta}=\boldsymbol{Y}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\qquad\text{or}\qquad\boldsymbol{\beta}'=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} \tag{762}$$

****

Derivative of Eigen Value

回転 直交行列が回転であること。

基底の変換

射影と回転

****

A　補足

A.1　実対称行列の固有値は実数である

$n \times n$ の実対称行列 $\boldsymbol{A}$ の固有値と固有ベクトルは

$$\boldsymbol{A}\boldsymbol{p} = \boldsymbol{\lambda}\boldsymbol{p} \tag{A.1}$$

を満たすが、この両辺の共役転置（要素の複素数をその共役複素数に置き換えた上で転置する）は $\bar{\boldsymbol{A}}' = \boldsymbol{A}$ であるから、

$$\bar{\boldsymbol{p}}'\boldsymbol{A} = \bar{\lambda}\bar{\boldsymbol{p}}' \tag{A.2}$$

である。したがって、最初の式の両辺の左から $\bar{\boldsymbol{p}}'$ を掛けると

$$\bar{\boldsymbol{p}}'\boldsymbol{A}\boldsymbol{p} = \boldsymbol{\lambda}\bar{\boldsymbol{p}}'\boldsymbol{p} = \lambda||\boldsymbol{p}||^2 \tag{A.3}$$

２番目の式の両辺の右から $\boldsymbol{p}$ を掛けると

$$\bar{\boldsymbol{p}}'\boldsymbol{A}\boldsymbol{p} = \bar{\lambda}\bar{\boldsymbol{p}}'\boldsymbol{p} = \bar{\lambda}||\boldsymbol{p}||^2 \tag{A.4}$$

を得る。したがって、

$$(\lambda - \bar{\lambda})||\boldsymbol{p}||^2 = 0 \tag{A.5}$$

であり、$\lambda - \bar{\lambda}$ が示される。

なお、上記の展開で、共役転置に関して

$$\bar{\boldsymbol{p}}'\boldsymbol{p} = ||\boldsymbol{p}||^2 \geq 0 \tag{A.6}$$

の関係を用いた。

A.2　実対称行列の階数と 0 でない固有値の数

$n \times n$ の実対称行列 $\boldsymbol{A}$ の固有値分解は

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}', \quad \boldsymbol{P}'\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P}' = \boldsymbol{I}_n \tag{A.7}$$

であるが、$\boldsymbol{P}$ 行列は直交行列であるのでフルランクであるため、右辺の階数は対角行列 $\boldsymbol{\Lambda}$ の階数に等しいが、それは $\boldsymbol{\Lambda}$ の対角要素の 0 でない値の数に等しい。したがって、実対称行列の階数は 0 でない固有値の数に等しい。

A.3　正定値行列の固有値は正である。

全ての $\boldsymbol{x}$ に関して

$$\boldsymbol{x}'\boldsymbol{A}\,\boldsymbol{x} > 0 \tag{A.8}$$

を満たす対称行列を正定値行列と呼ぶ。$\boldsymbol{A}$ の固有値分解を使うと

$$\boldsymbol{x}'\boldsymbol{A}\,\boldsymbol{x} \;=\; \boldsymbol{x}'\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'\boldsymbol{x} \;=\; \boldsymbol{y}'\boldsymbol{\Lambda}\boldsymbol{y} \;= \sum_{i=1}^{n} \lambda_i y_i^2 > 0 \tag{A.9}$$

となる。ただし $\boldsymbol{y} = \boldsymbol{P}'\boldsymbol{x}$ と置いた。$\boldsymbol{y}$ の要素を順に $\boldsymbol{y} = (1, 0, 0, \ldots)$, $\boldsymbol{y} = (0, 1, 0, \ldots)$ とすることにより、全ての固有値が正であることと同値であることが示せる。すなわち、正定値行列の固有値は全て正である。

全ての $\boldsymbol{x}$ に関して

$$\boldsymbol{x}'\boldsymbol{A}\ \boldsymbol{x} \geq 0 \tag{A.10}$$

を満たす対称行列を半正定値行列（非負定値行列）と呼ぶ。すなわち、半正定値行列の固有値は全て非負である。

このことから、$\boldsymbol{A}$ が正定値の場合には $\boldsymbol{A}^{-1}$ もまた正定値であること、ならびに、$|\boldsymbol{A}|$ が正であることがわかる。

A.4 $\boldsymbol{A}\boldsymbol{A}'$ が正定値であること

$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ と置くと

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{y}'\boldsymbol{y} = ||\boldsymbol{y}||^2 > 0 \tag{A.11}$$

$\boldsymbol{A}\boldsymbol{A}'$ の場合は $\boldsymbol{y} = \boldsymbol{A}'\boldsymbol{x}$ と置く。

A.5 $\boldsymbol{A}\boldsymbol{A}'$ と $\boldsymbol{A}'\boldsymbol{A}$ の固有値分解の関係

$\boldsymbol{A}$ を階数が r の $n \times p$ の実行列とする。$\boldsymbol{A}\boldsymbol{A}'$ の固有値分解は

$$\boldsymbol{A}\boldsymbol{A}'\boldsymbol{U} = \boldsymbol{U}\boldsymbol{\Lambda}, \quad \boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}_r \tag{A.12}$$

とすると、この両辺の左から $\boldsymbol{A}'$ を掛けると

$$\boldsymbol{A}'\boldsymbol{A} \quad \boldsymbol{A}'\boldsymbol{U} = \boldsymbol{A}'\boldsymbol{U} \quad \boldsymbol{\Lambda} \tag{A.13}$$

となり、$\boldsymbol{\Lambda}$ と $\boldsymbol{A}'\boldsymbol{U}$ が $\boldsymbol{A}'\boldsymbol{A}$ の固有値行列、固有値ベクトル行列であることが解る。そして、$\boldsymbol{V}'\boldsymbol{V} = \boldsymbol{I}_r$ を満たす固有値ベクトル行列は $\boldsymbol{A}'\boldsymbol{U}$ の各列に比例するため、対角行列 $\boldsymbol{D}$ を用いて

$$\boldsymbol{V} = \boldsymbol{A}'\boldsymbol{U}\boldsymbol{D} \tag{A.14}$$

として求められるが、

$$\boldsymbol{V}'\boldsymbol{V} = \boldsymbol{D}\boldsymbol{U}'\boldsymbol{A}\boldsymbol{A}'\boldsymbol{U}\boldsymbol{D} = \boldsymbol{D}\boldsymbol{\Lambda}\boldsymbol{D} = \boldsymbol{I}_r \tag{A.15}$$

であるためには

$$\boldsymbol{D} = \boldsymbol{\Lambda}^{-\frac{1}{2}} \tag{A.16}$$

でなければならない。したがって、

$$\boldsymbol{V} = \boldsymbol{A}'\boldsymbol{U}\boldsymbol{\Lambda}^{-\frac{1}{2}} \tag{A.17}$$

である。

同様に

$$\boldsymbol{U} = \boldsymbol{A}\boldsymbol{V}\boldsymbol{\Lambda}^{-\frac{1}{2}} \tag{A.18}$$

となる。

A.5.1 特異値分解と行列の階数

コンパクトな特異値分解から、$\boldsymbol{X}$ は階数が r の３つの行列の積で表されることが示された。このことより、$\boldsymbol{X}$ の階数が r であることが以下の様にして示される。

行列の階数の性質

$$\mathrm{rank}\left(\boldsymbol{A}'\right) = \mathrm{rank}\left(\boldsymbol{A}\right) \qquad \text{and} \qquad \mathrm{rank}\left(\boldsymbol{AB}\right) \leq \min(\mathrm{rank}\left(\boldsymbol{A}\right), \mathrm{rank}\left(\boldsymbol{B}\right)) \tag{A.19}$$

より

$$\mathrm{rank}\left(\boldsymbol{X}\right) = \mathrm{rank}\left(\boldsymbol{UDV}'\right) \leq \min(\mathrm{rank}\left(\boldsymbol{U}\right), \mathrm{rank}\left(\boldsymbol{D}\right), \mathrm{rank}\left(\boldsymbol{V}'\right)) = r \tag{A.20}$$

を得るが、これより $\min(r, \mathrm{rank}\left(\boldsymbol{X}\right)) = \mathrm{rank}\left(\boldsymbol{X}\right)$ である。また、

$$\mathrm{rank}\left(\boldsymbol{U}'\boldsymbol{X}\right) = \mathrm{rank}\left(\boldsymbol{DV}'\right) = r \leq \min(\mathrm{rank}\left(\boldsymbol{U}'\right), \mathrm{rank}\left(\boldsymbol{X}\right)) = \min(r, \mathrm{rank}\left(\boldsymbol{X}\right)) \tag{A.21}$$

であるから、$r \leq \mathrm{rank}\left(\boldsymbol{X}\right) \leq r$ を得る。

参考文献

Grayvill, F. A. (1976) *Theory and Application of the Linear Model.* Wadsworth Publishing Company, Inc., Belmont, California

Harville,D.A. (1997) *Matrix Algebra from a Statistician's Perspective.* Springer-Verlag New York, Inc.

Seber, G. A. F. (2008) *A Matrix Handbook for Statistician.* John Wiley & Sons, Inc., Hoboken, NJ.

Searl, S. R. (1971) *Linear Models.* John Wiley & Sons, Inc.