

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, 20230405

目次

1	ベクトル	2
1.1	ベクトルの演算	3
1.1.1	基本演算	3
1.1.2	ベクトルの内積	4
2	行列	4
2.1	行列の転置	5
2.2	行列の種類	6
2.3	行列の演算	6
2.3.1	基本演算	6
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	9
2.3.5	行列の積の一般形	10
2.3.6	演算規則	13
2.4	簡易表記と添字演算子	13
2.5	ベクトルの2次形式	14
2.6	行列のトレース	14
2.7	diag 演算子	14
2.8	並べ替え行列	15
2.9	等号成立条件	15
2.10	逆行列	16
2.11	連立方程式の解	16
3	ベクトルの線形独立と行列の階数	16
4	行列の利用	17
4.1	座標軸や図形の回転	17
4.2	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）	17
4.3	デザイン行列	18
4.4	距離行列	19
4.5	推移確率行列	20
4.6	ソシオメトリの行列	20
4.7	データ行列の要約統計量	20
4.8	確率変数と確率変数ベクトル	22
4.9	因子分析	23
4.10	多変量正規分布の密度関数	23
4.11	多項分布の確率	24
5	その他の演算子	24
5.1	分割行列の演算	24
5.2	ssq 演算子	26
5.3	Kronecker 積	27
5.4	vec 演算子	28
5.5	逆行列の演算規則といくつかの公式	29
6	スワイプ演算子と逆行列ならびに行列式	30
6.0.1	スワイプ演算子の定義	30
6.0.2	分割行列のスワイプ	31
6.0.3	スワイプ演算子を利用した様々な公式	32

7 齊次方程式と一般逆行列	33
7.1 齊次方程式	33
7.2 一般逆行列について	34
8 射影子行列***	35
9 行列の固有値と特異値分解	36
10 行列の微分	37
11 線形モデルと最小2乗解	39
11.1 線形モデルと最小2乗解	39
11.2 微分による解法	42

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、

$$\mathbf{a} = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 6 \end{pmatrix} \quad \mathbf{b} = (2, 4, 5) \tag{1}$$

の様なものであり、 \mathbf{a} を列ベクトル (または縦ベクトル)、 \mathbf{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 c 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 \mathbf{a} は 4×1 の列ベクトル、 \mathbf{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

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

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 \mathbf{a} と書くこ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、

$$\mathbf{a}' = (1, 3, 5, 6) \quad \mathbf{b}' = \begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix} \tag{3}$$

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

$$\mathbf{a} = (1, 3, 5, 6)' \tag{4}$$

または

$$\mathbf{a}' = (1, 3, 5, 6) \tag{5}$$

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

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{pmatrix} \quad \begin{array}{l} \mathbf{a} = (a_1, a_2, \dots, a_r)' \\ \text{または} \\ \mathbf{a}' = (a_1, a_2, \dots, a_r) \end{array} \tag{6}$$

と書ける。なお、ベクトル \mathbf{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

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

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。サイズが明らかの場合には単に $\mathbf{1}$ 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 $\mathbf{m} = (h, w, s)'$ 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 $\mathbf{x} = (x, y)'$ 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

$$\mathbf{a} = (a_1, a_2, \dots, a_r)' \quad \mathbf{b} = (b_1, b_2, \dots, b_r)' \quad (8)$$

の和もしくは差は

$$\mathbf{a} \pm \mathbf{b} = (a_1 \pm b_1, a_2 \pm b_2, \dots, a_r \pm b_r)' \quad (9)$$

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。

$$\lambda \mathbf{a} = (\lambda a_1, \lambda a_2, \dots, \lambda a_r)' \quad (10)$$

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。

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \quad (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}), \quad (11)$$

$$\lambda \mathbf{a} = \mathbf{a} \lambda \quad (\lambda + \mu)\mathbf{a} = \lambda \mathbf{a} + \mu \mathbf{a}, \quad \lambda(\mathbf{a} + \mathbf{b}) = \lambda \mathbf{a} + \lambda \mathbf{b}, \quad \lambda(\mu \mathbf{a}) = (\lambda \mu)\mathbf{a} \quad (12)$$

また、スカラーとベクトルの加減算は、 $r \times 1$ のベクトル \mathbf{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

$$\mathbf{a} \pm \lambda \mathbf{1}_r \quad \text{or} \quad \mathbf{a} \pm \mathbf{1}_r \lambda \quad (13)$$

と書くことが出来る。

1.1.1.1 例

$$(1, 2, 3) + (4, 5, 6) = (5, 7, 9) \quad (14)$$

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

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

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

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

$$m. = \frac{1}{p} \sum_{i=1}^p \mathbf{m}_i \quad (17)$$

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

$$\mathbf{m}_i = (h_i, w_i, s_i)' \quad (18)$$

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 \mathbf{a}' に $r \times 1$ の縦ベクトル \mathbf{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、

$$\mathbf{a}'\mathbf{b} = (a_1, a_2, \dots, a_r) \times \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \end{pmatrix} = \sum_{i=1}^r a_i b_i \quad (19)$$
$$= a_1 \times b_1 + a_2 \times b_2 + \dots + a_r \times b_r \quad (20)$$

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、

$$\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a} \quad (21)$$

である。(内積を (\mathbf{a}, \mathbf{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

$$\mathbf{a}'\mathbf{1} = \mathbf{1}'\mathbf{a} = \sum_{i=1}^r a_i \times 1 = \sum_{i=1}^r a_i \quad (22)$$

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

$$\mathbf{a}'\mathbf{a} = \sum_{i=1}^r a_i \times a_i = \sum_{i=1}^r a_i^2 \quad (23)$$

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、 $\|\mathbf{a}\|$ と書く。

1.1.2.1 例

$$(4 \ 5) \times \begin{pmatrix} 3 \\ 7 \end{pmatrix} = 4 \times 3 + 5 \times 7 = 47 \quad (24)$$

ベクトル \mathbf{a} の要素の平均 m と分散 s^2 は

$$m = \frac{1}{r} \mathbf{a}'\mathbf{1}, \quad s^2 = \frac{1}{r} \mathbf{a}'\mathbf{a} - m^2 \quad (25)$$

と書ける。

内積に関しては以下の演算規則が成立する。

$$\mathbf{a}'(\mathbf{b} + \mathbf{c}) = \mathbf{a}'\mathbf{b} + \mathbf{a}'\mathbf{c} \quad (26)$$

$$(\lambda\mathbf{a})'\mathbf{b} = \lambda(\mathbf{a}'\mathbf{b}) \quad (27)$$

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

$$\cos \theta = \frac{\mathbf{a}'\mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \quad (28)$$

となる。

2 行列

同じサイズ $r \times 1$ の c 個の列ベクトル

$$\mathbf{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{rj} \end{pmatrix}, \quad j = 1, 2, \dots, c \quad (29)$$

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

$$\mathbf{b}_i = (b_{i1}, b_{i2}, \dots, b_{ic})', i = 1, 2, \dots, r \quad (30)$$

を縦に並べたもの

$$\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_c), \quad \mathbf{B} = \begin{pmatrix} \mathbf{b}'_1 \\ \mathbf{b}'_2 \\ \vdots \\ \mathbf{b}'_r \end{pmatrix} \quad (31)$$

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 $a_{ij}, i = 1, 2, \dots, r, j = 1, 2, \dots, c$ もしくは $b_{ij}, i = 1, 2, \dots, r, j = 1, 2, \dots, c$ 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、 \mathbf{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、

$$\mathbf{A} = \{a_{ij}\}, i = 1, 2, \dots, r, j = 1, 2, \dots, c \quad (32)$$

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。

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

最後の表現に付いている括弧付きの添え字は $\mathbf{a}_{(j)}$ 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 $x_i, i = 1, 2, \dots, r$ 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 $y'_i, i = 1, 2, \dots, n$ を縦に並べると、以下に示す $n \times p$ の座標行列となる。

$$\mathbf{X} = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_r \end{pmatrix} \quad \mathbf{Y} = \begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_r \end{pmatrix} \quad (34)$$

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

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

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 \mathbf{B}) の ij 要素を $(\mathbf{B})_{ij}$ と書くとすれば、 $(\mathbf{A}')_{ij} = a_{ji}$ 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** $a_{11}, a_{22}, a_{33}, \dots$ を対称軸として、180 度回転した形となっている。

定義から、

$$(\mathbf{A}')' = \mathbf{A} \quad (36)$$

である。

2.1.0.1 例

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

2.2 行列の種類

行列は、その形状や性質により様々な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
 $a_{ij} = a_{ji}, i = 1, 2, \dots, r, j = 1, 2, \dots, c, \mathbf{A}' = \mathbf{A}$
- 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 \mathbf{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 $\text{diag}(\mathbf{A})$ と書く場合がある。この $\text{diag}()$ 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 \mathbf{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 \mathbf{O} と書く ($\mathbf{A} - \mathbf{A} = \mathbf{O}$ 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。($\mathbf{A}' = \mathbf{A}$) 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。

$$\mathbf{A} \pm \mathbf{B} = \{a_{ij} \pm b_{ij}\}, i = 1, 2, \dots, r, j = 1, 2, \dots, c \quad (38)$$

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。

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

行列とスカラーの積は以下のように定義される。

$$\lambda \mathbf{A} = \mathbf{A} \lambda = \{\lambda a_{ij}\}, i = 1, 2, \dots, r, j = 1, 2, \dots, c \quad (40)$$

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 \mathbf{A} と $c \times 1$ のベクトル \mathbf{b} との間には積が定義され、以下ようになる。

$$\mathbf{Ab} = \left\{ \sum_{k=1}^c a_{ik} b_k \right\}, i = 1, 2, \dots, r \quad (41)$$

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。

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \vdots \\ \mathbf{a}'_a \end{pmatrix} \quad (42)$$

ならば

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

$$= \begin{pmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \vdots \\ \mathbf{a}'_i \\ \vdots \\ \mathbf{a}'_a \end{pmatrix} \times \mathbf{b} = \begin{pmatrix} \mathbf{a}'_1 \mathbf{b} \\ \mathbf{a}'_2 \mathbf{b} \\ \vdots \\ \mathbf{a}'_i \mathbf{b} \\ \vdots \\ \mathbf{a}'_a \mathbf{b} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^c a_{1k} b_k \\ \sum_{k=1}^c a_{2k} b_k \\ \vdots \\ \sum_{k=1}^c a_{ik} b_k \\ \vdots \\ \sum_{k=1}^c a_{rk} b_k \end{pmatrix} \quad (44)$$

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

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

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

$$\mathbf{Ax} = \mathbf{c} \quad (46)$$

2.3.2.1 例

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

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 \mathbf{A} 行列と $b \times c$ の \mathbf{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。

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

これは、右側の行列 \mathbf{B} を c 個の列ベクトル $\mathbf{b}_{(j)}, j = 1, 2, \dots, c$ 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 $\mathbf{A}\mathbf{b}_{(j)}$ を行い（その結果は $a \times 1$ のベクトル）、その結果を横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。

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

ならば

$$\mathbf{AB} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1b} \\ a_{21} & a_{22} & \dots & a_{2b} \\ \vdots & \vdots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ib} \\ \vdots & \vdots & \dots & \vdots \\ a_{a1} & a_{a2} & \dots & a_{ab} \end{pmatrix} \times \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1j} & \dots & b_{1c} \\ b_{21} & b_{22} & \dots & b_{2j} & \dots & b_{2c} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ b_{b1} & b_{b2} & \dots & b_{bj} & \dots & b_{bc} \end{pmatrix} = \left\{ \sum_{k=1}^b a_{ik} b_{kj} \right\} \quad (50)$$

$$= \mathbf{A} \times (\mathbf{b}_{(1)}, \mathbf{b}_{(2)}, \dots, \mathbf{b}_{(j)}, \dots, \mathbf{b}_{(c)}) = (\mathbf{A}\mathbf{b}_{(1)}, \mathbf{A}\mathbf{b}_{(2)}, \dots, \mathbf{A}\mathbf{b}_{(j)}, \dots, \mathbf{A}\mathbf{b}_{(c)}) \quad (51)$$

$$= \begin{pmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \vdots \\ \mathbf{a}'_i \\ \vdots \\ \mathbf{a}'_a \end{pmatrix} \times (\mathbf{b}_{(1)}, \mathbf{b}_{(2)}, \dots, \mathbf{b}_{(j)}, \dots, \mathbf{b}_{(c)}) \quad (52)$$

$$= \begin{pmatrix} \left(\begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{b1} \end{pmatrix} \right), \left(\begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{b2} \end{pmatrix} \right), \dots, \left(\begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{bj} \end{pmatrix} \right), \dots, \left(\begin{pmatrix} b_{1c} \\ b_{2c} \\ \vdots \\ b_{bc} \end{pmatrix} \right) \\ \times \\ \left(\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1b} \\ a_{21} & a_{22} & \dots & a_{2b} \\ \vdots & \vdots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ib} \\ \vdots & \vdots & \dots & \vdots \\ a_{a1} & a_{a2} & \dots & a_{ab} \end{pmatrix} \right) \end{pmatrix} = \left\{ \mathbf{a}'_i \mathbf{b}_{(j)} \right\} \quad (53)$$

$$(54)$$

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

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & -4 & 6 \end{pmatrix} \times \begin{pmatrix} 1 & 4 & 7 & -1 \\ 2 & 5 & 8 & -2 \\ 3 & 6 & 9 & -3 \end{pmatrix} = \begin{pmatrix} 14 & 32 & 50 & -14 \\ 12 & 24 & 36 & -12 \end{pmatrix} \quad (55)$$

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

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

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 \mathbf{D} を対角行列とすれば、

$$\mathbf{AD} = (\mathbf{a}_{(1)}d_{11}, \mathbf{a}_{(2)}d_{22}, \dots, \mathbf{a}_{(j)}d_{jj}, \dots, \mathbf{a}_{(b)}d_{bb}) = \{\mathbf{a}_{ij}d_{jj}\}, \quad (57)$$

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

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。

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A} \quad (59)$$

自分自身の転置を左から掛けたものは

$$\mathbf{A}'\mathbf{A} = \left\{ \sum_{k=1}^b a_{ki}a_{kj} \right\} \quad (60)$$

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

$$\mathbf{AA}' = \left\{ \sum_{k=1}^b a_{ik}a_{jk} \right\} \quad (61)$$

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、 \mathbf{AA}' の対角要素には各行の2乗和 $\sum_{k=1}^b a_{ik}^2$ が入る。 \mathbf{A} がどのような行列であっても、 $\mathbf{A}'\mathbf{A}$ と \mathbf{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、 \mathbf{A} の各列は直交するという。また、

$$\mathbf{A}'\mathbf{A} = \mathbf{AA}' = \mathbf{I} \quad (62)$$

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 \mathbf{a} に右から $1 \times c$ の行ベクトル \mathbf{b}' を掛けると

$$\mathbf{ab}' = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_i \\ \vdots \\ a_r \end{pmatrix} \times (b_1, b_2, \dots, b_j, \dots, b_c) = \{a_i b_j\} \quad (63)$$

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、 $\mathbf{ab}' \neq \mathbf{ba}'$ 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。

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

2.3.4.1 例

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

2.3.5 行列の積の一般形

\mathbf{a} と \mathbf{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 \mathbf{c} とすると、

$$\mathbf{c} = \mathbf{a} \pm \mathbf{b}, \quad c_i = a_i \pm b_i \quad (66)$$

となる。二つのベクトルの内積を c とすると

$$c = \mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a}, \quad c = \sum_{i=1}^r a_i b_i \quad (67)$$

である。

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

$$\mathbf{C} = \mathbf{A} \pm \mathbf{B}, \quad c_{ij} = a_{ij} \pm b_{ij} \quad (68)$$

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

$$\mathbf{C} = \mathbf{AB}, \quad c_{ij} = \sum_{k=1}^{ca} a_{ik} b_{kj}, \quad i = 1, 2, \dots, ra, \quad j = 1, 2, \dots, cb \quad (69)$$

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

$$\mathbf{C} = \mathbf{BA}, \quad c_{ij} = \sum_{k=1}^{ra} b_{ik} a_{kj}, \quad i = 1, 2, \dots, rb, \quad j = 1, 2, \dots, ca \quad (70)$$

で、 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

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

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

$$C = A'B, \quad c_{ij} = \sum_{k=1}^{ra} a_{ki} b_{kj}, \quad i = 1, 2, \dots, ca, \quad j = 1, 2, \dots, cb \quad (72)$$

であり、 C 行列のサイズは $ca \times cb$ である。すなわち、(69) 式のパターンに於いて、転置となっている行列 A の添字を交換すればよい。また、 $rb = ra$ の場合、積 $B'A$ が定義でき、

$$C = B'A, \quad c_{ij} = \sum_{k=1}^{rb} b_{ki} a_{kj} = \sum_{k=1}^{rb} a_{kj} b_{ki}, \quad i = 1, 2, \dots, cb, \quad j = 1, 2, \dots, ca \quad (73)$$

であり、 C 行列のサイズは $cb \times ca$ である。これも、(70) 式の B の添字を交換してもものとなっている。自分自身の転置との積は常に定義可能で、

$$C = A'A, \quad c_{ij} = \sum_{k=1}^{ra} a_{ki} a_{kj}, \quad i = 1, 2, \dots, ca, \quad j = 1, 2, \dots, ca \quad (74)$$

$$C = AA', \quad c_{ij} = \sum_{k=1}^{ca} a_{ik} a_{jk}, \quad i = 1, 2, \dots, ra, \quad j = 1, 2, \dots, ra \quad (75)$$

となるが、 $A'A$ のサイズは $ca \times ca$ 、 AA' 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、 A 行列が対角行列の場合は、 AB は B の各行を A の対角要素倍したものであり、同様に、 B 行列が対角の場合は、 AB は A の各列 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、(69) 式は以下ようになる。なお、対角行列を強調するために $\text{diag}()$ の記号を用いている。

$$C = \text{diag}(A)B, \quad c_{ij} = a_{ii} b_{ij} \quad (76)$$

$$C = A \text{diag}(B), \quad c_{ij} = a_{ij} b_{jj} \quad (77)$$

$$C = \text{diag}(A) \text{diag}(B), \quad c_{ij} = \begin{cases} a_{ii} b_{ii} & i = j \\ 0 & i \neq j \end{cases} \quad (78)$$

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 $rb = 1$ 、であると考えるとよい。例えば、(69) 式に於いて B が縦ベクトルの場合、 $cb = 1$ であるから、

$$c = Ab, \quad c_i = c_{i1} = \sum_{k=1}^{ca} a_{ik} b_{k1} = \sum_{k=1}^{ca} a_{ik} b_k \quad (79)$$

となり、 c は $ra \times 1$ の縦ベクトルとなる。(72) 式に於いて A が列ベクトルの場合は $ca = 1$ であるから

$$c' = a'B, \quad c_i = c_{i1} = \sum_{k=1}^{ca} a_{k1} b_{kj} = \sum_{k=1}^{ca} a_k b_{kj} \quad (80)$$

となり 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、

$$c = A\mathbf{1}, \quad c_i = c_{i1} = \sum_{k=1}^{ca} a_{ik} \mathbf{1} = \sum_{k=1}^{ca} a_{ik} \quad (81)$$

$$c' = \mathbf{1}'B, \quad c_i = c_{i1} = \sum_{k=1}^{ca} \mathbf{1} b_{kj} = \sum_{k=1}^{ca} b_{kj} \quad (82)$$

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

$$D = A(BC), \quad d_{ij} = \sum_{k=1}^{ca} a_{ik}[BC]_{kj} = \sum_{k=1}^{ca} a_{ik} \left(\sum_{\ell=1}^{cb} b_{k\ell} c_{\ell j} \right) \quad (83)$$

$$= \sum_{k=1}^{ca} a_{ik} \sum_{\ell=1}^{cb} b_{k\ell} c_{\ell j} = \sum_{k=1}^{ca} \sum_{\ell=1}^{cb} a_{ik} b_{k\ell} c_{\ell j} \quad (84)$$

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

$$D = (AB)C, \quad d_{ij} = \sum_{k=1}^{cb} [AB]_{ik} c_{kj} = \sum_{k=1}^{cb} \left(\sum_{\ell=1}^{ca} a_{i\ell} b_{\ell k} \right) c_{kj} \quad (85)$$

$$= \sum_{k=1}^{cb} \sum_{\ell=1}^{ca} a_{i\ell} b_{\ell k} c_{kj} = \sum_{\ell=1}^{ca} a_{i\ell} \sum_{k=1}^{cb} b_{\ell k} c_{kj} = \sum_{k=1}^{ca} a_{ik} \sum_{\ell=1}^{cb} b_{k\ell} c_{\ell j} \quad (86)$$

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

$$F = ABCDE \quad (87)$$

$$f_{ij} = \sum_{i1=1}^{ca} \sum_{i2=1}^{cb} \sum_{i3=1}^{cc} \sum_{i4=1}^{cd} a_{i,i1} b_{i1,i2} c_{i2,i3} d_{i3,i4} e_{i4,j} \quad (88)$$

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

に示される様なパター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、

$$F = AB'CD'E \quad (90)$$

$$f_{ij} = \sum_{i1=1}^{ca} \sum_{i2=1}^{rb} \sum_{i3=1}^{cc} \sum_{i4=1}^{rd} a_{i,i1} b_{i2,i1} c_{i2,i3} d_{i4,i3} e_{i4,j} \quad (91)$$

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

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、 A 行列と D 行列が対角行列の場合は、まず、 $\text{diag}(bA)B$ をまとめることにより、総和記号が一つ減る。次に、 D 行列であるが、これを E 行列の左からかかっていると見るか、 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。

$$F = \text{diag}(A)BC\text{diag}(D)E = (\text{diag}(A)B)C(\text{diag}(D)E) \quad (93)$$

$$f_{ij} = \sum_{i1=1}^{ca} \sum_{i2=1}^{cc} (a_{i,i} b_{i,i1}) c_{i1,i2} (d_{i2,i2} e_{i2,j}) \quad (94)$$

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

$$F = \text{diag}(A)BC\text{diag}(D)E = (\text{diag}(A)B)(C\text{diag}(D))E \quad (96)$$

$$f_{ij} = \sum_{i1=1}^{ca} \sum_{i2=1}^{cc} (a_{i,i} b_{i,i1}) (c_{i1,i2} d_{i2,i2}) e_{i2,j} \quad (97)$$

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

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できる。例えば、以下の例で、 C 以外が列ベクトルであれば、 $B'CD$ はスカラーとなるため、それを前に出すことができる。

$$F = a(b'Cd)e' = (b'Cd)ab' \quad (99)$$

$$f_{ij} = \sum_{i1=1}^1 \sum_{i2=1}^{rb} \sum_{i3=1}^{cc} \sum_{i4=1}^1 a_{i,i1} b_{i2,i1} c_{i2,i3} d_{i3,i4} e_{j,i4} \quad (100)$$

$$= \sum_{i1=1}^1 a_{i,i1} \sum_{i2=1}^{rb} b_{i2,i1} \sum_{i3=1}^{cc} c_{i2,i3} \sum_{i4=1}^1 d_{i3,i4} e_{j,i4} \quad (101)$$

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

$$= a_i e_j \quad (103)$$

$$= \left(\sum_{i2=1}^{rb} b_{i2,i1} \sum_{i3=1}^{cc} c_{i2,i3} d_{i3,i4} \right) a_i e_j \quad (104)$$

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。

$$A + B + C = (A + B) + C = A + (B + C) = B + (A + C) \quad (105)$$

$$A(B + C) = AB + AC \quad \text{and} \quad (B + C)A = BA + CA \quad (106)$$

$$ABC = (AB)C = A(BC) \quad (107)$$

$$(A + B)' = A' + B', \quad (AB)' = B'A', \quad (A')' = A \quad (108)$$

$r \times c$ の行列 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、

$$A \pm \lambda \mathbf{1}_r \mathbf{1}'_c \quad \text{or} \quad A \pm \mathbf{1}_r \mathbf{1}'_c \lambda \quad \text{or} \quad A \pm \mathbf{1}_r \lambda \mathbf{1}'_c \quad (109)$$

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ェアでは、以下の様な簡易表記が用いられている。 A を $r \times c$ の行列、 r を $r \times 1$ の列ベクトル、 c を $1 \times c$ の行ベクトル、 λ をスカラーとする。

$$r \pm \lambda \mathbf{1}_r \rightarrow r \pm \lambda, \quad c \pm \lambda \mathbf{1}'_c \rightarrow c \pm \lambda \quad (110)$$

$$A \pm \lambda \mathbf{1}_r \mathbf{1}'_c \rightarrow A \pm \lambda \quad (111)$$

また、以下の様な添字演算子も利用されることがある。

$$A_{+j} = \sum_{i=1}^r a_{ij}, \quad A_{i+} = \sum_{j=1}^c a_{ij} \quad (112)$$

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、 $(A+B)C$ の様な演算結果の第 ij 要素を

$$\text{演算式の } ij \text{ 要素} = [\text{演算式}]_{ij} \quad (113)$$

と略記する。

2.5 ベクトルの2次形式

$b'Ab$ の形をベクトル b の **2次形式**と呼ぶ。この計算は以下ようになる。

$$b'Ab = b'(Ab) = \sum_{i=1}^r b_i \sum_{j=1}^r a_{ij} b_j \quad (114)$$

$$= \sum_{i=1}^r a_{ii} b_i^2 + 2 \sum_{i < j} a_{ij} b_i b_j \quad (\mathbf{A} \text{ が対称の場合}) \quad (115)$$

なお、 \mathbf{A} が対称行列の場合、どのような b , ただし $b \neq \mathbf{0}$, を選んでも2次形式が正となる行列 \mathbf{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 $\mathbf{A}'\mathbf{A}$ 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

$$\text{tr}(\mathbf{A}) = \sum_{j=1}^r a_{jj} \quad (116)$$

を**トレース (trace)**と呼び $\text{tr}(\mathbf{A})$ と書く。

トレース演算子は以下の様な性質を持っている。

$$\text{tr}(\lambda) = \lambda, \quad \text{tr}(\lambda\mathbf{A}) = \lambda\text{tr}(\mathbf{A}), \quad \text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}') \quad (117)$$

$$\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B}) \quad (118)$$

$$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA}) = \text{tr}(\mathbf{B}'\mathbf{A}') = \text{tr}(\mathbf{A}'\mathbf{B}'), \quad (119)$$

$$\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA}) = \text{tr}(\mathbf{CAB}) \quad (120)$$

$$= \text{tr}(\mathbf{C}'\mathbf{B}'\mathbf{A}') = \text{tr}(\mathbf{B}'\mathbf{A}'\mathbf{C}') = \text{tr}(\mathbf{A}'\mathbf{C}'\mathbf{B}') \quad (121)$$

$$\mathbf{x}'\mathbf{Ax} = \text{tr}(\mathbf{x}'\mathbf{Ax}) = \text{tr}(\mathbf{Ax}\mathbf{x}') \quad (122)$$

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

$$\text{tr}(\mathbf{A}'\mathbf{A}) = \text{tr}(\mathbf{AA}') = \sum_{i=1}^r \sum_{j=1}^c a_{ij}^2 \quad (123)$$

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

$$\text{RSS} = \sum_{i=1}^r (y_i - \hat{y}_i)^2 = (\mathbf{y} - \hat{\mathbf{y}})'(\mathbf{y} - \hat{\mathbf{y}}) \quad (124)$$

$$= \text{tr}((\mathbf{y} - \hat{\mathbf{y}})'(\mathbf{y} - \hat{\mathbf{y}})) = \text{tr}((\mathbf{y} - \hat{\mathbf{y}})(\mathbf{y} - \hat{\mathbf{y}})') \quad (125)$$

$$\text{RSS} = \sum_{i=1}^r \sum_{j=1}^c (y_{ij} - \hat{y}_{ij})^2 = \text{tr}((\mathbf{Y} - \hat{\mathbf{Y}})'(\mathbf{Y} - \hat{\mathbf{Y}})) \quad (126)$$

$$= \text{tr}((\mathbf{Y} - \hat{\mathbf{Y}})(\mathbf{Y} - \hat{\mathbf{Y}})') \quad (127)$$

と書くことができる。

2.7 diag 演算子

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

なお、 $\text{Diag}()$ を使わずに $\text{diag}()$ 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 $\text{Diag}()$ 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 $\mathbf{A} = [\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \dots, \mathbf{a}_{(c)}]$ の場合、 \mathbf{D} を d_i を対角要素として持つ対角行列とすると、

$$\mathbf{AD} = [\mathbf{a}_{(1)}d_1, \mathbf{a}_{(2)}d_2, \dots, \mathbf{a}_{(c)}d_c] \quad (128)$$

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 $A = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_c]'$ の場合、

$$DA = [d_1\mathbf{a}_1, d_2\mathbf{a}_2, \dots, d_r\mathbf{a}_c]' \quad (129)$$

となる。

2.8 並べ替え行列

$r \times c$ の行列 X の第 i 行と第 j 行を入れ替えたものを 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 P_{ij} を用いて $Y = P_{ij}X$ と書くことができる。以下に $r = 4, i = 1, j = 4$ と $r = 4, i = 2, j = 3$ の場合の例を示す。

$$P_{34} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad P_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (130)$$

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

$$Z = P_{i'j'}Y = P_{i'j'}P_{ij}X = PX \quad (P = P_{i'j'}P_{ij}) \quad (131)$$

である。例えば、上記の二つの行列の積は

$$P = P_{34} \times P_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (132)$$

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 $X = P'Z$ 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 P を定義して、 XP とすればよい。

2.9 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。

$A = B$ と $A - B = 0$ は同値である。

1 : 等号成立条件 1

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

なぜなら、 $\text{tr}(A'A)$ は 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、

2 : 等号成立条件 2

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

なぜなら、

$$(AB - AC)'(AB - AC) = (B' - C')(A'AB - A'AC) \quad (133)$$

であるが、 $A'AB = A'AC$ の場合、上式の右辺、したがって左辺は O となる。したがって、先の等号成立条件により、 $AB = AC$ となる。

2.10 逆行列

正方行列 A にある行列を右または左から掛けたときその結果が単位行列となるような行列を A の **逆行列** (inverse) と呼び、 A^{-1} で示す。すなわち

$$AA^{-1} = A^{-1}A = I \quad (134)$$

である。特に、 A がスカラー数の場合は $a^{-1} = 1/a$ となる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。

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \quad (135)$$

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**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**な行列という。正則な行列の逆行列は一意に定まる。

なお、サイズが 2 より大きな行列の逆行列の求め方はスウィープ演算子の節を参照のこと。

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

(正規) 直交行列の逆行列はその転置行列である。

$$A^{-1} = A' \quad \text{ただし} \quad AA' = A'A = I \quad (136)$$

2.10.0.1 例

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/3 \end{pmatrix} \quad (137)$$

逆行列は以下の様な演算規則を持つ。

$$\lambda^{-1} = \frac{1}{\lambda}, \quad (\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}, \quad I^{-1} = I \quad (138)$$

$$(A^{-1})^{-1} = A, \quad (A')^{-1} = (A^{-1})', \quad (AB)^{-1} = B^{-1}A^{-1} \quad (139)$$

2.11 連立方程式の解

連立方程式の係数行列 A が正則の場合、連立方程式の等号の両辺に左から係数行列の逆行列を掛けると解が求まる。

$$Ax = c \quad \leftarrow \rightarrow \quad = A^{-1}Ax = A^{-1}c \quad \leftarrow \rightarrow \quad x = A^{-1}c \quad (140)$$

解の存在条件

3 ベクトルの線形独立と行列の階数

$r \times c$ の行列 $A = [a_{(1)} \ a_{(2)} \ \dots \ a_{(c)}]$ を構成する列ベクトルが**線形独立** (1次独立, linearly independent) であるとは、 $Ax = 0$ を満たすような x ベクトルが $x = 0$ 以外に存在しないことをいう。列ベクトルが線形独立でない場合、すなわち、線形従属の場合には、 A_j を A からその第 j 列を含むいくつかの列を除いて出来る行列 (最大 $r \times (c-1)$) として、 $a_{(j)} = A_j b$ を満たすような列が少なくとも一つ存在する。 A の各列のうち q 個の列を用いて線形独立なベクトルの組を作ることは出来るが、 $q+1$ 個の線形独立な組を作ることが出来ない場合、その行列の**階数** (rank) は q であるという。なお、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線形独立な行ベクトルの組の最大個数と一致する。したがって、階数 $\leq r, c$ である。また、 $r \geq c$ の場合、行列 A の階数がその列数 c に等しい場合、その行列は列フルランク (full column rank) であるという。また、列フルランクである正方行列を正則行列 (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、 $A'A$ と AA' の階数は 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 A^{-1} が存在する。また、 A が列フルランクの場合 $(A'A)^{-1}$ が存在する。

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

$$A = PQ' \quad (141)$$

ただし、このような P, Q は一位には定まらない。

sweep を使う、see sweep3.tex 固有値分解を使う

3 : 行列の階数

- A のサイズが $r \times c$ の場合、 $\text{rank}(A) \leq \min(r, c)$
- $\text{rank}(A) = \text{rank}(A') = \text{rank}(A'A) = \text{rank}(AA')$
- $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$
- B のサイズが $c \times q$, $q \leq c$ で $\text{rank}(B) = q$ の場合 $\text{rank}(AB) = q$
- P と Q が正則の場合、 $\text{rank}(PAQ) = \text{rank}(A)$
- A が対称の場合 $\text{rank}(A) = A$ の 0 でない固有値の数

4 行列の利用

4.1 座標軸や図形の回転

平面上の一点の座標を $x = (x_1, x_2)'$ 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

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \times \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = T x \tag{142}$$

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 $x = T^{-1} y = T' y$ とすればよい。また、座標系を角度 θ' だけ回転する行列を S とすれば、 y をさらに θ' だけ回転したものは

$$z = S y = S(T x) = (S T) x \tag{143}$$

となるが、これは、行列 $S T$,

$$S T = \begin{bmatrix} \cos \theta' \cos \theta - \sin \theta' \sin \theta & -\cos \theta' \sin \theta - \sin \theta' \cos \theta \\ \sin \theta' \cos \theta + \cos \theta' \sin \theta & -\sin \theta' \sin \theta + \cos \theta' \cos \theta \end{bmatrix} \tag{144}$$

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 X ($p \times 2$) とすると、

$$Y = X T' \tag{145}$$

となる。

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

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \times \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = T' x \tag{146}$$

となる。

4.2 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 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

$$f_{r_i} = \sum_{j=1}^c f_{ij}, \quad \text{and} \quad f_{c_j} = \sum_{i=1}^r f_{ij}, \quad \text{and} \quad n = \sum_{i=1}^r \sum_{j=1}^c f_{ij} \tag{147}$$

$$p_{ij} = \frac{1}{n} f_{ij}, \quad \text{and} \quad p_{r_i} = \frac{1}{n} \sum_{j=1}^c f_{ij}, \quad \text{and} \quad p_{c_j} = \frac{1}{n} \sum_{i=1}^r f_{ij} \tag{148}$$

となるが、これを行列を用いて表すと

$$\mathbf{f}_r = \mathbf{F}\mathbf{1}_c \quad \text{and} \quad \mathbf{f}_c = (\mathbf{1}'_r \mathbf{F})' \quad \text{and} \quad n = \mathbf{1}'_r \mathbf{F}\mathbf{1} \quad (149)$$

$$\mathbf{P} = \frac{1}{n} \mathbf{F} \quad \text{and} \quad \mathbf{p}_r = \frac{1}{n} \mathbf{f}_r \quad \text{and} \quad \mathbf{p}_c = \frac{1}{n} \mathbf{f}_c \quad (150)$$

となる。二つの変数が独立であれば、

$$\mathbf{P} = \mathbf{p}_r \mathbf{p}'_c \quad (151)$$

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 \mathbf{C} 、ただし、 \mathbf{C} は $\sum_{i=1}^r c_{ij} = 1$, $j = 1, 2, \dots, c$ 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 \mathbf{R} 、ただし、 \mathbf{R} は $\sum_{j=1}^c r_{ij} = 1$, $i = 1, 2, \dots, r$ を満たす、は

$$\mathbf{C} = \mathbf{P} \text{Diag}(\mathbf{f}_c)^{-1} \quad \text{and} \quad \mathbf{R} = \text{Diag}(\mathbf{f}_r)^{-1} \mathbf{P} \quad (152)$$

となる。ただし、 $\text{Diag}(\mathbf{a})$ はベクトル \mathbf{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 \mathbf{C} と列に配置されたカテゴリの周辺確率 \mathbf{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

$$\mathbf{f}_r = \mathbf{C} \mathbf{f}_c \quad (= \mathbf{P} \text{Diag}(\mathbf{f}_c)^{-1} \mathbf{f}_c = \mathbf{P}\mathbf{1}) \quad (153)$$

となる。

4.3 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 $n_j, j = 1, 2, \dots, A$ とすると、

$$y_{ij} = \hat{y}_{ij} + \varepsilon_{ij}, \quad \hat{y}_{ij} = \mu + \alpha_j, \quad i = 1, 2, \dots, n_j, \quad j = 1, 2, \dots, A \quad (154)$$

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

$$\begin{pmatrix} \hat{y}_{11} \\ \hat{y}_{21} \\ \vdots \\ \hat{y}_{n_1 1} \\ \hat{y}_{12} \\ \hat{y}_{22} \\ \vdots \\ \hat{y}_{1A} \\ \vdots \\ \hat{y}_{n_{AA}} \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 0 & 0 & \dots & 1 \end{bmatrix} \times \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_A \end{pmatrix} \quad (155)$$

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、 $(\mathbf{X}'\mathbf{X})^{-1}$ 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 とい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

$$\begin{pmatrix} \hat{y}_{11} \\ \hat{y}_{21} \\ \vdots \\ \hat{y}_{n_1 1} \\ \hat{y}_{12} \\ \hat{y}_{22} \\ \vdots \\ \hat{y}_{1A} \\ \vdots \\ \hat{y}_{n_{AA}} \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & -1 & -1 & \dots & -1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & -1 & -1 & \dots & -1 \end{bmatrix} \times \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{A-1} \end{pmatrix} \quad (156)$$

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

$$y_{ijk} = \hat{y}_{ijk} + \varepsilon_{ijk}, \quad \hat{y}_{ijk} = \mu + \alpha_j + \beta_k, \quad i = 1, 2, \dots, n_{ij}, \quad j = 1, 2, \dots, A, \quad k = 1, 2, \dots, B \quad (157)$$

であるが、全ての n_{ij} を 1 とした場合のデザイン行列は

$$\begin{pmatrix} \hat{y}_{111} \\ \hat{y}_{112} \\ \vdots \\ \hat{y}_{11B} \\ \hat{y}_{121} \\ \hat{y}_{122} \\ \vdots \\ \hat{y}_{1A1} \\ \vdots \\ \hat{y}_{1AB} \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 1 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 0 & 0 & \dots & 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 1 \end{bmatrix} \times \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_A \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_B \end{pmatrix} \quad (158)$$

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 $X = \{\mathbf{1}, \mathbf{A}, \mathbf{B}\}$ 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 F は

$$F = AB' \quad (159)$$

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

$$\begin{pmatrix} \hat{y}_{111} \\ \hat{y}_{112} \\ \vdots \\ \hat{y}_{11B} \\ \hat{y}_{121} \\ \hat{y}_{122} \\ \vdots \\ \hat{y}_{1A1} \\ \vdots \\ \hat{y}_{1AB} \end{pmatrix} = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 0 & \dots & 0 & -1 & -1 & \dots & -1 \\ 1 & 0 & 1 & \dots & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & -1 & -1 & \dots & -1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & -1 & -1 & \dots & -1 & -1 & -1 & \dots & -1 \end{bmatrix} \times \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{A-1} \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{B-1} \end{pmatrix} \quad (160)$$

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 X の2つの列 $\mathbf{x}_{(1+j)}, \mathbf{x}_{(1+A+k)}$ 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。

4.4 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 $\mathbf{x}_j, j = 1, 2, \dots, N$ とすれば、

$$d_{ij}^2 = \sum_{a=1}^q (x_{ia} - x_{ja})^2 = (\mathbf{x}_i - \mathbf{x}_j)'(\mathbf{x}_i - \mathbf{x}_j) \quad (161)$$

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 D^2 とすれば、

$$D^2 = \text{diag}(\mathbf{X}\mathbf{X}')\mathbf{1}\mathbf{1}' - 2\mathbf{X}\mathbf{X}' + \mathbf{1}\mathbf{1}'\text{diag}(\mathbf{X}\mathbf{X}') \quad (162)$$

と書くことが出来る。

4.5 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

$$P = \begin{bmatrix} .5 & .1 & .1 \\ .3 & .6 & .1 \\ .2 & .3 & .8 \end{bmatrix} \quad (163)$$

で与えられているとする。この推移確率行列 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、 $\mathbf{1}'P = \mathbf{1}'$ 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 p_0 、 $\mathbf{1}'p_0 = 1$ であったとすると、時点を重ねるに連れその割合は

$$p_1 = Pp_0, \quad p_2 = Pp_1 = P^2p_0, \quad \dots \quad (164)$$

と変化していき、最終的には

$$P^\infty = \begin{bmatrix} .16667 & .16667 & .16667 \\ .26667 & .26667 & .26667 \\ .56667 & .56667 & .56667 \end{bmatrix} \quad (165)$$

を得る。

4.6 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、 S^ℓ の (i, j) 要素が 0 であるということは、間に $\ell-1$ 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 $\ell-1$ 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 $d_{ij}^{(\ell)}$ で表すと以下ようになる。

$$D^1 = S, \quad d_{ij}^{(\ell)} = \sum_{k \neq i, j} s_{ik} d_{kj}^{(\ell-1)}, \quad (166)$$

$$D = SD^{(\ell-1)} - \text{diag}(S)D^{(\ell-1)} - S\text{diag}(D^{(\ell-1)}) + \text{diag}(S)\text{diag}(D^{(\ell-1)}) \quad (167)$$

4.7 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 \mathbf{x} の要素の平均（スカラー量）は

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} = \frac{1}{N} \mathbf{1}' \mathbf{x} = \frac{1}{N} \mathbf{x}' \mathbf{1} \quad (168)$$

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 $\{x_i - \bar{x}\}$ は

$$\mathbf{x} - \mathbf{1} \left(\frac{1}{N} \mathbf{1}' \mathbf{x} \right) = \mathbf{x} - \frac{1}{N} \mathbf{1} \mathbf{1}' \mathbf{x} = \left(\mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}' \right) \mathbf{x} = \mathbf{J} \mathbf{x} \quad (169)$$

と書ける。この $N \times N$ の行列 $\mathbf{J} = \mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}'$ は **中心化演算子** (column centering operator) と呼ばれ $\mathbf{J} \mathbf{J} = \mathbf{J}$ 、 $\mathbf{J} \mathbf{1} = \mathbf{0}$ 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。

$$\text{var}(\mathbf{x}) = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N} = \frac{1}{N} (\mathbf{J} \mathbf{x})' \mathbf{J} \mathbf{x} = \frac{1}{N} \mathbf{x}' \mathbf{J} \mathbf{x} = \frac{1}{N} \|\mathbf{J} \mathbf{x}\|^2 = [\text{std}(\mathbf{x})]^2 \quad (170)$$

と定義する。また、 \mathbf{y} を \mathbf{x} と同じサイズのベクトルとすれば、 \mathbf{x} と \mathbf{y} の共分散を

$$\text{cov}(\mathbf{x}, \mathbf{y}) = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{N} = \frac{1}{N} \mathbf{x}' \mathbf{J} \mathbf{y} \quad (171)$$

相関係数を

$$\text{corr}(\mathbf{x}, \mathbf{y}) = \frac{\text{cov}(\mathbf{x}, \mathbf{y})}{\text{std}(\mathbf{x}) \text{std}(\mathbf{y})} \quad (172)$$

と定義する。すなわち、 $var(\cdot)$, $cov(\cdot, \cdot)$ 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。

$$\overline{a + b\mathbf{x}} = a + b\bar{\mathbf{x}} \quad (173)$$

$$var(a + b\mathbf{x}) = b^2 var(\mathbf{x}) \quad (174)$$

$$cov(a + b\mathbf{x}, c + d\mathbf{y}) = bd cov(\mathbf{x}, \mathbf{y}) \quad (175)$$

$$cov(\mathbf{x}, \mathbf{x}) = var(\mathbf{x}) \quad (176)$$

$$corr(a + b\mathbf{x}, c + d\mathbf{y}) = \text{sign}(bd) corr(\mathbf{x}, \mathbf{y}) \quad (177)$$

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 \mathbf{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。

$$\bar{\mathbf{X}} = \left\{ \frac{1}{N} \sum_{i=1}^N x_{ij} \right\} = \{\bar{\mathbf{x}}_{(j)}\} = \frac{1}{N} (\mathbf{1}' \mathbf{X})' = \frac{1}{N} \mathbf{X}' \mathbf{1} \quad (178)$$

従って、各列ごとの平均偏差ベクトルからなる行列は \mathbf{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

$$\mathbf{D}(\mathbf{X}) = \left\{ \frac{1}{N} \sum_{i=1}^N (x_{ij} - \bar{\mathbf{x}}_{(j)})(x_{ik} - \bar{\mathbf{x}}_{(k)}) \right\} = \frac{1}{N} (\mathbf{JX})' \mathbf{JX} = \frac{1}{N} \mathbf{X}' \mathbf{JX} \quad (179)$$

と書ける。次に、 $\text{diag}(\mathbf{A})$ で、正方行列 \mathbf{A} の非対角部分をゼロとおいた対角行列を、また、 \mathbf{A} が非負の要素を持つ対角行列の場合、 $(\mathbf{A})^{\frac{1}{2}}$ で $\mathbf{A} = (\mathbf{A})^{\frac{1}{2}} (\mathbf{A})^{\frac{1}{2}}$ 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、

$$\left[(\text{diag}(\mathbf{D}(\mathbf{X})))^{\frac{1}{2}} \right]^{-1} = (\text{diag}(\mathbf{D}(\mathbf{X})))^{-\frac{1}{2}} \quad (180)$$

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 $\mathbf{Z} = \{(x_{ij} - \bar{\mathbf{x}}_{(j)}) / \text{std}(\mathbf{x}_{(j)})\}$ は

$$\mathbf{Z} = \left(\mathbf{X} - \frac{1}{N} \mathbf{1} \bar{\mathbf{X}}' \right) (\text{diag}(\mathbf{D}(\mathbf{X})))^{-\frac{1}{2}} \quad (181)$$

と書ける。

また、 \mathbf{Y} を $N \times q$ の行列とすれば、 \mathbf{X} と \mathbf{Y} の共分散行列 ($p \times q$) は

$$\mathbf{C}(\mathbf{X}, \mathbf{Y}) = \left\{ \frac{1}{N} \sum_{i=1}^N (x_{ij} - \bar{\mathbf{x}}_{(j)})(y_{ik} - \bar{\mathbf{y}}_{(k)}) \right\} = \frac{1}{N} (\mathbf{JX})' \mathbf{JY} = \frac{1}{N} \mathbf{X}' \mathbf{JY} \quad (182)$$

相関係数行列は

$$\mathbf{R}(\mathbf{X}, \mathbf{Y}) = \left\{ \frac{1}{N} \frac{\sum_{i=1}^N (x_{ij} - \bar{\mathbf{x}}_{(j)})(y_{ik} - \bar{\mathbf{y}}_{(k)})}{\text{std}(\mathbf{x}_{(j)}) \text{std}(\mathbf{x}_{(k)})} \right\} \quad (183)$$

$$= \frac{1}{N} (\mathbf{JX}) \left[(\text{diag}(\mathbf{D}(\mathbf{X})))^{\frac{1}{2}} \right]^{-1} \mathbf{JY} \left[(\text{diag}(\mathbf{D}(\mathbf{Y})))^{\frac{1}{2}} \right]^{-1} \quad (184)$$

$$= \left[(\text{diag}(\mathbf{D}(\mathbf{X})))^{\frac{1}{2}} \right]^{-1} \mathbf{C}(\mathbf{X}, \mathbf{Y}) \left[(\text{diag}(\mathbf{D}(\mathbf{Y})))^{\frac{1}{2}} \right]^{-1} \quad (185)$$

と書くことが出来る。なお、

$$\mathbf{C}(\mathbf{X}, \mathbf{X}) = \mathbf{D}(\mathbf{X}) \quad (186)$$

$\mathbf{D}(\cdot)$, $\mathbf{C}(\cdot, \cdot)$ および $\mathbf{R}(\cdot, \cdot)$ は以下のような性質を持つ。

$$\overline{\mathbf{1a}' + \mathbf{XB}'} = \mathbf{a} + \mathbf{B}\bar{\mathbf{X}} \quad (187)$$

$$\mathbf{D}(\mathbf{1a}' + \mathbf{XB}') = \mathbf{BD}(\mathbf{X})\mathbf{B}' \quad (188)$$

$$\mathbf{C}(\mathbf{1a}' + \mathbf{XB}', \mathbf{1c}' + \mathbf{YD}') = \mathbf{BC}(\mathbf{X}, \mathbf{Y})\mathbf{D}' \quad (189)$$

$$\overline{\mathbf{X} + \mathbf{Y}} = \bar{\mathbf{X}} + \bar{\mathbf{Y}} \quad (190)$$

$$\mathbf{C}(\mathbf{X}, \mathbf{Y} + \mathbf{Z}) = \mathbf{C}(\mathbf{X}, \mathbf{Y}) + \mathbf{C}(\mathbf{X}, \mathbf{Z}) \quad (191)$$

$$\mathbf{D}(\mathbf{Y} + \mathbf{X}) = \mathbf{D}(\mathbf{Y}) + \mathbf{D}(\mathbf{X}) + \mathbf{C}(\mathbf{Y}, \mathbf{X}) + \mathbf{C}(\mathbf{X}, \mathbf{Y}) \quad (192)$$

4.8 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 $\{x_1, x_2, \dots, x_m\}$ とする。ただし、 $x_k, k = 1, 2, \dots, m$ は数値である。また、 X がそれぞれの値を取る確率を $\Pr(X = x_k) = \pi_k, k = 1, 2, \dots, m$ 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

$$\mathcal{E}(X) = \sum_{k=1}^m x_k \pi_k = \mathbf{x}' \boldsymbol{\pi} \quad (193)$$

$$\text{var}(X) = \sum_{k=1}^m (x_k - \mathcal{E}(X))^2 \pi_k = \sum_{k=1}^m x_k^2 \pi_k + \left(\sum_{k=1}^m x_k \pi_k \right)^2 \quad (194)$$

$$= \mathbf{x}' \text{Diag}(\boldsymbol{\pi}) \mathbf{x} - (\mathbf{x}' \boldsymbol{\pi})^2 = (\mathbf{x} \odot \mathbf{x})' \boldsymbol{\pi} - (\mathbf{x}' \boldsymbol{\pi})^2 \quad (195)$$

と書ける。ただし、 \mathbf{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、

$$\mathcal{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx \quad (196)$$

$$\text{var}(X) = \int_{-\infty}^{+\infty} (x - \mathcal{E}(X))^2 f(x) dx \quad (197)$$

$$= \int_{-\infty}^{+\infty} x^2 f(x) dx - \mathcal{E}(X)^2 \quad (198)$$

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 $\{x_1, x_2, \dots, x_{mx}\}$ Y の値域を $\{y_1, y_2, \dots, y_{my}\}$ とし、 $\Pr(X = x_k, Y = y_j) = \pi_{kj}, k = 1, 2, \dots, mx, j = 1, 2, \dots, my,$
 $\sum_{k=1}^{mx} \sum_{j=1}^{my} \pi_{kj} = 1$ とする。

$$\text{cov}(X, Y) = \sum_{k=1}^{mx} \sum_{j=1}^{my} (x_k - \mathcal{E}(X))(y_j - \mathcal{E}(Y)) \pi_{kj} \quad (199)$$

$$= \sum_{k=1}^{mx} \sum_{j=1}^{my} x_k y_j \pi_{kj} - \mathcal{E}(X) \mathcal{E}(Y) \quad (200)$$

$$= \mathbf{x}' \boldsymbol{\Pi} \mathbf{y} - \mathbf{x}' \boldsymbol{\pi} \mathbf{y}' \boldsymbol{\pi} \quad (201)$$

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、

$$\Pr(X = x_i = \boldsymbol{\Pi}[i+]), \quad \text{and} \quad \Pr(Y = y_j = \boldsymbol{\Pi}[+j]) \quad (202)$$

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、

$$\text{cov}(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mathcal{E}(X))(y - \mathcal{E}(Y)) f(x, y) dy dx \quad (203)$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f(x, y) dy dx - \mathcal{E}(X) \mathcal{E}(Y) \quad (204)$$

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 \mathbf{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

$$\boldsymbol{\mathcal{E}}(\mathbf{x}) \quad (r_x \times 1) \quad \text{and} \quad \boldsymbol{\mathcal{D}}(\mathbf{x}) = \boldsymbol{\mathcal{E}}(\mathbf{x} \mathbf{x}') - \boldsymbol{\mathcal{E}}(\mathbf{x}) \boldsymbol{\mathcal{E}}(\mathbf{x})' \quad (r_x \times r_x) \quad (205)$$

と記す。ただし、 $\boldsymbol{\mathcal{E}}(\mathbf{x})$ 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、 $\boldsymbol{\mathcal{D}}(\mathbf{x})$ はその ij 要素として x_i と x_j の共分散 $\text{cov}(x_i, x_j)$ 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 \mathbf{a} と $r_y \times r_x$ の行列 \mathbf{B} を用いて

$$\mathbf{y} = \mathbf{a} + \mathbf{B} \mathbf{x} \quad (206)$$

と変換したものの \mathbf{y} の期待値と分散共分散行列は以下ようになる。

$$\mathcal{E}(\mathbf{y}) = \mathbf{a} + \mathbf{B}\mathcal{E}(\mathbf{x}) \quad \text{and} \quad \mathcal{D}(\mathbf{y}) = \mathbf{B}\mathcal{D}(\mathbf{x})\mathbf{B}' \quad (207)$$

また、別の確率変数ベクトル \mathbf{y}, \mathbf{z} との間の共分散行列を $\mathcal{Cov}(\mathbf{x}, \mathbf{y}), \mathcal{Cov}(\mathbf{x}, \mathbf{z})$ とすれば、

$$\mathcal{E}(\mathbf{B}\mathbf{x} + \mathbf{A}\mathbf{y}) = \mathbf{B}\mathcal{E}(\mathbf{x}) + \mathbf{A}\mathcal{E}(\mathbf{y}), \quad \mathcal{Cov}(\mathbf{B}\mathbf{x}, \mathbf{A}\mathbf{y}) = \mathbf{B}\mathcal{Cov}(\mathbf{x}, \mathbf{y})\mathbf{A}' \quad (208)$$

$$\mathcal{E}(\mathbf{x} + \mathbf{y}) = \mathcal{E}(\mathbf{x}) + \mathcal{E}(\mathbf{y}), \quad \mathcal{Cov}(\mathbf{x}, \mathbf{y} + \mathbf{z}) = \mathcal{Cov}(\mathbf{x}, \mathbf{y}) + \mathcal{Cov}(\mathbf{x}, \mathbf{z}) \quad (209)$$

$$\mathcal{D}(\mathbf{x} + \mathbf{y}) = \mathcal{D}(\mathbf{x}) + \mathcal{D}(\mathbf{y}) + \mathcal{Cov}(\mathbf{x}, \mathbf{y}) + \mathcal{Cov}(\mathbf{y}, \mathbf{x}) \quad (210)$$

なお、期待値の場合は確率変数を要素として持つ行列 \mathbf{X} に関してベクトルの場合と同様に

$$\mathcal{E}(\mathbf{A} + \mathbf{B}\mathbf{X}) = \mathbf{A} + \mathbf{B}\mathcal{E}(\mathbf{X}) \quad (211)$$

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 $\mathbf{x}'\mathbf{A}\mathbf{x}$ の期待値は、トレースも線形演算子であることから、

$$\mathcal{E}(\mathbf{x}'\mathbf{A}\mathbf{x}) = \mathcal{E}(\text{tr}(\mathbf{A}\mathbf{x}\mathbf{x}')) = \text{tr}(\mathcal{E}(\mathbf{A}\mathbf{x}\mathbf{x}')) = \text{tr}(\mathbf{A}\mathcal{E}(\mathbf{x}\mathbf{x}')) = \text{tr}(\mathbf{A}(\mathcal{D}(\mathbf{x}) + \mathcal{E}(\mathbf{x})\mathcal{E}(\mathbf{x}')))) \quad (212)$$

$$= \text{tr}(\mathbf{A}\mathcal{D}(\mathbf{x})) + \text{tr}(\mathbf{A}\mathcal{E}(\mathbf{x})\mathcal{E}(\mathbf{x}')) = \mathcal{E}(\mathbf{x}')\mathbf{A}\mathcal{E}(\mathbf{x}) + \text{tr}(\mathbf{A}\mathcal{D}(\mathbf{x})) \quad (213)$$

となる。

4.9 因子分析

因子分析モデルは

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f} + \boldsymbol{\varepsilon} \quad (214)$$

$$\mathcal{E}(\mathbf{f}) = \mathbf{0}, \quad \mathcal{D}(\mathbf{f}) = \boldsymbol{\Phi}, \quad \mathcal{E}(\boldsymbol{\varepsilon}) = \mathbf{0}, \quad \mathcal{D}(\boldsymbol{\varepsilon}) = \boldsymbol{\Psi}, \quad \mathcal{Cov}(\mathbf{f}, \boldsymbol{\varepsilon}) = \mathbf{O} \quad (215)$$

と書くことができるが、これから \mathbf{y} の分散共分散行列

$$\mathcal{D}(\mathbf{y}) = \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Psi} \quad (216)$$

が導ける。

4.10 多変量正規分布の密度関数

\mathbf{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

$$f(\mathbf{y} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(\sqrt{2\pi})^p \sqrt{|\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})\right) \quad (217)$$

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 $\mathbf{y}_i, i = 1, 2, \dots, n$ 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

$$L = \prod_{i=1}^n f(\mathbf{y}_i | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (218)$$

$$= \left(\frac{1}{\sqrt{|\boldsymbol{\Sigma}|}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}_i - \boldsymbol{\mu})\right) \quad (219)$$

$$\propto \exp\left(-\frac{1}{2}\sum_{i=1}^n \text{tr}\left(\boldsymbol{\Sigma}^{-1}(\mathbf{y}_i - \boldsymbol{\mu})(\mathbf{y}_i - \boldsymbol{\mu})'\right)\right) \quad (220)$$

$$= \exp\left(-\frac{1}{2}\text{tr}\left(\boldsymbol{\Sigma}^{-1}\sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\mu})(\mathbf{y}_i - \boldsymbol{\mu})'\right)\right) \quad (221)$$

$$= \exp\left(-\frac{1}{2}\text{tr}\left(\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')'(\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}')\right)\right) \quad (222)$$

$$= \exp\left(-\frac{1}{2}\text{tr}\left(\boldsymbol{\Sigma}^{-1}\left(n(\boldsymbol{\mu} - \bar{\mathbf{Y}})(\boldsymbol{\mu} - \bar{\mathbf{Y}})' + (\mathbf{Y} - \mathbf{1}\bar{\mathbf{Y}})'(\mathbf{Y} - \mathbf{1}\bar{\mathbf{Y}})\right)\right)\right) \quad (223)$$

$$= \exp\left(-\frac{n}{2}\text{tr}\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \bar{\mathbf{Y}})(\boldsymbol{\mu} - \bar{\mathbf{Y}})'\right)\right) \times \exp\left(-\frac{n}{2}\text{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{C}(\mathbf{Y})\right)\right) \quad (224)$$

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。

4.11 多項分布の確率

$U_k, k = 0, 1, 2, \dots, m$ をたがいに背反な $m + 1$ 個の事象 $A_k, k = 0, 1, 2, \dots, m$ 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 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 $\pi = \{\pi_k\}$ で表すと、 $\sum_{k=0}^m \pi_k = 1$ である。このと

$$\mathcal{E}(U_k) = \pi_k, \text{var}(U_k) = \pi_k^2 - \pi_k, \text{cov}(U_k, U_j) = -\pi_k \pi_j \quad (225)$$

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

$$\mathcal{E}(U) = \pi, \mathcal{D}(U) = \text{Diag}(\pi) - \pi \pi' \quad (226)$$

と書ける。

5 その他の演算子

5.1 分割行列の演算

行列 A を分割し小行列 $A_{ik} r_{A_{ik}} \times c_{A_{ik}}, i = 1, 2, \dots, r_A, k = 1, 2, \dots, c_A$ を作る。同様に、行列 B を分割し小行列 $B_{kj} r_{B_{kj}} \times c_{B_{kj}}, k = 1, 2, \dots, r_B, j = 1, 2, \dots, c_B$ 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 A は $r_A \times c_A$ の行列、行列 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、 $r_A = r_B, c_A = c_B$ 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、

$$(A + B)_{ik} = A_{ik} + B_{ik} \quad (227)$$

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 A と 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。

$$AB = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1c_A} \\ A_{21} & A_{22} & \dots & A_{2c_A} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & A_{i2} & \dots & A_{ic_A} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r_A 1} & A_{r_A 2} & \dots & A_{r_A c_A} \end{bmatrix} \times \begin{bmatrix} B_{11} & B_{12} & \dots & B_{1j} & \dots & B_{1c_B} \\ B_{21} & B_{22} & \dots & B_{2j} & \dots & B_{2c_B} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ B_{r_B 1} & B_{r_B 2} & \dots & B_{r_B j} & \dots & B_{r_B c_B} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1j} & \dots & C_{1c_C} \\ C_{21} & C_{22} & \dots & C_{2j} & \dots & C_{2c_C} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C_{i1} & C_{i2} & \dots & C_{ij} & \dots & C_{ic_C} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ C_{r_C 1} & C_{r_C 2} & \dots & C_{r_C j} & \dots & C_{r_C c_C} \end{bmatrix} \quad (228)$$

ただし

$$(AB)_{ij} = C_{ij} = \sum_{k=1}^{c_{A_{ik}}} A_{ik} B_{kj} \quad (229)$$

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

$$A' = \begin{bmatrix} A'_{11} & A'_{21} & \dots & A'_{r_A 1} \\ A'_{12} & A'_{22} & \dots & A'_{r_A 2} \\ \vdots & \vdots & \ddots & \vdots \\ A'_{1j} & A'_{2j} & \dots & A'_{r_A j} \\ \vdots & \vdots & \ddots & \vdots \\ A'_{1c_A} & A'_{2c_A} & \dots & A'_{r_A c_A} \end{bmatrix} \quad (230)$$

分割行列のトレース

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^r \text{tr}(\mathbf{A}_{ii}) \quad (231)$$

分割行列の逆行列

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \\ -\mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22.1}^{-1} \end{bmatrix} \quad (232)$$

$$= \begin{bmatrix} \mathbf{A}_{11.2}^{-1} & -\mathbf{A}_{11.2}^{-1} \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{A}_{11.2}^{-1} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{A}_{11.2}^{-1} \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \end{bmatrix} \quad (233)$$

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

$$\mathbf{A}_{ii,j} = \mathbf{A}_{ii} - \mathbf{A}_{ij} \mathbf{A}_{jj}^{-1} \mathbf{A}_{ji} \quad (234)$$

と置いている。

分割行列の逆行列の公式の証明 (232) 式を用いて $\mathbf{A}^{-1} \mathbf{A}$ と $\mathbf{A} \mathbf{A}^{-1}$ を計算し、それらが単位行列になることを示す。
 $\mathbf{A}^{-1} \mathbf{A}$ の各要素：

$$\left[\mathbf{A}^{-1} \mathbf{A} \right]_{11} = \left(\mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \right) \mathbf{A}_{11} - \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \quad (235)$$

$$= \mathbf{A}_{11}^{-1} \mathbf{A}_{11} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{11} - \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \quad (236)$$

$$= \mathbf{I} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} - \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} = \mathbf{I} \quad (237)$$

$$\left[\mathbf{A}^{-1} \mathbf{A} \right]_{12} = \left(\mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \right) \mathbf{A}_{12} - \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{22} \quad (238)$$

$$= \mathbf{A}_{11}^{-1} \mathbf{A}_{12} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} - \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{22} \quad (239)$$

$$= \mathbf{A}_{11}^{-1} \mathbf{A}_{12} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \left(\mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} - \mathbf{A}_{22} \right) \quad (240)$$

$$= \mathbf{A}_{11}^{-1} \mathbf{A}_{12} - \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{22.1} \quad (241)$$

$$= \mathbf{A}_{11}^{-1} \mathbf{A}_{12} - \mathbf{A}_{11}^{-1} \mathbf{A}_{12} = \mathbf{O} \quad (242)$$

$$\left[\mathbf{A}^{-1} \mathbf{A} \right]_{21} = \left(-\mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \right) \mathbf{A}_{11} + \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \quad (243)$$

$$= -\mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{11} + \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \quad (244)$$

$$= -\mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} + \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} = \mathbf{O} \quad (245)$$

$$\left[\mathbf{A}^{-1} \mathbf{A} \right]_{22} = \left(-\mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \right) \mathbf{A}_{12} + \mathbf{A}_{22.1}^{-1} \mathbf{A}_{22} \quad (246)$$

$$= -\mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} + \mathbf{A}_{22.1}^{-1} \mathbf{A}_{22} \quad (247)$$

$$= \mathbf{A}_{22.1}^{-1} \left(-\mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} + \mathbf{A}_{22} \right) \quad (248)$$

$$= \mathbf{A}_{22.1}^{-1} \mathbf{A}_{22.1} = \mathbf{I} \quad (249)$$

$\mathbf{A} \mathbf{A}^{-1}$ の各要素：

$$\left[\mathbf{A} \mathbf{A}^{-1} \right]_{11} = \mathbf{A}_{11} \left(\mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \right) - \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \quad (250)$$

$$= \mathbf{A}_{11} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} - \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \quad (251)$$

$$= \mathbf{I} + \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} - \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} = \mathbf{I} \quad (252)$$

$$[\mathbf{A}\mathbf{A}^{-1}]_{12} = -\mathbf{A}_{11} \left(\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \right) + \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \quad (253)$$

$$= -\mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} + \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} = \mathbf{O} \quad (254)$$

$$[\mathbf{A}\mathbf{A}^{-1}]_{21} = \mathbf{A}_{21} \left(\mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \right) - \mathbf{A}_{22} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \quad (255)$$

$$= \mathbf{A}_{21} \mathbf{A}_{11}^{-1} + \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} - \mathbf{A}_{22} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \quad (256)$$

$$= \mathbf{A}_{21} \mathbf{A}_{11}^{-1} + \left(\mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} - \mathbf{A}_{22} \right) \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \quad (257)$$

$$= \mathbf{A}_{21} \mathbf{A}_{11}^{-1} + (-\mathbf{A}_{22.1}) \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \quad (258)$$

$$= \mathbf{A}_{21} \mathbf{A}_{11}^{-1} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} = \mathbf{O} \quad (259)$$

$$[\mathbf{A}\mathbf{A}^{-1}]_{22} = -\mathbf{A}_{21} \left(\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \right) + \mathbf{A}_{22} \mathbf{A}_{22.1}^{-1} \quad (260)$$

$$= \left(-\mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} + \mathbf{A}_{22} \right) \mathbf{A}_{22.1}^{-1} \quad (261)$$

$$= \mathbf{A}_{22.1} \mathbf{A}_{22.1}^{-1} = \mathbf{I} \quad (262)$$

証明終わり

(232) 式 に関しても同様に、(234) 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。

5.2 ssq 演算子

$ssq(\mathbf{A})$ で行列もしくはベクトル \mathbf{A} の要素の2乗和を表す。

$$ssq(\mathbf{A}) = \text{tr} [\mathbf{A}'\mathbf{A}] = \|\mathbf{A}\|^2 \quad \text{and} \quad ssq(\mathbf{a}) = \mathbf{a}'\mathbf{a} = \|\mathbf{a}\|^2 \quad (263)$$

例えば、回帰分析の残差の2乗和は

$$ssq(\mathbf{Y} - \mathbf{X}\mathbf{B}) = \text{tr} [(\mathbf{Y} - \mathbf{X}\mathbf{B})'(\mathbf{Y} - \mathbf{X}\mathbf{B})] \quad (264)$$

また、重み付きの残差の2乗和は

$$ssq(\mathbf{W}^{\frac{1}{2}}(\mathbf{Y} - \mathbf{X}\mathbf{B})) = \text{tr} ((\mathbf{Y} - \mathbf{X}\mathbf{B})'\mathbf{W}(\mathbf{Y} - \mathbf{X}\mathbf{B})) \quad (265)$$

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）。

$$(\mathbf{Y} - \mathbf{X}\mathbf{B})'\mathbf{W}(\mathbf{Y} - \mathbf{X}\mathbf{B}) = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'\mathbf{W}(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}) + (\mathbf{B} - \hat{\mathbf{B}})'\mathbf{X}'\mathbf{W}\mathbf{X}(\mathbf{B} - \hat{\mathbf{B}}) \quad (266)$$

ただし、 $\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{Y}$ 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

$$ay^2 + by + c = a \left(y + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a} \quad (267)$$

$$A(x - a)^2 + B(x - b)^2 = (A + B) \left(x - \frac{aA + bB}{A + B} \right)^2 + \frac{AB}{A + B} (a - b)^2 \quad (268)$$

のベクトル・行列版は、

$$\mathbf{x}'\mathbf{S}\mathbf{x} - 2\mathbf{c}'\mathbf{x} = (\mathbf{x} - \mathbf{d})'\mathbf{S}(\mathbf{x} - \mathbf{d}) - \mathbf{c}'\mathbf{S}\mathbf{c} \quad (269)$$

ただし、

$$\mathbf{d} = \mathbf{S}^{-1}\mathbf{c} \quad (270)$$

または

$$\text{tr}(\mathbf{X}'\mathbf{S}\mathbf{X}) - 2\text{tr}(\mathbf{C}'\mathbf{X}) = \text{tr}((\mathbf{X} - \mathbf{D})'\mathbf{S}(\mathbf{X} - \mathbf{D})) - \text{tr}(\mathbf{C}'\mathbf{S}\mathbf{C}) \quad (271)$$

ただし、

$$\mathbf{D} = \mathbf{S}^{-1}\mathbf{C} \quad (272)$$

となる。また、この平方完成の公式を使うと二つの \mathbf{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。

$$\begin{aligned} (\mathbf{x} - \mathbf{a})'\mathbf{C}(\mathbf{x} - \mathbf{a}) + (\mathbf{x} - \mathbf{b})'\mathbf{D}(\mathbf{x} - \mathbf{b}) &= (\mathbf{x} - (\mathbf{C} + \mathbf{D})^{-1}(\mathbf{C}\mathbf{a} + \mathbf{D}\mathbf{b}))'(\mathbf{C} + \mathbf{D})(\mathbf{x} - (\mathbf{C} + \mathbf{D})^{-1}(\mathbf{C}\mathbf{a} + \mathbf{D}\mathbf{b})) \\ &\quad + (\mathbf{a} - \mathbf{b})'\mathbf{C}(\mathbf{C} + \mathbf{D})^{-1}\mathbf{D}(\mathbf{a} - \mathbf{b}) \end{aligned} \quad (273)$$

$$\begin{aligned} &= (\mathbf{x} - (\mathbf{C} + \mathbf{D})^{-1}(\mathbf{C}\mathbf{a} + \mathbf{D}\mathbf{b}))'(\mathbf{C} + \mathbf{D})(\mathbf{x} - (\mathbf{C} + \mathbf{D})^{-1}(\mathbf{C}\mathbf{a} + \mathbf{D}\mathbf{b})) \\ &\quad + (\mathbf{a} - \mathbf{b})'(\mathbf{C}^{-1} + \mathbf{D}^{-1})^{-1}(\mathbf{a} - \mathbf{b}) \end{aligned} \quad (274)$$

なお、第2の形は \mathbf{C} 行列と \mathbf{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。

$$\begin{aligned} &(\mathbf{Y}_1 - \mathbf{X}_1\mathbf{B})'\mathbf{W}_1(\mathbf{Y}_1 - \mathbf{X}_1\mathbf{B}) + (\mathbf{Y}_2 - \mathbf{X}_2\mathbf{B})'\mathbf{W}_2(\mathbf{Y}_2 - \mathbf{X}_2\mathbf{B}) = \\ &(\mathbf{B} - \mathbf{A}^{-1}(\mathbf{X}'_1\mathbf{W}_1\mathbf{Y}_1 + \mathbf{X}'_2\mathbf{W}_2\mathbf{Y}_2))'\mathbf{A}(\mathbf{B} - \mathbf{A}^{-1}(\mathbf{X}'_1\mathbf{W}_1\mathbf{Y}_1 + \mathbf{X}'_2\mathbf{W}_2\mathbf{Y}_2)) \\ &\quad + (\mathbf{Y}_1 - \mathbf{X}_1\mathbf{B}_1)'\mathbf{X}'_1\mathbf{W}_1\mathbf{X}_1(\mathbf{Y}_1 - \mathbf{X}_1\mathbf{B}_1) \\ &\quad + (\mathbf{Y}_2 - \mathbf{X}_2\mathbf{B}_2)'\mathbf{X}'_2\mathbf{W}_2\mathbf{X}_2(\mathbf{Y}_2 - \mathbf{X}_2\mathbf{B}_2) \\ &\quad + (\mathbf{B}_1 - \mathbf{B}_2)'\mathbf{X}'_1\mathbf{W}_1\mathbf{X}_1\mathbf{A}^{-1}\mathbf{X}'_2\mathbf{W}_2\mathbf{X}_2(\mathbf{B}_1 - \mathbf{B}_2) \end{aligned} \quad (275)$$

$$\begin{aligned} &= \\ &(\mathbf{B} - \mathbf{A}^{-1}(\mathbf{X}'_1\mathbf{W}_1\mathbf{Y}_1 + \mathbf{X}'_2\mathbf{W}_2\mathbf{Y}_2))'\mathbf{A}(\mathbf{B} - \mathbf{A}^{-1}(\mathbf{X}'_1\mathbf{W}_1\mathbf{Y}_1 + \mathbf{X}'_2\mathbf{W}_2\mathbf{Y}_2)) \\ &\quad + (\mathbf{Y}_1 - \mathbf{X}_1\mathbf{B}_1)'\mathbf{X}'_1\mathbf{W}_1\mathbf{X}_1(\mathbf{Y}_1 - \mathbf{X}_1\mathbf{B}_1) \\ &\quad + (\mathbf{Y}_2 - \mathbf{X}_2\mathbf{B}_2)'\mathbf{X}'_2\mathbf{W}_2\mathbf{X}_2(\mathbf{Y}_2 - \mathbf{X}_2\mathbf{B}_2) \\ &\quad + (\mathbf{B}_1 - \mathbf{B}_2)'\mathbf{C}^{-1}(\mathbf{B}_1 - \mathbf{B}_2) \end{aligned} \quad (276)$$

ただし、

$$\mathbf{A} = \mathbf{X}'_1\mathbf{W}_1\mathbf{X}_1 + \mathbf{X}'_2\mathbf{W}_2\mathbf{X}_2 \quad (277)$$

$$\mathbf{B}_1 = (\mathbf{X}'_1\mathbf{W}_1\mathbf{X}_1)^{-1}\mathbf{X}'_1\mathbf{W}_1\mathbf{Y}_1 \quad (278)$$

$$\mathbf{B}_2 = (\mathbf{X}'_2\mathbf{W}_2\mathbf{X}_2)^{-1}\mathbf{X}'_2\mathbf{W}_2\mathbf{Y}_2 \quad (279)$$

$$\mathbf{C} = (\mathbf{X}'_1\mathbf{W}_1\mathbf{X}_1)^{-1} + (\mathbf{X}'_2\mathbf{W}_2\mathbf{X}_2)^{-1} \quad (280)$$

である。

5.3 Kronecker 積

\mathbf{A} を $ra \times ca$ の行列 \mathbf{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、

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \quad (281)$$

であるが、 \mathbf{A} と \mathbf{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。

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} \end{bmatrix} \quad (282)$$

すなわち、 \otimes 記号の左側にある(第1引数) \mathbf{A} 行列の各要素に、右側にある(第2引数) \mathbf{B} 行列を掛け合わせたものを、 \mathbf{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。

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} & a_{12} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \\ a_{21} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} & a_{22} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{11}b_{31} & a_{11}b_{32} & a_{12}b_{31} & a_{12}b_{32} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \\ a_{21}b_{31} & a_{21}b_{32} & a_{22}b_{31} & a_{22}b_{32} \end{bmatrix} \quad (283)$$

クロネッカー積は (282) 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、 $\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}$ であり、

$$\mathbf{B} \otimes \mathbf{A} = \begin{bmatrix} b_{11}\mathbf{A} & b_{12}\mathbf{A} \\ b_{21}\mathbf{A} & b_{22}\mathbf{A} \\ b_{31}\mathbf{A} & b_{32}\mathbf{A} \end{bmatrix} \quad (284)$$

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、 $\phi = [\phi_1, \phi_2, \phi_3]$ の時

$$\mathbf{I}_3 \odot \phi = \begin{bmatrix} \phi & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \phi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \phi \end{bmatrix} \quad \text{and} \quad \phi \odot \mathbf{I} = \begin{bmatrix} \phi_1 \mathbf{I}_3 \\ \phi_2 \mathbf{I}_3 \\ \phi_3 \mathbf{I}_3 \end{bmatrix} \quad (285)$$

$$\mathbf{1}_3 \otimes \phi = \begin{bmatrix} \phi \\ \phi \\ \phi \end{bmatrix} \quad \text{and} \quad \phi \otimes \mathbf{1}_3 = \begin{bmatrix} \phi_1 \mathbf{1}_3 \\ \phi_2 \mathbf{1}_3 \\ \phi_3 \mathbf{1}_3 \end{bmatrix} \quad (286)$$

また、この拡張として、 \mathbf{A} 行列が対角行列の場合、 $\mathbf{A} \otimes \mathbf{B}$ は \mathbf{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。

$$\mathbf{a} \otimes \mathbf{b}' = \mathbf{b}' \otimes \mathbf{a} = \mathbf{ab}' \quad (287)$$

$$(\lambda \mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (\lambda \mathbf{B}) = \lambda (\mathbf{A} \otimes \mathbf{B}) \quad (288)$$

$$(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C} \quad (289)$$

$$\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C} \quad (290)$$

$$\mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} \quad (291)$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BC}) \quad (292)$$

$$(\mathbf{A} \otimes \mathbf{B})' = \mathbf{A}' \otimes \mathbf{B}' \quad (293)$$

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} \quad (294)$$

$$(\mathbf{A} \otimes \mathbf{B})^- = \mathbf{A}^- \otimes \mathbf{B}^- \quad (295)$$

$$\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{A})\text{tr}(\mathbf{B}) \quad (296)$$

$$\text{rank}(\mathbf{A} \otimes \mathbf{B}) = \text{rank}(\mathbf{A})\text{tr}(\mathbf{B}) \quad (297)$$

\mathbf{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。

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} \mathbf{A}_{11} \otimes \mathbf{B} & \mathbf{A}_{12} \otimes \mathbf{B} & \dots & \mathbf{A}_{1c_A} \otimes \mathbf{B} \\ \mathbf{A}_{21} \otimes \mathbf{B} & \mathbf{A}_{22} \otimes \mathbf{B} & \dots & \mathbf{A}_{2c_A} \otimes \mathbf{B} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{A}_{i1} \otimes \mathbf{B} & \mathbf{A}_{i2} \otimes \mathbf{B} & \dots & \mathbf{A}_{ic_A} \otimes \mathbf{B} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{A}_{r_A1} \otimes \mathbf{B} & \mathbf{A}_{r_A2} \otimes \mathbf{B} & \dots & \mathbf{A}_{r_Ac_A} \otimes \mathbf{B} \end{bmatrix} \quad (298)$$

5.4 vec 演算子

$\text{vec}()$ 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、

$$\mathbf{A} = [\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \mathbf{a}_{(3)}] = \begin{pmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \end{pmatrix} \quad (299)$$

の場合、

$$\text{vec}(\mathbf{A}) = \begin{pmatrix} \mathbf{a}_{(1)} \\ \mathbf{a}_{(2)} \\ \mathbf{a}_{(3)} \end{pmatrix} \quad (300)$$

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

$$\text{vec}(\mathbf{A}') = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix}' \quad (301)$$

とすればよい。ただし、 \mathbf{a} が列ベクトルの場合は

$$\text{vec}(\mathbf{a}) = \text{vec}(\mathbf{a}') = \mathbf{a} \quad (302)$$

であり、またもう一つの列ベクトル \mathbf{b} を用いて

$$\text{vec}(\mathbf{ba}') = \mathbf{a} \otimes \mathbf{b} \quad (303)$$

という関係がある。

$r \times c$ の行列 \mathbf{A} から作られる二つのベクトル $\text{vec}(\mathbf{A})$ と $\text{vec}(\mathbf{A}')$ との間には転換行列 (commutation matrix) $\mathbf{K}(r, c)$ を用いて

$$\text{vec}(\mathbf{A}) = \mathbf{K}(c, r)\text{vec}(\mathbf{A}') \quad (304)$$

$$\text{vec}(\mathbf{A}') = \mathbf{K}(r, c)\text{vec}(\mathbf{A}) \quad (305)$$

$$(306)$$

という関係がある。たとえば、

$$\mathbf{K}(2, 3) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{K}(3, 2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (307)$$

であるが、見て分かる様に $\mathbf{K}(r, c) = \mathbf{K}(c, r)' = \mathbf{K}(c, r)^{-1}$ である。

$\text{vec}()$ 演算子は以下の性質を持つ。

$$\text{vec}(\lambda\mathbf{A}) = \lambda\text{vec}(\mathbf{A}) \quad \text{and} \quad \text{vec}(\mathbf{A} + \mathbf{B}) = \text{vec}(\mathbf{A}) + \text{vec}(\mathbf{B}) \quad (308)$$

行列の積に $\text{vec}()$ 演算子を作用させると以下の様になる。

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\text{vec}(\mathbf{B}) \quad (309)$$

$$\text{vec}(\mathbf{AB}) = (\mathbf{I} \otimes \mathbf{A})\text{vec}(\mathbf{B}) = (\mathbf{B}' \otimes \mathbf{A})\text{vec}(\mathbf{I}) = (\mathbf{B}' \otimes \mathbf{I})\text{vec}(\mathbf{A}) \quad (310)$$

$\text{vec}()$ 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。

$$\text{tr}(\mathbf{ABCD}) = \text{vec}(\mathbf{D}')'(\mathbf{C}' \otimes \mathbf{A})\text{vec}(\mathbf{B}) \quad (311)$$

$$= \text{vec}(\mathbf{D}')'(\mathbf{A} \otimes \mathbf{C}')\text{vec}(\mathbf{B}') \quad (312)$$

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。

$$\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCIA}) = \text{vec}(\mathbf{A}')'(\mathbf{I} \otimes \mathbf{B})\text{vec}(\mathbf{C}) \quad (313)$$

二つの行列の積のトレース (トレースが取れるためには \mathbf{A}' と \mathbf{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、

$$\text{tr}(\mathbf{AB}) = \text{vec}(\mathbf{A}')'\text{vec}(\mathbf{B}) \quad (314)$$

を得る。

5.5 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. $(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}$
2. $(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1}$
3. $(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{A}^{-1}$
4. $(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{B}^{-1} - \mathbf{B}^{-1}(\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}\mathbf{B}^{-1}$
5. $(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}$

6. $(A + B)^{-1} = A^{-1} - (A + B)^{-1}BA^{-1}$
7. $(A + B)^{-1} = B^{-1} - B^{-1}A(A + B)^{-1}$
8. $(A + B)^{-1} = B^{-1} - (A + B)^{-1}AB^{-1}$
9. $(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$
10. $(A + BCD)^{-1}B = A^{-1}B(C^{-1} + DA^{-1}B)^{-1}C^{-1}$
11. $D(A + BCD)^{-1} = C^{-1}(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$
12. $D(A + BCD)^{-1}B = C^{-1} - C^{-1}(C^{-1} + DA^{-1}B)^{-1}C^{-1}$

(註)

公式 1-4 は A と B を A^{-1} と B^{-1} で置き換える。

公式 5-8 は $(A + B)^{-1}$ を残す。

公式 11-14 は A と C を A^{-1} と C^{-1} で置き換える。

公式 1-4 は、たとえば、

$$(A^{-1} + B^{-1})^{-1} = A(A + B)^{-1}B \quad (315)$$

という形で用いる場合が多い。

公式 5-8 は、 $A(A + B)^{-1}$ の様な形の式を単純化するとき利用されることが多い。

$$A(A + B)^{-1} = I - B(A + B)^{-1} \quad \text{formula5} \quad (316)$$

$$(A + B)^{-1}A = I - (A + B)^{-1}B \quad \text{formula6} \quad (317)$$

$$B(A + B)^{-1} = I - A(A + B)^{-1} \quad \text{formula7} \quad (318)$$

$$(A + B)^{-1}B = I - (A + B)^{-1}A \quad \text{formula8} \quad (319)$$

である。

6 スイープ演算子と逆行列ならびに行列式

6.0.1 スイープ演算子の定義

スイープ演算子 (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 A が

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{ni} & \cdots & a_{nn} \end{bmatrix} \quad \begin{matrix} i \text{ 行} \\ \\ \\ \\ \\ i \text{ 列} \end{matrix} \quad (320)$$

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

$$\text{Swp}[A | i] = \begin{bmatrix} a_{11} - a_{1i}a_{i1}/a_{ii} & a_{12} - a_{1i}a_{i2}/a_{ii} & \cdots & -a_{1i}/a_{ii} & \cdots & a_{1n} - a_{1i}a_{in}/a_{ii} \\ a_{21} - a_{2i}a_{i1}/a_{ii} & a_{22} - a_{2i}a_{i2}/a_{ii} & \cdots & -a_{2i}/a_{ii} & \cdots & a_{2n} - a_{2i}a_{in}/a_{ii} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1}/a_{ii} & a_{i2}/a_{ii} & \cdots & 1/a_{ii} & \cdots & a_{in}/a_{ii} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} - a_{ni}a_{i1}/a_{ii} & a_{n2} - a_{ni}a_{i2}/a_{ii} & \cdots & -a_{ni}/a_{ii} & \cdots & a_{nn} - a_{ni}a_{in}/a_{ii} \end{bmatrix} \quad \begin{matrix} i \text{ 行} \\ \\ \\ \\ \\ i \text{ 列} \end{matrix} \quad (321)$$

すなわち行列 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。

このスワイプ演算子は $a_{ii} \neq 0$ の場合、可逆的

$$\text{Swp}[\text{Swp}[\mathbf{A} \mid i] \mid i] = \mathbf{A} \quad (322)$$

であり、また、各スワイ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

$$\text{Swp}[\text{Swp}[\mathbf{A} \mid i] \mid j] = \text{Swp}[\text{Swp}[\mathbf{A} \mid j] \mid i] \quad (323)$$

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ワイプを

$$\text{Swp}[\mathbf{A} \mid \{i, j, \dots, k\}] = \text{Swp}[\dots \text{Swp}[\text{Swp}[\mathbf{A} \mid i] \mid j] \dots \mid k] \quad (324)$$

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

スワイ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ワイプを行った結果

$$\text{Swp}[\mathbf{A} \mid \{1, 2, \dots, n\}] = \text{Swp}[\dots \text{Swp}[\text{Swp}[\mathbf{A} \mid 1] \mid 2] \dots \mid n] = \mathbf{A}^{-1} \quad (325)$$

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

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \times (a_{11}a_{22} - a_{12}a_{21})/a_{11} = a_{11}a_{22} - a_{21}a_{12} \quad (326)$$

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 \mathbf{Q} , \mathbf{R} とすれば、

$$(\mathbf{QAR})^{-1} = \mathbf{R}^{-1} \mathbf{A}^{-1} \mathbf{Q}^{-1} \quad (327)$$

であるが、 $\mathbf{Q} = \mathbf{P}$, $\mathbf{R} = \mathbf{P}'$ として並べ替え行列 \mathbf{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。

6.0.2 分割行列のスワイプ

$n \times n$ の行列 \mathbf{A} の分割を、

$$\mathbf{A} = \begin{matrix} & \begin{matrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1i} & \cdots & \mathbf{A}_{1G} \end{matrix} \\ \begin{matrix} \mathbf{A}_{21} \\ \vdots \\ \mathbf{A}_{i1} \\ \vdots \\ \mathbf{A}_{G1} \end{matrix} & \begin{matrix} \mathbf{A}_{22} & \cdots & \mathbf{A}_{2i} & \cdots & \mathbf{A}_{2G} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{A}_{i2} & \cdots & \mathbf{A}_{ii} & \cdots & \mathbf{A}_{iG} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{A}_{G2} & \cdots & \mathbf{A}_{Gi} & \cdots & \mathbf{A}_{GG} \end{matrix} \end{matrix} \quad \begin{matrix} \\ \\ \\ i \text{ 行} \\ \\ \\ \\ \\ i \text{ 列} \end{matrix} \quad (328)$$

とする。ただし、行列 \mathbf{A}_{ij} , $i = 1, 2, \dots, G$, $j = 1, 2, \dots, G$ 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 \mathbf{A}_{ii} によるスワイプ $\text{Swp}[\mathbf{A} \mid (i)]$ を

$$\text{Swp}[\mathbf{A} \mid (i)] = \begin{matrix} & \begin{matrix} \mathbf{A}_{11} - \mathbf{A}_{1i} \mathbf{A}_{ii}^{-1} \mathbf{A}_{i1} & \mathbf{A}_{12} - \mathbf{A}_{1i} \mathbf{A}_{ii}^{-1} \mathbf{A}_{i2} & \cdots & -\mathbf{A}_{1i} \mathbf{A}_{ii}^{-1} & \cdots & \mathbf{A}_{1G} - \mathbf{A}_{1i} \mathbf{A}_{ii}^{-1} \mathbf{A}_{iG} \end{matrix} \\ \begin{matrix} \mathbf{A}_{21} - \mathbf{A}_{2i} \mathbf{A}_{ii}^{-1} \mathbf{A}_{i1} \\ \vdots \\ \mathbf{A}_{i1}^{-1} \mathbf{A}_{i1} \\ \vdots \\ \mathbf{A}_{G1} - \mathbf{A}_{Gi} \mathbf{A}_{ii}^{-1} \mathbf{A}_{i1} \end{matrix} & \begin{matrix} \mathbf{A}_{22} - \mathbf{A}_{2i} \mathbf{A}_{ii}^{-1} \mathbf{A}_{i2} \\ \vdots \\ \mathbf{A}_{ii}^{-1} \mathbf{A}_{i2} \\ \vdots \\ \mathbf{A}_{G2} - \mathbf{A}_{Gi} \mathbf{A}_{ii}^{-1} \mathbf{A}_{i2} \end{matrix} & \cdots & \begin{matrix} -\mathbf{A}_{2i} \mathbf{A}_{ii}^{-1} \\ \vdots \\ \mathbf{A}_{ii}^{-1} \\ \vdots \\ -\mathbf{A}_{Gi} \mathbf{A}_{ii}^{-1} \end{matrix} & \cdots & \begin{matrix} \mathbf{A}_{2G} - \mathbf{A}_{2i} \mathbf{A}_{ii}^{-1} \mathbf{A}_{iG} \\ \vdots \\ \mathbf{A}_{ii}^{-1} \mathbf{A}_{iG} \\ \vdots \\ \mathbf{A}_{GG} - \mathbf{A}_{Gi} \mathbf{A}_{ii}^{-1} \mathbf{A}_{iG} \end{matrix} \end{matrix} \quad \begin{matrix} \\ \\ \\ i \text{ 行} \\ \\ \\ \\ \\ i \text{ 列} \end{matrix} \quad (329)$$

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weepに等しい。

$$\text{Swp}[\mathbf{A} \mid (i)] = \text{Swp} \left[\mathbf{A} \mid \left\{ \sum_{j=1}^{i-1} n_j + 1, \sum_{j=1}^{i-1} n_j + 2, \dots, \sum_{j=1}^i n_j \right\} \right] \quad (330)$$

たとえば、行列が

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \quad (331)$$

の様に 2 分割されている場合には

$$\text{Swp}[\mathbf{A} \mid (1)] = \begin{bmatrix} \mathbf{A}_{11}^{-1} & \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \\ -\mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \end{bmatrix} \quad (332)$$

$$\text{Swp}[\mathbf{A} \mid (2)] = \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21} & -\mathbf{A}_{12} \mathbf{A}_{22}^{-1} \\ \mathbf{A}_{22}^{-1} \mathbf{A}_{21} & \mathbf{A}_{22}^{-1} \end{bmatrix} \quad (333)$$

であり、行列が 3 分割されている場合には

$$\text{Swp}[\mathbf{A} \mid (1)] = \begin{bmatrix} \mathbf{A}_{11}^{-1} & \mathbf{A}_{11}^{-1} \mathbf{A}_{12} & \mathbf{A}_{11}^{-1} \mathbf{A}_{13} \\ -\mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} & \mathbf{A}_{23} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{13} \\ -\mathbf{A}_{31} \mathbf{A}_{11}^{-1} & \mathbf{A}_{32} - \mathbf{A}_{31} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} & \mathbf{A}_{33} - \mathbf{A}_{31} \mathbf{A}_{11}^{-1} \mathbf{A}_{13} \end{bmatrix} \quad (334)$$

$$\text{Swp}[\mathbf{A} \mid (2)] = \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21} & -\mathbf{A}_{12} \mathbf{A}_{22}^{-1} & \mathbf{A}_{13} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{23} \\ \mathbf{A}_{22}^{-1} \mathbf{A}_{21} & \mathbf{A}_{22}^{-1} & \mathbf{A}_{22}^{-1} \mathbf{A}_{23} \\ \mathbf{A}_{31} - \mathbf{A}_{32} \mathbf{A}_{22}^{-1} \mathbf{A}_{21} & -\mathbf{A}_{32} \mathbf{A}_{22}^{-1} & \mathbf{A}_{33} - \mathbf{A}_{32} \mathbf{A}_{22}^{-1} \mathbf{A}_{23} \end{bmatrix} \quad (335)$$

$$\text{Swp}[\mathbf{A} \mid (3)] = \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{13} \mathbf{A}_{33}^{-1} \mathbf{A}_{31} & \mathbf{A}_{12} - \mathbf{A}_{13} \mathbf{A}_{33}^{-1} \mathbf{A}_{32} & -\mathbf{A}_{13} \mathbf{A}_{33}^{-1} \\ \mathbf{A}_{21} - \mathbf{A}_{23} \mathbf{A}_{33}^{-1} \mathbf{A}_{31} & \mathbf{A}_{22} - \mathbf{A}_{23} \mathbf{A}_{33}^{-1} \mathbf{A}_{32} & -\mathbf{A}_{23} \mathbf{A}_{33}^{-1} \\ \mathbf{A}_{33}^{-1} \mathbf{A}_{31} & \mathbf{A}_{33}^{-1} \mathbf{A}_{32} & \mathbf{A}_{33}^{-1} \end{bmatrix} \quad (336)$$

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、

$$\text{Swp}[\text{Swp}[\mathbf{A} \mid (i)] \mid (i)] = \mathbf{A} \quad (337)$$

$$\text{Swp}[\text{Swp}[\mathbf{A} \mid (i)] \mid (j)] = \text{Swp}[\text{Swp}[\mathbf{A} \mid (j)] \mid (i)] \quad (338)$$

$$\text{Swp}[\mathbf{A} \mid \{(i), (j), \dots, (k)\}] = \text{Swp}[\dots \text{Swp}[\text{Swp}[\mathbf{A} \mid (i)] \mid (j)] \dots \mid (k)] \quad (339)$$

である。

6.0.3 sweep演算子を利用した様々な公式

sweep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

$$\mathbf{A}^{-1} = \text{Swp}[\mathbf{A} \mid \{(1), (2)\}] = \text{Swp}[\text{Swp}[\mathbf{A} \mid (1)] \mid (2)] = \text{Swp}[\text{Swp}[\mathbf{A} \mid (2)] \mid (1)] \quad (340)$$

という 3 通りの方法が考えられるが、この最後の二つのsweepの結果 $\text{Swp}[\text{Swp}[\mathbf{A} \mid (1)] \mid (2)]$ と $\text{Swp}[\text{Swp}[\mathbf{A} \mid (2)] \mid (1)]$ を書き下せば

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{A}_{22.1}^{-1} \\ -\mathbf{A}_{22.1}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22.1}^{-1} \end{bmatrix} \quad (341)$$

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11.2}^{-1} & -\mathbf{A}_{11.2}^{-1} \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{A}_{11.2}^{-1} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{A}_{11.2}^{-1} \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \end{bmatrix} \quad (342)$$

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 (residual sum of squares and cross-product matrix) を

$$\mathbf{A}_{ii,j} = \mathbf{A}_{ii} - \mathbf{A}_{ij} \mathbf{A}_{jj}^{-1} \mathbf{A}_{ji} \quad (343)$$

とおいている。

(341) 式 と (342) 式 を比べることにより残差積和行列の逆行列

$$\mathbf{A}_{ii,j}^{-1} = \mathbf{A}_{ii}^{-1} + \mathbf{A}_{ii}^{-1} \mathbf{A}_{ij} \mathbf{A}_{jj,i}^{-1} \mathbf{A}_{ji} \mathbf{A}_{ii}^{-1} \quad (344)$$

を得る。また、残差積和行列の逆行列公式において、

$$\mathbf{A}_{ii} \rightarrow \mathbf{A}, \mathbf{A}_{ij} \rightarrow \mathbf{B}, -\mathbf{A}_{jj}^{-1} \rightarrow \mathbf{C}, \mathbf{A}_{ji} \rightarrow \mathbf{D} \quad (345)$$

とおけば、Woodburry-Lindley-Smith の公式

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1} + \mathbf{DA}^{-1} \mathbf{B})^{-1} \mathbf{DA}^{-1} \quad (346)$$

を得る。また、 $\mathbf{A}_{12} = \mathbf{O}$ とおくと 下三角行列の逆行列

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{O} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} & \mathbf{O} \\ -\mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22}^{-1} \end{bmatrix} \quad (347)$$

を得る。

また、スワイ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、(331) 式 で定義された分割行列の逆行列を (340) 式 の $\text{Swp}[\text{Swp}[\mathbf{A} \mid (1) \mid (2)]]$ を用いて計算する際には、最初のピボットが \mathbf{A}_{11} 、その結果として得られる行列が (332) 式 であるから、2 番目のピボットは $\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}$ となる。したがって、この行列の行列式は、

$$|\mathbf{A}| = |\mathbf{A}_{11}| \times |\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}| \quad (348)$$

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

$$|\mathbf{A}| = |\mathbf{A}_{22}| \times |\mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21}| \quad (349)$$

となり、この二つは同値である。したがって、 $r \times c$ の行列 \mathbf{P} と $c \times r$ の行列 \mathbf{Q} に関して以下の公式を得る。

$$|\mathbf{I}_r - \mathbf{PQ}| = |\mathbf{I}_c - \mathbf{QP}| \quad (350)$$

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。

$$|\mathbf{AB}| = |\mathbf{A}| \times |\mathbf{B}| = |\mathbf{BA}| \quad (351)$$

$$|\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|} \quad (352)$$

$$|\lambda \mathbf{A}| = \lambda^r |\mathbf{A}| \quad \text{and} \quad |\mathbf{A}| = |\mathbf{A}'| \quad (353)$$

7 斉次方程式と一般逆行列

7.1 斉次方程式

連立 1 次方程式

$$\mathbf{Ax} = \mathbf{c} \quad (354)$$

が解を持つためには、係数行列 \mathbf{A} が正則である (特異でない) ことが必要である、そのときには、

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{c} \quad (355)$$

が解となる。係数行列 \mathbf{A} の逆行列が存在しない場合 (\mathbf{A} 行列が特異行列の場合) には、解は、不定 (解は一意には定まらないが、方程式を満たす \mathbf{x} は存在する) もしくは不能 (方程式を満たす \mathbf{x} は存在しない) となが、本節では解が存在しない場合は除外して、不定の場合を扱う。係数行列が特異行列であり、解が不定となる場合とは、もしも \mathbf{A} の第 i 行 (\mathbf{a}'_i) が残りの行 $\mathbf{A}_{(i)}$ を用いて

$$\mathbf{a}'_i = \mathbf{q}' \mathbf{A}_{(i)} \quad (356)$$

と書けるなら、 \mathbf{c} の第 i 要素も、残りの要素 $\mathbf{c}_{(i)}$ を用いて

$$c_i = \mathbf{q}' \mathbf{c}_{(i)} \quad (357)$$

と書ける場合である。

解が不定の場合、それらの解の全体は、任意のベクトル y を用いて

$$x^* = A^-c + (I - A^-A)y \quad (358)$$

と書くことが出来る。ただし、 A^- は

$$A = AA^-A \quad (359)$$

を満たす行列で 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。これ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、

$$Ax^* = A(A^-c + (I - A^-A)y) = AA^-c + A(I - A^-A)y \quad (360)$$

$$= AA^-c + Ay - AA^-Ay = AA^-c \quad (361)$$

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 c は、任意のベクトル z を用いて $c = Az$ と書けなければならない。(c が A の列空間の中にある。) したがって、

$$Ax^* = AA^-c = AA^-Az = Az = c \quad (362)$$

となる。

なお、斉次 (同時) 方程式 $Ax = 0$ が $x = 0$ 以外の解を持つためには A が特異でなければならないが、その場合の上記の公式において $c = 0$ とし解を求めればよい。

7.2 一般逆行列について

通常は、前節で述べたように

$$A = AA^-A \quad (363)$$

を満たす行列 A^- を一般逆行列 (g1 inverse, conditional inverse, generalized inverse) とよぶ。また、この条件に加えて

$$A^- = A^-AA^- \quad (364)$$

を満たす行列を反射型一般逆行列 (g2-inverse, reflexive generalized inverse) と呼ぶ。また、反射型一般逆行列のなかで、 AA^- と A^-A が対称という条件を付与したものをムーアペンローズ逆行列 (Moor-Penrose inverse) と呼ぶが、この行列は一意に定まり、特に A^+ と記される。なお A が正則な場合は $A^+ = A^- = A^{-1}$ である。

一般逆行列に関しては以下の性質が知られている。

ある行列の一般逆行列を転置したものは元の行列の転置の一般逆行列となっている。

$$(A')^- = (A^-)' \quad (365)$$

A が $m \times n$ の行列であり、そのランクが行または列数に一致する場合、

$$A'(AA')^{-1} \quad \text{if } \text{rank}(A) = m \text{ (行フルランク、右逆行列)} \quad (366)$$

$$(A'A)^{-1}A' \quad \text{if } \text{rank}(A) = n \text{ (列フルランク、左逆行列)} \quad (367)$$

は A の一般逆行列となる。

以下に示す行列は

$$AA^-, A^-A, I - AA^-, I - A^-A \quad (368)$$

対称かつ幂等である。

以下の 3 つの表現はそれぞれ A の一般逆行列であるが、(Graville)

$$A^-AA^- \quad (369)$$

$$A'(AA')^- \quad (370)$$

$$(A'A)^-A' \quad (371)$$

最初のふたつは反射型一般逆行列となる。

行列 B が列フルランク（階数が列数に等しい）であり、行列 C が行フルランク（階数が行数に等しい）の場合、

$$(BC)^{-} = C^{-}B^{-} \quad (372)$$

である。

任意の一般逆行列から

$$A^{+} = A'(AA')^{-}A(A'A)^{-}A' \quad (373)$$

でムーアペンローズ一般逆行列を作ることが出来る。

ムーアペンローズ一般逆行列のムーアペンローズ一般逆行列は元の行列となる。

$$(A^{+})^{+} = A \quad (374)$$

また、

$$(A'A)^{+} = A^{+}(A')^{+} \quad (375)$$

$$(A'A)^{+} = A^{+}(A')^{+} \quad (376)$$

$$(AA^{+})^{+} = AA^{+} \quad (377)$$

$$(A^{+}A)^{+} = A^{+}A \quad (378)$$

$$A(A'A)^{-}A' = AA^{+} \quad (379)$$

スイープ演算子を用いて逆行列を計算する際に、どのように並べ替えを行っても、ピボットが 0 となる場合があるが、これは行列が特異行列の場合である。その場合、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、上記の (363) 式と

$$A^{-} = A^{-}AA^{-} \quad (380)$$

を満たす一般行列 A^{-} を得る。

8 射影子行列***

どのような $N \times 1$ のベクトル x も、ある $N \times a$ の行列 A の列の列の線形結合として書ける部分 x_1 とそれに直交する部分 x_2 に分解することが出来る。すなわち、 $x = x_1 + x_2$, $x_1'x_2 = 0$ と書ける。このようなふたつの直交するベクトルは

$$x_1 = P(A)x, \quad P(A) = A(A'A)^{-}A' \quad \text{and} \quad x_2 = Q(A)x, \quad Q(A) = I - P(A) \quad (381)$$

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

4: 射影子行列

$$P(A)A = A$$

これを示すには、左辺の左から $A'A$ を掛けたものを、一般逆行列の性質を用いて以下の変形すると

$$A'AP(A)A = A'AA(A'A)^{-}A'A = AA'A \quad (382)$$

すなわち

$$A'A P(A)A = AA' A \quad (383)$$

となるが、等号成立条件より、このことは $P(A)A = A$ を示している。

この射影子行列は対称であり冪等 (idempotent) $P(A)^2 = P(A)$ である。また、定義から、 $P(A)Q_A = O$ となる。射影子行列 $P(A)$ の階数は 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。なお、ベクトル $\mathbf{1}_N$ の直交補空間への射影子行列は

$$J = I_N - \mathbf{1}_N(\mathbf{1}_N'\mathbf{1}_N)^{-1}\mathbf{1}_N' = I_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N' \quad (384)$$

となるが、これは先に見た中心化演算子である。

9 行列の固有値と特異値分解

a 次の正方行列 A に関して

$$Ap = \lambda p \quad (385)$$

を満たすスカラー λ を A の**固有値** または固有根 (eigen value, characteristic root)、ベクトル p を λ に対応する**固有ベクトル** (eigen vector) と呼ぶ。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 ($\|p\|^2 = 1$) に規準化したものを用いる。一般に、 a 次の正方行列の固有値は重根も含めれば a 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 Λ_q 、それに対応する固有ベクトルを横に並べた行列を P_q とすれば、上記の関係はまとめて

$$AP_q = P_q \Lambda_q \quad (386)$$

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 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。また、対称行列の階数はその行列のゼロでない固有値の個数に等しい。対称行列の規準化された固有ベクトルのうち、ゼロでない単根に対応するものは互いに直交する。すなわち、 Λ_q が 0 を含まないならば、

$$P_q' P_q = I \quad (387)$$

である。したがって、

$$P_q' A P_q = \Lambda_q \quad (388)$$

である。すなわち、行列 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る。

また、 A が正則の場合、(A の固有値に 0 が含まれない場合) a 個の全ての固有ベクトルを並べたものを P とすると、

$$P' P = I \quad \text{and} \quad P P' = I \quad (389)$$

である。したがって、

$$A = P \Lambda P' \quad (390)$$

と書けるが、これを A の固有分解 (スペクトル分解) (eigen decomposition) と呼ぶ。行列の固有分解を用いると、 A の逆行列は

$$A^{-1} = P \Lambda^{-1} P' \quad (391)$$

と書ける。また、

$$\text{tr}(A) = \text{tr}(P \Lambda P') = \text{tr}(\Lambda P' P) = \text{tr}(\Lambda) = \sum_{i=1}^a \lambda_i \quad (392)$$

である。

なお、一般に、 $P_q, \Lambda_q, q < a$ を用いた場合には

$$A \neq P_q \Lambda_q P_q' \quad (393)$$

であるが、左辺は右辺の階数 q の最小 2 乗近似、すなわち

$$\text{RSS} = \text{tr} \left[(A - B)' (A - B) \right], \text{rank}(B) = q \quad (394)$$

を最小とするものとなっている。

A の階数が r 、すなわち、 A のゼロでない固有値が全部で r 個しかない場合、

$$A = P_r \Lambda_r P_r', \quad P_r' P_r = I \quad (395)$$

と書くことが出来るが、これを対称行列 A の縮小された固有値分解と呼ぶ。なお、この場合、

$$AA = P_r \Lambda_r P_r' P_r' P_r \Lambda_r P_r' = P_r \Lambda_r^2 P_r' \quad (396)$$

であるから、対称行列の自分自身との積 (2 乗) AA の固有ベクトルは A の固有ベクトルに等しく、固有値はもとの固有値の 2 乗に等しくなる。

また、

$$\mathbf{A}^+ = \mathbf{P}_r \mathbf{\Lambda}_r^{-1} \mathbf{P}_r' \quad (397)$$

は \mathbf{A} の一般化逆行列 (ムーア・ペンローズ逆行列) になる。
 どのようなベクトル \mathbf{x} を用いても

$$\mathbf{x}' \mathbf{A} \mathbf{x} > 0 \quad (398)$$

であるような対称行列を **正定値行列** (または正値行列 positive definite 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、 $\mathbf{A}'\mathbf{A}$ 行列や $\mathbf{A}\mathbf{A}'$ 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。

今、 \mathbf{X} を $a \times b, b \leq a$ の階数 r の行列とし、 $\mathbf{X}'\mathbf{X}$ ならびに $\mathbf{X}\mathbf{X}'$ の縮小された固有分解を

$$\mathbf{X}'\mathbf{X} = \mathbf{P}\mathbf{\Lambda}^2\mathbf{P}', \mathbf{X}\mathbf{X}' = \mathbf{Q}\mathbf{\Lambda}^2\mathbf{Q}' \quad (399)$$

とすれば、

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{P}' \quad (400)$$

と書きあらわすことができる。これを行列 \mathbf{X} の特異値分解 (singular value decomposition) と呼ぶ。なお、 \mathbf{P} 等として、階数が $q, q \leq r$ のもの、 \mathbf{P}_q 等、を用いた場合、

$$\hat{\mathbf{X}} = \mathbf{Q}_q \mathbf{\Lambda}_q \mathbf{P}_q' \quad (401)$$

は

$$\text{RSS} = \sum_{i=1}^a \sum_{j=1}^b (x_{ij} - \hat{x}_{ij})^2 \quad (402)$$

を最小とする階数 q の行列となることが知られている。

階数 r の非負定値対称行列 \mathbf{A} はそれを階数 r の列フルランク行列 \mathbf{B} の積として

$$\mathbf{A} = \mathbf{B}\mathbf{B}' \quad (403)$$

の形に分解することが出来る。ただし、

$$\mathbf{B} = \mathbf{P}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{T} \quad (404)$$

であり、 $\mathbf{\Lambda}^{\frac{1}{2}}$ は \mathbf{A} の正の固有値の平方根を要素として持つ行列、 \mathbf{P} はそれに対応する固有ベクトル、また、 \mathbf{T} は任意の直交行列である。なお、この分解は任意に \mathbf{T} を選べるため一意ではない。また、

$$\mathbf{B} = \mathbf{P}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{P}' \quad (405)$$

で定義される行列は行列 \mathbf{A} の対称な平方根行列 (symmetric square root matrix) と呼ばれることがある。

10 行列の微分

いま、 $\mathbf{X} = [x_{ij}]$ を、 $a \times b$ の行列とし、 \mathbf{U} や \mathbf{V} の要素が $u_{ij} = g_{ij}(x_{ij}), v_{ij} = h_{ij}(x_{ij})$ という形で \mathbf{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 \mathbf{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、

$$\frac{\partial f}{\partial \mathbf{X}} = \left[\frac{\partial f}{\partial x_{ij}} \right] \quad (406)$$

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

$$\frac{\partial f}{\partial \mathbf{X}'} = \left(\frac{\partial f}{\partial \mathbf{X}} \right)' \quad (407)$$

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

$$\frac{\partial f}{\partial \mathbf{X}} = \text{diag} \left(\frac{\partial f}{\partial \mathbf{X}} \right) \quad \text{ただし } \mathbf{X} \text{ は対角行列} \quad (408)$$

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

$$\frac{\partial \text{tr}(\mathbf{UV})}{\partial \mathbf{X}} = \frac{\partial \text{tr}(\mathbf{UV}_c)}{\partial \mathbf{X}} + \frac{\partial \text{tr}(\mathbf{U}_c \mathbf{V})}{\partial \mathbf{X}} \quad (409)$$

ただし、 $\frac{\partial \text{tr}(\mathbf{UV}_c)}{\partial \mathbf{X}}$ は \mathbf{V} を \mathbf{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

$$\frac{\partial f}{\partial \mathbf{X}} = \frac{\partial \text{tr} \left(\left[\frac{\partial f}{\partial \mathbf{U}} \right]_c \right)' \mathbf{U}}{\partial \mathbf{X}} \quad (410)$$

ただし、 $\left[\frac{\partial f}{\partial \mathbf{U}} \right]_c$ は、 f を \mathbf{U} の要素で微分したものを、 \mathbf{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、 \mathbf{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、 \mathbf{A} および \mathbf{B} は \mathbf{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。

$$\frac{\partial \text{tr}(\mathbf{AX})}{\partial \mathbf{X}} = \mathbf{A}', \quad \mathbf{X} \text{ が非対称の場合} \quad (411)$$

$$= \mathbf{A} + \mathbf{A}' - \text{diag}(\mathbf{A}), \quad \mathbf{X} \text{ が対称の場合} \quad (412)$$

$$\frac{\partial \text{tr}(\mathbf{AX}')}{\partial \mathbf{X}} = \mathbf{A}, \quad \mathbf{X} \text{ が非対称の場合} \quad (413)$$

$$= \mathbf{A} + \mathbf{A}' - \text{diag}(\mathbf{A}), \quad \mathbf{X} \text{ が対称の場合} \quad (414)$$

$$\frac{\partial \text{tr}(\mathbf{X}'\mathbf{AX})}{\partial \mathbf{X}} = (\mathbf{A} + \mathbf{A}')\mathbf{X}, \quad \mathbf{X} \text{ が非対称の場合} \quad (415)$$

$$= (\mathbf{A} + \mathbf{A}')\mathbf{X} + \mathbf{X}(\mathbf{A} + \mathbf{A}') - \text{diag}((\mathbf{A} + \mathbf{A}')\mathbf{X}), \quad \mathbf{X} \text{ が対称の場合} \quad (416)$$

$$\frac{\partial \text{tr}(\mathbf{XAX}')}{\partial \mathbf{X}} = \mathbf{X}(\mathbf{A} + \mathbf{A}'), \quad \mathbf{X} \text{ が非対称の場合} \quad (417)$$

$$= (\mathbf{A} + \mathbf{A}')\mathbf{X} + \mathbf{X}(\mathbf{A} + \mathbf{A}') - \text{diag}((\mathbf{A} + \mathbf{A}')\mathbf{X}), \quad \mathbf{X} \text{ が対称の場合} \quad (418)$$

$$\frac{\partial \text{tr}(\mathbf{X}'\mathbf{AXB})}{\partial \mathbf{X}} = \mathbf{AXB} + \mathbf{A}'\mathbf{XB}', \quad \mathbf{X} \text{ が非対称の場合} \quad (419)$$

$$= \mathbf{AXB} + \mathbf{BXA} + \mathbf{A}'\mathbf{XB}' + \mathbf{B}'\mathbf{XA}' - \text{diag}(\mathbf{AXB}) - \text{diag}(\mathbf{BXA}), \quad (420)$$

$$\mathbf{X} \text{ が対称の場合} \quad (421)$$

$$\frac{\partial \text{tr}(\mathbf{XAXB})}{\partial \mathbf{X}} = \mathbf{B}'\mathbf{X}'\mathbf{A}' + \mathbf{A}'\mathbf{X}'\mathbf{B}', \quad \mathbf{X} \text{ が非対称の場合} \quad (422)$$

$$= \mathbf{B}'\mathbf{X}'\mathbf{A}' + \mathbf{A}'\mathbf{X}'\mathbf{B}' + \mathbf{BXA} + \mathbf{AXB} - \text{diag}(\mathbf{BXA}) - \text{diag}(\mathbf{AXB}), \quad (423)$$

$$\mathbf{X} \text{ が対称の場合} \quad (424)$$

$$\frac{\partial \text{tr}(\mathbf{X}^{-1}\mathbf{A})}{\partial \mathbf{X}} = -(\mathbf{X}^{-1}\mathbf{AX}^{-1})', \quad \mathbf{X} \text{ が非対称の場合} \quad (425)$$

$$= -(\mathbf{X}^{-1}\mathbf{AX}^{-1})' - \mathbf{X}^{-1}\mathbf{AX}^{-1} + \text{diag}(\mathbf{X}^{-1}\mathbf{A}'\mathbf{X}^{-1}), \quad \mathbf{X} \text{ が対称の場合} \quad (426)$$

$$\frac{\partial |\mathbf{X}|^a}{\partial \mathbf{X}} = a|\mathbf{X}|^a(\mathbf{X}^{-1})', \quad \mathbf{X} \text{ が非対称の場合} \quad (427)$$

$$= a|\mathbf{X}|^a(2\mathbf{X}^{-1} - \text{diag}(\mathbf{X}^{-1})), \quad \mathbf{X} \text{ が対称の場合} \quad (428)$$

$$\frac{\partial |\mathbf{X}\mathbf{X}'|}{\partial \mathbf{X}} = 2|\mathbf{X}\mathbf{X}'|(\mathbf{X}\mathbf{X}')^{-1}\mathbf{X}, \quad \mathbf{X} \text{ が非対称の場合} \quad (429)$$

$$= 2|\mathbf{X}\mathbf{X}'| \left((\mathbf{X}\mathbf{X}')^{-1}\mathbf{X} + ((\mathbf{X}\mathbf{X}')^{-1}\mathbf{X})' - \text{diag} \left((\mathbf{X}\mathbf{X}')^{-1}\mathbf{X} \right) \right), \quad \mathbf{X} \text{ が対称の場合} \quad (430)$$

$$\frac{\partial |\mathbf{X}'\mathbf{X}|}{\partial \mathbf{X}} = 2|\mathbf{X}'\mathbf{X}|\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}, \quad \mathbf{X} \text{ が非対称の場合} \quad (431)$$

$$= 2|\mathbf{X}'\mathbf{X}| \left(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1})' - \text{diag} \left(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \right) \right), \quad \mathbf{X} \text{ が対称の場合} \quad (432)$$

$$\frac{\partial |\mathbf{A}\mathbf{X}\mathbf{B}|}{\partial \mathbf{X}} = |\mathbf{A}\mathbf{X}\mathbf{B}|\mathbf{A}'(\mathbf{B}'\mathbf{X}'\mathbf{A}')^{-1}\mathbf{B}', \quad \mathbf{X} \text{ が非対称の場合} \quad (433)$$

$$= |\mathbf{A}\mathbf{X}\mathbf{B}| \left(\mathbf{A}'(\mathbf{B}'\mathbf{X}'\mathbf{A}')^{-1}\mathbf{B}' + \mathbf{B}(\mathbf{A}\mathbf{X}\mathbf{B})^{-1}\mathbf{A} - \text{diag} \left(\mathbf{B}(\mathbf{A}\mathbf{X}\mathbf{B})^{-1}\mathbf{A} \right) \right), \quad \mathbf{X} \text{ が対称の場合} \quad (434)$$

$$\frac{\partial \ln |\mathbf{X}|}{\partial \mathbf{X}} = (\mathbf{X}^{-1})', \quad \mathbf{X} \text{ が非対称の場合} \quad (435)$$

$$= 2\mathbf{X}^{-1} - \text{diag} \left(\mathbf{X}^{-1} \right), \quad \mathbf{X} \text{ が対称の場合} \quad (436)$$

なお、 $\frac{\partial \mathbf{X}'\mathbf{A}}{\partial \mathbf{X}}$ や $\frac{\partial \mathbf{A}\mathbf{X}\mathbf{X}'}{\partial \mathbf{X}}$ 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、 \mathbf{Y} が \mathbf{X} の関数である場合、

$$\frac{\partial \text{tr}(\mathbf{Y}^{-1}\mathbf{A})}{\partial \mathbf{X}} = -\frac{\partial \text{tr} \left(\left[\mathbf{Y}^{-1}\mathbf{A}'\mathbf{Y}^{-1} \right]_c \mathbf{Y} \right)}{\partial \mathbf{X}} \quad (437)$$

また、

$$\frac{\partial \ln |\mathbf{Y}|}{\partial \mathbf{X}} = \frac{\partial \text{tr} \left(\left(\mathbf{Y}^{-1} \right)_c \mathbf{Y} \right)}{\partial \mathbf{X}} \quad (438)$$

である。

11 線形モデルと最小2乗解

11.1 線形モデルと最小2乗解

$n \times p$ の基準変数行列 \mathbf{Y} , $n \times q$ の説明変数行列 \mathbf{X} , $q \times p$ の係数行列 \mathbf{B} , $n \times p$ の誤差行列 \mathbf{E} を用いて作られる線形モデル

$$\mathbf{Y} = \hat{\mathbf{Y}} + \mathbf{E} \quad \text{ただし} \quad \hat{\mathbf{Y}} = \mathbf{X}\boldsymbol{\beta}' \quad (439)$$

$$\mathbf{y}_{(j)} = \hat{\mathbf{y}}_{(j)} + \boldsymbol{\varepsilon}_{(j)} \quad \text{ただし} \quad \hat{\mathbf{y}}_{(j)} = \mathbf{X}\boldsymbol{\beta}_j \quad (440)$$

$$\mathbf{y}_i = \hat{\mathbf{y}}_i + \boldsymbol{\varepsilon}_i \quad \text{ただし} \quad \hat{\mathbf{y}}_i = \boldsymbol{\beta}\mathbf{x}_i \quad (441)$$

$$y_{ij} = \hat{y}_{ij} + \varepsilon_{ij} \quad \text{ただし} \quad \hat{y}_{ij} = \sum_{k=1}^q x_{ik}\beta_{jk} \quad (442)$$

を考える。

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1q} \\ x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2q} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{iq} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nj} & \dots & x_{nq} \end{pmatrix} = \begin{pmatrix} \mathbf{x}'_1 \\ \mathbf{x}'_2 \\ \vdots \\ \mathbf{x}'_i \\ \vdots \\ \mathbf{x}'_q \end{pmatrix} = (\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \dots, \mathbf{x}_{(n)}) \quad (443)$$

$$\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} & \dots & y_{1j} & \dots & y_{1p} \\ y_{21} & y_{22} & \dots & y_{2j} & \dots & y_{2p} \\ & & & \vdots & & \\ y_{i1} & y_{i2} & \dots & y_{ij} & \dots & y_{ip} \\ & & & \vdots & & \\ y_{n1} & y_{n2} & \dots & y_{nj} & \dots & y_{np} \end{pmatrix} = \begin{pmatrix} \mathbf{y}'_1 \\ \mathbf{y}'_2 \\ \vdots \\ \mathbf{y}'_i \\ \vdots \\ \mathbf{y}'_p \end{pmatrix} = (\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \dots, \mathbf{y}_{(n)}) \quad (444)$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1k} & \dots & \beta_{1q} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2k} & \dots & \beta_{2q} \\ & & & \vdots & & \\ \beta_{j1} & \beta_{j2} & \dots & \beta_{jk} & \dots & \beta_{jq} \\ & & & \vdots & & \\ \beta_{p1} & \beta_{p2} & \dots & \beta_{pk} & \dots & \beta_{pq} \end{pmatrix} = \begin{pmatrix} \beta'_1 \\ \beta'_2 \\ \vdots \\ \beta'_j \\ \vdots \\ \beta'_p \end{pmatrix} = (\boldsymbol{\beta}_{(1)}, \boldsymbol{\beta}_{(2)}, \dots, \boldsymbol{\beta}_{(q)}) \quad (445)$$

$$\boldsymbol{\beta}' = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1j} & \dots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2j} & \dots & \beta_{2p} \\ & & & \vdots & & \\ \beta_{k1} & \beta_{k2} & \dots & \beta_{kj} & \dots & \beta_{kp} \\ & & & \vdots & & \\ \beta_{q1} & \beta_{q2} & \dots & \beta_{qj} & \dots & \beta_{qp} \end{pmatrix} = \begin{pmatrix} \beta'_{(1)} \\ \beta'_{(2)} \\ \vdots \\ \beta'_{(k)} \\ \vdots \\ \beta'_{(q)} \end{pmatrix} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_p) \quad (446)$$

データ行列 \mathbf{Y} の期待値 $\widehat{\mathbf{Y}}$ 、すなわち、行列 \mathbf{X} に行列 $\boldsymbol{\beta}'$ を掛けたものは

$$\widehat{\mathbf{Y}} = \mathbf{X}\boldsymbol{\beta}' = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1q} \\ x_{21} & x_{22} & \dots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i1} & x_{i2} & \dots & x_{iq} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nq} \end{pmatrix} \times \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1j} & \dots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2j} & \dots & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \dots & \beta_{qj} & \dots & \beta_{qp} \end{pmatrix} = \left\{ \sum_{k=1}^b x_{ik}\beta_{kj} \right\} \quad (447)$$

$$= \mathbf{X} \times (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_j, \dots, \boldsymbol{\beta}_p) = (\mathbf{X}\boldsymbol{\beta}_1, \mathbf{X}\boldsymbol{\beta}_2, \dots, \mathbf{X}\boldsymbol{\beta}_j, \dots, \mathbf{X}\boldsymbol{\beta}_p) \quad (448)$$

$$= \begin{pmatrix} \mathbf{x}'_1 \\ \mathbf{x}'_2 \\ \vdots \\ \mathbf{x}'_i \\ \vdots \\ \mathbf{x}'_n \end{pmatrix} \times (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_j, \dots, \boldsymbol{\beta}_p) \quad (449)$$

$$= \begin{pmatrix} \begin{pmatrix} \beta_{11} \\ \beta_{21} \\ \vdots \\ \beta_{q1} \end{pmatrix}, \begin{pmatrix} \beta_{12} \\ \beta_{22} \\ \vdots \\ \beta_{q2} \end{pmatrix}, \dots, \begin{pmatrix} \beta_{1j} \\ \beta_{2j} \\ \vdots \\ \beta_{qj} \end{pmatrix}, \dots, \begin{pmatrix} \beta_{1p} \\ \beta_{2p} \\ \vdots \\ \beta_{qp} \end{pmatrix} \end{pmatrix} \times \begin{pmatrix} \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1q} \\ x_{21} & x_{22} & \dots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i1} & x_{i2} & \dots & x_{iq} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nq} \end{pmatrix} \end{pmatrix} = \left\{ \mathbf{x}'_i\boldsymbol{\beta}_j \right\} \quad (450)$$

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

$$\text{RSS} = \sum_{j=1}^p \sum_{i=1}^n (y_{ij} - \widehat{y}_{ij})^2 = \sum_{j=1}^p (\mathbf{y}_{(j)} - \mathbf{X}\boldsymbol{\beta}_j)'(\mathbf{y}_{(j)} - \mathbf{X}\boldsymbol{\beta}_j) = \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\beta}\mathbf{x}_i)'(\mathbf{y}_i - \boldsymbol{\beta}\mathbf{x}_i) \quad (451)$$

$$= \text{tr}[(\mathbf{Y} - \widehat{\mathbf{Y}})'(\mathbf{Y} - \widehat{\mathbf{Y}})] = \text{tr}[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}')'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}')] \quad (452)$$

を最小とする最小2乗解で、それは

$$\widehat{\boldsymbol{\beta}'} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \quad (453)$$

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

$$\text{tr}[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}')'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}')] = \text{tr}((\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})\mathbf{X}'\mathbf{X}(\boldsymbol{\beta}' - \widehat{\boldsymbol{\beta}}')) + \text{tr}[(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}')] \quad (454)$$

$$= \text{tr}((\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})\mathbf{X}'\mathbf{X}(\boldsymbol{\beta}' - \widehat{\boldsymbol{\beta}}')) - \text{tr}(\widehat{\boldsymbol{\beta}}\mathbf{X}'\mathbf{X}\widehat{\boldsymbol{\beta}}') + \text{tr}(\mathbf{Y}'\mathbf{Y}) \quad (455)$$

$$= \text{tr}((\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})\mathbf{X}'\mathbf{X}(\boldsymbol{\beta}' - \widehat{\boldsymbol{\beta}}')) - \text{tr}(\mathbf{Y}'\mathbf{X}\widehat{\boldsymbol{\beta}}') + \text{tr}(\mathbf{Y}'\mathbf{Y}) \quad (456)$$

と展開すれば、 $\mathbf{X}'\mathbf{X}$ 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、 $X'X$ が特異な場合は

$$\widehat{\beta}' = (X'X)^{-} X'Y \quad (457)$$

で与えられるが、この場合、 $\widehat{\beta}'$ は一意に定まらない。ただし、その場合でも、 \widehat{Y} の部分の最小2乗推定値 $X\widehat{\beta}'$ は一意に定まる。なぜなら、

$$X\widehat{\beta}' = X(X'X)^{-} X'Y \quad (458)$$

となるが、この $X(X'X)^{-} X'$ 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

$$\hat{\beta} = (X'X)^{-1} X'y \quad (459)$$

となる。

11.2 微分による解法

$$\text{RSS} = \sum_{j=1}^p \sum_{i=1}^n (y_{ij} - \widehat{y}_{ij})^2 = \sum_{j=1}^p (\mathbf{y}_{(j)} - X\beta_j)' (\mathbf{y}_{(j)} - X\beta_j) = \sum_{i=1}^n (\mathbf{y}_i - \beta \mathbf{x}_i)' (\mathbf{y}_i - \beta \mathbf{x}_i) \quad (460)$$

$$= \text{tr}[(Y - \widehat{Y})'(Y - \widehat{Y})] = \text{tr}[(Y - X\beta)'(Y - X\beta)] \quad (461)$$

$$= \text{tr}(Y'Y) - \text{tr}(\beta X'Y) + \text{tr}(Y'X\beta') + \text{tr}(\beta X'X\beta') \quad (462)$$

$$= \text{tr}(Y'Y) - 2\text{tr}(Y'X\beta') + \text{tr}(\beta X'X\beta') \quad (463)$$

$$(464)$$

$$\frac{\partial \text{RSS}}{\partial \beta} = -2Y'X + 2\beta X'X \quad (465)$$

$$Y'X = \beta X'X \quad (466)$$

$$\beta = Y'X(X'X)^{-1} \quad \text{or} \quad \beta' = (X'X)^{-1} X'Y \quad (467)$$

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
