

行列の列空間と階数、連立方程式の解、行列の分解、その他

前川真一

20080207, 20200512,0612,18,23dnc,20200702,20,06,09,0804dnc,06,07,08,10c,11,12,13,14,
20220515,16,19,26

目次

1	行列の列空間・基底・直交化	2
1.1	行列の列空間	2
1.2	回帰分析と射影子	3
1.2.1	射影子の幾何学的解釈	6
1.3	行列の直交基底	8
1.3.1	グラム・シュミットの直交化 The Gram-Schmidt Orthogonalization	8
1.3.2	列ブロック単位での直交化	11
1.4	ベクトルの線形独立と行列の階数	18
2	連立一次方程式の一般解と一般逆行列	19
2.1	連立一次方程式の一般解	19
2.1.1	回帰分析との関係	20
3	一般逆行列について	22
3.1	一般逆行列	23
3.1.1	一般逆行列の性質	23
3.1.2	一般逆行列の生成	24
3.2	ムーアペンローズ逆行列	26
3.2.1	ムーアペンローズ逆行列の生成	26
3.2.2	ムーアペンローズ逆行列の性質	27
4	行列の固有値と特異値分解	28
4.1	固有値分解	28
4.2	特異値分解	30
4.2.1	特異値分解の存在	30
4.2.2	特異値分解	32
4.2.3	補足	33

1 行列の列空間・基底・直交化

1.1 行列の列空間

$n \times q$ の行列 X の各列を

$$X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_q] \quad (1)$$

と記すとき、行列 X の列空間とは、 n 個の列ベクトルの線形結合

$$\mathbf{b} = \sum_{j=1}^q (\beta_j \times \mathbf{x}_j) = X\boldsymbol{\beta} \quad (2)$$

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

あるベクトル \mathbf{b} が行列 X の列空間に含まれることを

$$\mathbf{b} \in \mathcal{C}(X) \quad (3)$$

と書く。このとき、(2) 式を満たすベクトル $\boldsymbol{\beta}$ が存在する。

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

$$Z = XA \quad (4)$$

を定義すると、

$$\mathbf{b} = X\boldsymbol{\beta} = XAA^{-1}\boldsymbol{\beta} = (XA)(A^{-1}\boldsymbol{\beta}) = Z\boldsymbol{\beta}^* \quad (5)$$

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

$$\mathbf{b} \in \mathcal{C}(X) \leftrightarrow \mathbf{b} \in \mathcal{C}(XA) \quad (6)$$

である。すなわち、

1 : 行列の列空間

X の列空間と $Z=XA$ の列空間は、 A が正則の場合等しくなる。

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

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

1.2 回帰分析と射影子

$$\mathbf{y} = \hat{\mathbf{y}} + \varepsilon \quad (7)$$

$$\hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} \quad (8)$$

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

$$\text{RSS} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \text{tr} \left[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right] \quad (9)$$

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

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (10)$$

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

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \quad (11)$$

$$= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (12)$$

$$= \mathbf{P}(\mathbf{X})\mathbf{y} \quad (13)$$

ただし、

$$\mathbf{P}(\mathbf{X}) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \quad (14)$$

は、 \mathbf{X} の列空間への射影子と呼ばれる、対称かつ冪等

$$\mathbf{P}(\mathbf{X}) = \mathbf{P}(\mathbf{X})', \mathbf{P}(\mathbf{X})^2 = \mathbf{P}(\mathbf{X}) \quad (15)$$

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

$$\mathbf{P}(\mathbf{X})\mathbf{X} = \mathbf{X} \quad (16)$$

である。

特に、 $q = 1$ の場合、

$$\mathbf{P}(\mathbf{x}) = \mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}' = \frac{1}{\mathbf{x}'\mathbf{x}}\mathbf{x}\mathbf{x}' \quad (17)$$

であるから、

$$\mathbf{P}(\mathbf{x})\mathbf{x} = \frac{1}{\mathbf{x}'\mathbf{x}}\mathbf{x}\mathbf{x}'\mathbf{x} = \frac{\mathbf{x}'\mathbf{x}}{\mathbf{x}'\mathbf{x}}\mathbf{x} = \mathbf{x} \quad (18)$$

である。

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

$$\mathbf{P}(\mathbf{X})\mathbf{b} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{b} \quad (19)$$

$$= \mathbf{X}\boldsymbol{\beta} \quad (20)$$

ただし

$$\beta = (X'X)^{-1}X'b \quad (21)$$

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

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

$$P(XA) = XA(A'X'XA)^{-1}A'X' \quad (22)$$

$$= XAA^{-1}(X'X)^{-1}A'^{-1}A'X' \quad (23)$$

$$= X(X'X)^{-1}X' \quad (24)$$

である。

また、

$$Q(X) = I - P(X) = I - X(X'X)^{-1}X' \quad (25)$$

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

$$P(X)Q(X) = Q(X)P(X) = O \quad (26)$$

したがって、

$$Q(X)X = Q(X)P(X)X = O \quad (27)$$

である。なお、回帰分析に於ける残差ベクトルは (13) 式より

$$y - \hat{y} = y - P(X)y = (I - P(X))y = Q(X)y \quad (28)$$

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

いま、 $Q(X)$ の定義から

$$P(X) + Q(X) = I \quad (29)$$

であるから、このふたつの射影子を用いれば、ベクトル y を

$$y = (P(X) + Q(X))y = P(X)y + Q(X)y \quad (30)$$

$$= X(X'X)^{-1}X'y + (I - X(X'X)^{-1}X')y \quad (31)$$

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

$$Y = P(X)Y + Q(X)Y \quad (32)$$

$$= X(X'X)^{-1}X'Y + (I - X(X'X)^{-1}X')Y \quad (33)$$

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

最後に、行列 X が $X = [X_1, X_2]$ と二つに分割されていて、しかも、そのふたつの行列 X_1 と X_2 が直交する場合、すなわち、

$$X_1' X_2 = \mathbf{O} \quad (34)$$

の場合には、 X の列空間への射影子は

$$P(X) = P([X_1, X_2]) \quad (35)$$

$$= \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} X_1' X_1 & X_1' X_2 \\ X_2' X_1 & X_2' X_2 \end{bmatrix}^{-1} \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \quad (36)$$

$$= \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} (X_1' X_1)^{-1} & \mathbf{O} \\ \mathbf{O} & (X_2' X_2)^{-1} \end{bmatrix} \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \quad (37)$$

$$= X_1 (X_1' X_1)^{-1} X_1' + X_2 (X_2' X_2)^{-1} X_2' \quad (38)$$

$$= P(X_1) + P(X_2) \quad (39)$$

と、 X_1 の列空間への射影子と X_2 への射影子行列の和として表現できる。同様に、回帰係数行列は

$$\hat{\beta} = \begin{bmatrix} X_1' X_1 & X_1' X_2 \\ X_2' X_1 & X_2' X_2 \end{bmatrix}^{-1} \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} y \quad (40)$$

$$= \begin{bmatrix} (X_1' X_1)^{-1} & \mathbf{O} \\ \mathbf{O} & (X_2' X_2)^{-1} \end{bmatrix} \begin{bmatrix} X_1' y \\ X_2' y \end{bmatrix} \quad (41)$$

$$= \begin{bmatrix} (X_1' X_1)^{-1} X_1' y \\ (X_2' X_2)^{-1} X_2' y \end{bmatrix} \quad (42)$$

となり、これは、 y を別々にそれぞれ X_1 と X_2 へ回帰した時に得られる回帰係数行列を縦に並べたものと等しくなる。

以下の回帰モデル

$$\mathbf{y} = \mathbf{X}_1 \mathbf{b}_1 + \varepsilon_1, \quad \mathbf{y} = \mathbf{X}_2 \mathbf{b}_2 + \varepsilon_2 \quad (43)$$

$$\mathbf{y} = \mathbf{X}_1 \boldsymbol{\beta}_1 + \mathbf{X}_2 \boldsymbol{\beta}_2 + \varepsilon \quad (44)$$

$$\mathbf{y} = \mathbf{X}_1 \boldsymbol{\gamma}_1 + \mathbf{Q}(\mathbf{X}_1) \mathbf{X}_2 \boldsymbol{\gamma}_2 + \varepsilon \quad (45)$$

$$\mathbf{y} = \mathbf{Q}(\mathbf{X}_2) \mathbf{X}_1 \boldsymbol{\lambda}_1 + \mathbf{X}_2 \boldsymbol{\lambda}_2 + \varepsilon \quad (46)$$

$$\mathbf{y} = \mathbf{Q}(\mathbf{X}_2) \mathbf{X}_1 \boldsymbol{\delta}_1 + \varepsilon_3, \quad \mathbf{y} = \mathbf{Q}(\mathbf{X}_1) \mathbf{X}_2 \boldsymbol{\delta}_2 + \varepsilon_4 \quad (47)$$

$$\mathbf{y} = \mathbf{Q}(\mathbf{X}_2) \mathbf{X}_1 \boldsymbol{\kappa}_1 + \mathbf{Q}(\mathbf{X}_1) \mathbf{X}_2 \boldsymbol{\kappa}_2 + \varepsilon \quad (48)$$

におけるパラメタ間には

$$\boldsymbol{\beta}_1 = \boldsymbol{\delta}_1 = \boldsymbol{\lambda}_1 \quad (49)$$

$$\boldsymbol{\beta}_2 = \boldsymbol{\delta}_2 = \boldsymbol{\gamma}_2 \quad (50)$$

$$\mathbf{b}_1 = \boldsymbol{\gamma}_1, \quad \mathbf{b}_2 = \boldsymbol{\lambda}_2 \quad (51)$$

$$(52)$$

という関係がある。

1.2.1 射影子の幾何学的解釈

行列 \mathbf{X} の列空間への射影子は (17) 式 で定義されるが、 \mathbf{X} は n 次元のベクトルが q 個集まったものであると考えられる。

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

$$\mathbf{u} = \mathbf{X} \boldsymbol{\beta} \quad (53)$$

と書けるため、

$$\mathbf{P}(\mathbf{X}) \mathbf{u} = \mathbf{P}(\mathbf{X}) \mathbf{X} \boldsymbol{\beta} = \mathbf{X} \boldsymbol{\beta} = \mathbf{u} \quad (54)$$

となる。これは $n = q > 2$ の場合にも同様である。

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

$$\mathbf{P}(\mathbf{X}) = \frac{1}{\mathbf{v}_1' \mathbf{v}_1} \mathbf{v}_1 \mathbf{v}_1' \quad (55)$$

となるが、 $\mathbf{v} = \mathbf{P}(\mathbf{X}) \mathbf{u}$ は、ベクトル \mathbf{u} の頂点からこの \mathbf{v}_1 に向かって下ろした垂線の足の 2 次元空間における座標を表すことになる。

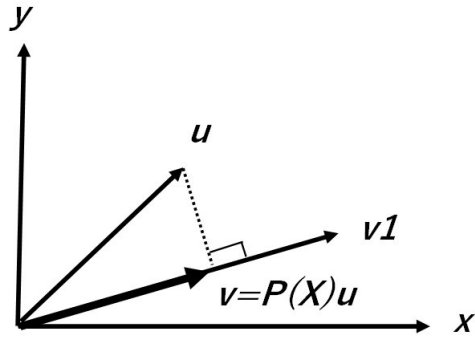


図 1: 直線への射影

$n = 3$ の場合にも同様で、 X の列数が 1 の場合には、 $P(X)u$ はベクトル u から x へ下ろした垂線の足の 3 次元空間における座標となる。また、 X の階数が $q = 2$ でフルランクの場合、すなわち、 $X = [v_1, v_2]$ の場合には、 X の列空間はベクトル v_1 と v_2 を基底として持つ平面を形成するため、 $v = P(X)u$ はベクトル u からその平面に下ろした垂線の足の 3 次元空間における座標となる。

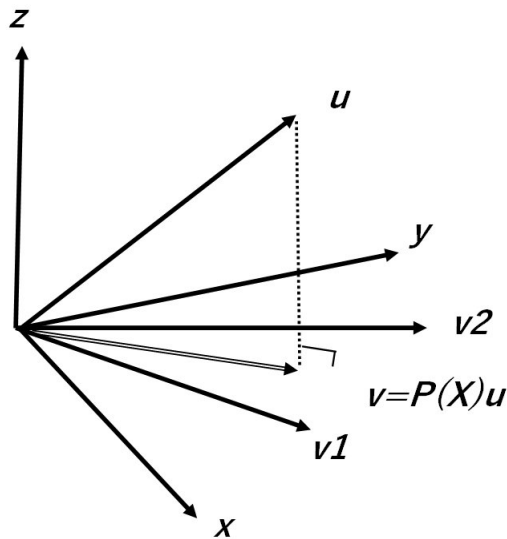


図 2: 2次元平面への射影

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

 座標行列との関係の補足

1.3 行列の直交基底

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

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

$$X = [x_1, x_2, \dots, x_q] \quad (56)$$

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

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

- 第 1 列はそのまま
 $u_1 = x_1$
- 第 2 列は x_2 を x_1 へ回帰した残差とする。
 $u_2 = Q(x_1) x_2$
- 第 3 列は x_3 を x_1 と x_2 へ回帰した残差とする。
 $u_3 = Q(x_1, x_2) x_3$
 \vdots
- 第 j 列は x_j を x_1, x_2, \dots, x_{j-1} へ回帰した残差とする。
 $u_j = Q(x_1, x_2, \dots, x_{j-1}) x_j$
 \vdots
- 第 q 列は x_q を x_1, x_2, \dots, x_{q-1} へ回帰した残差とする。
 $u_q = Q(x_1, x_2, \dots, x_{q-1}) x_q$

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

$$Q = U \text{diag}(U'U)^{-\frac{1}{2}} \quad (57)$$

が X 行列の直交基底行列となる。(この Q は射影子を表す Q ではないことに注意。単なる行列で、伝統的に Q と呼ばれるものである。)

そして、この直交基底は

$$U = X A \quad (58)$$

ただし、 A は $q \times q$ の上三角行列、という形で書き表される。なぜなら、 x_j を x_1, x_2, \dots, x_{j-1} へ回帰したときの回帰係数ベクトルを $\beta_j = [\beta_{1j}, \beta_{2j}, \dots, \beta_{j-1,j}]$ とすれば、

$$u_j = x_j - \sum_{k=1}^{j-1} \beta_{kj} x_k \quad (59)$$

であるから、まとめて

$$[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \dots, \mathbf{u}_q] = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \dots, \mathbf{x}_q] \times \begin{bmatrix} 1 & -\beta_{12} & -\beta_{13} & -\beta_{14} & \cdots & -\beta_{1q} \\ 0 & 1 & -\beta_{23} & -\beta_{24} & \cdots & -\beta_{2q} \\ 0 & 0 & 1 & -\beta_{34} & \cdots & -\beta_{3q} \\ 0 & 0 & 0 & 1 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (60)$$

と書けるが、この上三角行列が (58) 式の \mathbf{A} 行列となる。したがって、

$$\mathbf{Q} = \mathbf{X} \mathbf{A} \text{diag}(\mathbf{A}' \mathbf{X}' \mathbf{X} \mathbf{A})^{-\frac{1}{2}} = \mathbf{X} \mathbf{A}^* \quad (61)$$

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

また、上式を \mathbf{X} に関して解けば

$$\mathbf{X} = \mathbf{Q} \mathbf{A}^{*-1} = \mathbf{Q} \mathbf{R} \quad (62)$$

ただし、

$$\mathbf{R} = \mathbf{A}^{*-1} = \text{diag}(\mathbf{A}' \mathbf{X}' \mathbf{X} \mathbf{A})^{\frac{1}{2}} \mathbf{A}^{-1} \quad (63)$$

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

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

表 1 に 8×5 の行列 \mathbf{X} のグラム・シュミットの直交化の結果を示す。ただし、この結果は $\mathbf{X} = \mathbf{U} \mathbf{R}$ かつ $\mathbf{U} = \mathbf{X} \mathbf{A}$ を与えるもので、 $\mathbf{Q}' \mathbf{Q} = \mathbf{I}$ となる \mathbf{Q} 行列はこの \mathbf{U} 行列の各列を (61) 式で正規化したもの、それに対応する \mathbf{R} は (63) 式として与えられる。

表 1: グラム・シュミットの直交化：列フルランクの場合

\mathbf{X}						\mathbf{U}					
	X1	X2	X3	X4	X5		U1	U2	U3	U4	U5
1)	7	3	1	6	9	1)	7	-3.560	-2.280	-0.683	4.687
2)	6	9	5	9	2	2)	6	3.377	-1.014	1.460	-0.561
3)	6	9	7	9	0	3)	6	3.377	0.986	2.926	-3.710
4)	3	6	5	7	8	4)	3	3.189	1.246	3.931	5.900
5)	5	9	8	0	2	5)	5	4.315	2.241	-4.605	-0.447
6)	7	1	7	1	8	6)	7	-5.560	4.717	0.220	-0.371
7)	7	8	4	5	4	7)	7	1.440	-1.771	-3.255	1.207
8)	7	4	1	9	0	8)	7	-2.560	-2.778	1.563	-4.072

\mathbf{R}						\mathbf{A}					
	U1	U2	U3	U4	U5		U1	U2	U3	U4	U5
U1	1	0.937	0.722	0.914	0.639	X1	1	-0.937	-0.255	-0.736	-0.692
U2	0	1	0.498	0.389	-0.288	X2	0	1	-0.498	-0.754	0.623
U3	0	0	1	-0.733	0.481	X3	0	0	1	0.733	-0.575
U4	0	0	0	1	0.127	X4	0	0	0	1	-0.127
U5	0	0	0	0	1	X5	0	0	0	0	1

なお、上記の手続きを行う際に、ある時点で回帰の残差 u_j が $\mathbf{0}$ となった場合、それは x_j が $[x_1, x_2, \dots, x_{j-1}]$ の列空間に含まれることを示しているが、この場合には、以降の回帰係数の計算には、 x_j を省いて行い、 $\beta_{i,j}, i = 1, 2, \dots$ と $\beta_{j,i}, i = j+1, j+2, \dots$ を全て 0 として先に進んで行く。すると、最終的にいくつかの全ての要素が 0 である列を持つ U 行列と全ての要素が 0 である行を持つ A 行列が生成されるが、この、 U 行列の中の全てが 0 でない列の数を、行列の階数 (rank) と呼び rank (X) と記す。また、行列の階数とその列数に等しい場合を列フルランク、行数に等しい場合を行フルランクと呼ぶ。行フルランクな行列の各列は互いに線形独立である。そして、 $u_j = \mathbf{0}$ に対応した x_j は、 $[x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_p]$ に線形従属であるという。

表 2: グラム・シュミットの直交化：列フルランクではない場合

X						U					
	X1	X2	X3	X4	X5		U1	U2	U3	U4	U5
1)	7	14	1	8	9	1)	7	0	-4.053	0	4.955
2)	6	12	5	11	2	2)	6	0	0.669	0	-1.905
3)	6	12	7	13	0	3)	6	0	2.669	0	-4.117
4)	3	6	5	8	8	4)	3	0	2.834	0	5.783
5)	5	10	8	13	2	5)	5	0	4.391	0	-1.660
6)	7	14	7	14	8	6)	7	0	1.947	0	3.321
7)	7	14	4	11	4	7)	7	0	-1.053	0	-0.362
8)	7	14	1	8	0	8)	7	0	-4.053	0	-4.045

R						A					
	X1	X2	X3	X4	X5		U1	U2	U3	U4	U5
U1	1	2	0.722	1.722	0.639	X1	1	0	-0.722	0	-0.563
U2	0	0	0	0	0	X2	0	0	0	0	0
U3	0	0	1	1	0.106	X3	0	0	1	0	-0.106
U4	0	0	0	0	0	X4	0	0	0	0	0
U5	0	0	0	0	1	X5	0	0	0	0	1

表 3: 縮小化されたグラム・シュミットの直交化

U				R					
	U1	U3	U5		X1	X2	X3	X4	X5
1)	7	-4.053	4.955	U1	1	2	0.722	1.722	0.639
2)	6	0.669	-1.905	U3	0	0	1	1	0.106
3)	6	2.669	-4.117	U5	0	0	0	0	1
4)	3	2.834	5.783						
5)	5	4.391	-1.660	A					
6)	7	1.947	3.321		U1	U3	U5		
7)	7	-1.053	-0.362	X1	1	-0.722	-0.563		
8)	7	-4.053	-4.045	X2	0	0	0		
				X3	0	1	-0.106		
				X4	0	0	0		
				X5	0	0	1		

表 2 に X 行列の第 2 列と 4 列がそれ以前の列の線形結合で表される場合、すなわち、行列 X の階数が $\text{rank}(X) = 5 - 2 = 3$ である場合の例を示した。この例からも明らかのように、行列が列フルランクでない場合、すなわち $\text{rank}(X) \leq q$ の場合には $U'U$ は対角行列ではあるが、対角要素に $q - \text{rank}(X)$ 個の 0 を持ち、それを正規化した Q 行列も同様である。

そこで、 U 行列の全てが 0 である列を取り除いた $n \times \text{rank}(X)$ の行列 U^\dagger とそれを (61) 式で正規化して得られる Q^\dagger ならびに A 行列の全てが 0 の行を取り除いた $q \times \text{rank}(X)$ の A^\dagger それらから (63) 式を用いて計算される R^\dagger 行列を定義すると、

$$X = Q^\dagger R^\dagger, \quad U^\dagger = X A^\dagger, \quad Q^{\dagger'} Q^\dagger = I \quad (64)$$

となる。表 2 に対応するこれらの行列を表 3 に示した。

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

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

$$X = Q^* R^*, \quad \text{ただし } R^* = (Q^{*'} Q^*)^{-1} Q^{*'} X \quad (65)$$

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

1.3.2 列ブロック単位での直交化

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

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

$$X = [X_1, X_2, \dots, X_G] \quad (66)$$

$q_j, j = 1, 2, \dots, q$ をそれぞれのブロックの列数とする。このとき、第 j ブロック X_j を第 1 ブロックから第 $j-1$ ブロックまで、すなわち $X_{1:(j-1)} = [X_1, X_2, \dots, X_{j-1}]$ に回帰した時の $\sum_{k=1}^{j-1} q_k \times n_j$ の回帰係数行列を B_j とし、その行列を X 行列のブロックに対応した形で分割したものを

$$B_j = \begin{bmatrix} X_1' X_1 & X_1' X_2 & \dots & X_1' X_{j-1} \\ X_2' X_1 & X_2' X_2 & \dots & X_2' X_{j-1} \\ \vdots & \vdots & & \\ X_{j-1}' X_1 & X_{j-1}' X_2 & \dots & X_{j-1}' X_{j-1} \end{bmatrix}^{-1} \times \begin{bmatrix} X_{1j}' X_j \\ X_{2j}' X_j \\ \vdots \\ X_{j-1,j}' X_j \end{bmatrix} = \begin{bmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{j-1,j} \end{bmatrix}, \quad j = 2, 3, \dots, G \quad (67)$$

とすれば、 X_j を $X_{1:(j-1)}$ へ回帰した時の $n \times q_j$ の残差行列 U_j は

$$U_j = X_j - X_{1:(j-1)} B_j = X_j - \sum_{k=1}^{j-1} X_k B_{kj} \quad (68)$$

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

$$U = [U_1, U_2, \dots, U_G] = X A \quad (69)$$

という形で U 行列を作成することが出来る。

$$A = \begin{bmatrix} I & -B_{12} & \cdots & -B_{1,j-1} & -B_{1,j} & \cdots & -B_{1,G} \\ O & I & \cdots & -B_{2,i-1} & -B_{2,j} & \cdots & -B_{2,G} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ O & O & \cdots & I & -B_{j-1,j} & \cdots & -B_{j-1,G} \\ O & O & \cdots & O & I & \cdots & -B_{j,G} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ O & O & \cdots & O & O & \cdots & I \end{bmatrix} \quad (70)$$

j 列

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

表 4 に先ほどの X 行列を $X_1 = [x_1, x_2]$, $X_2 = [x_3, x_4]$, $X_3 = x_5$ としたときの例を示す。

表 4: 列ブロック単位の直交化

U	A					$U'U$	A					
	Q1	Q2	Q3	Q4	Q5		Q1	Q2	Q3	Q4	Q5	
1)	7	3	-2.280	0.987	4.687	X1	1	0	-0.255	-0.550	-0.692	
2)	6	9	-1.014	2.203	-0.561	X2	0	1	-0.498	-0.389	0.623	
3)	6	9	0.986	2.203	-3.710	X3	0	0	1	0	-0.575	
4)	3	6	1.246	3.018	5.900	X4	0	0	0	1	-0.127	
5)	5	9	2.241	-6.247	-0.447	X5	0	0	0	0	1	
6)	7	1	4.717	-3.236	-0.371							
7)	7	8	-1.771	-1.957	1.207							
8)	7	4	-2.778	3.598	-4.072							
							Q1	Q2	Q3	Q4	Q5	
							Q1	302	283	0	0	0
							Q2	283	369	0	0	0
							Q3	0	0	47	-34	0
							Q4	0	0	-34	86	0
							Q5	0	0	0	0	89

なお、以下の記述では、各列ブロックを構成する行列が列フルランクでない場合を考慮して、逆行列の代わりに一般逆行列を用いる。一般逆行列には様々なものがあるが、ここでは正方行列の一般逆行列のみを考えれば良く、スweep演算子を用いれば、線形従属な列に対応する回帰係数が 0 となるような回帰係数行列が得られる。一般逆行列の選択により回帰係数行列そのものは一意に定まらないが、 $P()$ や $Q()$ と表記される射影子は一意に定まるため、最終的に得られる Z 行列は一意に定まる。

以下に、(70) 式の別の表現を与える。

前節の Gram-Schmidt の直交化を踏まえると、行列 X が $\begin{bmatrix} X_1 & X_2 \end{bmatrix}$ と 2 分割されている場合、 A 行列を、以下のように、単位行列の右上部分に X_2 を X_1 へ回帰したときの回帰係数行列の符号を逆にしたものを持つ様に定義すると

$$A = \begin{bmatrix} I & -(X_1'X)^- X_1'X_2 \\ O & I \end{bmatrix} \quad (71)$$

X は、 X_1 と X_1 に直交する部分の二つに再編成することが出来る。すなわち、 A 行列を用いると

$$XA = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \times \begin{bmatrix} I & -(X_1'X_1)^{-1}X_1'X_2 \\ O & I \end{bmatrix} \quad (72)$$

$$= \begin{bmatrix} X_1 & Q(X_1)X_2 \end{bmatrix} \quad (73)$$

である。この、結果として得られる $XA = \begin{bmatrix} X_1 & Q(X_1)X_2 \end{bmatrix}$ は、もとの行列 X と同じ列空間を持つ。

なお、このことは、以下のようにして確認することが出来る。 X の列空間に含まれるベクトル y は

$$y = X\beta = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = X_1\beta_1 + X_2\beta_2 \quad (74)$$

と書けるが、

$$X_2 = P(X_1)X_2 + Q(X_1)X_2 \quad (75)$$

であるから、

$$y = X_1\beta_1 + X_2\beta_2 \quad (76)$$

$$= X_1\beta_1 + (P(X_1)X_2 + Q(X_1)X_2)\beta_2 \quad (77)$$

$$= X_1\beta_1 + X_1 \left((X_1'X_1)^{-1}X_1'X_2 \right) \beta_2 + Q(X_1)X_2\beta_2 \quad (78)$$

$$= X_1 \left(\beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 \right) + Q(X_1)X_2\beta_2 \quad (79)$$

$$= X_1\beta_3 + Q(X_1)X_2\beta_2 \quad (80)$$

$$= \begin{bmatrix} X_1 & Q(X_1)X_2 \end{bmatrix} \begin{bmatrix} \beta_3 \\ \beta_2 \end{bmatrix} \quad (81)$$

となり、 y は $[X_1, Q(X_1)X_2]$ の列空間にも含まれていることが分かる。

したがって、行列 X が $X = [X_1, X_2]$ と分割されている時に、上記の方法で $[X_1, Q(X_1)X_2]$ という基底を作った場合、(39) 式によれば、 X の列空間への射影子は、互いに直交する二つの行列 X_1 と $Q(X_1)X_2$ への射影子の和

$$P(X) = P([X_1, X_2]) \quad (82)$$

$$= P(X_1) + P(Q(X_1)X_2) \quad (83)$$

$$= X_1(X_1'X_1)^{-1}X_1' + Q(X_1)X_2(X_2'Q(X_1)X_2)^{-1}X_2'Q(X_1) \quad (84)$$

に等しくなる。なお、 $P(Q(X_1)X_2)$ は、 X_2 の列空間のうち、 X_1 に直交する部分、すなわち、 $Q(X_1)X_2$ の列空間、への射影子である。また、任意のベクトル y を X に回帰した時の回帰係数行列も、(42) 式によれば、 y を X_1 に回帰して得られる回帰係数行列と、 y を $Q(X_1)X_2$ に回帰して得られる回帰係数行列とを縦に並べたものとなる。

同様に、 X が $\begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}$ と 3 分割されている場合、 $A^{(1)}$ 行列を

$$A^{(1)} = \begin{bmatrix} I & -(X_1'X_1)^{-1}X_1'X_2 & -(X_1'X_1)^{-1}X_1'X_3 \\ O & I & O \\ O & O & I \end{bmatrix} \quad (85)$$

とすると、 X は、 X_1 と X_1 に直交する二つの部分、 $Q(X_1)X_2$ と $Q(X_1)X_3$ 、の合計三つの部分に再編成することが出来る。すなわち、

$$\begin{aligned} Z^{(1)} = XA^{(1)} &= \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} \times \begin{bmatrix} I & -(X_1'X_1)^{-1}X_1'X_2 & -(X_1'X_1)^{-1}X_1'X_3 \\ O & I & O \\ O & O & I \end{bmatrix} \\ &= \begin{bmatrix} X_1 & Q(X_1)X_2 & Q(X_1)X_3 \end{bmatrix} \end{aligned} \quad (86)$$

が得られる。(2分割の場合の X_2 がさらに X_2 と X_3 に分割されていると考えても同じ。) この $Z^{(1)}$ 行列の第2列ブロックには X_1 に直交する X_2 が、第3列ブロックには X_1 に直交する X_3 が入っている。

次に、この $Z^{(1)}$ 行列の右から係数行列を掛け、 X_1 にも $Q(X_1)X_2$ にも直交するブロックを $Q(X_1)X_3$ から作ることを考える。

そのためには、 $Q(X_1)X_3$ を $[X_1, Q(X_1)X_2]$ へ回帰をした残差を求めれば良いが、 X_1 と $Q(X_1)X_2$ が直交するため、回帰係数は、 $Q(X_1)X_3$ の X_1 への回帰と $Q(X_1)X_3$ の $Q(X_1)X_2$ への回帰の二つを独立に行って求めることが出来る。また、また、 $Q(X_1)X_3$ を X_1 へ回帰した回帰係数は O である。

したがって、 $Q(X_1)X_3$ を $Q(X_1)X_2$ へ回帰した際の回帰係数行列は

$$((Q(X_1)X_2)'Q(X_1)X_2)^{-1}(Q(X_1)X_2)'Q(X_1)X_3 \quad (87)$$

$$= (X_2'Q(X_1)Q(X_1)X_2)^{-1}(Q(X_1)X_2)'Q(X_1)X_3 \quad (88)$$

$$= (X_2'Q(X_1)X_2)^{-1}(Q(X_1)X_2)'Q(X_1)X_3 \quad (89)$$

となり、第(2,3)列ブロックにこの回帰係数行列の符号を反転させたものを持つ

$$A^{(2)} = \begin{bmatrix} I & O & O \\ O & I & -(X_2'Q(X_1)X_2)^{-1}(Q(X_1)X_2)'Q(X_1)X_3 \\ O & O & I \end{bmatrix} \quad (90)$$

と定義される $A^{(2)}$ 行列を $Z^{(1)}$ の右から掛けると、結果として得られる $Z^{(2)}$ 行列は

$$\begin{aligned} Z^{(2)} = Z^{(1)}A^{(2)} &= \begin{bmatrix} X_1 & Q(X_1)X_2 & Q(X_1)X_3 \end{bmatrix} \times \\ &\begin{bmatrix} I & O & O \\ O & I & -(X_2'Q(X_1)X_2)^{-1}(Q(X_1)X_2)'Q(X_1)X_3 \\ O & O & I \end{bmatrix} \\ &= \begin{bmatrix} X_1 & Q(X_1)X_2 & Q(Q(X_1)X_2)Q(X_1)X_3 \end{bmatrix} \end{aligned} \quad (91)$$

となる。ただし、上記の結果としてえられる $Z^{(2)}$ 行列の第3列ブロックは

$$\text{第3列ブロック} = Q(X_1)X_3 - Q(X_1)X_2(X_2'Q(X_1)X_2)^{-1}(Q(X_1)X_2)'Q(X_1)X_3 \quad (92)$$

となるが、この式の第2項の $Q(X_1)X_3$ にかかる部分を

$$P(Q(X_1)X_2) = Q(X_1)X_2(X_2'Q(X_1)X_2)^{-1}(Q(X_1)X_2)' \quad (93)$$

すなわち、 $Q(X_1)X_2$ の列空間への射影子と置いた。したがって、

$$\text{第3列ブロック} = Q(X_1)X_3 - P(Q(X_1)X_2)Q(X_1)X_3 \quad (94)$$

$$= (I - P(Q(X_1)X_2))Q(X_1)X_3 = Q(Q(X_1)X_2)Q(X_1)X_3 \quad (95)$$

と書ける。

このとき、この第 3 列ブロックである、 $Q(Q(X_1)X_2)Q(X_1)X_3$ は X_1 にも X_2 にも直交することは以下のようにして確かめられる。まず X_1 とは

$$X_1' Q(Q(X_1)X_2)Q(X_1)X_3 = X_1'(I - P(Q(X_1)X_2))Q(X_1)X_3 \quad (96)$$

$$= (X_1' - X_1'P(Q(X_1)X_2))Q(X_1)X_3 \quad (97)$$

$$(98)$$

であるが、(93) 式より、 $X_1'P(Q(X_1)X_2) = O$ かつ $X_1'Q(X_1) = O$ であるため、この二つは直交している。次に X_2 であるが、

$$X_2' Q(Q(X_1)X_2)Q(X_1)X_3 = X_2'(I - P(Q(X_1)X_2))Q(X_1)X_3 \quad (99)$$

$$= X_2'Q(X_1)X_3 - X_2'P(Q(X_1)X_2) Q(X_1)X_3 \quad (100)$$

$$= X_2'Q(X_1)X_3 -$$

$$X_2'Q(X_1)X_2(X_2'Q(X_1)X_2)^{-1}(Q(X_1)X_2)'Q(X_1)X_3$$

であるが、最後の第 2 項は

$$X_2'Q(X_1)X_2(X_2'Q(X_1)X_2)^{-1}(Q(X_1)X_2)' X_3 = \quad (101)$$

$$X_2'Q(X_1) Q(X_1)X_2(X_2'Q(X_1)X_2)^{-1}(Q(X_1)X_2)' X_3 = \quad (102)$$

$$(Q(X_1)X_2)' Q(X_1)X_2(X_2'Q(X_1)X_2)^{-1}(Q(X_1)X_2)' X_3 = (Q(X_1)X_2)' X_3 \quad (103)$$

であるため、この二つもまた直交する。したがって、この $Z^{(2)}$ 行列の第 3 列ブロックには X_1 と X_2 に直交する X_3 が入っていることになるが、これは (91) 式の第 3 列ブロックに注目すれば

$$Q([X_1, X_2])X_3 = Q(Q(X_1)X_2) Q(X_1)X_3 \quad (104)$$

すなわち、 X_3 を $[X_1, X_2]$ へ回帰した残差は、 $Q(X_1)X_3$ を $Q(X_1)X_2$ へ回帰した残差に等しい、ということである。このことは (104) 式の右辺の X_3 に掛かる部分が

$$Q(Q(X_1)X_2) Q(X_1) = (I - P(Q(X_1)X_2)) (I - P(X_1)) \quad (105)$$

$$= (I - P(X_1) - P(Q(X_1)X_2) + P(Q(X_1)X_2)P(X_1))$$

$$= (I - P(X_1) - P(Q(X_1)X_2)) \quad (106)$$

であり、左辺が、 $[X_1, X_2]$ と $[X_1, Q(X_1)X_2]$ の列空間が等しいことと後者の各列が互いに直交していること、ならびに、(83) 式により

$$Q([X_1, X_2])X_3 = (I - P([X_1, X_2])) \quad (107)$$

$$= (I - P([X_1, Q(X_1)X_2])) \quad (108)$$

$$= (I - P(X_1) - P(Q(X_1)X_2)) \quad (109)$$

となることから確認することが出来る。

なお、(86) 式と (91) 式より

$$Z^{(2)} = Z^{(1)}A^{(2)} = X A^{(1)}A^{(2)} \quad (110)$$

であるから、係数行列として以下に示す $A^{(12)}$ を用いれば X から直接 $Z^{(2)}$ を求めることが出来る。いま、

$$A^{(12)} = A^{(1)} A^{(2)} = \begin{bmatrix} I & -(X_1' X_1)^{-1} X_1' X_2 & (X_1' X_1)^{-1} X_1' X_2 (X_2' Q(X_1) X_2)^{-1} X_2' Q(X_1) X_3 - (X_1' X_1)^{-1} X_1' X_3 \\ O & I & -(X_2' Q(X_1) X_2)^{-1} X_2' Q(X_1) X_3 \\ O & O & I \end{bmatrix} \quad (111)$$

とし、これを X の右から掛けたものを

$$X A^{(12)} = [X_1, X_1 A_{12}^{(12)} + X_2, X_1 A_{13}^{(12)} + X_2 A_{23}^{(12)} + X_3] = [Z_1^{(2)}, Z_2^{(2)}, Z_3^{(2)}] = Z^{(2)} \quad (112)$$

とすれば、 $Z_1^{(2)} = X_1$ であり、

$$Z_2^{(2)} = -X_1 (X_1' X_1)^{-1} X_1' X_2 + X_2 = Q(X_1) X_2 \quad (113)$$

また

$$Z_3^{(2)} = X_1 (X_1' X_1)^{-1} X_1' X_2 (X_2' Q(X_1) X_2)^{-1} X_2' Q(X_1) X_3 - X_1 (X_1' X_1)^{-1} X_1' X_3 - X_2 (X_2' Q(X_1) X_2)^{-1} X_2' Q(X_1) X_3 + X_3 \quad (114)$$

$$= (I - Q(X_1)) X_2 (X_2' Q(X_1) X_2)^{-1} X_2' Q(X_1) X_3 - (I - Q(X_1)) X_3 - X_2 (X_2' Q(X_1) X_2)^{-1} X_2' Q(X_1) X_3 + X_3 \quad (115)$$

$$= -Q(X_1) X_2 (X_2' Q(X_1) X_2)^{-1} X_2' Q(X_1) X_3 + Q(X_1) X_3 \quad (116)$$

$$= -Q(X_1) X_2 (X_2' Q(X_1) X_2)^{-1} X_2' Q(X_1) Q(X_1) X_3 + Q(X_1) X_3 \quad (117)$$

$$= -P(Q(X_1) X_2) Q(X_1) X_3 + Q(X_1) X_3 \quad (118)$$

$$= Q(Q(X_1) X_2) Q(X_1) X_3 \quad (119)$$

となり、(91) 式 と一致する。

まとめると、 X が三つの列グループに分割されている場合、

$$Z^{(2)} = [X_1, Q(X_1) X_2, Q([X_1, X_2]) X_3] \quad (120)$$

という互いに直交する列ブロックを持つ $Z^{(2)}$ は、(91) 式 ならびに (110) 式 のふた通りの方法で求めることが出来る。また、(39) 式 を再帰的に利用することにより

$$P([X_1, X_2, X_3]) = P([X_1, Q(X_1) X_2, Q([X_1, X_2]) X_3]) \quad (121)$$

$$= P(X_1) + P(Q(X_1) X_2) + P(Q([X_1, X_2]) X_3) \quad (122)$$

と分解されるが、

$$P(Q([X_1, X_2]) X_3) = P(Q(Q(X_1) X_2) Q(X_1) X_3) \quad (123)$$

である。

なお、 X が四つの列グループに分割されている場合には、(91) 式 の第 3 列ブロックが以下のように分割されていることになるが、

$$Z^{(2)} = [X_1, Q(X_1) X_2, Q([X_1, X_2]) X_3, Q([X_1, X_2]) X_4] \quad (124)$$

これの第 1 列ブロックと第 2 列ブロックをまとめると

$$\mathbf{Z}^{(2)} = [[\mathbf{X}_1, \mathbf{Q}(\mathbf{X}_1)\mathbf{X}_2], \mathbf{Q}([\mathbf{X}_1, \mathbf{X}_2])\mathbf{X}_3, \mathbf{Q}([\mathbf{X}_1, \mathbf{X}_2])\mathbf{X}_4] \quad (125)$$

となり、第 1 列ブロックには $[\mathbf{X}_1, \mathbf{X}_2]$ と同じ列空間を持つ行列が、そして第 2 ブロックと第 3 ブロックにはそれに直交する \mathbf{X}_3 と \mathbf{X}_4 が含まれる形となり、(86) 式と同じ形をしていることが分かる。したがって、(125) 式の第 1 列ブロック $[\mathbf{X}_1, \mathbf{Q}(\mathbf{X}_1)\mathbf{X}_2]$ を \mathbf{X}_1 、第 2 列ブロック $\mathbf{Q}([\mathbf{X}_1, \mathbf{X}_2])\mathbf{X}_3$ を \mathbf{X}_2 、そして第 3 列ブロック $\mathbf{Q}([\mathbf{X}_1, \mathbf{X}_2])\mathbf{X}_4$ を \mathbf{X}_3 と読み替えて上記のプロセスを繰り返せば、 $\mathbf{Z}^{(2)}$ として

$$\mathbf{Z}^{(3)} = [[\mathbf{X}_1, \mathbf{Q}(\mathbf{X}_1)\mathbf{X}_2], \mathbf{Q}([\mathbf{X}_1, \mathbf{X}_2])\mathbf{X}_3, \mathbf{Q}([\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3])\mathbf{X}_4] \quad (126)$$

$$= [\mathbf{X}_1, \mathbf{Q}(\mathbf{X}_1)\mathbf{X}_2, \mathbf{Q}([\mathbf{X}_1, \mathbf{X}_2])\mathbf{X}_3, \mathbf{Q}([\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3])\mathbf{X}_4] \quad (127)$$

を得る。列ブロック数が 5 以上の場合も同様である。

一般に、 \mathbf{X} が

$$\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_G] \quad (128)$$

と G 群に分割されている場合、直交化のための $G \times G$ の係数行列 \mathbf{A} を

$$\mathbf{A}^{(i)} = \begin{bmatrix} \mathbf{I} & \mathbf{O} & \dots & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{I} & \dots & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{I} & \mathbf{A}_{i,i+1} & \dots & \mathbf{A}_{iG} \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} & \mathbf{I} & \dots & \mathbf{O} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} & \mathbf{O} & \dots & \mathbf{I} \end{bmatrix}, \quad i \text{ 行 } i = 1, 2, \dots, G-1$$

i 列

(129)

用いて

$$\mathbf{A} = \prod_{i=1}^{G-1} \mathbf{A}_i \quad (130)$$

と定義する。ただし \mathbf{A}_{ij} , $i < j$ は $\mathbf{Q}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{i-2})\mathbf{X}_j$ を

$$\mathbf{Z}^{(i-1)} = \left[\mathbf{X}_1 \quad \mathbf{Q}(\mathbf{X}_1)\mathbf{X}_2 \quad \mathbf{Q}(\mathbf{X}_1, \mathbf{X}_2)\mathbf{X}_3 \dots \mathbf{Q}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{i-2})\mathbf{X}_{i-1} \right] \quad (131)$$

に回帰したときの $\mathbf{Q}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{i-2})\mathbf{X}_{i-1}$ に掛かる回帰係数行列の符号を変えたものである。

一般に、 $\mathbf{Z}^{(i)}$ 行列の各列ブロックは $i-1$ 列ブロックまでは互いに直交し、 i 列ブロック以降は $i-1$ 列ブロックと直交する。すると、

$$\mathbf{Z}^{(1)} = \mathbf{X}\mathbf{A}^{(1)} \quad (132)$$

$$\mathbf{Z}^{(2)} = \mathbf{Z}^{(1)}\mathbf{A}^{(2)} = \mathbf{X}\mathbf{A}^{(1)}\mathbf{A}^{(2)} \quad (133)$$

$$\mathbf{Z}^{(3)} = \mathbf{Z}^{(2)}\mathbf{A}^{(3)} = \mathbf{X}\mathbf{A}^{(1)}\mathbf{A}^{(2)}\mathbf{A}^{(3)} \quad (134)$$

$$\vdots \quad \vdots \quad (135)$$

$$\mathbf{Z}^{(G-1)} = \mathbf{Z}^{(G-2)}\mathbf{A}^{(G-1)} = \mathbf{X} \prod_{i=1}^{G-1} \mathbf{A}^{(i)} = \mathbf{X}\mathbf{A} \quad (136)$$

である。

1.4 ベクトルの線形独立と行列の階数

$a \times b$ の行列 $A = [a_{(1)} \ a_{(2)} \ \dots \ a_{(b)}]$ のグラム・シュミットの直交化を行った場合、途中で残差が 0 となる列が出現する場合がある。そのような場合は u を $\mathbf{0}$ として先に進むと、最終的に完成する Q 行列のいくつかの列が $\mathbf{0}$ となっている場合がある。このような場合、行列 A の列は線形独立では無いという。また、 A の列数 b から Q 行列の全てが 0 である列の数を引いたものが行列の階数 (rank) と呼ばれる。 Q 行列が $\mathbf{0}$ を一つも含まない場合、すなわち、階数が列数に等しい場合、 A の b 個の列ベクトルは線形独立であるという。また、階数が b より小さい場合は、いくつかのベクトルが他のベクトルの線形結合で書き表されるという意味で、線形従属であると呼ぶ。

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

ある行列に正則な行列を掛けた行列の階数は元の行列の階数に等しい。また、 $A'A$ 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、 A が行フルランクの場合、 $(AA')^{-1}$ が存在する。

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

$$A = PQ' \tag{137}$$

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

3 : 行列の階数

- A のサイズが $a \times b$ の場合、 $\text{rank}(A) \leq \min(a, b)$
- $\text{rank}(A) = \text{rank}(A') = \text{rank}(A'A) = \text{rank}(AA')$
- $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$
- B のサイズが $b \times r$, $r \leq b$ で $\text{rank}(B) = r$ の場合 $\text{rank}(AB) = r$
- P と Q が正則の場合、 $\text{rank}(PAQ) = \text{rank}(A)$
- A が特異の場合、行列式の値は 0 となる。
- A が対称の場合 $\text{rank}(A) = A$ の 0 でない固有値の数

2 連立一次方程式の一般解と一般逆行列

2.1 連立一次方程式の一般解

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

$$Ax = c \quad (138)$$

が解を持つためには、 A の階数が、 A の横に c を付け加えたものの階数に等しい

$$\text{rank}(A) = \text{rank}([A, c]) \quad (139)$$

ことが必要である。

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

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

が唯一の解となる。また、 $b \times b$ の A 行列が正則な場合はどのような $b \times 1$ のベクトルも A の列空間に含まれるため、(139) 式が満たされる。

係数行列 A の逆行列が存在しない場合（ A 行列が特異行列の場合）には、(139) 式は c が A の列空間に含まれることを意味し、その場合には、解は一意には定まらないが方程式を満たす x は存在する、すなわち、解は不定 (indeterminate) ということになる。そして、(139) 式が満たされていないならばこの方程式は不能 (inconsistent)、すなわち、方程式を満たす x は存在しない場合となる。

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

$$x^* = A^-c + (I - A^-A)y \quad (141)$$

と書くことが出来る。ただし、 A^- は

$$A = AA^-A \quad (142)$$

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。なお、上式の $(I - A^-A)$ の部分は、行列 A の零空間 (null space) とか核空間 (Kernel) と呼ばれるもので、

$$A(I - A^-A) = A - AA^-A = A - A = O \quad (143)$$

となるものである。

なお、 A がフルランクの場合 $A^- = A^{-1}$ と書けるから、この一般解は (140) 式と一致する。また、 A が列フルランクの場合、階数 b と未知数 x の数は等しく

$$A^- = (A'A)^{-1}A' \quad (144)$$

であるから、

$$(I - A^-A) = (I - (A'A)^{-1}A'A) = O \quad (145)$$

となり、解は

$$\boldsymbol{x} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{c} \quad (146)$$

と一意に定まる。

この一般解が確かに解となっていることを示すためにもとの方程式に代入すると、

$$\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{A}(\boldsymbol{A}^{-}\boldsymbol{c} + (\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A})\boldsymbol{y}) = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{c} + \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A})\boldsymbol{y} \quad (147)$$

$$= \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{c} + \boldsymbol{A}\boldsymbol{y} - \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{y} = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{c} \quad (148)$$

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

$$\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{c} = \boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{z} = \boldsymbol{A}\boldsymbol{z} = \boldsymbol{c} \quad (149)$$

となる。

なお、(139) 式が満たされていない場合はこの方程式は解を持たず、(141) 式は計算できるもののそれは解では無いことに注意すべきである。

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

$$\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{A}^{-}\boldsymbol{A})\boldsymbol{y} \quad (150)$$

となる。したがって斉次連立方程式の自明でない解は不定である。

また、 $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots$ を斉次方程式 $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ の解であるとすれば、

$$\sum_i \boldsymbol{A}\boldsymbol{x}_i = \boldsymbol{A} \sum_i \boldsymbol{x}_i = \mathbf{0} \quad (151)$$

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

2.1.1 回帰分析との関係

重回帰分析モデルは、データの数を n 、説明変数の数を q とすれば

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e} \quad (152)$$

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

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} \quad (153)$$

を $\boldsymbol{\beta}$ に関する連立方程式と見做して

$$\boldsymbol{\beta} = \boldsymbol{X}^{-}\boldsymbol{y} + (\boldsymbol{I} - \boldsymbol{X}^{-}\boldsymbol{X})\boldsymbol{c} \quad (154)$$

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

$$\mathbf{X}^- = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \quad (155)$$

を用いれば、この方程式の見かけ上の解は

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (156)$$

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

$\hat{\boldsymbol{\beta}}$ は

$$\text{RSS} = \text{tr} \left[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right] \quad (157)$$

を最小とする $\boldsymbol{\beta}$ の推定値として

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (158)$$

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

$$(\mathbf{X}'\mathbf{X})\boldsymbol{\beta} = \mathbf{X}'\mathbf{y} \quad (159)$$

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

3 一般逆行列について

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

$$g1) \quad A = AA^-A \quad (160)$$

$$g2) \quad A^- = A^-AA^- \quad (161)$$

$$g3) \quad AA^- = (AA^-)' \quad (162)$$

$$g4) \quad A^-A = (A^-A)' \quad (163)$$

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

また、条件 g1 に加えて、条件 g2 を満たす行列を反射型一般逆行列 (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 (g3 と g4) をも満たすものをムーア・ペンローズ逆行列 (Moor-Penrose inverse, the pseud inverse, 疑似逆行列) と呼ぶが、この行列は一意に定まり、特に A^+ と記される。なお A が正則な場合は $A^+ = A^- = A^{-1}$ である。

数値例として、以下の対称な正方行列 A

$$A = \begin{bmatrix} 8 & 8 & 0 \\ 8 & 9 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad (164)$$

のいくつかの一般逆行列とムーア・ペンローズ逆行列を示すが、 B_1 は条件 g1 を、 B_2 は条件 g1 と g2 を、 B_3 は条件 g1, g2, と g3 を、 B_4 は条件 g1, g2, と g4 を満たす。

$$\begin{aligned} B_1 &= \begin{bmatrix} 65/8 & -8 & 7 \\ 6 & -6 & 7 \\ 3 & -3 & 3 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 9/8 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 5/12 & -7/24 & -17/24 \\ -1/3 & 1/3 & 2/3 \\ 0 & 0 & 0 \end{bmatrix} \\ B_4 &= \begin{bmatrix} 5/12 & -1/3 & 0 \\ -7/24 & 1/3 & 0 \\ -17/24 & 2/3 & 0 \end{bmatrix}, \quad A^+ = \begin{bmatrix} 1/6 & -1/12 & -1/4 \\ -1/12 & 1/8 & 5/24 \\ -1/4 & 5/24 & 11/24 \end{bmatrix} \end{aligned} \quad (165)$$

また、矩形行列

$$A = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad (166)$$

のいくつかの一般逆行列とムーア・ペンローズ逆行列を示すが、 B_1 は条件 g1 を、 B_2 は条

件 g_1 と g_2 を、 B_3 は条件 $g_1, g_2,$ と g_3 を満たす。

$$\begin{aligned}
 B_1 &= \begin{bmatrix} -5/2 & 3 & 7 & -1 \\ -5 & 5 & 7 & 1 \\ -7 & 7 & 3 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1/2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 1/4 & 1/4 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
 A^+ &= \begin{bmatrix} 1/6 & 1/6 & 0 & -1/3 \\ 1/12 & 1/12 & 0 & 1/3 \\ -1/12 & -1/12 & 0 & 2/3 \end{bmatrix} \tag{167}
 \end{aligned}$$

3.1 一般逆行列

3.1.1 一般逆行列の性質

一般逆行列に関しては以下の性質が知られている。

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。

$$(A')^- = (A^-)' \tag{168}$$

証明:

A' の一般逆行列は

$$A'(A')^-A' = A' \tag{169}$$

を満たさなければならないが、これに (168) 式 を代入すると

$$A'(A')^-A' = A'(A^-)'A' = (AA^-A)'' = A' \tag{170}$$

を得る。

証明終わり

以下に示す行列は冪等である。

$$AA^-, A^-A \tag{171}$$

(証明略)

T を正則行列とすれば

$$(AT)^- = T^{-1}A^- \quad \text{and} \quad (TA)^- = A^-T^{-1} \tag{172}$$

である。

証明:

AT の一般逆行列は条件

$$AT(AT)^-AT \tag{173}$$

を満たさなければならないが、これに (172) 式 を代入すると

$$AT(AT)^-AT = AT T^{-1}A^-AT \tag{174}$$

$$= AA^-AT = AT \tag{175}$$

を得る。

証明終わり

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、

$$B^- B = I \quad \text{and} \quad C C^- = I \quad (176)$$

であり、

$$(BC)^- = C^- B^- \quad (177)$$

である。

証明:

$B^- B$ を K と置き、両辺の左から B を掛けることにより

$$B B^- B = B = B K \quad (178)$$

を得る。また、この最後の等式の両辺の左から B' を掛けると

$$B' B = B' B K \quad (179)$$

を得るが $B' B$ は逆行列を持つため $K = I$ を得る。

証明終わり

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

$$B_1 = A^- + Y(I - AA^-), \quad B_2 = A^- + (I - A^-A)Y \quad (180)$$

$$B_3 = A^- + YQ(A), \quad B_4 = A^- + Q(A)Y \quad (181)$$

証明:

$$(I - AA^-)A = A(I - A^-A) = Q(A)A = AQ(A) = O \quad (182)$$

となり、常に条件 g1 が満たされるからである。

証明終わり

3.1.2 一般逆行列の生成

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

$$PAQ = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = B \quad (183)$$

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

$$B^- = \begin{bmatrix} B_{11}^{-1} & O \\ O & O \end{bmatrix} \quad (184)$$

として得られることから、

$$A^- = QB^-P \quad (185)$$

として得られる。

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

$$BB^-B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{21}B_{11}^{-1}B_{12} \end{bmatrix} \quad (186)$$

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

$$\begin{pmatrix} B_{12} \\ B_{22} \end{pmatrix} = \begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix} C \quad (187)$$

と書けるが、これは

$$B_{12} = B_{11}C, \quad \text{and} \quad B_{22} = B_{21}C \quad (188)$$

を意味し、 $C = B_{11}^{-1}B_{12}$ を得るため、(186) 式の第 2-2 要素は B_{22} となり、一般逆行列が満たすべき条件 (160) 式が満たされることを示している。また、並べ替え行列は自分自身が自分自身の逆行列となっているため (183) 式より

$$A = PBQ \quad (189)$$

となるが、これと (185) 式より、

$$A A^- A = PBQ QB^-P PBQ \quad (190)$$

$$= P BB^-B Q = PBQ = A \quad (191)$$

を得る。

(160) 式と (161) 式を共に満たす一般逆行列 g2-inverse はスweep演算子を用いて簡単に計算することが出来る。基本的にスweep演算子は、正方行列の中の正則部分の逆行列を計算し、それ以外の部分を 0 と置く操作をするため、行列が矩形の場合は、正方となるように全てが 0 の列や行を付け加えてからスweep演算子を用いたのちに、元の矩形行列に戻せば良い。

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

$$A = QR \quad (192)$$

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

$$A^- = R^- Q' \quad (193)$$

である。(証明略)

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

$$A^- = R^{*-} Q^{*-} \quad (194)$$

である。

証明:

階数分解と QR 分解の間には、 T を正則行列として、

$$Q^* = QT \quad \text{and} \quad R^* = T^{-1}R \quad (195)$$

という関係があるため、

$$AA^{-}A = AR^*Q^*A = ARTT^{-1}Q^{-}A \quad (196)$$

$$= QRR^{-}Q^{-}QR = QRR^{-}R = QR \quad (197)$$

を得る。

証明終わり

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)

$$A^{-}AA^{-}, A'(AA')^{-}, (A'A)^{-}A' \quad (198)$$

最後のふたつは反射型一般逆行列となる。

証明:

$A^{-} = A^{-}AA^{-}$ の場合

$$AA^{-}A = A A^{-}AA^{-} A \quad (199)$$

$$= AA^{-}A A^{-}A = A A^{-}A = A \quad (200)$$

$A^{-} = (A'A)^{-}A'$ の場合

$$AA^{-}A = A (A'A)^{-}A' A \quad (201)$$

$$= A(A'A)^{-}A' A = A \quad (202)$$

$$A^{-}AA^{-} = (A'A)^{-}A' A (A'A)^{-}A' \quad (203)$$

$$= (A'A)^{-}A' A(A'A)^{-}A' = (A'A)^{-}A' = A^{-} \quad (204)$$

証明終わり

(198) 式の第 2 式と第 3 式の特例として A が行フルランクもしくは列フルランクの場合には

$$A^{-} = A'(AA')^{-1} \quad \text{or} \quad A^{-} = (A'A)^{-1}A' \quad (205)$$

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

3.2 ムーアペンローズ逆行列

3.2.1 ムーアペンローズ逆行列の生成

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

$$A = U_r D_r V_r', \quad U_r' U_r = V_r' V_r = I_r, \quad D \text{ は対角} \quad (206)$$

とすると

$$A^+ = V_r D_r^{-1} U_r' \quad (207)$$

で与えられる。(証明略)

また、 A の階数分解を (65) 式 とすると

$$A^+ = R^*(R^*R^*)^{-1}(Q^*Q^*)^{-1}Q^* \quad (208)$$

である。なお、 L が列フルランク、 R が行フルランクの場合

$$L_L^{-1} = (L'L)^{-1}L' \quad \text{and} \quad R_R^{-1} = R'(RR')^{-1} \quad (209)$$

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

また、ムーアペンローズ逆行列は任意の一般逆行列を用いて

$$A^+ = A'A(A'AA'A)^-A' \quad (210)$$

$$A^+ = AA'(AA'AA')^-A \quad (211)$$

$$A^+ = A'(AA')^-A(A'A)^- \quad (212)$$

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

なお、

$$A^+AA^+, \quad A'(AA')^+, \quad (A'A)^+A' \quad (213)$$

は A 行列のムーアペンローズ逆行列となる。(証明略)

3.2.2 ムーアペンローズ逆行列の性質

ムーアペンローズ一般逆行列のムーアペンローズ逆行列は元の行列となる。

$$(A^+)^+ = A \quad (214)$$

また、以下のことが言える。

$$(A'A)^+ = A^+(A')^+ \quad (215)$$

$$(AA^+)^+ = AA^+ \quad (216)$$

$$(A^+A)^+ = A^+A \quad (217)$$

$$A(A'A)^-A' = AA^+ \quad (218)$$

4 行列の固有値と特異値分解

a 次の正方行列 A に関して

$$A p = \lambda p \quad (219)$$

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。

上式を変形すれば、

$$(A - \lambda I) p = 0 \quad (220)$$

となるが、これは p に関する斉次方程式であり、 $p = 0$ 以外の解が存在するためには係数行列 $(A - \lambda I)$ の行列式が 0

$$|A - \lambda I| = 0 \quad (221)$$

でなければならぬが、これを λ の方程式とみたものを固有方程式と呼ぶ。例えば、 $a = 2$ の場合

$$|A - \lambda I| = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12} * a_{21} = 0 \quad (222)$$

という λ に関する2次方程式となる。

上記の定義から明らかなように固有ベクトルの長さは一意に定まらず、通常はそのノルムを 1 ($\|p\|^2 = 1$) 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

$$A P_q = P_q \Lambda_q \quad (223)$$

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。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、

$$P_q' P_q = I \quad (224)$$

であるため、

$$P_q' A P_q = \Lambda_q \quad (225)$$

である。

4.1 固有値分解

対称行列 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、

$$P' P = I \quad \text{and} \quad P P' = I \quad (226)$$

である。したがって、

$$P' A P = \Lambda \quad (227)$$

であり、

$$A = P \Lambda P' \quad (228)$$

と書けるが、これを A の固有値分解（スペクトル分解）（eigen decomposition, spectral decomposition）と呼ぶ。

すなわち、対称行列 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り、その階数はその行列のゼロでない固有値の個数に等しい。

行列の固有値分解を用いると、 A の逆行列は

$$A^{-1} = P \Lambda^{-1} P' \quad (229)$$

と書ける。また、行列の冪乗は

$$A^n = P \Lambda^n P' \quad (230)$$

となる。すなわち、行列をべき乗したものの固有値は元の固有値をべき乗したものとなり、固有ベクトルは変わらない。

$$\text{tr}(A) = \text{tr}(P \Lambda P') = \text{tr}(\Lambda P' P) = \text{tr}(\Lambda) = \sum_{i=1}^a \lambda_i \quad (231)$$

である。

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、それらに対角要素を持つ対角行列を Λ_r とし、それらに対応する固有ベクトルを列に持つ $a \times r$ の行列を P_r とすれば

$$A = P_r \Lambda_r P_r', \quad P_r' P_r = I, \quad P_r P_r' \neq I \quad (232)$$

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

$$A \neq P_q \Lambda_q P_q' \quad (233)$$

であるが、右辺を \hat{A} と置けば、それは A 行列の階数 q の最小2乗近似、すなわち

$$\text{RSS} = \text{tr} \left[(A - \hat{A})' (A - \hat{A}) \right], \quad \text{rank}(\hat{A}) = q \quad (234)$$

を最小とするものとなっている。

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

$$A = B B' \quad (235)$$

の形に分解することが出来る。ただし、

$$B = P_r \Lambda_r^{\frac{1}{2}} T \quad (236)$$

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

$$B = P_r \Lambda_r^{\frac{1}{2}} P_r' \quad (237)$$

で定義される行列は

$$BB = P_r \Lambda_r^{\frac{1}{2}} P_r' P_r \Lambda_r^{\frac{1}{2}} P_r' = P_r \Lambda_r^{\frac{1}{2}} P_r' P_r \Lambda_r^{\frac{1}{2}} P_r' = P_r \Lambda_r^{\frac{1}{2}} \Lambda_r^{\frac{1}{2}} P_r' = A \quad (238)$$

となるため、行列 A の対称な平方根行列 (symmetric square root matrix) と呼ばれることがある。

また、

$$A^+ = P_r \Lambda_r^{-1} P_r' \quad (239)$$

は

$$AA^+ A = P_r \Lambda_r P_r' P_r \Lambda_r^{-1} P_r' P_r \Lambda_r P_r' = P_r \Lambda_r P_r' = A \quad (240)$$

等を示せるため、 A のムーア・ペンローズ逆行列になる。

$\mathbf{0}$ では無いどのようなベクトル x を用いても2次形式が正

$$x' A x > 0 \quad (241)$$

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、 $A'A$ 行列や AA' 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。

4.2 特異値分解

4.2.1 特異値分解の存在

$a \times b$ の行列 X は

$$X = U D V' \quad (242)$$

という形に分解することが出来る。ただし、 U は $a \times a$ の直交行列、 D は $a \times b$ の行列で $a > b$ の場合は $b \times b$ の対角行列の下に $(a - b) \times b$ の 0 を加えたもの、 $a < b$ の場合は $a \times a$ の対角行列の右に $a \times (b - a)$ の 0 を加えたもの、 V は $b \times b$ の直交行列である。

このような分解が可能なおよそ以下のようにして示せる。まず、 V が直交行列であるとすると、(242) 式の両辺の右から VD^{-} を掛けると

$$X V D^{-} = U D V' V D^{-} \quad (243)$$

$$= U D D^{-} \quad (244)$$

を得るが、 DD^{-} は単位行列の下や右に 0 を加えた $b \times a$ の行列である。したがって、 U 行列は

$$U = XVD^{-} \quad (245)$$

と書くことが出来るが、この行列の D^{-} の 0 の対角要素に対応する列には 0 が入っている。
次に、この U 行列が直交するという事より

$$U'U = D^{-'}V'X'XVD^{-} = I \text{ に 0 を付加したもの} \quad (246)$$

である。ただし、(245) 式 で定義される U の 0 部分を適当に補えば、完全に U を直交化することも可能である。

ここで、 V を $X'X$ の固有ベクトルであるとし、その固有値を Λ とすれば

$$X'XV = V\Lambda \quad (247)$$

より

$$V'X'XV = \Lambda \quad (248)$$

であるため、上記の直交条件は

$$U'U = D^{-'}V'X'XVD^{-} \quad (249)$$

$$= D^{-'}\Lambda D^{-} \quad (250)$$

となるが、ここで

$$D = \Lambda^{\frac{1}{2}} \text{ に 0 を付加したもの} \quad (251)$$

と定義すれば上記の直交条件は満たされる。

また、(247) 式 の左から X を掛けると

$$XX'XV = XV\Lambda \quad (252)$$

となるが、これは XV が、したがって、その各列の定数倍である U が XX' の固有ベクトルであること、ならびに $X'X$ と XX' の固有値は等しいことを示している。

以上より

- U は XX' の固有ベクトル
- V は $X'X$ の固有ベクトル
- D は $X'X$ もしくは XX' の固有値の平方根からなる対角行列

となる。

4.2.2 特異値分解

今、 X を $a \times b$ の階数 r の行列とし、 $X'X$ ならびに XX' の縮小された固有値分解を

$$X'X = V_r D_r^2 V_r', \quad XX' = U_r D_r^2 U_r' \quad (253)$$

ただし、 V_r は $b \times r$ の行列、 U_r は $a \times r$ の行列、 D_r は $r \times r$ の対角行列、とすれば、

$$X = U_r D_r V_r' \quad (254)$$

と書きあらわすことができる。これを行列 X の特異値分解 (SVD, Singular Value Decomposition) と呼ぶ。一般的に

$$V_r' V_r = I_r, \quad \text{and} \quad U_r' U_r = I_r \quad (255)$$

であるが、 X の階数が $b = r$ の場合は

$$V_r' V_r = V_r V_r' = I_r \quad (256)$$

であり、 X の階数が $a = r$ の場合は

$$U_r' U_r = U_r U_r' = I_r \quad (257)$$

である。

また、 D_r をさらに分解したものをいれれば

$$X = U_r D_r V_r' = U_r D_r^{\frac{1}{2}} D_r^{\frac{1}{2}} V_r' = U^* (V^*)' \quad (258)$$

ただし、

$$U^* = U_r D_r^{\frac{1}{2}}, \quad \text{and} \quad V^* = V_r D_r^{\frac{1}{2}} \quad (259)$$

と書ける。

なお、 P 等として、階数が $q, q \leq r$ のものを用いた場合、

$$\hat{X} = U_q D_q V_q' \quad (260)$$

は

$$\text{RSS} = \sum_{i=1}^a \sum_{j=1}^b (x_{ij} - \hat{x}_{ij})^2 \quad (261)$$

を最小とする階数 q の行列となることが知られている。これを行列 X の Eckart-Young Decomposition と呼ぶ。

また、これを

$$\hat{X} = F A' \quad (262)$$

ただし、

$$F = a U_q \quad \text{and} \quad A = \frac{1}{a} V_q D_q \quad (263)$$

と書くことも可能であるが、これは主成分分析に他ならない。

4.2.3 補足

4.2.3.1 正定値行列

全ての x に関して

$$x'Ax > 0 \quad (264)$$

を満たす行列を正定値行列と呼ぶ。 A の固有値分解を使うと

$$x'Ax = x'P\Lambda P'x = y'\Lambda y > 0 \quad (265)$$

y の要素を順に $y = (1, 0, 0, \dots)$, $y = (0, 1, 0, \dots)$ とすることにより、全ての固有値が正であることと同値であることが示せる。すなわち、正定値行列の固有値は全て正、半正定値行列の固有値は全て非負である。

4.2.3.2 応用例

4 :

A が $n \times r$, $n \geq r$ の行列、 D が $r \times r$ の対称行列の時

$$B = AT \quad (266)$$

で作られる行列 B が

$$B'B = D \quad (267)$$

となるような $r \times r$ の T 行列を求めよ。

解

$A'A$ は半正定値の対称行列であるのでその固有値分解を

$$A'A = P\Lambda P', \quad P'P = I_r \quad (268)$$

また、 D の固有値分解を

$$D = Q\Delta Q', \quad Q'Q = I_r \quad (269)$$

とすると、 R を任意の $r \times r$ の直交行列として

$$T = P\Lambda^{-\frac{1}{2}}R\Delta^{\frac{1}{2}}Q' \quad (270)$$

が求める解である。(一意には定まらない。)

なぜなら、

$$B'B = T'A'AT = T'P\Lambda P'T \quad (271)$$

$$= Q\Delta^{\frac{1}{2}}R'\Lambda^{-\frac{1}{2}}P'P\Lambda P'P\Lambda^{-\frac{1}{2}}R\Delta^{\frac{1}{2}}Q' \quad (272)$$

$$= Q\Delta^{\frac{1}{2}}R'\Lambda^{-\frac{1}{2}}\Lambda\Lambda^{-\frac{1}{2}}R\Delta^{\frac{1}{2}}Q' \quad (273)$$

$$= Q\Delta^{\frac{1}{2}}R'R\Delta^{\frac{1}{2}}Q' \quad (274)$$

$$= Q\Delta^{\frac{1}{2}}\Delta^{\frac{1}{2}}Q' = Q\Delta Q' = D \quad (275)$$

である。

直交行列の自由な要素

R を $r \times r$ として

$$R'R = I \quad (276)$$

は対角要素が全て 1、非対角要素が全て 0 という制約であるが、 $R'R$ は対称であるため、実質的に 0 の制約は $r(r-1)/2$ 個である。したがって制約の数は $r + r(r-1)/2 = r(r+1)/2$ であり、自由に決められる要素の数は $r(r-1)/2$ である。

$B'B = D$ という制約条件は対称であるため実質的に $r(r+1)/2$ 個の制約である。しかし T は $r \times r$ の行列であるため $r^2 - r(r+1)/2 = r(r-1)/2$ 個の自由な要素があると考えられるが、これが R の自由な要素に対応している。

$B'B = \text{対角行列}$ という制約

制約条件の数は $r(r-1)/2$ であり、 $B'B = D$ よりも弱い。上記の解で D を対角行列とし、 $Q = I, \Delta = D$ と置き、 D の対角要素は未知とする。したがって、 $B'B = \Delta^{\frac{1}{2}} R' R \Delta^{\frac{1}{2}}$ となるが、 R が $R'R = \text{対角行列}$ という $r(r-1)/2$ 個の制約を満たせばこの制約は満たされることになるが、このような R は無限に存在する。

T が直交行列の場合

直交行列 T の自由な要素の数は $r(r-1)/2$ である。制約条件 $B'B = \text{対角行列}$ が持つ制約の数と一致するため、この条件で直交行列 T が一意に定まる。すなわち、 $B'B = D$ の制約は強すぎることになる。

制約条件 $B'B = \text{対角行列}$ を満たす直交行列 T は

$$T = P \quad (277)$$

である。なぜなら、

$$B'B = P'A'AP = P' P \Lambda P' P = \Lambda = \text{対角行列} \quad (278)$$

となるからである。(厳密には、対角要素の順序の入れ替えという不定性が残っている。)

\mathbf{X} が $a \times b$ の階数 r の行列の時

$$\mathbf{X} = \mathbf{U}_r \mathbf{D}_r \mathbf{V}_r' \quad (279)$$

を行列 \mathbf{X} の (縮小された compact) 特異値分解と呼ぶ。ただし、 $\mathbf{X}'\mathbf{X}$ ならびに $\mathbf{X}\mathbf{X}'$ の縮小された固有値分解を

$$\mathbf{X}'\mathbf{X} = \mathbf{V}_r \mathbf{D}_r^2 \mathbf{V}_r', \quad \mathbf{X}\mathbf{X}' = \mathbf{U}_r \mathbf{D}_r^2 \mathbf{U}_r' \quad (280)$$

ただし、 \mathbf{V}_r は $b \times r$ の行列、 \mathbf{U}_r は $a \times r$ の行列、 \mathbf{D}_r は $r \times r$ の対角行列で

$$\mathbf{V}_r' \mathbf{V}_r = \mathbf{I}_r, \quad \text{and} \quad \mathbf{U}_r' \mathbf{U}_r = \mathbf{I}_r \quad (281)$$

満たすものとする。