数理統計学における射影

関　本　年　彦

はじめに

k 個の説明変数 $x_1, ..., x_k$ と1個の目的変数 y からなる線形モデル

$$y = x_1\beta_1 + \cdots + x_k\beta_k \tag{i}$$

を資料に当てはめ，最適な係数 $\beta_1, \cdots, \beta_k$ を求めるのに最もよく用いられる手法が最小二乗法である。この場合，線形性とは $\beta_1, \cdots, \beta_k$ に関していうものであって，解析しようとする資料の特性に合わせて，$x_1, ..., x_k$ はそれらを変数とする適当な関数 $g_1(x_1, \ldots, x_k), \ldots, g_k(x_1, \ldots, x_k)$ に置き換えることは往々行われるが，それでも $\beta_1, \ldots, \beta_k$ に関して線形であることには変わりはない。また，指数関数や対数関数などを用いて簡単に (i) の形に還元できるような非線形モデルを用いることもよくある。しかしながら，最小二乗法の **数理** を考究する場面では，よりよい見通しを得るため (i) の線形モデルをその対象にするのが妥当である。

与えられた資料

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}, \qquad \begin{pmatrix} y_1 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$

を少し加工した

$$X=\begin{pmatrix} x_{11}-\bar{x}_1 & x_{12}-\bar{x}_2 & \cdots & x_{1k}-\bar{x}_k \\ x_{21}-\bar{x}_1 & x_{22}-\bar{x}_2 & \cdots & x_{2k}-\bar{x}_k \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}-\bar{x}_1 & x_{n2}-\bar{x}_2 & \cdots & x_{nk}-\bar{x}_k \end{pmatrix},\quad y=\begin{pmatrix} y_1-\bar{y} \\ y_2-\bar{y} \\ \vdots \\ y_n-\bar{y} \end{pmatrix} \qquad \text{(ii)}$$

を用いた方が議論が多少簡単になる[1)]。ここで，$\bar{x}_i=\dfrac{1}{n}\sum_{j=1}^{n}x_{ji}(1\le i\le k)$，$\bar{y}=\dfrac{1}{n}\sum_{j=1}^{n}y_j$ である。

モデル (i) に資料 (ii) を当てはめたとき，ε_j を誤差項として

$$y_j-\bar{y}=(x_{j1}-\bar{x}_1)\beta_1+\cdots+(x_{jk}-\bar{x}_k)\beta_k+\varepsilon_j \qquad 1\le j\le n \qquad \text{(iii)}$$

となるが，これを (ii) の行列を用いて書き直すと

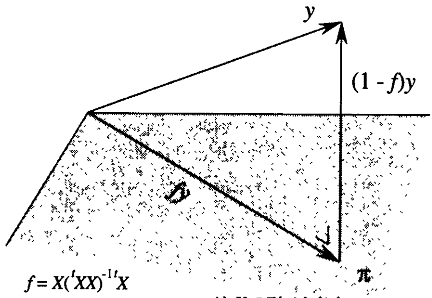
$$y=X\beta+\varepsilon \qquad \beta={}^t(\beta_1,\ldots,\beta_k),\ \varepsilon={}^t(\varepsilon_1,\ldots,\varepsilon_n) \qquad \text{(iv)}$$

となる。最小二乗法は，係数 β として

$$\|\varepsilon\|^2=\|y-X\beta\|^2 \qquad \text{(v)}$$

の値を最小とするような $b={}^t(b_1,\cdots,b_k)$ を採用するのであるが，このような b を見出すのに，行列 $X({}^tXX)^{-1}\,{}^tX$ で表される1次変換が X の列ベクトルから生成される部分空間への正射影であることを利用するところが，最小二乗法数理の核心部分である。

図1　ベクトル y の正射影 fy



1)　(ii) の y は，(i) のそれとは別物であり，以降の y は (ii) の意味である。

実際，行列 $X({}^tXX)^{-1}\,{}^tX$ で表される1次変換が X の列ベクトルから生成される部分空間への正射影であることを容認すれば，任意の $\xi\in R^k$ に対して $y-X({}^tXX)^{-1}\,{}^tXy\perp X\xi$ であるから，

$$\begin{aligned}\|y-X\beta\|^2 &= \|(y-X({}^tXX)^{-1}\,{}^tXy)+(X({}^tXX)^{-1}\,{}^tXy-X\beta)\|^2\\ &= \|y-X({}^tXX)^{-1}\,{}^tXy\|^2\\ &\quad +2\langle y-X({}^tXX)^{-1}\,{}^tXy,\ X(({}^tXX)^{-1}\,{}^tXy-\beta)\rangle\\ &\quad +\|X({}^tXX)^{-1}\,{}^tXy-X\beta\|^2\\ &= \|y-X({}^tXX)^{-1}\,{}^tXy\|^2+\|(X({}^tXX)^{-1}\,{}^tXy-X\beta\|^2\\ &\geq \|y-X({}^tXX)^{-1}\,{}^tXy\|^2\end{aligned}$$

この等号は $\beta=({}^tXX)^{-1}\,{}^tXy$ のとき成り立つのであるから，求める b は，

$$b=({}^tXX)^{-1}\,{}^tXy \tag{vi}$$

である。

1. 準備

1. 1 ユークリッド空間

本論は，実ユークリッド空間 R^n で議論を進める。R は実数の集合，R^n は n 次元列ベクトル全体からなるベクトル空間である。以下では R^n を E とも書くことにする。E の要素は，

$$\begin{pmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{pmatrix},\quad a_i\in R\,(1\leq i\leq n)$$

で，転置行列の記法を用いてこれを ${}^t(a_1,a_2,\ldots,a_n)$ とも書く。a_i $(1\leq$

$i \le n$) を前記ベクトルの第 i 座標という。E の要素で，第 i 座標が 1 で他の座標はすべて 0 であるようなベクトルを e_i で表すと，$e_1, \ldots, e_n$ は E の基底となるが，この基底を E の **自然基底** という。

$x = {}^t(x_1, \ldots, x_n)$, $y = {}^t(y_1, \ldots, y_n) \in E$ に対する **内積** は

$$\langle x, y\rangle = {}^t xy = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \tag{1}$$

で定義され，$\|x\| = \sqrt{\langle x, x\rangle}$ はベクトル x の長さまたは**ノルム**とよばれる。ノルムは以下の性質（**ノルムの公理**）を満たす：

N1. $\|ax\| = |a|\|x\|$ （α は実数），

N2. $\|x+y\| \le \|x\| + \|y\|$ （三角不等式），

N3. $\|x\| = 0$ と $x = 0$ とは同値である。

E から別のユークリッド空間 F への線形写像全体の集合を $\boldsymbol{L}(E;F)$ で表すことにする。$\boldsymbol{L}(E;F)$ は実数上のベクトル空間の構造をもつ。E から E への線形写像を E の **1 次変換**といい，E の 1 次変換全体の集合を $\boldsymbol{L}(E)$ と記すことにする。E に基底を定めると $\boldsymbol{L}(E)$ と実数要素の n 次正方行列全体の集合 $\boldsymbol{M}_n$ との間に 1 対 1 の対応が生じる。$\boldsymbol{L}(E)$，$\boldsymbol{M}_n$ 両者ともベクトル空間であるが，$\boldsymbol{M}_n$ において (i, j) 要素 $(1 \le i, j \le n)$ が 1 で他の要素はすべて 0 であるような行列を考えると，それらは基底を成すから $\dim \boldsymbol{L}(E) = \dim \boldsymbol{M}_n = n^2$ である。さらに，$\boldsymbol{L}(E)$ では $f, g \in \boldsymbol{L}(E)$ に対して合成変換 $g \circ f$ を考えることにより，また $\boldsymbol{M}_n$ では行列の積を考えることにより，$\boldsymbol{L}(E)$，$\boldsymbol{M}_n$ はともに多元環の構造をもつ。

1. 2　双対空間

ベクトル空間 E の 1 次形式（E から R への線形写像）全体からなる集合はベクトル空間となるが，これを E の **双対空間** といい E^* と記す。1

次形式 $\varphi \in E^*$ に限って，$x \in E$ に対する値の通常の記法 $\varphi(x)$ の代わりに $\varphi \cdot x$ とも書くことにする。E の基底 $\{x_1, x_2, \ldots, x_n\}$ に対して，各 $i\,(1 \leq i \leq n)$ について1次形式 x_i^* を $x_i^* \cdot x_j = \delta_{ij}$（Kronecker のデルタ）$(1 \leq j \leq n)$ によって定義すると，$\{x_1^*, x_2^*, \ldots, x_n^*\}$ は E^* の基底になるが[2]これを $\{x_1, x_2, \ldots, x_n\}$ の **双対基底** という。

系　$\dim E = \dim E^*$ である。

註　E^* の双対空間 E^{**} については，$\{x_1^*, x_2^*, \ldots, x_n^*\}$ の双対基底を $\{\tilde{x}_1^*, \tilde{x}_2^*, \ldots, \tilde{x}_n^*\}$ とすると $\tilde{x}_i \mapsto x_i$ によって E^{**} は E と同形となるので，以降混同の恐れがない限り E^{**} を E と同一視する。

$y \in E$ に対して，$\mathbf{\Phi}(y) \in E^*$ を（ベクトル空間 E を明示する場合には $\mathbf{\Phi}_E$ と記す）$\mathbf{\Phi}(y) \cdot x = \langle x, y \rangle\ (x \in E)$ によって定義すると，$\mathbf{\Phi} : E \to E^*$ は線形写像であり，$\mathbf{\Phi}(y) = 0$ ならば $\mathbf{\Phi}(y) \cdot y = \langle y, y \rangle = \|y\|^2 = 0$ であるから $y = 0$，すなわち $\mathbf{\Phi}$ は単射である。さらに，$\dim E = \dim E^* = n$ より $\mathbf{\Phi}$ は全射であり，したがって線形同形であることがわかる。そこで，$\mathbf{\Phi}$ の逆写像を $\mathbf{\Psi}$（ベクトル空間 E を明示する場合には $\mathbf{\Psi}_E$）で表すことにする。

1. 3　転置行列と転置写像

行列 $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ の行と列を入れ替えた行列は $\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$ であるが，一般に，$n \times m$ 行列 A の行と列を入れ替えた $m \times n$ 行列を A の **転置行列** といい，${}^t\!A$ と記す。

---

2)　一般に，E からもう一つの線形空間への線形写像 f は，基底に対する値 $f(x_1), \cdots, f(x_n)$ によって定まるから，任意の $y^* \in E^*$ は　$y^* = \sum_{j=1}^{n} (y^* \cdot x_j) x_j^*$ と表される。

転置行列に関する基本性質　(1) $m \times n$ 行列に関する操作 $A \to {}^tA$ の線形性：${}^t(A+B) = {}^tA + {}^tB$，${}^t(\alpha A) = \alpha\,{}^tA$，(2) 積 AB が定義できるとき ${}^t(AB) = {}^tB\,{}^tA$，(3) 正則行列については $({}^tA)^{-1} = {}^t(A^{-1})$

命題 1．1　A, B を n 次正方行列とする。すべての $x, y \in E$ に対して

$$\langle Ax, y\rangle = \langle x, By\rangle$$

が成り立つならば，$B = {}^tA$ である。

証明　$Ae_j = \sum_{i=1}^{n} e_i \langle e_i, Ae_j\rangle$ $1 \le j \le n$ であるから，A の第 j 列ベクトルは ${}^t(\langle e_1, Ae_j\rangle, \ldots, \langle e_n, Ae_j\rangle)$，すなわち A の (i,j) 要素は $\langle e_i, Ae_j\rangle$ である。一方，$Be_j = \sum_{i=1}^{n} e_i \langle e_i, Be_j\rangle$ であるから，B の (i,j) 要素は $\langle e_i, Be_j\rangle = \langle Ae_i, e_j\rangle$ である。これから $B = {}^tA$ であることがわかる。□

E の 1 次変換 $f \in \boldsymbol{L}(E)$ の E の自然基底に関する表現行列は，その第 j 列ベクトルが ${}^t(\langle e_1, f(e_j)\rangle, \ldots, \langle e_n, f(e_j)\rangle)$ であるような行列であり，この対応は $\boldsymbol{L}(E)$ と $\boldsymbol{M}_n$ とのあいだの線形同形である（多元環としての同形でもある）。

$f \in \boldsymbol{L}(E)$ を E の 1 次変換，$x, y \in E$ とする。$\varphi(y) \in E^*$ を $\varphi(y) \cdot x = \langle f(x), y\rangle$ によって定義すると，φ は E から E^* への線形写像である。そこで，$g = \Psi \circ \varphi$ とおくと g は E の 1 次変換で

$$\langle f(x), y\rangle = \langle x, g(y)\rangle \qquad x, y \in E \tag{2}$$

が成り立つ。任意の 1 次変換 g' が $\langle x, g'(y)\rangle = \langle x, g(y)\rangle$ を満たすと，$\langle x, g'(y) - g(y)\rangle = 0$ で，$x = g'(y) - g(y)$ とおけば $\|g'(y) - g(y)\|^2 = 0$，すなわち $g'(y) = g(y)$ を得るから 1 次変換 g は (2) によって一意的に定まる。

定義 1. 1 （**転置写像**） $f \in \boldsymbol{L}(E)$ を E の１次変換とするとき，(2) によって定義される１次変換を f の **転置写像** または **随伴変換** といい，${}^t\!f$ と記す。この記号を用いて (2) を書き直すと，

$$\langle f(x), y\rangle = \langle x, {}^t\!f(y)\rangle \qquad x, y \in E \tag{2'}$$

である。

１次変換 f の自然基底に関する表現行列を A とすると，転置写像 ${}^t\!f$ の表現行列は ${}^tA$ となる。

転置写像の基本性質 $f, g \in \boldsymbol{L}(E)$, $\alpha \in \boldsymbol{R}$ に対して以下が成り立つ：(1) ${}^{tt}\!f = f$, (2) ${}^t(f+g) = {}^t\!f + {}^tg$, ${}^t(\alpha f) = \alpha\, {}^t\!f$, (3) ${}^t(g \circ f) = {}^t\!f \circ {}^tg$, (4) f が同形ならば $({}^t\!f)^{-1} = {}^t(f^{-1})$。

1. 4 直交性

$x, y \in E$ について，$\langle x, y\rangle = 0$ であるとき x と y は **直交** するといい $x \perp y$ と書く。E の部分集合 S, T について，S の任意の要素が T の任意の要素と直交するとき，S と T とは直交するといい $S \perp T$ と書く。さらに，E の部分集合 S と直交する E の要素全体を $S^{\perp}$ と記す。

基本性質 E をベクトル空間，S, T をその部分集合，$(S_\lambda)_{\lambda\in\Lambda}$ を部分集合族とするとき以下が成り立つ：(1) $S^{\perp}$ は E の部分空間である，(2) $S \subset T$ ならば $S^{\perp} \supset T^{\perp}$，(3) $S^{\perp} = (\boldsymbol{R}S)^{\perp}$，(4) $(\bigcup_{\lambda\in\Lambda} S_\lambda)^{\perp} = \bigcap_{\lambda\in\Lambda} S_\lambda^{\perp}$，(5) $S^{\perp\perp} \supset \boldsymbol{R}S$，$S^{\perp\perp\perp} = S^{\perp}$。

E の要素 $u_1, \ldots, u_k\,(1 \le k \le n)$ について，$\langle u_i, u_j\rangle = \delta_{ij}$ $\;1 \le i, j \le k$ が成り立つとき，$u_1, \ldots, u_k$ を **正規直交系** といい，正規直交系である基底を **正規直交基底** という。正規直交系は１次独立である。

Schmidt の直交化　$u_1, \ldots, u_k$ が正規直交系で $x, u_1, \ldots, u_k$ は1次独立であるとする。

$$v = x - \sum_{i=1}^{k} u_i \langle u_i , x \rangle$$

とおくと $\langle v, u_i \rangle = 0 \quad 1 \leq i \leq k$ であるから，$u = \dfrac{v}{\|v\|}$ とおくと $u_1, \cdots, u_k, u$ も正規直交系となる。

ユークリッド空間においては，与えられた基底に対して Schmidt の直交化を繰り返し適用することにより，正規直交基底を作ることが出来る。

1. 5　直和

この節に限り，とくに断らない場合，かならずしもユークリッド空間ではない任意の実ベクトル空間を対象とし次元の有限・無限を問わない。

直積　$(E_\iota)_{\iota \in I}$ をベクトル空間の族とする。I は任意の集合（添字集合）である。$(E_\iota)_{\iota \in I}$ の積集合は，その要素 $(x_\iota)_{\iota \in I}, (y_\iota)_{\iota \in I}, \cdots$ に対して演算 $\alpha(x_\iota) = (\alpha x_\iota), \alpha \in \boldsymbol{R}$，$(x_\iota) + (y_\iota) = (x_\iota + y_\iota)$ を定義することによりベクトル空間（**直積**）となり，$\prod_{\iota \in I} E_\iota$ と表記される。要素 $x = (x_\iota) \in \prod E_\iota$ を x_ι へ写す線形写像を pr_ι と記すことにする $(pr_\iota(x) = x_\iota)$。次の命題は直積の基本性質である。

命題 1. 2　$(E_\iota)_{\iota \in I}$ をベクトル空間の族，F をベクトル空間，各 $\iota \in I$ に対して $f_\iota : F \to E_\iota$ を線形写像とすると，各 $\iota \in I$ について $pr_\iota \circ f = f_\iota$ となるような線形写像 $f : F \to E = \prod_{\iota \in I} E_\iota$ がただ一つ存在する。

$$\begin{tikzcd}
F \arrow[r, "f_\iota"] \arrow[dr, "f"'] & E_\iota \\
& \prod E_\iota \arrow[u, "\mathrm{pr}_\iota"']
\end{tikzcd}$$

直和　ベクトル空間族 $(E_\iota)_{\iota\in I}$ の直積 $E=\prod_{\iota\in I}E_\iota$ の要素 x で，有限個を除くすべての $\iota\in I$ に対して $pr_\iota(x)=0$ であるようなものからなる集合を族 (E_ι) の**外部直和**といい，$\bigoplus_{\iota\in I}E_\iota$ と表記する。外部直和は直積 $\prod_{\iota\in I}E_\iota$ の部分空間である。x_κ を $pr_\iota(x)=0\,(\iota\neq\kappa),=x_\kappa(\iota=\kappa)$ であるような $x\in\bigoplus E_\iota$ へ写す線形写像 $j_\kappa:E_\kappa\to\bigoplus E_\iota$ を **標準単射** といい，

$$x=\sum_{\iota\in I}j_\iota(pr_\iota\, x)\qquad x\in E=\bigoplus_{\iota\in I}E_\iota \tag{3}$$

が成り立つ。**命題 1．2** に対応して，外部直和について次の命題が成り立つ。

命題 1．3　$(E_\iota)_{\iota\in I}$ をベクトル空間の族，F をベクトル空間，各 $\iota\in I$ に対して $f_\iota:E_\iota\to F$ を線形写像とすると，各 $\iota\in I$ に対して

$$g\circ j_\iota=f_\iota$$

$$\begin{tikzcd} E_\iota \arrow[r, "f_\iota"] \arrow[d, "j_\iota"'] & F \\ \bigoplus E_\iota \arrow[ur, "g"'] & \end{tikzcd}$$

が成り立つような線形写像 $g:\bigoplus E_\iota\to F$ がただ一つ存在する。

記号　ここで定義された $g=\sum_{\iota\in I}f_\iota\circ\mathrm{pr}_\iota:\bigoplus_{\iota\in I}E_\iota\to F$ を単に $\sum_{\iota\in I}f_\iota$ と表すことにする。

部分空間の和と直和　$(E_\iota)_{\iota\in I}$ をベクトル空間 E の部分空間族とするとき，集合 $\bigcup_{\iota\in I}E_\iota$ から生成される E の部分空間（有限個の E_ι からの要素 x_ι の和 $\sum x_\iota$ 全体の集合で，これは E の部分空間となる）を族 $(E_\iota)_{\iota\in I}$ の **和** といい，$\sum_{\iota\in I}E_\iota$ と記す。とくに，有限個の部分空間族 $E_1,\ldots,E_n$ の場合には，その和を $E_1+\cdots+E_n$，$\sum_{i=1}^{n}E_i$ などと記す。

定義 1．2 （**部分空間族の直和**） $(E_\iota)_{\iota\in I}$ をベクトル空間 E の部分空間族とする。**命題 1．3** の f_ι として E_ι から E への標準単射 j_ι ($x_\iota \in E_\iota$ に対して $j_\iota(x_\iota) = x_\iota \in E$) をとって $\sum_{\iota\in I} j_\iota : \bigoplus_{\iota\in I} E_\iota \to E$ が同形となる場合，E は部分空間族 $(E_\iota)_{\iota\in I}$ の **直和** であるという。これは，E の要素 x が E_ι の要素の和 $x = \sum x_\iota$ として一意的に表せることである。

記法　ベクトル空間の部分空間族 $(E_\iota)_{\iota\in I}$ の和 $\sum_{\iota\in I} E_\iota$ が直和であるとき，外部直和との混同の恐れがなければ以降，これを $\bigoplus_{\iota\in I} E_\iota$ と記すことにする。また，有限個の部分空間族 $E_1,\ldots,E_n$ の場合には，$E_1 \oplus \cdots \oplus E_n$ とも記すことにする。

命題 1．4 $(E_\iota)_{\iota\in I}$ がベクトル空間 E の部分空間族であるとき，以下は同値である：

a)　部分空間 $\sum_{\iota\in I} E_\iota$ は，族 $(E_\iota)_{\iota\in I}$ の直和である。
b)　$\sum_{\iota\in I} x_\iota = 0\ (x_\iota \in E_\iota)$ ならば，各 $\iota \in I$ に対して $x_\iota = 0$ である。
c)　各 $\kappa \in I$ に対して，$E_\kappa \cap \sum_{\iota\neq\kappa} E_\iota = \{0\}$ である。

証明　a) $\Leftrightarrow$ b): a) $\Rightarrow$ b) は明らかである。$\sum x_\iota = \sum y_\iota$ とすると，$\sum (x_\iota - y_\iota) = 0$ であるから，仮定 b) より各 $\iota \in I$ に対して $x_\iota = y_\iota$。これは，a) が成り立っていることを示す。
b) $\Rightarrow$ c): $x \in E_\kappa \cap \sum_{\iota\neq\kappa} E_\iota$ とすると，$x - x = 0$ で，第1項の x を $x - x$ の E_κ 成分と見れば，b) より $x = 0$ でなければならない。したがって，$E_\kappa \cap \sum_{\iota\neq\kappa} E_\iota = \{0\}$。
c) $\Rightarrow$ b): $\sum x_\iota = 0\ (x_\iota \in E_\iota)$ とすると，$x_\kappa = -\sum_{\iota\neq\kappa} x_\iota \in E_\kappa \cap \sum_{\iota\neq\kappa} E_\iota$。したがって，$x_\kappa = 0$。□

命題 1．5 E を n 次元ユークリッド空間，F をその部分空間とすると，

$$E = F \oplus F^{\perp} \tag{4}$$

が成り立つ。

証明　F が E と一致したり，零空間である場合は明らかなので，そうでないと仮定する。F の基底を $x_1, \dots, x_\kappa (1 \leq k < n)$ とし，これらを拡張した $x_1, \dots, x_\kappa, x_{\kappa+1}, \dots, x_n$ を E の基底とする。この基底に順次 Schmidt の直交化を施して E の正規直交基底 $u_1, \dots, u_\kappa, u_{\kappa+1}, \dots, u_n$ を作る。このとき，$u_1, \dots, u_\kappa$ は F の正規直交基底となる。$u_{k+1}, \dots, u_n$ から生成される部分空間を G とすると，$G \subset F^{\perp}$ であるから $E = F + F^{\perp}$ であり，$x \in F \cap F^{\perp}$ とすると $\|x\|^2 = \langle x, x \rangle = 0$ であるから $x = 0$，すなわち $F \cap F^{\perp} = \{0\}$ である。したがって，(4) が成り立つ。□

2.　射影

記法　ここで，以降用いる二つの記法を定義しておく：$a_1, \cdots, a_\kappa$ をベクトルとして

$[a_1, \dots, a_\kappa]$ は $a_1, \dots, a_\kappa$ で生成される部分空間，
$(a_1, \dots, a_\kappa)$ は $a_1, \dots, a_\kappa$ を列ベクトルとする行列。

また，以降，1 次変換を表す行列　というとき，とくに断らない限り自然基底に関する表現である。

2. 1　射影の定義

定義 2. 1　ベクトル空間 E 上の 1 次変換 $f \in \boldsymbol{L}(E)$ が

$$f \circ f (= f^2) = f \tag{5}$$

を満たすとき E の $F = f(E)$ 上への **射影** という。

この射影の定義は，数学特有のその対象をきわめて抽象化したものであるが，それゆえにまた数学を展開していく上できわめて強力な効果を発揮するのである。はじめに，この定義によれば，E 上の零変換 0（すなわち，各 $x \in E$ に対して $0(x) = 0 \in E$ である変換）および，恒等変換 1（すなわち，各 $x \in E$ に対して $1(x) = x$ である変換）は，零空間 $\{0\}$ および E 上への射影であることに注意しておく。つぎに，$f \in \boldsymbol{L}(E)$ が射影ならば，$1-f$ も射影である。

(5) と固有値の定義[3] から，$f \in \boldsymbol{L}(E)$ が射影であるときは，その固有値は 1 または 0 である，という射影の顕著な性質が得られる。またこの事実から，f の階数はトレースに等しい，すなわち，

$$\operatorname{rank} f = \operatorname{Tr} f$$

であることもただちにわかる[4]。

2. 2　射影と直和

命題 2. 1　$f \in \boldsymbol{L}(E)$ が射影ならば，$E = f(E) + \operatorname{Ker} f$ で，これは直和となる。すなわち，

$$E = f(E) \oplus \operatorname{Ker} f \tag{6}$$

が成り立つ。

証明　$x \in E$ とすると，$x - f(x) \in \operatorname{Ker} f$ であるから $E = f(E) + \operatorname{Ker} f$ を得る。つぎに，$f(E) \cap \operatorname{Ker} f$ から任意の要素 $f(x)$ をとると，$f(x) =$

3)　1 次変換 f の固有値は，ゼロベクトルでない x について $f(x) = \lambda x$ を満たすスカラー λ。このとき，x を λ に属する固有ベクトルという。

4)　$\operatorname{rank} f$ は部分空間 $f(E)$ の階数（次元)。また，$\operatorname{Tr} f$ は f を表す行列の対角要素の和で，この値は f を表すすべての行列について等しい。

$f^2(x)=0$ であるから $f(E)\cap \mathrm{Ker}\, f=\{0\}$，すなわち $E=f(E)+\mathrm{Ker}\, f$ は直和である。□

E が二つの部分空間 F, G の直和であるとき，E の各要素 z は F と G の要素の和 $x+y$ として表すことができ，これら $x\in F, y\in G$ は z に対して一意的に定まる。したがって，E から F への写像 $z\mapsto x$ が定義できるが，この写像は E の F 上への射影である。このように，射影は部分空間が一つ与えられただけでは定まらず，部分空間とその補空間が与えられてはじめて決定される。

例　太陽による地上への射影は，地球から太陽へのベクトル（太陽の位置）次第で種々の異なったものになる。

上述の結果，つぎの定理が得られる。

定理 2. 1　ベクトル空間の射影全体からなる集合と，そのベクトル空間の二部分空間の直和の組全体からなる集合との間には1対1の対応が存在する。

2. 3　射影を表す行列の一般形

前節で述べたように，射影には二つの部分空間が関係する。この節では，E はこれまでと同じく n 次元ユークリッド空間とし，$x_1,x_2,\ldots,x_\kappa$，y_1，$y_2,\ldots,y_\kappa\in E$ $(0\le k\le n)$ とする。さらに，二つの部分空間 F, G，および二つの $n\times k$ 行列 X, Y を以下のように定めておく。

$$F=[x_1,x_2,\ldots,x_\kappa],\qquad G=[y_1,y_2,\ldots,y_\kappa],$$
$$X=(x_1,x_2,\ldots,x_\kappa),\qquad Y=(y_1,y_2,\ldots,y_\kappa).$$

命題 2. 2　$\mathrm{rank}\, X=\mathrm{rank}\, Y=k$ $(0<k\le n)$[5] で同時に $F\cap G^\perp=\{0\}$

であることは，正方行列 ${}^tYX$ が正則であることと同値である。

証明　はじめに，${}^tYX$ が正則であることを仮定する。もしも $\mathrm{rank}\ X < k$ であるとすると，$x_1,\ldots,x_k$ は1次従属であるから $X\xi = x_1\xi_1 + \cdots + x_\kappa\xi_\kappa = 0$ であるようなゼロでないベクトル $\xi = {}^t(\xi_1,\ldots,\xi_\kappa) \in R^k$ が存在するが，この ξ に対して ${}^tYX\xi = 0$ である。これは ${}^tYX$ が正則であることに矛盾するから $\mathrm{rank}\ X = k$ でなければならない。同様に，$\mathrm{rank}\ Y < k$ を仮定すると，${}^tYX$ が正則でないという矛盾が生じ[6]，$\mathrm{rank}\ Y = k$ でなければならないことが解る。$\mathrm{rank}\ X = k$ であることが解ったので，F の任意の要素は $X\xi = x_1\xi_1 + \cdots + x_k\xi_k$ のような形に書ける。そこで，$X\xi \in F \cap G^\perp$ とすると，${}^tYX\xi = 0$ であるから $\xi = 0$，したがって，$X\xi = 0$，すなわち $F \cap G^\perp = \{0\}$ である。

つぎに，逆を証明するには，${}^tYX\xi = 0$ のとき $\xi = 0$ であることをいえばよい。実際，${}^tYX\xi = 0$ より $X\xi \in G^\perp$ であるから $X\xi \in F \cap G^\perp$，したがって，$X\xi = 0$ でなければならず，さらに，$\mathrm{rank}\ X = k$ であるから $\xi = 0$ である。□

系　行列 ${}^tXX$ が正則であることと，$\mathrm{rank}\ X = k \quad (0 < k \le n)$ であることとは同値である。

定理 2. 2　${}^tYX$ が正則ならば，

$$X({}^tYX)^{-1}\,{}^tY \tag{7}$$

---

5)　ここでは，行列の階数 (rank) を列ベクトルのうち1次独立なものの最大個数と定義する。これは，行ベクトルのうち1次独立なものの最大個数にも等しい。

6)　一般に，正方行列 A については，(1) A が正則であることと $\det A \neq 0$ であることとは同値である。また (2) $\det A = \det {}^tA$ の二つの事実を用いている。

は E の F 上への射影を表す行列である。また，この射影の核は $G^{\perp}$ である。

証明　$\{X({}^tYX)^{-1}\,{}^tY\}\{X({}^tYX)^{-1}\,{}^tY\} = X({}^tYX)^{-1}\,{}^tY$ であるから，(7) は射影を表す行列である。つぎに，この定理の仮定に**命題 2. 2** を適用して $E = F \oplus G^{\perp}$ である。そして，行列 (7) が表す射影を f とすると，各 $i(1 \leq i \leq k)$ について $X({}^tYX)^{-1}\,{}^tYx_i = Xe_i = x_i$ より $F \subset f(E)$，また $G^{\perp}$ の任意の要素 y に対して ${}^tYy = 0$ より $G^{\perp} \subset \mathrm{Ker}\, f$ であるから，(6) によって $F(E) = F$, $\mathrm{Ker}\, f = G^{\perp}$ でなければならない。□

別証　$y_{k+1}, \ldots, y_n$ を部分空間 $G^{\perp}$ の基底とする。このとき，ベクトル $y \in \boldsymbol{R}^n$ を E の基底 $x_1, \ldots, x_k, y_{k+1}, \ldots, y_n$ を用いて

$$y = x_1\xi_1 + \cdots + x_k\xi_k + y_{k+1}\xi_{k+1} + \cdots + y_n\xi_n \qquad (\xi_1, \ldots, \xi_n \in \boldsymbol{R}) \quad (8)$$

と表すと，その核が $G^{\perp}$ である F 上への射影は y を $x_1\xi_1 + \cdots + x_k\xi_k = X\xi'$　$(\xi' = {}^t(\xi_1, \ldots, \xi_k) \in R^k)$ へ写す写像であるから，

$$X\xi' = X({}^tYX)^{-1}\,{}^tYy \quad (9)$$

が成り立つことをいえばよい。実際，各 $i(k+1 \leq i \leq n)$ に対して ${}^tYy_i = 0$ であることに注意して，(8) の両辺に左側から ${}^tY$ を乗じると ${}^tYy_i = {}^tYX\xi'$ を得るから，$\xi' = ({}^tYX)^{-1}\,{}^tYy$ であり，したがって，(9) が成り立つ。□

2. 4　射影とその特異値分解

任意の $m \times n$ 行列 A は，適当な m 次元直交行列 U と n 次元直交行列 V によって，

$$A = \begin{cases} U\begin{pmatrix} D & 0 \end{pmatrix} {}^tV & (m < n \text{ のとき}), \\ U\begin{pmatrix} D \\ 0 \end{pmatrix} {}^tV & (m \geq n \text{ のとき}) \end{cases} \qquad \begin{aligned} D &= \mathrm{diag}(\sigma_1, \ldots, \sigma_L), \\ &\sigma_1 \geq \cdots \geq \sigma_L \geq 0 \end{aligned} \tag{10}$$

と分解できる[7)]。ここで，$L = \min\{m, n\}$ である。ここに現れる $\sigma_1, \ldots, \sigma_L$ を行列 A の **特異値** といい，上の等式の右辺を行列 A の特異値分解という。以下の証明からわかるように，行列 A の特異値は行列 ${}^tAA$ の固有値から定まるので A に対し一意的に定まる。

特異値分解の略証　$m < n$ の場合について考える。半正値対称行列 ${}^tAA$ の固有値を $\sigma_1^2 \geq \cdots \geq \sigma_n^2 \geq 0$[8)]，$v_i\,(1 \leq i \leq n)$ を固有値 σ_i^2 に属するたがいに直交する単位ベクトル[9)]，さらに，$V = (v_1, \ldots, v_n)$ を n 次直交行列とすると，

$${}^tAA = V\mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2)\,{}^tV$$

が成り立つ。行列 ${}^tAA$ の階数は高々 m であるから，$\sigma_{m+1} = \cdots = \sigma_n = 0$ であり，さらに，$\sigma_1 \geq \cdots \geq \sigma_k > \sigma_{k+1} = \cdots = \sigma_m = 0$ であるとする。

$$u_i = \sigma_i^{-1} A\, v_i, \quad 1 \leq i \leq k \tag{11}$$

とおくと，$1 \leq i, j \leq k$ ならば $\langle u_i, u_j \rangle = \langle \sigma_i^{-1} v_i, \sigma_j^{-1}\,{}^tAA\, v_j \rangle = \langle \sigma_i^{-1} v_i, \sigma_j v_j \rangle = \delta_{ij}$，すなわち $u_1, \ldots, u_k$ は正規直交系であるが，これらに適当な

---

7) $\mathrm{diag}(\sigma_1, \ldots, \sigma_L)$ は，対角要素が $(\sigma_1, \ldots, \sigma_L)$ で他の要素がすべてゼロである L 次対角行列を表す。

8) 対称行列の固有値は実数で，相異なる固有値に属する固有ベクトルは直交する。また，対称行列 B が半正値であるとは，任意のベクトル x に対して $\langle Bx, x \rangle \geq 0$ であることで，x としてとくに固有値 λ に属する固有ベクトルをとると，$\lambda\|x\|^2 \geq 0$，したがって $\lambda \geq 0$ である。

9) 同じ固有値に属する複数の固有ベクトルについては **Schmidt の直交化**を適用する

たがいに直交する単位ベクトル $u_{k+1}, \ldots, u_m$ を付け加え，m 次正方行列 $U = (u_1, \ldots, u_m)$ が直交行列となるようにする。$k+1 \leq i \leq n$ である i に対しては $0 = \langle v_i, {}^t\!AA\, v_i \rangle = \langle A\, v_i, A\, v_1 \rangle$ より $A\, v_i = 0$ であるから，(11) とあわせて

$$A\, v_i = U\begin{pmatrix} D & 0 \end{pmatrix} {}^t V v_i, \quad 1 \leq i \leq n$$

が成り立つ。したがって (10) が得られる。$m \geq n$ の場合も同様に証明できる。□

$m < n$ のとき，

$$\sigma_1 \geq \cdots \geq \sigma_k > \sigma_{k+1} = \cdots = \sigma_n = 0$$

を仮定し，あらためて $U = (u_1, \ldots, u_k)$ を $m \times k$ 行列，$V = (v_1, \ldots, v_\kappa)$ を $n \times k$ 行列とし，$D = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$ とおくと，$A\, v_i = UD\, {}^t V v_i$ $(1 \leq i \leq n)$ であるから

$$A = UD\, {}^t V \tag{12}$$

が成り立つ。

行列 A の特異値分解が (12) で与えられた場合，A で表現される線形写像を f とすると，

$$f(\boldsymbol{R}^n) = [u_1, \ldots, u_k], \qquad \mathrm{Ker}(f) = [v_1, \ldots, v_k]^\perp \tag{13}$$

である。以下では，とくに断らないかぎり (12) を行列 A の特異値分解ということにする。

命題 2．3 f をベクトル空間 E 上の零写像でない射影，f の表現行列 A の特異値分解が (12) で与えられるとすると，

$$ {}^tVU = {}^tUV = D^{-1} \tag{14} $$

が成り立つ。

証明 f が射影であるから $AA = A$，したがって $UD\,{}^tVUD\,{}^tV = UD\,{}^tV$ である。この両辺に，左から $D^{-1}\,{}^tU$ を，右から VD^{-1} を掛けて ${}^tVU = D^{-1}$ であり，これは対角行列で対称であるから (14) を得る。□

2. 5 線形写像のノルム

E, F をユークリッド空間とするとき，E から F への線形写像 $f \in \boldsymbol{L}(E;F)$ のノルム $\|f\|$ を

$$ \|f\| = \max\{\|f(x)\| : x \in E, \|x\| = 1\} \tag{15} $$

と定義する[10)]。このノルムは，ノルムの公理 (i) $\|f\| = 0$ と $f = 0$ とは同値，(ii) $\|\alpha f\| = |\alpha|\|f\|$，(ii) $\|f + g\| \leq \|f\| + \|g\|$，$(\alpha \in \boldsymbol{R}, f, g \in \boldsymbol{L}(E;F))$ を満たしている。

$m \times n$ 行列 $A \in \boldsymbol{M}(m,n)$ に対するノルムも同様に定義する：

$$ \|A\| = \max\{\|Ax\| : x \in \boldsymbol{R}^n, \|x\| = 1\} \tag{16} $$

E, F にそれぞれ正規直交基底を定め，これらの基底に関する線形写像 $f \in \boldsymbol{L}(E;F)$ の表現行列が A であるとすると $\|f\| = \|A\|$ である。これらのノルムは以下の性質を持っている：

(1) $\|f(x)\| \leq \|f\|\,\|x\|$ $(x \in E)$

(2) g を F から第三のユークリッド空間 G への線形写像とすると，

10) $\|x\|$ は E のノルムを，$\|f(x)\|$ は F のノルムを用いており，一般にはこれら二つのノルムは別物であるが，ここでは混乱の恐れはないので同じ記号を用いている。

$\|g \circ f\| \leq \|g\| \|f\|$

(3) $\|{}^t f\| = \|f\|$,

(4) $\|{}^t f f\| = \|f {}^t f\| = \|f\|^2$,

(5) 正規直交ベクトル系を列ベクトルとする行列 P に対して,

$\|PAx\| = \|Ax\|$ $(x \in \boldsymbol{R}^n)$, $\|A\,{}^t P\| = \|A\|$。

命題 2.4 f が零写像でない射影ならば,

$$\|f\| \geq 1 \tag{17}$$

証明 $\|f\| = \|f^2\| \leq \|f\| \|f\|$ の両辺を $\|f\|$ で割ればよい。□

2.6 正射影

この節でも, E を n 次元ユークリッド空間, F, G を部分空間とする。

定義 2.2 E の F 上への射影 f が

$${}^t f = f \tag{18}$$

であるとき, **正射影** という。

定義 2.3 f がユークリッド空間 E の零写像でない射影であるとき, 以下は同値である。

a) f は正射影である。

b) $f(E) \perp \mathrm{Ker}(f)$。

c) $\|f\| = 1$。

証明 a) $\Rightarrow$ b): $f(x) \in f(E), y \in \mathrm{Ker}(f)$ を任意にとると, $\langle f(x), y \rangle = \langle x, f(y) \rangle = \langle x, 0 \rangle = 0$ より b) が成り立つ。

b) $\Rightarrow$ c): 任意の単位ベクトル $x \in E$ に対して，$x = f(x) + (1-f)(x)$ より $1 = \|x\|^2 = \|f(x)\|^2 + 2\langle f(x), (1-f)(x)\rangle + \|(1-f)(x)\|^2$ であり，b) より $\langle f(x), (1-f)(x)\rangle = 0$ であるから $\|f(x)\| \leq 1$ である。f が射影であることから $\|f\| \geq 1$ なので c) が成り立つ。

c) $\Rightarrow$ a):

$$UD\,{}^tV (D = \mathrm{diag}(\sigma_1, \ldots \sigma_k),\ \ \sigma_1 \geq \cdots \geq \sigma_k > 0 \tag{19}$$

を f の表現行列の特異値分解とすると，(14) より ${}^tVU = {}^tUV = D^{-1}$ で，${}^tVU$ の各要素は単位ベクトルどうしの内積であるから，$1 \geq \sigma_k^{-1} \geq \cdots \geq \sigma_1^{-1}$，すなわち $1 \leq \sigma_k \leq \cdots \leq \sigma_1$ でなければならない。一方，$\|f\| = \|UD\,{}^tV\| = \|D\| = \sigma_1 = 1$ であるから $\sigma_1 = \cdots = \sigma_k = 1$ でなければならず，

$${}^tVU = {}^tUV = D^{-1} = D = I_k \tag{20}$$

を得る。つぎに，V は，${}^tff$ の固有値 $\sigma_1^2, \ldots, \sigma_k^2$ に属してたがいに直交する単位固有ベクトルを列ベクトルとする行列であるが，このうち固有値 0 に属して正規直交系をなす固有ベクトルを列ベクトルとする行列を W と記すことにすると，$(V\,W)$ は直交行列であるから

$$(V\,W)\begin{pmatrix} {}^tV \\ {}^tW \end{pmatrix} = V\,{}^tV + W\,{}^tW = I_n$$

2 番目の等号で結ばれた両辺に左から ${}^tU$ を，右から U を乗じると，

$$\begin{aligned} {}^tUV\,{}^tVU + {}^tUW\,{}^tWU &= D^{-2} + {}^tUW\,{}^tWU \\ &= I_k + {}^tUW\,{}^tWU = {}^tUI_n\,U = I_k \end{aligned}$$

が得られ，これから ${}^tUW\,{}^tWU = 0$ であることがわかる。$\|{}^tUW\,{}^tWU\| = \|{}^tUW\,{}^t({}^tUW)\| = \|{}^tUW\|^2 = 0$ より ${}^tUW = 0$ を得るが，これを用いて

$$U = (V\,W)\begin{pmatrix} {}^tV \\ \\ {}^tW \end{pmatrix} U = (V\,W)\begin{pmatrix} I_k \\ \\ 0 \end{pmatrix} = V$$

が成り立つことがわかる。この結果，f の表現行列 (19) について ${}^t(UD\,{}^tV) = {}^t(UD\,{}^tU) = UD\,{}^tU$ が成り立つから ${}^tf = f$ を得る。□

系　ユークリッド空間の正射影全体からなる集合と，そのユークリッド空間の部分空間全体からなる集合との間には 1 対 1 の対応が存在する。

証　定理の b) から，ユークリッド空間 E の正射影 f は，E の直和 $E = f(E) \oplus f(E)^{\perp}$ を与えるが，この直和は部分空間 $f(E)$ だけで決定されるからである。□

系　X を列ベクトルが 1 次独立である $n \times k$ 行列 $(1 \leq k \leq n)$ とすると，$X({}^tXX)^{-1}\,{}^tX$ は正射影を表す行列で，その正射影を f とすると $f(E)$ は X の列ベクトルが生成する部分空間に等しい。

2. 7　決定係数

以下は　**はじめに**　の続きである。$X({}^tXX)^{-1}\,{}^tX$ で表される正射影を f と記すことにする。$fy \perp (1-f)y$ であるから

$$\|y\|^2 = \|fy\|^2 + \|(1-f)y\|^2 \tag{21}$$

が成り立ち，右辺の第 2 項 $\|(1-f)y\|^2$ は誤差の平方和である。この式から明らに，誤差の割合 $\dfrac{\|(1-f)y\|^2}{\|y\|^2}$ は $\dfrac{\|fy\|^2}{\|y\|^2}$ の値が大きいほど小さい。

すなわち，後者の値は，モデルの精度の目安となるもので **決定係数** とよばれ，通常 R^2 と記される。図 1 からも明らかなように，R^2 は二つのベクトル y, yf がなす角の余弦 (cosine) の平方であり，したがって，内積を

用いて $\left\langle \frac{y}{\|y\|}, \frac{fy}{\|fy\|} \right\rangle^2$ とも書ける：

$$R^2 = \frac{\|fy\|^2}{\|y\|^2} = \left\langle \frac{y}{\|y\|}, \frac{fy}{\|fy\|} \right\rangle^2 \tag{22}$$

この式から，決定係数は，また，資料 y と，それをベクトルと見なして平面 π へ正射影して得られる資料 fy との相関係数の平方であることもわかる。

3. 多重正規分布

基本的定積分

次の定積分が基本的である。

$$\int_{-\infty}^{+\infty} \exp(-x^2)dx = \sqrt{\pi} \tag{23}$$

ガンマ関数

$$\Gamma(p) = \int_0^{+\infty} t^{p-1}e^{-t}dt,\ p > 0 \tag{24}$$

の記法を用いると，変数変換 $t = x^2$ により，

$$\int_{-\infty}^{+\infty} \exp(-x^2)dx = \Gamma(\frac{1}{2}) \tag{25}$$

が得られ，ガンマ関数の再帰公式 $\Gamma(p+1) = p\Gamma(p)$ により一般に

$$\int_{-\infty}^{+\infty} x^{2n}\exp(-x^2)dx = \Gamma(n+\frac{1}{2})$$

$$= \left(n-\frac{1}{2}\right)\left(n-\frac{3}{2}\right)\cdots\frac{1}{2}\Gamma(\frac{1}{2}),\ n = 1, 2, \cdots \tag{26}$$

が得られる。

3. 1　確率ベクトル

$\mathbf{X}_1, \cdots, \mathbf{X}_n$ を確率変数とするとき，これらを座標とする確率ベクトルを $\mathbf{X} = {}^t(\mathbf{X}_1, \cdots, \mathbf{X}_n)$ と書くことにする。各 $\mathbf{X}_i\,(1 \le i \le n)$ が期待値 $\mathbf{E}(\mathbf{X}_i)$ をもつとき，${}^t(\mathbf{E}(\mathbf{X}_1), \cdots, \mathbf{E}(\mathbf{X}_n))$ を $\mathbf{E}(\mathbf{X})$ と書く。$\mathbf{M}$ を (j, k) 要素が確率変数 $\mathbf{M}_{jk}$ であるような行列とし，各 $\mathbf{M}_{jk}$ が期待値をもつとき $\mathbf{E}(\mathbf{M})$ を (j, k) 要素が $\mathbf{E}(\mathbf{M}_{jk})$ である行列とする。

定義 3. 1　$\mathbf{E}(\mathbf{X}) = 0$ であるとき，$\mathbf{X}$ の共分散行列を

$$\mathrm{Var}(\mathbf{X}) = \mathbf{E}(\mathbf{X}\,{}^t\mathbf{X}) \tag{27}$$

と定義する。$\mathbf{E}(\mathbf{X}\,{}^t\mathbf{X})$ は (j, k) 要素が $\mathbf{E}(\mathbf{X}_j\,\mathbf{X}_k)$ である $n \times n$ 行列である。$\mathbf{E}$ は線形作用素であるから，$m \times n$ 行列 A に対して

$$\mathrm{Var}(A\,\mathbf{X}) = A\,\mathrm{Var}(\mathbf{X})\,{}^tA \tag{28}$$

が成り立つ。

3. 2　正規確率ベクトル

定義 3. 2　Q を n 次正値対称行列とするとき，

$$\varphi(x) = \gamma^{-1} \cdot \exp\left(-\frac{1}{2}\,{}^t x Q\, x\right) \quad x \in \boldsymbol{R}^n,\ \gamma \in \boldsymbol{R} \tag{29}$$

のような形の密度関数をもつ n 次確率ベクトルは **原点を中心とする正規分布** をするという。

定理 3. 1　密度関数に (29) をもつ原点を中心とする n 次元正規確率ベクトル $\mathbf{X} = {}^t(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ は，以下の性質を持つ：

(1)　適当な直交行列 C を用いて $\mathbf{Z} = C\mathbf{X} = {}^t(\mathbf{Z}_1, \ldots, \mathbf{Z}_n)$ と変換することにより各要素 $\mathbf{Z}_i$ が互いに独立である正規確率ベクトルとするこ

とができる，

(2)　$Q^{-1} = \mathrm{Var}(\mathbf{X})$,

(3)　$\gamma = \sqrt{(2\pi)^n \det \mathrm{Var}(\mathbf{X})}.$

証明　$\mathbf{X}$ の密度関数が (29) であるとすると，Q は正値対称行列であるから，行列式が 1 である適当な直交行列 C を用いて $CQ\,{}^tC$ を対角要素が $\sigma_1^{-2}, \ldots, \sigma_n^{-2}$ $(\sigma_i > 0)$ であるような対角行列とすることができる。$\mathbf{Z} = C\mathbf{X}$ とおくと，$\mathbf{Z}$ は，その密度関数が

$$\begin{aligned} \varphi({}^tCz)\det{}^tC &= \varphi({}^tCz) = \gamma^{-1}\exp(-\tfrac{1}{2}\,{}^tzDz) \\ &= \gamma^{-1}\exp\left\{-\frac{1}{2}\left(\frac{z_1}{\sigma_1}\right)^2\right\}\cdots\exp\left\{-\frac{1}{2}\left(\frac{z_n}{\sigma_n}\right)^2\right\} \end{aligned} \tag{30}$$

ただし，$D = \mathrm{diag}(\sigma_1^{-2}, \ldots, \sigma_n^{-2})$ であり，定積分が 1 であることから

$$\gamma = \sqrt{(2\pi)^n}\,\sigma_1 \cdots \sigma_n \tag{31}$$

でなければならない。これは，$Z_1, \ldots, Z_n$ がたがいに独立な平均 0，分散がそれぞれ $\sigma_1^2, \ldots, \sigma_n^2$ である正規確率変数であることを示しており，(1) が証明された。つぎに，$D^{-1} = \mathrm{Var}(\mathbf{Z}) = C\,\mathrm{Var}(\mathbf{X})\,{}^tC$ であるから，$\mathrm{Var}(\mathbf{X}) = ({}^tCDC)^{-1} = Q^{-1}$ が得られる (2)。さらに，$\sigma_1^2 \cdots \sigma_n^2 = \det D^{-1} = \det(C\,\mathrm{Var}(\mathbf{X})\,{}^tC) = \det \mathrm{Var}(\mathbf{X})$ であるから，(31) と併せて (3) が得られる。□

系　$\mathbf{X}_1, \mathbf{X}_2$ を 1 次元確率変数とするとき，${}^t(\mathbf{X}_1, \mathbf{X}_2)$ が原点を中心とする正規分布をするならば，$\mathbf{E}(\mathbf{X}_1\mathbf{X}_2) = 0$ は $\mathbf{X}_1$ と $\mathbf{X}_2$ が独立であることの必要十分条件である。

3. 3　条件付多重正規分布

$\mathbf{Z} = {}^t(\mathbf{X}_1, \ldots, \mathbf{X}_k, \mathbf{Y}_{k+1}, \ldots, \mathbf{Y}_n)$ は

$$\varphi(x,y) = \gamma^{-1}\exp\left\{-\frac{1}{2}({}^tx, {}^ty)Q\begin{pmatrix}x\\ y\end{pmatrix}\right\} \qquad x \in \boldsymbol{R}^k, \tag{32}$$

$$y \in \boldsymbol{R}^{n-k} \quad \gamma^2 = (2\pi)^n \det \mathrm{Var}(\mathbf{Z})$$

を密度関数とする正規確率ベクトルとし，$\mathbf{X} = {}^t(\mathbf{X}_1, \ldots, \mathbf{X}_k)$, $\mathbf{Y} = {}^t(\mathbf{Y}_{k+1}, \ldots, \mathbf{Y}_n)$ とおく。

$\mathbf{Y}$ の周辺密度関数を $\psi(y)$ とすると，$\mathbf{X}$ の $\mathbf{Y} = y$ のもとでの条件付密度関数は

$$\varphi(x|y) = \frac{\varphi(x,y)}{\psi(y)}$$

である。はじめに，これらを求める準備をしておく。

B を $k \times (n-k)$ 行列として，

$$\mathbf{W} = \begin{pmatrix}\mathbf{U}\\ \mathbf{V}\end{pmatrix} = \begin{pmatrix}I_k & B\\ 0 & I_{n-k}\end{pmatrix}\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix} = \begin{pmatrix}\mathbf{X} - B\mathbf{Y}\\ \mathbf{Y}\end{pmatrix} \tag{33}$$

とおくと，$\mathbf{E}(\mathbf{U}\,{}^t\mathbf{V}) = \mathbf{E}(\mathbf{X}\,{}^t\mathbf{Y}) - B\,\mathbf{E}(\mathbf{Y}\,{}^t\mathbf{Y}) = \mathbf{E}(\mathbf{X}\,{}^t\mathbf{Y}) - B\ \mathrm{Var}(\mathbf{Y})$ であるから，$\mathbf{E}(\mathbf{U}\,{}^t\mathbf{V}) = 0$ となるために，すなわち $\mathbf{U}$ と $\mathbf{V}$ が独立であるために B が満たすべき条件は，

$$B = \mathbf{E}(\mathbf{X}\,{}^t\mathbf{Y})\mathrm{Var}(\mathbf{Y})^{-1} \tag{34}$$

である。以下では (34) が成り立っているものとする。

定理 3. 2　$\mathbf{Y}$ の周辺密度関数は

$$\psi(y) = \{(2\pi)^{n-k}\det\mathbf{E}(\mathbf{Y}\,{}^t\mathbf{Y})\}^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}{}^ty\mathbf{E}(\mathbf{Y}\,{}^t\mathbf{Y})^{-1}y\right\} \tag{35}$$

である。

証明　$\mathbf{W}$ の密度関数 $\eta(u,v)$ は

$$\{(2\pi)^n \det \mathbf{E}(\mathbf{W}\,{}^t\mathbf{W})\}^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}({}^tu\mathbf{E}(\mathbf{U}\,{}^t\mathbf{U})^{-1}u + {}^tv\mathbf{E}(\mathbf{V}\,{}^t\mathbf{V})^{-1}v\right\} \quad (36)$$

で，$\det \mathbf{E}(\mathbf{W}\,{}^t\mathbf{W}) = \det \mathbf{E}(\mathbf{U}\,{}^t\mathbf{U}) \cdot \det \mathbf{E}(\mathbf{V}\,{}^t\mathbf{V})$ であるから，

$$\int_{-\infty}^{+\infty} \eta(u,v)dx = \{(2\pi)^{n-k} \det \mathbf{E}(\mathbf{V}\,{}^t\mathbf{V})\}^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\,{}^tv\mathbf{E}(\mathbf{V}\,{}^t\mathbf{V})^{-1}v\right\}$$

を得るが，$v = y$, $\mathbf{V} = \mathbf{Y}$ であるから右辺は (35) に等しい。□

定理　3．3　条件 $\mathbf{Y} = y$ のもとでの $\mathbf{X}$ の条件付確率の密度関数は

$$\begin{aligned}&\varphi(x|y)\\ &\quad = \{(2\pi)^k \det \mathbf{E}(\mathbf{X}\,{}^t\mathbf{X})\}^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\,{}^t(x - By)\mathbf{E}(\mathbf{X}\,{}^t\mathbf{X})^{-1}(x - By)\right\} \quad (37)\end{aligned}$$

であり，条件 $\mathbf{Y} = y$ のもとでの $\mathbf{X}$ の条件付期待値は

$$\mathbf{E}(\mathbf{X}|\mathbf{Y} = y) = \mathbf{E}(\mathbf{X}\,{}^t\mathbf{Y})\mathrm{Var}(\mathbf{Y})^{-1}y \quad (38)$$

である。

証明

$$\varphi(x|y) = \frac{\eta(u,v)}{\psi(y)}$$

を計算すると式 (37) となる。つぎに，

$$\mathbf{E}(\mathbf{X}|\mathbf{Y} = y) = \int_{-\infty}^{+\infty} x\varphi(x|y)dx$$

$$= \int_{-\infty}^{+\infty} (x - By)\varphi(x|y)\, d(x - By) + By \int_{-\infty}^{+\infty} \varphi(x|y)\, dx$$

である。もともと $\mathbf{E}(\mathbf{Z}) = 0$ であったから (33) よりと $\mathbf{E}(\mathbf{W}) = 0$ であり，したがって，$\mathbf{E}(\mathbf{U}) = 0$ である。したがって，第 1 項の積分はゼロであり，また，第 2 項は明らかに By となるから (38) が得られる。□

3. 4　χ^2 分布の自由度

$\mathbf{X}_1, \ldots, \mathbf{X}_n$ がたがいに独立な平均 0，分散 1 の正規確率変数とするとき，確率変数

$$\mathbf{X}_1^2 + \cdots + \mathbf{X}_n^2$$

は**自由度** n の χ^2 **分布**をなすという。

標本分散　上と同じく，$\mathbf{X}_1, \ldots, \mathbf{X}_n$ がたがいに独立な平均 0，分散 1 の正規確率変数であるとき，

$$\overline{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i, \qquad \frac{1}{n}\sum_{i=1}^{n} (\mathbf{X}_i - \overline{\mathbf{X}})^2$$

は，それぞれ，**標本平均**，**標本分散**とよばれる。

命題 3. 1　$\mathbf{X}_1, \ldots, \mathbf{X}_n$ をたがいに独立な平均 0，分散 1 の正規確率変数とすると，確率変数

$$\mathbf{Y} = \sum_{i=1}^{n} (\mathbf{X}_i - \overline{\mathbf{X}})^2$$

は自由度 $n - 1$ の χ^2 分布をなす。

証明　$\mathbf{X} = {}^t(\mathbf{X}_1, \ldots, \mathbf{X}_n)$ とおいて $\mathbf{Y}$ を書き直すと，

$$\mathbf{Y} = \sum_{i=1}^{n} \mathbf{X}_1^2 - n\overline{\mathbf{X}}^2 = \sum_{i=1}^{n} \mathbf{X}_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} \mathbf{X}_i\right)^2$$

$$= {}^t\mathbf{X} I_n \mathbf{X} - \frac{1}{n}\, {}^t\mathbf{X} \mathbf{1}_n \mathbf{X} = {}^t\mathbf{X}\left(I_n - \frac{1}{n}\mathbf{1}_n\right)\mathbf{X},$$

ここで，I_n は n 次単位行列，$\mathbf{1}_n$ はすべての要素が 1 である n 次正方行列である。行列 $A = I_n - \frac{1}{n}\mathbf{1}_n$ は $AA = A$ を満たすから，その固有値は 1 または 0 であって，その階数は

$$\mathrm{Tr}\, A = n\left(1 - \frac{1}{n}\right) = n - 1$$

に等しい。さらに，対称行列であるから適当な n 次直交行列 C により

$$A = {}^tC \operatorname{diag}(1, \ldots, 1, 0) C$$

とすることが出来る。したがって，$D = \operatorname{diag}(1, \ldots, 1, 0)$, $\mathbf{Z} = {}^t(\mathbf{Z}_1, \ldots, \mathbf{Z}_n) = C\mathbf{X}$ とおくと，

$$\mathbf{Y} = {}^t\mathbf{X} A \mathbf{X} = {}^t\mathbf{X}\, {}^tCDC\mathbf{X} = {}^t\mathbf{Z} D \mathbf{Z} = \mathbf{Z}_1^2 + \cdots + \mathbf{Z}_{n-1}^2$$

となる。これらの $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ については，

$$\mathbf{E}(\mathbf{Z}) = C\mathbf{E}(\mathbf{X}) = 0 \in \boldsymbol{R}^n,$$

$$\mathrm{Var}(\mathbf{Z}) = C \,\mathrm{Var}(\mathbf{X})\, {}^tC = C I_n\, {}^tC = C\, {}^tC = I_n$$

が成り立つから，たがいに独立な平均 0，分散 1 である正規確率変数であることが分かる。以上から，確率変数 $\mathbf{Y}$ は自由度 $n-1$ の χ^2 分布をなすことが証明された。□

参　考　文　献

[1] 弥永昌吉，布川正巳編「代数学（現代数学演習叢書 1)」(岩波書店，1972 年)

[2] William Feller. *An Introduction to Probability Theory and Its Applications,* Volume II. John Wiley & Sons, 1971.

[3] N. Bourbaki. *Algèbre* Chapitre 2, *Algèbre Linéaire,* Hermann, 1962.

[4] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis,* John Wiley & Sons, 1984.