論　文

ビジネスデータ分析の理論的基礎（Ⅱ）
―遡及Cholesky分解と多変量分布―

Theoretical Foundation of Business Data Analysis（Ⅱ）
－Retrospective Cholesky Decomposition and Multivariate Distribution－

松　縄　　規
MATSUNAWA Tadashi

概要： Part （I） に引き続き、ビジネスデータ分析等の分野の多変量データ解析に登場するいくつかの代表的な多変量分布の、遡及 （そきゅう）コレスキー分解を利用した誘導を行う。すなわち，正値対称行列**A**に対して、対角線要素が全て正の実下三角行列**T**を構成し平方根分解**A=T'T**を利用する。本稿では、この分解の有効性を更に知るために、Wishart分布と密接に関連するHotellingの T^2統計量の分布及びWilksの $\boldsymbol{\Lambda}$ 統計量の分布の誘導を行う。本稿における章および数式等の番号については、前稿との内容の連続性を考慮して、通し番号を用いることにする。

キーワード： 多変量解析、遡及コレスキー分解、Hotellingの T^2統計量の分布、Wilks の $\boldsymbol{\Lambda}$ 統計量の分布。

5. Wishart 分布関連の多変量分布の誘導

第3章におけるWishart分布の誘導過程および、よく知られた一変量でのカイ二乗分布、t 分布、F 分布、ベータ分布の相互関係を考慮すると，Hotellingの T^2統計量やWilksの $\boldsymbol{\Lambda}$ 統計量の密度関数の導出も，比較的容易に行える。以下にこのことを示す。

5.1. Hotelling の統計量の分布

一変量の統計学で、正規母集団の平均に関する推測において t 統計量がよく用いられる。 以下のように、この統計量の二乗を多変量の場合に拡張したものがHotellingの T^2 統計量と呼ばれるものである：

$\mathbf{z}' = \left(z_1, \cdots, z_p\right)$ を p -次元標準正規 $N_p\left(\mathbf{0}, \mathbf{I}_p\right)$ に従う確率ベクトル、$\mathbf{A}\left(p \times p\right)$ をWishart分布 $W_p\left(n, \mathbf{I}_p\right)$ に従うランダム行列とし、両者は統計的に独立に分布するものとする。この時、Hotellingの統計量 $T^2 := n \cdot \mathbf{z}'\mathbf{A}^{-1}\mathbf{z}$ に対し、次の規格化した統計量は、

$$\frac{T^2/p}{n/\left(n-p+1\right)} \sim F^{p}_{n-p+1} \tag{5.1.1}$$

まつなわ　ただし（経営情報学科）

に従って分布する。ただし F^{p}_{n-p+1}、は自由度 $\left(p, n-p+1\right)$ の F 分布に従う確率変数である。

誘導. $\mathbf{z}'/\|\mathbf{z}\|$ を p-次元実空間 R^p の第 p 座標軸の単位ベクトルとする。すなわち、

$$\mathbf{z}'/\|\mathbf{z}\| = \left(z_1/\|\mathbf{z}\|, z_2/\|\mathbf{z}\|, \cdots, z_{p-1}/\|\mathbf{z}\|, z_p/\|\mathbf{z}\|\right) \tag{5.1.2}$$

とする。さらに、$\left\{\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_{p-1}, \mathbf{z}'/\|\mathbf{z}\|\right\}$ を R^p の任意の正規直行基底とし、

$$\boldsymbol{\Gamma}_{p\times p} := \begin{pmatrix} \mathbf{z}'/\|\mathbf{z}\| \\ \boldsymbol{\gamma}'_{p-1} \\ \vdots \\ \boldsymbol{\gamma}'_2 \\ \boldsymbol{\gamma}'_1 \end{pmatrix}_{p\times p} \tag{5.1.3}$$

を考える。Part (I) にならい、$\boldsymbol{\gamma}'_i =: \left(\gamma_{i,1}, \gamma_{i,2}, \cdots, \gamma_{i,p}\right)$ と記すと、

$$\boldsymbol{\Gamma}\mathbf{z} = \begin{pmatrix} z_1/\|\mathbf{z}\| & z_2/\|\mathbf{z}\| & \cdots & z_{p-1}/\|\mathbf{z}\| & z_p/\|\mathbf{z}\| \\ \gamma_{p-1,1} & \gamma_{p-1,2} & \cdots & \cdots & \gamma_{p-1,p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{2,1} & \gamma_{2,2} & \cdots & \cdots & \gamma_{2,p} \\ \gamma_{1,1} & \gamma_{1,2} & \cdots & \cdots & \gamma_{1,p} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_{p-1} \\ z_p \end{pmatrix} = \begin{pmatrix} \|\mathbf{z}\| \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}_{p\times 1}. \tag{5.1.4}$$

よって、

$$\frac{T^2}{n} = \mathbf{z}'\left(\boldsymbol{\Gamma}'\boldsymbol{\Gamma}'^{-1}\right)\mathbf{A}^{-1}\left(\boldsymbol{\Gamma}'\boldsymbol{\Gamma}'^{-1}\right)\mathbf{z} = \left(\boldsymbol{\Gamma}\mathbf{z}\right)'\left(\boldsymbol{\Gamma}\mathbf{A}\boldsymbol{\Gamma}'\right)^{-1}\left(\boldsymbol{\Gamma}\mathbf{z}\right)$$

$$= \begin{pmatrix} \|\mathbf{z}\| & 0 & \cdots & 0 & 0 \end{pmatrix}_{1\times p}\left(\boldsymbol{\Gamma}\mathbf{A}\boldsymbol{\Gamma}'\right)^{-1} \begin{pmatrix} \|\mathbf{z}\| \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}_{p\times 1} \tag{5.1.5}$$

と表せる。ところで、

(5.1.6) $$\|\mathbf{z}\|^2 = z_1^2 + z_2^2 + \cdots + z_p^2 \sim \chi^2(p),\ \left(\because z_j \sim N(0,1),\ j = 1,\cdots,p\right)$$

であり、$\mathbf{\Gamma A \Gamma}' = \mathbf{A} \sim W_p\left(n, \mathbf{I}_p\right)$だから、遡及Cholesky分解により、$\mathbf{A}^{-1} = \left(\mathbf{\Gamma A \Gamma}'\right)^{-1}$の（1,1）要素の逆数は、$|\mathbf{A}|/|\mathbf{A}_{p-1}|$である。このことと、Part Iの（3.14）式等を参照して

(5.1.7) $$a_{1,1} = t_{1,1}^2 = |\mathbf{A}|/|\mathbf{A}_{p-1}| = \left\| P_{\mathcal{L}\left(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_{p-1}\right)^{\perp}} \mathbf{z} \right\|^2 \sim \chi^2(n-p+1)$$

となる。ここに、$\left\| P_{\mathcal{L}\left(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_{p-1}\right)^{\perp}} \mathbf{z} \right\|^2$は線形部分空間$\mathcal{L}\left(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_{p-1}\right)$の直交補空間上への正射影ベクトルのノルムを表す。$\mathbf{z}$と$P_{L\left(\gamma_1, \gamma_2, \cdots, \gamma_{p-1}\right)^{\perp}} \mathbf{z}$は独立に分布するから、

(5.1.8) $$\frac{T^2}{n} = \|\mathbf{z}\|^2 \Big/ \left\| P_{\mathcal{L}\left(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \cdots, \boldsymbol{\gamma}_{p-1}\right)^{\perp}} \mathbf{z} \right\|^2 \sim \frac{\chi^2(p)}{\chi^2(n-p+1)}$$

が従う。自由度で規格化すれば、所要のF分布と関連づけられる:

(5.1.9) $$\frac{T^2/p}{n/(n-p+1)} \sim F_{n-p+1}^{p}$$

なお、T^2そのものの密度関数は、（5.1.9）の関係から、次のように与えられる：

(5.1.10) $$f_p\left(t^2 \mid n\right) = \frac{1}{B\left(\frac{p}{2}, \frac{n-p+1}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \left(\frac{t^2}{n}\right)^{\frac{p}{2}-1} \frac{1}{n},$$

ただし、$B(\alpha, \beta)$はパラメータα, βの一変量ベータ関数を表す:

(5.1.11) $$B(\alpha, \beta) = \int_0^1 u^{\alpha-1} (1-u)^{\beta-1}\, du,\ \ (\alpha > 0, \beta > 0).$$

5.2. Wilksの$\mathbf{\Lambda}$-統計量の分布

この統計量は、一変量の場合の、ベータ統計量、F統計量あるいは分散分析の拡張ともみなせる。多変量分散分析で馴染みのある統計量であり、その使用事例を詳しく扱っている書籍は少なくない。しかし、その分布の誘導については、Wishart分布の場合と同様、これまでの多変量解析の教科書では、省略されていたり、与えられていても、間接的なものが散見される程度のように見受けられる（cf. Wilks（1962））。そこで以下に、遡及Cholesky分解を利用したWishart分布の導出を考慮して、Wilksの$\mathbf{\Lambda}$－統計量の

分布の誘導を試みる：

命題　5.2.1　$\mathbf{E}_{p\times p} \sim W_p(n-q, \boldsymbol{\Sigma}), (n-q \geq p)$とし、$\mathbf{z}_i(p\times 1) \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}), (i=1,\cdots,n)$に対し、$\mathbf{H}_{p\times p} := \mathbf{Z}_{p\times q}\mathbf{Z}'_{q\times p} = \Sigma_{i=1}^{q}\mathbf{z}_i\mathbf{z}'_i \sim W_p(q, \boldsymbol{\Sigma}), (p \leq q)$とする。また、$\mathbf{E}$と$\mathbf{H}$は統計的に独立と仮定する。この時、次の行列式の比に基づく統計量

$$\Lambda := \frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|} \tag{5.2.1}$$

は、次のように、互いに独立に分布する確率変数の積に分解される：

$$\Lambda = \prod_{i=1}^{p} t_{i,i}^2,\ \ t_{i,i}^2 \sim B\left(\tfrac{n-q-i+1}{2}, \tfrac{q}{2}\right),\ \ (p \leq q) \tag{5.2.2}$$

又は、

$$\Lambda = \prod_{j=1}^{q} t_{j,j}^2,\ \ t_{j,j}^2 \sim B\left(\tfrac{n-p-j+1}{2}, \tfrac{p}{2}\right),\ \ (p > q). \tag{5.2.3}$$

証明　まず，$\boldsymbol{\Sigma} = \mathbf{I}_p$の場合を証明する。$\mathbf{y}_i \sim N(\mathbf{0}, \mathbf{I}_p)(i=1,\cdots,n-q)$とし、

$$\mathbf{E}_\eta := \begin{pmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_{n-q} \end{pmatrix}_{p\times(n-q)} =: \begin{pmatrix} \boldsymbol{\eta}'_1 \\ \boldsymbol{\eta}'_2 \\ \vdots \\ \boldsymbol{\eta}'_p \end{pmatrix}_{p\times(n-q)} \tag{5.2.4}$$

を考える。ここに、$\boldsymbol{\eta}_i((n-q)\times 1) \sim N(\mathbf{0}, \mathbf{I}_{n-q}), (i=1,\cdots,p)$を表す。同様に

$$\mathbf{H}_\xi := \begin{pmatrix} \mathbf{z}_1 & \mathbf{z}_2 & \cdots & \mathbf{z}_{n-q} \end{pmatrix}_{p\times q} =: \begin{pmatrix} \boldsymbol{\xi}'_1 \\ \boldsymbol{\xi}'_2 \\ \vdots \\ \boldsymbol{\xi}'_p \end{pmatrix}_{p\times q} \tag{5.2.5}$$

を考える。ここに、$\boldsymbol{\xi}_i((n-q)\times 1) \sim N(\mathbf{0}, \mathbf{I}_q), (i=1,\cdots,p)$を表す。上記の設定で、

$$\mathbf{E} = \mathbf{E}_\eta\mathbf{E}'_\eta = \left(\boldsymbol{\eta}'_i\boldsymbol{\eta}_j\right)_{p\times p},\ \ \mathbf{H} = \mathbf{H}_\xi\mathbf{H}'_\xi = \left(\boldsymbol{\xi}'_i\boldsymbol{\xi}_j\right)_{p\times p} \tag{5.2. 6 }$$

なるGramian表示を得る。この時、（5.2.1）式右辺の分母における行列は

(5.2.7) $$\mathbf{E}+\mathbf{H}=: \mathbf{E}_{\eta}\mathbf{E}_{\eta}'+\mathbf{H}_{\xi}\mathbf{H}'=\left(\boldsymbol{\eta}_i'\boldsymbol{\eta}_j+\boldsymbol{\xi}_i'\boldsymbol{\xi}_j\right)_{p\times p}$$

と表現できる。ここで合成列ベクトル$\begin{pmatrix}\boldsymbol{\eta}_i & \boldsymbol{\xi}_\iota\end{pmatrix}' =: \tilde{\mathbf{x}}_i\ (n\times 1), (i=1,\cdots,p)$を考えると、

$$\mathbf{E}+\mathbf{H}=\begin{pmatrix}\boldsymbol{\eta}_1' & \boldsymbol{\xi}_1' \\ \boldsymbol{\eta}_2' & \boldsymbol{\xi}_2' \\ \vdots & \vdots \\ \boldsymbol{\eta}_p' & \boldsymbol{\xi}_p'\end{pmatrix}_{p\times n}\begin{pmatrix}\boldsymbol{\eta}_1 & \boldsymbol{\eta}_2 & \cdots & \boldsymbol{\eta}_p \\ \boldsymbol{\xi}_1 & \boldsymbol{\xi}_2 & \cdots & \boldsymbol{\xi}_p\end{pmatrix}_{n\times p}=\left(\boldsymbol{\eta}_i'\boldsymbol{\eta}_j+\boldsymbol{\xi}_i'\boldsymbol{\xi}_j\right)_{p\times p}$$

(5.2.8) $$=: \begin{pmatrix}\mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_p'\end{pmatrix}_{p\times n}\begin{pmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_p\end{pmatrix}_{n\times p}=\left(\mathbf{x}_i'\mathbf{x}_j\right)_{p\times p}$$

と表すことができ、新たなGramian表示を得る。したがって、(5.2.1)は次のように行列式の比として表せる：

(5.2.9) $$\Lambda=\frac{\left|\left(\boldsymbol{\eta}_i'\boldsymbol{\eta}_j\right)_{p\times p}\right|}{\left|\left(\mathbf{x}_i'\mathbf{x}_j\right)_{p\times p}\right|}.$$

そこで、正値対称行列$\left(\boldsymbol{\eta}_i'\boldsymbol{\eta}_j\right)_{p\times p}$に対し、実空間$R^{n-q}$で基底$\left\{\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_p\right\}$を持つ遡及Cholesky分解を適用する。ただし、ここで

(5.2.10) $$\mathbf{e}_i'=\Bigl(\underbrace{0,\cdots,0}_{i-1},1,\underbrace{0,\cdots,0}_{n-q-i}\Bigr)_{1\times(n-q)},\quad (i=1,\cdots,p)$$

を意味する。

ところで、Gramian行列式$\left|\left(\boldsymbol{\eta}_i'\boldsymbol{\eta}_j\right)_{p\times p}\right|$は、$\boldsymbol{\eta}_p,\boldsymbol{\eta}_{p-1},\cdots,\boldsymbol{\eta}_1$を辺とする多次元平行体の体積の二乗である(cf.斉藤(1966))から、

(5.2.11) $$\left|\left(\boldsymbol{\eta}_i'\boldsymbol{\eta}_j\right)_{p\times p}\right| = \left\|\boldsymbol{\eta}_p\right\|^2 \cdot \left\|P_{\mathcal{L}_\eta(\mathbf{e}_1)^\perp}\boldsymbol{\eta}_{p-1}\right\|^2 \cdot\cdots\cdot \left\|P_{\mathcal{L}_\eta(\mathbf{e}_1,\cdots,\mathbf{e}_{p-1})^\perp}\boldsymbol{\eta}_1\right\|^2$$

と表せる。ただし、$P_{\mathcal{L}_\eta(\mathbf{e}_1)^\perp}\boldsymbol{\eta}_{p-1}$は線形部分空間$\mathcal{L}_\eta\left(\mathbf{e}_2\right)$に沿った、$\mathcal{L}_\eta\left(\mathbf{e}_1\right)$上への正射影ベクトルを、$\left\|P_{\mathcal{L}_\eta(\mathbf{e}_1)^\perp}\boldsymbol{\eta}_{p-1}\right\|$は線形部分空間$\mathcal{L}_\eta\left(\mathbf{e}_1\right)$の直行補空間上への正射影ベクトルのノルムを、それぞれ表す。その他も同様である。なお、ベクトル$\boldsymbol{\eta}_i = \begin{pmatrix}\eta_{i,1} & \eta_{i,2} & \cdots & \eta_{i,p}\end{pmatrix}'$は$N\left(\mathbf{0},\mathbf{I}_{n-q}\right),(i=1,\cdots,p)$に従って分布しているから、第3章の議論（p.57 in Part （I））を参照して、

(5.2.12) $$\left|\left(\boldsymbol{\eta}_i'\boldsymbol{\eta}_j\right)_{p\times p}\right| \sim \chi^2\left(n-q\right)\cdot\chi^2\left(n-q-1\right)\cdot\cdots\cdot\chi^2\left(n-q-p+1\right)$$

であることが分かる。（5.2.9）の分母$\left|\left(\mathbf{x}_i'\mathbf{x}_j\right)_{p\times p}\right|$についても、実空間$R^n$で基底$\left\{\mathbf{u}_1,\ \mathbf{u}_2,\ \cdots,\ \mathbf{u}_p\right\}$を持つ遡及Cholesky分解の適用を考える。ただし、

(5.2.13) $$\mathbf{u}_i' = \Bigl(\underbrace{0,\cdots,0}_{i-1},1,\underbrace{0,\cdots,0}_{n-i}\Bigr)_{1\times n} = \Bigl(\mathbf{e}_i',\underbrace{0,\cdots,0}_{q},\Bigr)_{1\times n},(i=1,\cdots,p)$$

を表す。前と同様、$\left|\left(\mathbf{x}_i'\mathbf{x}_j\right)_{p\times p}\right|$は、$\mathbf{x}_p,\mathbf{x}_{p-1},\cdots,\mathbf{x}_1$を辺とする平行体の体積の二乗で、

(5.2.14) $$\left|\left(\mathbf{x}_i'\mathbf{x}_j\right)_{p\times p}\right| = \left\|\mathbf{x}_p\right\|^2 \cdot \left\|P_{\mathcal{L}_\eta(\mathbf{u}_1)^\perp}\mathbf{x}_{p-1}\right\|^2 \cdot\cdots\cdot \left\|P_{\mathcal{L}_\eta(\mathbf{u}_1,\cdots,\mathbf{u}_{p-1})^\perp}\mathbf{x}_1\right\|^2$$

と表せる。再度第3章の議論を用いれば、

(5.2.15) $$\left|\left(\mathbf{x}_i'\mathbf{x}_j\right)_{p\times p}\right| \sim \chi^2\left(n\right)\cdot\chi^2\left(n-1\right)\cdot\cdots\cdot\chi^2\left(n-p+1\right).$$

結局、（5.2.9）,（5.2.11）および（5.2.14）から,次の表現を得る：

(5.2.16) $$\boldsymbol{\Lambda} = \frac{\left\|\boldsymbol{\eta}_p\right\|^2}{\left\|\mathbf{x}_p\right\|^2}\cdot\frac{\left\|P_{\mathcal{L}_\eta(\mathbf{e}_1)^\perp}\boldsymbol{\eta}_{p-1}\right\|^2}{\left\|P_{\mathcal{L}_\eta(\mathbf{u}_1)^\perp}\mathbf{x}_{p-1}\right\|^2}\cdot\cdots\cdot\frac{\left\|P_{\mathcal{L}_\eta(\mathbf{e}_1,\cdots,\mathbf{e}_{p-1})^\perp}\boldsymbol{\eta}_1\right\|^2}{\left\|P_{\mathcal{L}_\eta(\mathbf{u}_1,\cdots,\mathbf{u}_{p-1})^\perp}\mathbf{x}_1\right\|^2}.$$

ところで、仮定により、$\left\|\boldsymbol{\eta}_p\right\|^2$ と $\left\|\boldsymbol{\xi}_p\right\|^2$ は互いに独立に分布するから、右辺第1項の分母

$\left\|\mathbf{x}_p\right\|^2 = \left\|\boldsymbol{\eta}_p\right\|^2 + \left\|\boldsymbol{\xi}_p\right\|^2$ はと分解される。よって、

$$\frac{\left\|\boldsymbol{\eta}_p\right\|^2}{\left\|\mathbf{x}_p\right\|^2} \sim \frac{\chi^2\left(n-q\right)}{\chi^2\left(n-q\right)+\chi^2\left(q\right)} \sim B\left(\tfrac{n-q}{2}, \tfrac{q}{2}\right) \tag{5.2.17}$$

が従う（cf. Wilks（1962））。この結果と（5.2.12）および（5.2.15）から、（5.2.16）の第 2 項以降の各項

$$\frac{\left\|P_{\mathcal{L}_{\eta}\left(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1}\right)^{\perp}}\boldsymbol{\eta}_{p+1-i}\right\|^2}{\left\|P_{\mathcal{L}_{\eta}\left(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1}\right)^{\perp}}\mathbf{x}_{p+1-i}\right\|^2},\ \left(i=2,\cdots,p\right) \tag{5.2.18}$$

についても，（5.2.17）と同様、分母のカイ二乗確率変数が独立な二つのカイ二乗確率変数に分割できて、各項が互いに独立にベータ分布することが推察される。以下でこのことを確かめる。分子に関し，$i=2,\cdots,p$ に対し、

$$R^{n-q} = \mathcal{L}_{\eta}\left(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1}\right) + \mathcal{L}_{\eta}\left(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1}\right)^{\perp}, \tag{5.2.19}$$

$$\dim\mathcal{L}_{\eta}\left(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1}\right) = n-q-i+1 \tag{5.2.20}$$

$\left(i=2,\cdots,p\right)$である。分母の $\mathcal{L}_{\mathrm{x}}\left(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1}\right)$はで正規直交基底$\left\{\mathbf{u}_1,\cdots,\mathbf{u}_{i-1}\right\}$で張られる線形部分空間であるが、（5.2.13）を想起すれば、次の直和表現が可能である：

$$R^{n} = \left\{\mathcal{L}_{\mathrm{x}}\left(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1}\right) \oplus \mathcal{L}_{\mathrm{x}}\left({}_{+\mathbf{0}}\mathbf{e}_1,\cdots,{}_{+\mathbf{0}}\mathbf{e}_{i-1}\right)^{\perp}\right\} \oplus \mathcal{L}_{\mathrm{x}}\left(O_{i-1}\right), \left(i=2,\cdots,p\right) \tag{5.2.21}$$

と表せる。ただし、

$$\mathcal{L}_{\mathrm{x}}\left({}_{+\mathbf{0}}\mathbf{e}_1,\cdots,{}_{+\mathbf{0}}\mathbf{e}_{i-1}\right)^{\perp} := \mathcal{L}_{\mathrm{x}}\begin{pmatrix}\mathbf{e}_1,\cdots,\mathbf{e}_{i-1}\\ \mathbf{0}_{q\times(i-1)}\end{pmatrix}^{\perp} \tag{5.2.22}$$

を意味し，$\mathcal{L}_{\mathrm{x}}\left(O_{i-1}\right)$ は $\mathcal{L}_{\mathrm{x}}\left(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1}\right)$ と $\mathcal{L}_{\eta}\left(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1}\right)^{\perp}$ に直交する補空間を表す。よって、

$$\dim\mathcal{L}_{\mathrm{x}}\left(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1}\right) = i-1, \tag{5.2.23}$$

$$\dim\mathcal{L}_{\mathrm{x}}\left({}_{+\mathbf{0}}\mathbf{e}_1,\cdots,{}_{+\mathbf{0}}\mathbf{e}_{i-1}\right)^{\perp} = \dim\mathcal{L}_{\eta}\left(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1}\right)^{\perp} = n-q-i+1, \tag{5.2.24}$$

$$\dim\mathcal{L}_{\mathrm{x}}\left(O_{i-1}\right) = n-\left(i-1\right)-\left(n-q-i+1\right) = q \tag{5.2.25}$$

である。（5.2.21）により、

(5.2.26) $$\left\|\mathbf{x}_{p+1-i}\right\|^2 = \left\|P_{\mathcal{L}_{\mathbf{x}}(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1})}\mathbf{x}_{p+1-i}\right\|^2 + \left\|P_{\mathcal{L}_{\mathbf{x}}\left({}_{+\mathbf{0}}\mathbf{e}_1,\cdots,{}_{+\mathbf{0}}\mathbf{e}_{i-1}\right)^{\perp}}\mathbf{x}_{p+1-i}\right\|^2 + \left\|P_{\mathcal{L}_{\mathbf{x}}(O_{i-1})}\mathbf{x}_{p+1-i}\right\|^2 .$$

よって、

(5.2.27) $$\left\|P_{\mathcal{L}_{\mathbf{x}}(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1})^{\perp}}\mathbf{x}_{p+1-i}\right\|^2 = \left\|\mathbf{x}_{p+1-i}\right\|^2 - \left\|P_{\mathcal{L}_{\mathbf{x}}(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1})}\mathbf{x}_{p+1-i}\right\|^2$$

(5.2.28) $$= \left\|P_{\mathcal{L}_{\mathbf{x}}\left({}_{+\mathbf{0}}\mathbf{e}_1,\cdots,{}_{+\mathbf{0}}\mathbf{e}_{i-1}\right)^{\perp}}\mathbf{x}_{p+1-i}\right\|^2 + \left\|P_{\mathcal{L}_{\mathbf{x}}(O_{i-1})}\mathbf{x}_{p+1-i}\right\|^2 .$$

ところで、空間図形の垂線の性質から、

(5.2.29) $$\left\|P_{\mathcal{L}_{\mathbf{x}}(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1})^{\perp}}\mathbf{x}_{p+1-i}\right\|^2 = \left\|P_{\mathcal{L}_{\eta}(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1})^{\perp}}\eta_{p+1-i}\right\|^2 .$$

また、

(5.2.30) $$\mathcal{L}_{\mathbf{x}}\left({}_{+\mathbf{0}}\mathbf{e}_1,\cdots,{}_{+\mathbf{0}}\mathbf{e}_{i-1}\right)^{\perp} = \mathcal{L}_{\eta}\left(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1}\right)^{\perp} \quad \text{in } R^{n-q}$$

であるから、（5.2.27）～（5.2.30）および（5.2.20），（5.2.23），（5.2.25）より

(5.2.31) $$\left\|P_{\mathcal{L}_{\mathbf{x}}(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1})^{\perp}}\mathbf{x}_{p+1-i}\right\|^2 = \left\|P_{\mathcal{L}_{\eta}(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1})^{\perp}}\eta_{p+1-i}\right\|^2 + \left\|P_{\mathcal{L}_{\mathbf{x}}(O_{i-1})}\mathbf{x}_{p+1-i}\right\|^2$$

(5.2.32) $$\sim \chi^2\left(n-i+1\right) = \chi^2\left(n-q-i+1\right) + \chi^2\left(q\right).$$

故に、 $i = 2,\cdots,p$ に対し、

(5.2.33) $$\frac{\left\|P_{\mathcal{L}_{\eta}(\mathbf{e}_1,\cdots,\mathbf{e}_{i-1})^{\perp}}\eta_{p+1-i}\right\|^2}{\left\|P_{\mathcal{L}_{\mathbf{x}}(\mathbf{u}_1,\cdots,\mathbf{u}_{i-1})^{\perp}}\mathbf{x}_{p+1-i}\right\|^2} \sim \frac{\chi^2\left(n-q-i+1\right)}{\chi^2\left(n-q-i+1\right)+\chi^2\left(q\right)} \sim B\left(\tfrac{n-q-i+1}{2},\tfrac{q}{2}\right).$$

と表せる。したがって、（5.2.16），（5.2.17），（5.2.33）から、

(5.2.34) $$\Lambda =: t_{p,p}^2 \cdot t_{p-1,p-1}^2 \cdot\cdots\cdot t_{1,1}^2$$

と表現できる。ただし、（5.2.17）および（5.2.33）から

(5.2.35) $$t_{i,i}^2 \sim B\left(\tfrac{n-q-i+1}{2},\tfrac{q}{2}\right),\ \left(i = 1,\cdots,p\right).$$

よって、命題の前半（5.2.2）が証明された。

次に、この結果を基に、定理の後半の（5.2.3）をモーメント法により証明する。（5.2.35）のベータ確率変数について、そのr次の中心積率は次式で与えられる（cf. Johnson,et.al.（1995）,p.217）：

$$(5.2.36)\qquad E\left[\left(t_{i,i}^2\right)^r\right]=\frac{\Gamma\left(\frac{n-q-i+1}{2}+r\right)}{\Gamma\left(\frac{n-q-i+1}{2}\right)}\cdot\frac{\Gamma\left(\frac{n-i+1}{2}\right)}{\Gamma\left(\frac{n-i+1}{2}+r\right)},\ \text{for all } r.$$

確率変数$t_{i,i}^2\left(i=1,\cdots,p\right)$は互いに独立に分布しているから、（5.2.34）を参照して

$$(5.2.37)\qquad E\left[\Lambda^r\right]=\prod_{i=1}^{p}\frac{\Gamma\left(\frac{n-q-i+1}{2}+r\right)}{\Gamma\left(\frac{n-q-i+1}{2}\right)}\cdot\frac{\Gamma\left(\frac{n-i+1}{2}\right)}{\Gamma\left(\frac{n-i+1}{2}+r\right)},\ \text{for all } r$$

と与えられる。そこで、特に$p>q$の場合を考える。上式右辺の共通項を相殺して、

$$(5.2.38)\qquad E\left[\Lambda^r\right]=\prod_{j=1}^{q}\frac{\Gamma\left(\frac{n-j+1}{2}\right)}{\Gamma\left(\frac{n-j+1}{2}+r\right)}\cdot\frac{\Gamma\left(\frac{n-p-j+1}{2}+r\right)}{\Gamma\left(\frac{n-p-j+1}{2}\right)},\ \text{for all } r.$$

これと（5.2.36）を比較すると、モメーン法により、

$$\Lambda=\prod_{j=1}^{q}t_{j,j}^2,\ \ t_{j,j}^2\sim B\left(\frac{n-p-j+1}{2},\frac{p}{2}\right),\ \ p>q$$

であることが示された。

元のWishart分布$W_p\left(\bullet,\Sigma\right)$の分散共分散行列が$\Sigma>\mathbf{0}$の場合について考える。実は$\mathbf{\Lambda}$の分布は上で詳述した$\mathbf{\Sigma}=\mathbf{I}_p$の場合と結果は同じであることが以下のようにして示せる。遡及Cholesky分解により、$\mathbf{\Sigma}=\mathbf{C}'\mathbf{C}$となる下三角行列$\mathbf{C}_{p\times p}$が存在する。すなわち、$\mathbf{C}^{-1\prime}\mathbf{\Sigma}\mathbf{C}^{-1}=\mathbf{I}_p$が成立する。そこで、（5.2.4）の$\mathbf{y}_i$に対し次の変換

$$(5.2.39)\qquad \mathbf{C}^{-1\prime}\mathbf{y}_i=:\tilde{\mathbf{y}}_i,\ \ \mathbf{y}_i\sim \text{i.i.d.}\,N_p\left(\mathbf{0},\mathbf{I}_p\right),\left(i=1,\cdots,n-q\right)$$

を考える。（5.2.4）より、

$$(5.2.40)\qquad \mathbf{E}_{\eta}:=\mathbf{C}'_{p\times p}\begin{pmatrix}\tilde{\mathbf{y}}_1 & \tilde{\mathbf{y}}_2 & \cdots & \tilde{\mathbf{y}}_{n-q}\end{pmatrix}_{p\times(n-q)}=:\mathbf{C}'_{p\times p}\begin{pmatrix}\tilde{\boldsymbol{\eta}}'_1\\ \tilde{\boldsymbol{\eta}}'_2\\ \vdots\\ \tilde{\boldsymbol{\eta}}'_p\end{pmatrix}_{p\times(n-q)},$$

ここに、$\tilde{\boldsymbol{\eta}}'_i\left(\left(n-q\right)\times 1\right)\sim N\left(\mathbf{0},\mathbf{I}_{n-q}\right),\left(i=1,\cdots,p\right)$である。同様に、次の変換を考え得る：

$$(5.2.41)\qquad \mathbf{H}_{\xi} := \mathbf{C}'_{p\times p}\begin{pmatrix}\tilde{\mathbf{z}}_1 & \tilde{\mathbf{z}}_2 & \cdots & \tilde{\mathbf{z}}_{n-q}\end{pmatrix}_{p\times q} =: \mathbf{C}'_{p\times p}\begin{pmatrix}\tilde{\boldsymbol{\xi}}'_1\\ \tilde{\boldsymbol{\xi}}'_2\\ \vdots\\ \tilde{\boldsymbol{\xi}}'_p\end{pmatrix}_{p\times q},$$

ここに、$\tilde{\mathbf{z}}_i \sim N\left(\mathbf{0},\mathbf{I}_p\right);\tilde{\boldsymbol{\xi}}'_i\,(q\times 1) \sim N\left(\mathbf{0},\mathbf{I}_p\right),(i=1,\cdots,q)$を意味する。したがって、（5.2.6）に対応して、

$$(5.2.42)\qquad \mathbf{E} := \mathbf{E}_{\eta}\mathbf{E}'_{\eta} = \mathbf{C}'_{p\times p}\left(\tilde{\boldsymbol{\eta}}'_i\tilde{\boldsymbol{\eta}}_j\right)_{p\times p}\mathbf{C}_{p\times p},$$

$$(5.2.43)\qquad \mathrm{H} := \mathrm{H}_{\xi}\mathrm{H}'_{\xi} = \mathrm{C}'_{p\times p}\left(\tilde{\xi}'_i\tilde{\xi}_j\right)_{p\times p}\mathrm{C}_{p\times p},$$

と表現できる。また、ベクトル$\begin{pmatrix}\tilde{\boldsymbol{\eta}}_i & \tilde{\boldsymbol{\xi}}_\iota\end{pmatrix}' =: \tilde{\mathbf{x}}_i\,(n\times 1),(i=1,\cdots,p)$を考えると、（5.2.7）に対応して

$$(5.2.44)\qquad \mathbf{E}+\mathrm{H} = \mathbf{C}'_{p\times p}\left(\tilde{\mathbf{x}}'_i\tilde{\mathbf{x}}_j\right)_{p\times p}\mathbf{C}_{p\times p}$$

と表現できる。結局、

$$(5.2.45)\qquad \boldsymbol{\Lambda} = \frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|} = \frac{\left|\mathbf{C}'_{p\times p}\left(\tilde{\boldsymbol{\eta}}'_i\tilde{\boldsymbol{\eta}}_j\right)_{p\times p}\mathbf{C}_{p\times p}\right|}{\left|\mathbf{C}'_{p\times p}\left(\tilde{\mathbf{x}}'_i\tilde{\mathbf{x}}_j\right)_{p\times p}\mathbf{C}_{p\times p}\right|} = \frac{\left|\left(\tilde{\boldsymbol{\eta}}'_i\tilde{\boldsymbol{\eta}}_j\right)_{p\times p}\right|}{\left|\left(\tilde{\mathbf{x}}'_i\tilde{\mathbf{x}}_j\right)_{p\times p}\right|}$$

となる。これは、$\boldsymbol{\Sigma}_{p\times p} = \mathbf{I}_{p\times p}$の場合と同一である（証明了）。

注　5.2.1. $\boldsymbol{\Lambda}$統計量は

$$\boldsymbol{\Lambda} = \frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|} = \frac{1}{|\mathbf{E}|^{-1}\left|\mathbf{E}\left(\mathbf{I}+\mathbf{E}^{-1}\mathbf{H}\right)\right|} = \frac{1}{\left|\mathbf{I}+\mathbf{E}^{-1}\mathbf{H}\right|} = \prod_{i=1}^{s}\frac{1}{1+\boldsymbol{\lambda}_i}$$

とも表せる。ここに、$s = rank\mathbf{H}$, $\boldsymbol{\lambda}_i\,(1,\cdots,s)$は$\mathbf{E}^{-1}\mathbf{H}$の非負固有値を表す。$\mathbf{E}^{-1}\mathbf{H}$は一般に非対称であるから、直接には一般化固有値問題を解くことになるが、前稿の遡及Cholesky分解$\mathbf{E} = \mathbf{T}'\mathbf{T}$を用いれば、対称行列$\mathbf{T}'^{-1}\mathbf{H}\mathbf{T}^{-1}$に対する標準固有値問題を解くことになり、問題の簡素化が可能である。

6．むすび

本稿では、Part I に続いて、遡及的なCholesky分解に基づく多変量解析の分布論の一端としてHotellingのT^2統計量とWilksのΛ統計量の分布誘導を行った。前者は多変量解析の代表的分布で広く用いられていることは論をまたない。後者についてもその有用さは、経営学や心理学等の先端での応用事例の積み重ねから、疑いの無いところである。定評のある多変量解析のパッケージソフトにはこれらの分布を使用してデータ分析等を行えるようになっている。Part I でも触れたように、多変量解析を学ぶ際に分布の誘導はかなり厄介な部分である。本稿では多変量分布の理論面を新しい側面から整理した。本稿がこの分野の理論的基盤を知る上で多少なりとも役に立つならば幸いである。

参考文献

松縄　規（2008）ビジネスデータ分析の理論的基礎（I）、— 遡及Cholesky分解と多変量分布 —，富山短期大学紀要　第43巻，49－60.

斎藤正彦　（1966）．線型代数学入門、東京大学出版会.

Johnson, N. L., Kotz, S. and Balakrishnan, N. （1995）. *Continuous Univariate Distributions* Vol. 2 , John Wiley and Sons, Inc., New York.

Rencher, A. C. （1995）. *Methods of Multivariate Statistical Analysis*, John Wiley and Sons, Inc., New York.

Wilks, S. S. （1962）． *Mathematical Statistics*, Wiley, New York.

（平成20年10月8日受付、平成20年10月31日受理）