日本統計学会誌
第22巻，第3号（増刊号），1993年
313頁～356頁

多変量解析の現状と展望

藤越康祝*，柳井晴夫**

Recent Developments and Perspectives in Multivariate Analysis

Yasunori Fujikoshi* and Haruo Yanai**

Multivariate analysis is generically known as a collection of statistical notion, methods and theory on multivariate data of several variables. In general, there are two directions in the course of research in multivariate analysis. One lays stress on the research of statistical inference, assuming probability distributions for multivariate data. The other is free from the assumptions of probability distributions, and lays stress on reduced representations, etc. of multivariate data. In the latter case, various methods have been proposed for qualitative or multi-way data as well as conventional multivariate data. The former is called multivariate statistical analysis, and the latter is called multivariate data analysis. Strictly speaking, these two directions cannot be discriminated. On the other side, it is natural that one should create a system of multivariate statistical methods, making up for the two directions each other. However, at the present time great progress is being made in two directions, and two directions are rather independently developing. From these reasons we first divide multivariate statistical methods into methods related to statistical inference and those related to multivariate data analysis, and discuss current perspectives and future developments in some topics of methods. The contributor of Part I : Multivariate Statistical Inference is Yasunori Fujikoshi, and the contributor of Part II : Multivariate Data Analysis is Haruo Yanai. In the last part of Part III we list books and proceedings on multivariate analysis, discriminating these into various areas. The contents of this paper are as follows:

多変量解析とは，簡単にいえば，複数個の変量に関する多変量データを分析するための統計的諸概念，諸方法，並びに，それらに関連する統計理論の総称であるといえる．一般に，多変量解析の研究には，２つの異なる方向がある．１つの方向は，確率モデルを想定し，母集団の母数推測に重点をおく研究である．他は，確率モデルの想定を必ずしも前提とせずに，多変量データの縮約表現法等に重点をおく研究である．後者の場合には，多変量データが質的データで表される場合，さらには，多重配列データによって表される場合に対しても，多くの手法が開発されている．前者は多変量推測法とよばれ，後者は多変量データ解析法とよばれる．これら２つの研究の方向は，厳密には区別されるものではなく，また，一方，当然のことながら，２つの方向は互いに補い合って，１つの多変量統計的方法の体系が作り上げられていくべきものである．しかし，多変量解析の理論と応用は，膨大なものになってきており，また，これら２つの研究方向は独立に発展していることから，本稿ではこれを２つに分けて現状と展望を論じる．第Ｉ部；多変量推測法の執筆者は藤越康祝で，第II部；多変量データ解析法の執筆者は柳井晴夫である．なお，最後の第III部においては，多変量解析に関する書物を分野別に上げている．

論文受付：1993年２月　改訂受付：1993年４月　受理：1993年４月
* 広島大学　理学部数学教室，〒724 東広島市鏡山 1-3-1
** 大学入試センター，〒153 東京都目黒区駒場 2-19-23

Contents

第I部　多変量推測法

藤越康祝

1. はじめに—研究動向

多変量推測法の現状を理解する上で，Anderson による多変量解析に関する 2 冊の本が参考になる．1つは1958年に出版され，他はその改訂版で1984年に出版されたものである．最初の本は，多変量解析の初期の時代から50年代後半までの多変量推測法に関する主要な結果が体系的にまとめられたものであるといえる．多変量解析の初期における重要な進展としては，Fisher (1915，1928，1936)による標本相関係数の分布の導出；回帰係数，偏回帰係数，重相関係数に関する有意性検定；正準判別分析法の導入，Wishart (1928) による Wishart 分布の導出，Hotelling (1931，1933，1936) によるスチューデント t 統計量の多変量への拡張；主成分・正準相関分析法の開発，Wilks (1932)による平均ベクトルと共分散行列に関するいくつかの仮説に対する尤度比検定，Fisher (1939), Hsu (1939), Roy (1939) 等による 2 つの Wishart 行列から定まる固有値の分布の導出，等である．これに，Rao (1946, 1948)による判別関数の係数，および，付加情報に関する有意性検定，Box (1949) による Box 型モーメントをもつ統計量の分布の漸近展開，Anderson and Rubin (1956)に代表される因子分析モデルに関する推測理論，等を加えると，Anderson の最初の本が出版されるまでの多変量推測法の概要を知ることができる．これらの研究においては，基礎となる確率モデルは多変量正規分布であり，また，推定・検定の基準は最尤法に基づくものになっている．改訂版においても，基礎となる確率モデルとしては多変量正規分布が仮定され，また，枠組も因子分析が独立した章として扱われている以外は，ほぼ同様である．しかし，初版が発刊されてからの26年間の間における多変量推測法の新しい進展が大幅に追加されている．これらの新しい進展の主なものは，Anderson が改訂版の序文において述べているように，次の通りである．

A 1. 平均ベクトルと共分散行列の点推定問題において，ある種の損失関数に関して最尤推定量より良い推定量，例えば，Stein 推定量，Bayes 推定量等の縮小推定量が導入されている．

A 2. 検定問題において，尤度比検定の他に不変検定が提案され，これらの検出力関数の許容性，不偏性，単調性が研究されている．

A 3. 検定統計量の精密分布，極限分布，漸近展開に関して，新しい結果が与えられている．

A 4. 平均ベクトルと共分散行列に関する同時信頼領域が展開されている．

A 5. 同時方程式モデル，関数関係モデル等の新しい話題が紹介されている．

上記の A 1～A 5 以外にも重要な多変量推測法に関する研究がされていることはいうまでもない．多変量データ解析法を含めた多変量解析全般に関する解説論文としては，Dempster (1971), Rao (1972，1983)，塩谷 (1979), Sibson (1984), Schervish (1987)，等がある．これらにおいて強調されている問題点の1つは，推測理論のほとんどが多変量正規性の仮定のもとで展開されているという点である．これは，現実の多変量データには，その確率モデルとして多変量正規性の前提が適当でないものが多いことを意味している．Schervish (1987) は多変量解析の 2 冊の本 Anderson (1984), Dillon and Goldstein (1984) において扱われている内容を比較しながら，多変量解析全般の解説を行っている．これら 2 冊の本のうち，前者は多変量推測法に関するものであり，後者は，種々の方法とその応用に重点がおかれており，むしろ，多変量データ解析法に属するものである．この解説論文においては，Anderson の本では取り上げられていないが，Dillon and Goldstein において扱われている多くの有用な多変量解析法を紹

介し，多変量解析の理論と方法はこの２冊の本の合併として把えられるべきであることを主張している．Anderson の本で扱われていない解析法とは，クラスター分析，多次元尺度法，多変量グラフ解析法等の探索的方法とパス解析法であって，推測論的側面からの発展が十分でない分野である．これらの方法は，本論文の第II部で取り上げられる．この２冊の本によって，多変量解析に関する話題はかなり網羅される．しかし，個々の内容は膨大になってきており，この２冊の本でカバーされるものではなく，また，話題としても，変量選択法，成長曲線モデル，ブートストラップ法(Efron (1979))を含む多変量ノンパラメトリック推測法，射影追跡(Huber (1985))，多変量ランダムモデル等の重要なものが除かれている．Sen (1986) はこれら２冊の本を含む多変量解析に関する１６冊の本について，比較を中心とした書評を行っている．

一般に，推測理論においては，その目標として統計的方法の構築とその発展・評価が掲げられ，また，その展開においては，明解さ，首尾一貫性，並びに，現実性が要求される．これらの点は，多変量推測法の現状の発展をみる上で，とくに重要である．上述のように，多変量推測法は主として多変量正規性のもとで発展しており，従って，現実のデータ解析の適用においては多くの吟味されるべき問題点が残されている．また，推測的展開が十分でない多変量データ解析法も多い．すなわち，大雑把にいえば，現状の多変量推測法とデータ解析法との間にはかなりの乖離がみられ，この乖離をなくしていくことが当面の課題であろう．従って，今後の多変量推測法の研究においては，上記のＡ1〜Ａ5をさらに発展させることも必要であろうが，むしろ，大きな発展が望まれる下記の研究課題が重要であろう．

Ｂ1. 新しい多変量モデルの構築とその推測

Ｂ2. 非正規多変量モデル；(1) Elliptically Contoured モデル，(2) 多変量離散モデル，(3) 連続型・離散型混合モデル，(4) 与えられた周辺分布をもつ多変量モデル，等に関する推測

Ｂ3. 次元縮小法，変量選択法に関する推測

Ｂ4. クラスター分析，多次元尺度法，数量化法，等に対する推測的アプローチ

Ｂ5. 多変量ノンパラメトリック推測法

Ｂ6. 多変量統計的方法のロバストネス

Ｂ7. 多変量平均構造・共分散構造・不等式制約構造，等に関する推測

Ｂ8. 多変量ランダム効果モデル，多変量混合効果モデルにおける推測

Ｂ9. 欠測値がある場合の推測

Ｂ10. 高次元小標本の場合の推測

Ｂ11. 多変量経時データ，多変量寿命データ，方向性データ，等に対するモデルと推測

Ｂ12. 多変量予測法

Ｂ13. 漸近的近似法（漸近展開，ブートストラップ法等）の精度

多変量推測法の発展は多岐にわたり，また，分野によっては膨大な結果が得られている．従って，全般について個別に現状を解説することは，この論文の範囲を超えるものである．本論文においては，標本分布，多変量線形モデル，次元縮小法，及び，上記Ｂ1〜Ｂ13に関連したいくつかの話題に焦点をあてる．また，個々の結果については，できるだけ最近のサーベイ論文を引用文献として掲げることにする．なお，今後の多変量解析の研究の方向については，1992年5月にペンシルバニア州立大で開催された多変量解析—今後の方向—に関する国際会議におけるプロシーディング，C. R. Rao (ed.) (1993), Multivariate Analysis: Future Direction (North-Holland Series in Statistics and Probability Vol. 7) が参考になると思われる．

2. 検定統計量の標本分布

2.1 精密標本分布

Wilks (1932) の研究以来，多変量正規分布の平均ベクトルと共分散行列に関する種々の仮説に対して尤度比検定を含むいくつかの検定統計量が導入され，それらの標本分布が研究されている．この発展をみる上で，多くの研究者によって調べられてきた統計量

$$T_1=|S_e|/|S_e+S_h|$$

の分布を考える．ここに，S_e, S_h は互いに独立で，それぞれウィシャート分布 $W_p(\Sigma, n)$，非心ウィシャート分布 $W_p(\Sigma, c; \Omega)$ に従い，$n\geq p$ とする．一般性を失うことなく，$\Sigma=I_p$, $\Omega=diag(\omega_1,\cdots,\omega_p)$, $\omega_1\geq\cdots\geq\omega_p\geq 0$ としてよい．この統計量は仮説"$\Omega=0$"の検定に用いられるものであって，その仮説のもとでの分布 (null 分布) はラムダ分布 $\Lambda_{p,c,n}$ と呼ばれる．T_1 の null 分布を調べる上で，T_1 が互いに独立なベータ変量の積として表されること，すなわち $T_1=V_1\cdots V_p$ となる構造をもつことが基本となる．ここに，V_i はベータ分布 $\beta\left(\frac{1}{2}(n+1-i), \frac{1}{2}c\right)$ に従う．

検定統計量の null 分布が上のように独立なベータ変量の積になるものはかなりある．また，$|S_1|/|\Sigma|=U_1\cdots U_p$, $U_i\sim\chi^2(n-p+i)$ のように独立なカイ二乗分布の積として表されるものもある．Mathai (1973) はこのような統計量の精密分布を求める方法について解説している．そこでは，6つの方法；(1) 直接計算法，(2) 特性関数法，(3) たたみこみ法，(4) 逆メリン変換法，(5) 部分分数法，(6) 留数計算法 が紹介されている．一般に，$\min(p, c)$ が大きくなるにつれて，精密分布表示を求めることが困難になる．

検定統計量 T_1 は，多変量線形仮説に対する尤度比基準であり，この場合，T_1 の他に

Lawley-Hotelling 基準；$T_2=\mathrm{tr}S_hS_e^{-1}$

Bartlett-Nanda-Pillai 基準；$T_3=\mathrm{tr}S_h(S_e+S_h)^{-1}$

Roy 基準；$T_4=l_1=S_hS_e^{-1}$ の最大固有値

が提案されている．T_1 を含めたこれらの null 分布については，Krishnaiah (1978, 1980), Anderson (1984) を参照されたい．

検定統計量の null 分布に関する結果は，p，c が小さいときは早くから求められている．例えば，Wilks (1932) は $(p, c)=(1, 3)$, $(2, 3)$, $(3, 3)$, $(3, 4)$, $(4, 4)$ のとき T_1 の分布を具体的に与えている．これに比べ，対立仮説のもとでの分布 (nonnull 分布) を求める問題は長い間未解決のままであった．この突破口は，James (1960, 1964) の zonal 多項式の導入により開かれた．また，Constantine (1963) による zonal 多項式を用いた行列変数の超幾何級数の導入によるところも大きい．これにより，例えば，非心ウィシャート分布の確率密度関数が ${}_0F_1$ 型の超幾何級数を用いて表され，また，T_1 のモーメントが

$$E(T_1^h)=\frac{\Gamma_p(h+c/2)\Gamma_p((n+c)/2)}{\Gamma_p(c/2)\Gamma_p(h+(n+c)/2)}{}_1F_1(h;h+(n+c)/2;-\Omega)$$

として表される．ここに，$\Gamma_p(a)=\pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma((a-(i-1))/2)$．zonal 多項式，行列変数の超幾何級数については，Muirhead (1982)，竹村 (1991) に詳しくまとめらている．1960 年代から 1970 年代にかけて，多くの検定統計量の nonnull が zonal 多項式，とくに，行列変数の超幾何級数を用いて表された．Pillai (1976, 1977) は固有値を用いて表される検定統計量の null 分布および nonnull 分布について，409 編の参考論文を含むサーベイを行っている．Davis (1979,

1980）は1つの行列変数のzonal多項式を2つの行列変数のzonal多項式に拡張し，さらに，Chikuse (1980) は$r(\geq 3)$個の行列変数を含むzonal多項式を導入している．これらの拡張されたzonal多項式を用いて表される標本分布については，Chikuse and Davis (1986) を参照されたい．多変量標本分布の導出におけるzonal多項式の利用に焦点を当てた最近のサーベイ論文としては，Hayakawa (1989) を参照されたい．多くの多変量標本分布がzonal多項式を含む形で求められるようになってきたが，それらを用いた数値的確率計算への利用においてはいくつかの問題点が残されている．一つは，一般の場合のzonal多項式が具体的に求められていない点である．また，分布表示は全ての次数のzonal多項式を含む無限級数になっているが，その収束性が一般に遅いことである．2次元の場合にはzonal多項式に対する一般表示が知られており，Pillai and Jayachandran (1967) は検定統計量 $T_1,\cdots, T_4$ のnonnull分布を具体的に求め，これを利用してこれらの検出力を数値的に与えている．多変量の標本分布は，H-関数やG-関数を用いても表されているが，これについてはRathie (1989) のサーベイ論文を参照されたい．

2.2 漸近展開

多変量解析の標本分布の導出においては，与えられた統計量の標本分布を正確に求めることが困難であるとか，あるいはたとえ求められたとしても余りにも複雑で数値計算には利用できないことがしばしば生ずる．このため，何等かの近似が必要となり，とくに標本数を無限大としたときの極限分布，あるいは，極限分布を第1近似とする漸近展開近似の導出が早くから研究されている．多変量の場合，極限分布による近似が有効であるためには相当数の標本が必要となり，中程度の標本に対しても有効である漸近展開近似が重要な位置を占めている．標本の大きさnに依存する分布関数$F_n(x)$をその極限分布$G(x)$のまわりで展開したときの漸近展開の典型的な形は

$$G_{k,n}(x)=G(x)+\sum a_j(x)g(x)n^{-j}$$

または，nを$\sqrt{n}$で置き換えたものである．ここで，$g(x)$はGの確率密度関数，$a_j(x), j=1,\cdots, k-1$は適当な多項式を表す．漸近展開については，漸近展開式の導出，漸近展開の正当化，および，誤差限界の導出，に関する研究があり，それぞれ，独立に研究されている．漸近展開の正当化とは，$F_n(x)$を$G_{k,n}(x)$で近似したときの誤差$R_{k,n}(x)=F_n(x)-G_{k,n}(x)$が正しいオーダーでおさえられること，例えば，$R_{k,n}(x)=O(n^{-k})$を示すことである．漸近展開式の導出の過程においては，必ずしも誤差項の厳密な評価がなされない場合が多い．このような漸近展開式は形式的な漸近展開とよばれている．

多変量解析における漸近展開式の最初の2，3の項を具体的に求める研究は，非常に進んでいて，多くの結果が多変量解析の本に紹介されている．Andersonの本の初版には，平均ベクトルと共分散行列に関する種々の尤度比基準のnull分布の漸近展開が載せられている．これらの統計量はそのモーメントがBox型になっているものである．その後，種々の漸近展開を求める方法が導入され，尤度比基準を含む各種検定統計量のnull，および，nonnull分布の漸近展開が求められている．これらの多くは，Andersonの第2版，Muirhead (1982), Siotani, Hayakawa and Fujikoshi (1985) 等の本に紹介されている．漸近展開の導出については，日本の研究者が大きく貢献している．これらの結果については，塩谷 (1975, 1976), Muirhead (1978), Nagao (1989), 等のサーベイ論文，Sugiura (1976), Fujikoshi (1977), Konishi (1979) 等を参照されたい．塩谷 (1975, 1976) は，漸近展開式を求める5つの方法を詳しく説明するとともに，その方法を用いて得られた結果を列挙している．これらの5つの方法は，(1) テイラー展開に基づく方法，(2) 特性関数の正確な表示から導く方法，(3) 超幾何級数が満足する偏微分方程式

系に基づく方法，(4) 統計量の正規変量による展開を利用する方法，(5) 直交変量行列の歪対称行列表現を用いる方法，である．このうち，方法 (2)，(3) は，zonal 多項式を含む精密標本分布からより利用しやすい漸近展開近似を得るための方法であるといえる．これらの方法は，基礎になる確率モデルが多変量正規分布である場合の方法である．しかし，方法 (1)，(4) をより一般にした，いわゆる摂動展開法は一般の多変量連続型非正規モデルの場合にも適用できる方法である．この方法を用いて，Fujikoshi (1980) は適当な正則条件をみたす多変量非正規母集団からの標本に基づく標本共分散行列の固有値の分布の漸近展開を与えている．統計量の関数の漸近展開も求められるが，Konishi (1987) は統計量の漸近展開と正規化変換・分散安定化変換との関連をサーベイしている．

漸近展開の正当化の研究，すなわち誤差項の正しいオーダー評価を行うことは，漸近展開式の存在証明に相当していて，展開式を具体的に求める研究とはかなり異なっている．この研究における大きな進展は Bhattacharya and Ghosh (1978) によってなされた．この論文は，独立変量の和の関数として表される統計量について，その極限分布が正規分布になる場合の漸近展開の正当化を示したもので，広く応用されている．その後，Chandra and Ghosh (1979, 1980) は極限分布がカイ二乗分布，非心カイ二乗分布になる場合について，適当な正則条件のもとでの正当化を与えている．関連した最近の研究としては，正当化のための正則条件をゆるめることが，Bai and Rao (1991) 等によってなされている．誤差項のオーダー評価が示されることは，例えば

$$|R_{k,n}(x)|\leq C_k n^{-k},\quad (\text{または } C_k n^{-k/2})\quad n\geq n_0$$

を満たす十分大きな定数 C_k, n_0 の存在を保証するものであるが，しかし，具体的な C_k, n_0 は求められていないのがふつうである．従って，漸近展開による近似の精度を知るためには，誤差項のオーダー評価だけでは不十分であって，上のように誤差項のオーダー評価を含む形での誤差限界の導出，すなわち C_k を具体的に求めることが必要となる．これに関する結果としては，独立和の分布に対する正規近似の誤差限界である Berry-Esseen 限界が知られている．誤差限界に関する研究はあまり進んでいないが，最近，若干の進展がみられることを注意したい．Kariya and Maekawa (1982), Fujikoshi (1985 a) はそれぞれ，SUR (Seemingly Unrelated Regression) モデル，GMANOVA モデルにおける推定量の漸近展開に対して誤差限界を与えている．また，Shimizu (1987) は正規尺度変量について，同様な研究を行っている．これらを含む誤差限界に関するサーベイ論文として，Fujikoshi and Shimizu (1988), Fujikoshi (1993) がある．後者の論文には，ある種の Cornish-Fisher 展開に対する誤差限界が紹介されている．

3. 多変量線形モデルとその一般化

3.1 多変量線形モデル

N 個の固体の各々について p 次元変量 y が観測され，固体 i の観測変量を y_i とし，線形モデル

$$y_i=\Xi' a_i+e_i,\quad i=1,\cdots,N$$

を仮定する．ここに，Ξ は $k\times p$ の未知パラメータ行列，a_i は $k\times 1$ の既知ベクトルで，$A=[a_1,\cdots,a_N]'$ の階級は k とする．また，誤差ベクトル e_i は互いに独立で，それぞれ，$N_p(0,\Sigma)$ に従うとする．$Y=[y_1,\cdots,y_N]'$ とおくとき，$E(Y)=A\Xi$ であって，上のモデルは，多変量分散分析 (MANOVA) モデル，多変量共分散分析 (MANCOVA) モデル，多変量回帰モデル，等を含んでいる．行列 $C:c\times k$ を $\mathrm{rank}(C)=c$ の既知行列とするとき，検定問題“$C\Xi=O$” vs “$C\Xi\neq$

O"に対する標準形は次のようになる：確率行列 $Z:N\times p$ の各行は互いに独立でそれぞれ $N_p(\cdot,\Sigma)$ に従い，

$$E(Z)=E(\begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix})=\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ O \end{bmatrix}\begin{array}{l} c \\ k-c \\ n=N-k \end{array}$$

とするとき，仮説検定問題は"$\zeta_1=O$" vs "$\zeta_1\neq O$"として与えられる．尤度比基準は，$S_h=Z_1'Z_1$, $S_e=Z_3'Z_3$ とおくとき，$T_1=|S_e|/|S_e+S_h|$ となり第2節で考察された統計量と同じものであって，これを含めた4つの検定統計量 $T_1,\cdots,T_4$ が代表的な検定統計量である．不変検定は $S_hS_e^{-1}$ の固有値 $l_1\geq\cdots\geq l_p\geq 0$ に基づくものである．上の4つの検定法を含む不変検定の性質については，最近かなりのことがわかってきている．$p=1$ または $c=1$ のとき，不変検定は全て一致し，一様最強力不変検定となる．一般に，不変検定の検出力関数は p, c, n の他に $\Sigma^{-1}\zeta_1'\zeta_1$ の固有値 $\omega_1\geq\cdots\geq\omega_p\geq 0$ に依存している．Das Gupta, Anderson and Mudhaolkar (1964) は，検出力関数が各 ω_i の単調増加関数になるための十分条件を与えている．この結果を用いると，T_1, T_2, T_4 の単調性が示される．T_3 の単調性については，その棄却点が1以下であるという条件が必要になる．不変性は単調性からも得られる結果であるが，より一般的な不偏性条件は Perlman and Olkin (1980) によって与えられている．各検定法 $T_1,\cdots,T_4$ の許容性はそれぞれ個別に示されてきたが，Anderson and Takemura (1982) は不変検定の許容性について新しい統一的な証明法を与えている．検定法の比較については，検出力の数値的比較，漸近展開法，等による結果がある．とくに，検出力関数の漸近展開比較 ($O(n^{-2})$ を無視した) に関する結果として，T_1, T_2, T_3 の検出力は

$$\gamma=\frac{\sigma_\omega^2}{\bar{\omega}^2}-\frac{(p-1)(p+2)}{pc+2},\quad (\bar{\omega}=\frac{1}{p}\sum_{i=1}^{p}\omega_i,\ \sigma_\omega^2=\frac{1}{p}\sum_{i=1}^{p}(\omega_i-\bar{\omega})^2)$$

の値が正か負かによって順序づけられることが知られている (Anderson (1984), Fujikoshi (1988), 竹村(1991))．すなわち，$\gamma>0$ のとき，検出力は大きい順に $T_2>T_1>T_3$ となり，$\gamma<0$ のときには検出力の順番は逆となる．この結果は T_1, T_2, T_3 を含む検定法のクラスに対しても拡張されている (Fujikoshi (1988))．

上記の結果は，固定効果モデルに適用されるものである．ランダム効果を含むモデルとしては，例えば

$$y_i=\Xi' a_i+\beta_i+e_i,\ i=1,\cdots,N$$

が考えられる．ここに，新たに追加された項 β_i は，固体 i のランダム効果を表す変量であって，e_i とは独立で，$N_p(0,\Lambda)$ に従っている．また，共分散行列 Λ は半正定値である．このモデルにおける検定問題としては，"$\mathrm{rank}(\Lambda)=m$" vs "$\mathrm{rank}(\Lambda)>m$"に関心がある．k 個の処理またはグループからなる1元配置モデルにおいて，繰り返し数が等しい場合には尤度比基準が求められている (例えば，Anderson and Olkin (1986))．また，この尤度比基準，および，関連する統計量の漸近分布が研究されている (Anderson (1989), Anderson and Amemiya (1991))．Kuriki (1993) は Wishart 行列から出発して，その共分散行列 Σ についての階層的仮説 $H_0:\Sigma=\sigma^2I_p$, $H_1:\Sigma\geq\sigma^2I_p$, $H_2:\Sigma>O$ に関連して2つの検定問題"H_0 vs $H_1\cap H$, "H_1 vs $H_2\cap H_1$" の尤度比基準の性質を調べている．仮説 H_1 は，ランダム効果に関連して Rao (1965) によって導入されたモデルである．一般に，多変量ランダム効果モデルにおける統計的推測の発展はあまりみられず，多くの未解決問題が残されている．

3.2 多変量線形モデルの拡張

各個体について，ある変量 y が p 個の時点で測定される場合には，次の一般化線形モデルが重要となる．

$$y_i = X\Xi' a_i,\ i=1,\cdots,N$$

ここに，$X: p\times q$ は $\mathrm{rank}(X)=q$ の既知行列で，$\Xi: q\times p$ は未知パラメータ行列である．X は個体内計画行列とよばれている．この場合の観測行列 Y の平均構造は，$E(Y)=A\Xi X'$ である．このモデルは，Rao (1959), Potthoff and Roy (1964) によって導入され，成長データの解析に応用されている．仮説検定問題 "$C\Xi D=O$" vs "$C\Xi D\neq O$" ($C: c\times k, K: q\times d$ はそれぞれ階数 c，d の既知行列) に対する標準形は次のようになる：確率行列 $Z: N\times p$ の各行は互いに独立でそれぞれ $N_p(\cdot,\Sigma)$ に従い，かつ

$$E(Z)=E\left(\begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33}\end{bmatrix}\right)=\begin{array}{cc} \begin{bmatrix} \zeta_{11} & \zeta_{12} & O \\ \zeta_{21} & \zeta_{22} & O \\ O & O & O \end{bmatrix} & \begin{array}{l} c \\ k-c \\ n=N-k \end{array} \\ \begin{array}{ccc} q-d & d & p-q \end{array} & \end{array}$$

とするとき，仮説検定問題は $H_0: \zeta_{12}=O$ vs $H_1: \zeta_{12}\neq O$ として与えられる．この検定に対する尤度比基準は Khatri (1966), Gleser and Olkin (1970) によって求められた．$V=[Z_{31}\ Z_{32}\ Z_{33}]'[Z_{31}\ Z_{32}\ Z_{33}]$, $V_{ij}=Z_{3i}'Z_{3j}$, $U=(I+Z_{13}V_{33}^{-1}Z_{13}')^{-1/2}(Z_{12}-Z_{13}V_{33}^{-1}V_{32})$ とおくとき，尤度比統計量は $T_1=|S_e|/|S_e+S_h|$ と表せる．ここに，$S_e=V_{22\cdot3}=V_{22}-V_{23}V_{33}^{-1}V_{32}$, $S_h=U'U$ である．MANOVA の場合と同様に，他の検定統計量 T_2, T_3, T_4 も同様に定義される．S_e はウィシャート分布 $W_d(n-(p-q),\Sigma_{22\cdot3})$ に従い，また，S_h は，$W_1=Z_{13}V_{33}^{-1}Z_{13}'$ が与えられたとき，S_e とは独立に非心ウィシャート分布 $W_d(c,\Sigma_{22\cdot3};\Theta_{12}'(I+W_1)^{-1}\Theta_{12})$ に従っている．ここに，$\Sigma_{22\cdot3}=\Sigma_{22}-\Sigma_{23}\Sigma_{33}^{-1}\Sigma_{32}$, Σ_{ij} は Σ の分割行列である．この結果を利用して，検定統計量 $T_1,\cdots,T_4$ に基づく検定法について MANOVA の場合と同様な性質が得られている．しかし，許容性については，条件つき許容性から本来の許容性が得られないため十分な結果が得られていない．これらの検定法は統計量 $W=US_e^{-1}U'$ に基づいている．Kariya (1978) は不変検定の本質的完備クラスは W と W_1 に基づく検定であることを示し，さらに，局所最良不変検定の棄却域は $T_5=tr(I+W_1)^{-1}[(c+n-p+q)d^{-1}W(I+W)^{-1}-I]>t$ で与えられることを示している．検定問題については Kariya (1985) に詳しく解説されている．

一般化 MANOVA モデルについてはいくつかのサーベイ論文があるが，とくに von Rosen (1991) は 237 編の引用文献を挙げて解説している．その内容は，推定，検定，標本分布，共分散構造，信頼領域，拡張モデル，Ξ に関する線形制約，欠測値，ベイズ法，縦断的研究，等の多岐にわたり，それぞれについて文献を中心にした解説がなされている．なお，推定問題，予測問題については Rao (1987 a, b) を参照されたい．

一般化 MANOVA モデルとその関連モデルは成長データを初めとする経時データの解析へ応用されている．このとき，固体の変動を考慮したランダムモデル

$$y_i = X\Xi' a_i + X\beta_i + e_i,\quad i=1,\cdots,N$$

も提案されている (Rao (1965))．β_i は，MANOVA の場合のランダムモデルと同様に e_i とは独立で，$N_q(0,\Delta)$ に従う確率変量である．しかし，e_i には $Var(e_i)=\sigma^2 I_p$ が仮定され，全誤差項 $\varepsilon_i=X\beta_i+e_i$ については

$$E(\varepsilon_i)=0, \quad Var(\varepsilon_i)=\Sigma_\varepsilon=X\Delta X'+\sigma^2 I_p$$

となっている．共分散構造 Σ_ε は Rao の共分散構造と呼ばれている．半正定値行列 Δ の不変推定量は容易に求められるが，不偏推定量は必ずしも半正定値ではないという欠点がある．経時データにおいては，測定時点や測定条件が固体毎に異なる場合がしばしば生ずる．また，各固体の繰り返し測定数も固体によって異なったり，欠測値も生じがちである．このようなアンバランスなデータに対するモデルとして，固体 i の p_i 次元観測変量 y_i について，個体内計画行列 X を固体に依存した X_i にしたモデルが提案されている（Laird and Ware (1982)）．また，パラメータ Ξ, σ^2, Δ の MLE を数値的に求める方法として，EM または NR アルゴリズムが適用されている（Lindstrom and Bates (1988)）．Vonesh and Carter (1987) は拡張モデルのもとでの反復計算を必要としない推測法を提案している．最近，誤差項の共分散構造として自己回帰構造を想定する場合等を含め一般化 MANOVA およびその拡張モデルに関する論文は多い．これらに関しては，先に述べた von Rosen (1991) のサーベイ論文，および，Crowder and Hand (1990) の研究書を参照されたい．同一固体について複数個の変量が繰り返し測定される場合のデータは多変量経時データとよばれる．バランス型で共分散行列が一般の正定値行列である場合には，一般化 MANOVA モデルの理論を適用することができる．ランダム効果共分散構造をもつ場合には，新しい推測法が必要となる（Reinsel (1982, 1984), Boik (1988)）．

4. 次元推定と変量選択

4.1 問題の定式化

多変量解析法の多くは，多変量データを小数次元の空間に表現することと関係している．このとき，縮小表現空間の座標とその次元が重要となる．例えば，多群の判別分析において，群間を判別するためには何個の判別関数が必要であるか，2組の変量間の関係を記述するためには何個の正準相関変量が必要であるか，多次元のデータ集合の変動の大部分を説明するためには何個の主成分が必要であるか，等は新しい座標の次元を決定する問題である．また，多変量データの解析において，用いられている変量の中に冗長的なものがあるかどうかも問題になる．後者の問題は変量選択問題とよばれる．

上のような問題は，多変量解析の本，Kshirsagar (1972), Muirhead (1982), Anderson (1984), Siotani, Hayakawa and Fujikoshi (1985) においても扱われている．一般に，ある統計解析に対して p 次元変量 $\boldsymbol{x}=(x_1,\cdots,x_p)'$ から新しい変量

$$y_i=y_i(\boldsymbol{x}), \quad i=1,\cdots,q$$

が定義され，かつ，これらのうちの最初の m 個による縮小表現に関心があるとする．また，変量の集合 $\{y_1,\cdots,y_m\}$ の各々について，統計解析の目的に関する情報量 $I(\{y_1,\cdots,y_m\})\geq 0$ が与えられているとする．このとき，変換変量の次元は，通常 $I(\{y_1,\cdots,y_m\})=I(\{y_1,\cdots,y_q\})$ を満たす最小な m として定義される．一般には，与えられた比率 $g(0<g\leq 1)$ に対して，$I(\{y_1,\cdots,y_m\})\geq gI(\{y_1,\cdots,y_q\})$ を満たす最小な m にも関心がある．とくに，$g=1$ のときの m は前者の次元になる．後者のような次元は主成分分析に適用される．この定式化に従って，判別分析，正準相関分析の通常の次元が定義され，また，主成分分析の実際的な次元が定義される．

(1) 判別分析　k 個の p 次元正規母集団 $N_p(\mu_i, \Sigma)$, $i=1,\cdots,k$ からの大きさ N_i の標本に基づく群内変動行列，群間変動行列をそれぞれ A，B とする．このとき，A，B は互いに独立で，それぞれウィシャート分布 $W_p(n, \Sigma)$，非心ウィシャート分布 $W_p(k-1, \Sigma; N\Omega)$ に従う．ここに，$N=N_1+\cdots+N_k$, $n=N-k$, $\Omega=\Sigma_{i=1}^k(n_i/N)(\mu_i-\bar{\mu})(\mu_i-\bar{\mu})'$, $\bar{\mu}=(1/N)\Sigma_{i=1}^N N_i\bar{\mu}_i$．変換変量 y_i

$=\gamma_i'\boldsymbol{x}$, $i=1,\cdots,q(=\min(p,k-1)$ は判別関数であって，γ_i は $\Sigma^{-1}\Omega$ の第 i 固有値 w_i に対応する固有ベクトルである．$I(\{y_1,\cdots,y_m\})=\omega_1+\cdots+\omega_m$ と定義すれば，判別変量の次元はゼロでない ω_i の個数である．従って，次元がmであるとは$H_m:\mathrm{rank}(\Omega)=m$，または，$\omega_1\geq\cdots\geq\omega_m>\omega_{m+1}=\cdots=\omega_p=0$ を意味している．

(2) 正準相関分析　p 次元変量 $\boldsymbol{x}=(\boldsymbol{x}_1',\boldsymbol{x}_2')'$ は $N_p(\mu,\Sigma)$ に従うとし，$\boldsymbol{x}$ の分割 $\boldsymbol{x}_1:p_1\times 1$, $\boldsymbol{x}_2:p_2\times 1$ に対応して，μ,Σ を分割する．2 組の変量 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 間の関係は正準相関変量で記述される．第 i 正準相関変量，第 i 正準相関係数をそれぞれ $\boldsymbol{y}:2\times 1$, ρ_i, $i=1,\cdots,q(=\min(p_1,p_2))$ とし，$I(\{\boldsymbol{y}_1,\cdots,\boldsymbol{y}_m\})=\rho_1^2+\cdots+\rho_m^2$ とする．このとき，正準相関変量の次元はゼロでない ρ_i の個数である．また，次元がmであるとは，$H_m:\mathrm{rank}(\Sigma_{12})=m$，または，$\rho_1^2\geq\cdots\geq\rho_m^2>\rho_{m+1}^2=\cdots=\rho_q^2=0$ を意味している．

(3) 主成分分析　p 次元変量 $\boldsymbol{x}$ は $N_p(\mu,\Sigma)$ に従うとする．$\boldsymbol{x}$ の変動を少数個の変換変量で説明するとき，主成分 $y_i=\gamma_i'\boldsymbol{x}$, $i=1,\cdots,q(=p)$ が用いられる．ここに，$\boldsymbol{\gamma}_i$ は Σ の第 i 固有値 λ_i に対応する長さ 1 の固有ベクトルである．$y_1,\cdots,y_m$ を用いたときの情報量は $I(\{y_1,\cdots,y_m\})=\lambda_1+\cdots+\lambda_m$ と定義する．この場合の次元としては，与えられた比率 $g(0<g<1)$ に対して

$$(\lambda_1+\cdots+\lambda_m)/(\lambda_1+\cdots+\lambda_p)\geq g$$

となる最小なmを考えるのが実際的である．判別分析の場合におけるように，$g=1$ とすると，次元は $\lambda_m>\lambda_{m+1}=\cdots=\lambda_p=0$ となるmとして定義される．この場合には，退化分布の問題が生じ，推測問題が困難になる．

上の次元の定義において，正規性の仮定は必要でないが，検定等を考えるとき必要となるのであらかじめ仮定している．次に，p 次元変量 $\boldsymbol{x}=(x_1,\cdots,x_p)'$ についての統計解析において，変量 $x_1,\cdots,x_p$ の選択を問題にする．冗長的変量を除くことは，推測の効率の向上，データ収集に関する負担の軽減，さらには，データ解釈の簡単化が期待される．定式化は，変換変量 $y_1,\cdots,y_q$ の選択の場合とほぼ同様であるが，異なる点は $x_1,\cdots,x_p$ の場合には重要性に関する順序がついていないことである．従って，この場合，$x_1,\cdots,x_p$ の全ての部分集合について検討する必要がある．記号の簡単化のため，変量の集合と対応する添字の集合を同一視する．例えば，$\{x_1,x_2\}\equiv\{1,2\}$．一般に，変量の部分集合 $\{j_1,\cdots,j_m\}$ の各々に対して統計解析の目的に対する自然な情報量 $I(\{j_1,\cdots,j_m\})$ が定められているとする．このとき，変量選択問題は制約条件

$$M(\{j_1,\cdots,j_m\}):\quad I(\{j_1,\cdots,j_m\})=I(\{1,\cdots,p\})$$

をみたす最小な変量の部分集合を決定することとして定式化される．推測に関連して，$M(\{j_1,\cdots,j_m\})$ のもとでの最大尤度が基本的となり，最大尤度が計算しやすい $M(\{j_1,\cdots,j_m\})$ の表示が問題となる．モデル $M(\{1,\cdots,m\})$ は，$\boldsymbol{x}_1=(x_1,\cdots,x_m)'$ だけで十分であり $\boldsymbol{x}_2=(x_{m+1},\cdots,x_p)'$ は冗長であることを意味しており，$\boldsymbol{x}_2$ の冗長性仮説ともよばれる．Rao (1970) は 2 群の判別分析における冗長性仮説に対して種々の同値命題を与えている．多群の判別解析に関しては，$I(\{1,\cdots,m\})=\mathrm{tr}\,\Sigma_{11}^{-1}\Omega_{11}$ と定義すると

$$M(\{1,\cdots,m\}):\quad \mathrm{tr}\,\Sigma_{11}^{-1}\Omega_{11}=\mathrm{tr}\,\Sigma^{-1}\Omega$$

で，これは"$\mu_{21}-\Sigma_{21}\Sigma_{11}^{-1}\mu_{11}=\cdots=\mu_{2k}-\Sigma_{21}\Sigma_{11}^{-1}\mu_{1k}$"と同値である．種々のモデルにおける冗長性については Fujikoshi (1989) に解説されている．

4.2 推　　測

正準相関分析の場合は判別分析の場合とほぼ平行的な結果が得られるので，以下では判別分析の場合について述べる．固定されたmに対し，次元に関する検定問題 $H_m:\mathrm{rank}(\Omega)=m$ vs

$K_m:\mathrm{rank}(\Omega)>m$ は古くから考えられている．$m=0$ のとき，水準 α の尤度比検定の棄却域は $|A|/|A+B|=\prod_{j=1}^{p}(1+l_j)^{-1}<c_0(\alpha)$ で与えられる．ここに，$l_1\geq\cdots\geq l_p$ は$A^{-1}B$ の固有値である．一般のmに対する尤度比検定の棄却域は $W_m=\prod_{m+1}^{p}(1+l_j)^{-1}<c_m(\alpha)$ である．$-n\log W_m$ は仮説 H_m のもとで漸近的に自由度 $f_m=(p-m)(k-1-m)$ のカイ二乗分布に従うが，その近似の改良として $\tilde{W}_m=-[n+(1/2)(k-p-2)+\Sigma_{j=1}^{m}l_j^{-1}]\log W_m$ が提案されている（Lawley (1959), Fujikoshi (1977), Glynn and Muirhead (1978), Chou and Muirhead (1979)）．$m\geq 1$ の場合には相似検定とならないため，$c(\alpha)$ を仮説 H_m のもので $\sup P(W_m<c(\alpha))\leq\alpha$ となるように定める必要があるが，最適な $c(\alpha)$ については Schott (1984) を参照されたい．次元mを推定する方法として，逐次検定法による方法とモデル選択基準を適用する方法が考えられている．伝統的に用いられている検定法とは，仮説$H_0, H_1,\cdots$をこの順に検定して，$H_0,\cdots,H_{m-1}$ が棄却され H_m が選択されれば，次元はmであると推定する方法である．検定の順序としてはこの逆も考えられる．検定法による問題点の１つは，各段階での棄却点は厳密には条件つきで評価して決められるべきものであるが，非常に複雑になるため条件を無視した評価がなされていることである．モデル選択基準を適用する方法とは，例えば AIC（Akaike (1973)）を用いて，$\min(A_0,\cdots,A_q)=A_m$ のとき，次元をmと推定する方法である（Fujikoshi and Veitch (1978)）．ここに，$q=\min(p,k-1), A_q\equiv 0$,

$$A_m=-n\log W_m-2(p-m)(k-1-m),\ m=0,\cdots,k-1.$$

一般に AIC 選択法は一致性をもっていない．一致性をもつようにするためには，例えば，Schwartz (1978) に従って，$\tilde{A}_m=-n\log W-(p-m)(k-1-m)\log n$ と修正すればよい．なお，変量選択問題に対しても同様なモデル選択基準が求められている（Fujikoshi (1985 b)）．

主成分分析における主成分の次元の推定法として，上記の定式化に基づく逐次検定法を適用できる．この他の方法として，交差検証法（cross-validation method）による方法も提案されている（Wold (1978), Krzanowski (1987)）．

5.　その他の話題

この節では，主として第１節で述べた研究課題 B 1〜B 13 に関する最近の進展を取り上げる．これらについては，総合的なものではなく，断片的な紹介である．

5.1　多変量非正規モデル

最近，非正規連続型多変量分布として Elliptically Contoured Distribution（楕円等高面分布，単に EC 分布という）にかなりの関心がよせられている．この分布は，p 次元変量 $\boldsymbol{x}$ の pdf が

$$|\Lambda|^{-1}g[(\boldsymbol{x}-\mu)'\Lambda^{-1}(\boldsymbol{x}-\mu)]$$

として与えられる分布であって，記号として $EC(\mu,\Lambda)$ が用いられている．関数 g を適当に選ぶことにより，特別な場合として p 次元正規分布，多変量 t-分布が得られる．正規分布は平均ベクトル μ と共分散行列 Λ で規定されるが，EC 分布には，この他に尖度パラメータ κ が含まれている．EC 分布の基本的な性質は Kelker (1970) によって展開された．また，1980 年以前の EC 分布に関する論文は，Chmielewski (1980) によってサーベイされている．正規性の場合と同じ，または，類似の結果が得られることから，1980 年代にも多くの論文が発表されている．これらについては，Fang and Anderson (1990), Fang and Zhang (1990) を参照されたい．EC 分布に関しては多くの理論が得られ，また，ロバストネスの研究にも利用されている（例えば，Kariya and Sinha (1988)）．しかし，連続型多変量データの誤差モデルとしての有用性に関しては，今後，理論面のみならず実際面からの検討が必要である．

多変量分布族の一つ構成法として，周辺分布が指定された分布になるような分布族が考えられている．また，分布の基本的性質を利用してこれらの分布族の比較も行われている．これらに関する最近の研究については，Joe (1993) を参照されたい．

5.2 ロバストネス

MANOVA モデルにおける代表的検定法 $T_1 \sim T_4$ (2.1 節を参照) においては，2つの基本的仮定；(i) 正規性，(ii) 共分散行列の均一性，が前提になっている．これらの仮定に関するロバストネスの研究は，基本的には，仮定がくずれることによる検定統計量の棄却点および検出力への影響を調べることである．1980 年までの研究については，Ito (1980) によってサーベイされている．仮定 (i)，(ii) を前提としない一般的な状況のもとでの検定統計量の分布は非常に複雑であるため，何等かの近似または数値実験を用いて調べられている．1980 年代には，非正規性モデルとして EC 分布に限定したロバストネスが研究されている．この場合には，null 分布が正規分布の場合と同一であるとか，また，ある種の最適性が保存される，等の明確な結果が得られている．これらについての結果は Kariya and Sinha (1988) に系統的にまとめられている．一般に，正規性の仮定のもとで提案された推測法が，EC 分布のもとでどのような影響を受けるかの研究は現在も進行中である．推測法のロバストネスを調べる他の方法としては，はずれ値による影響が調べられている．この研究は，各観測値が分析結果にどのように影響するかという感度分析に関係しており，これについては田中 (1992) のサーベイ論文を参照されたい．

ロバストネスに関する研究の他の方向として，ロバスト推測法を探る研究が行われている．平均ベクトル，共分散行列についての代表的なロバスト推定として，M-推定量が知られている (Maronna (1976))．回帰モデルにおける M-推定に関する文献は多い．この中で，Bai, Rao and Wu (1992) は多変回帰モデルの M-推定に関する一般的な結果を与えている．M-推定量を用いた各種推測が考えられるが (例えば，Campbell (1980, 1982))，これらの性質に関する研究は今後の課題と思われる．

5.3 多変量欠測値

欠測値がある場合の推測問題は，欠測値が単調的 (または階層的) である場合には比較的容易である．一般に，p 次元母集団からの標本 $\boldsymbol{x}_\alpha=(x_{\alpha 1},\cdots,x_{\alpha p})'$, $\alpha=1,\cdots,N$ は，$x_{\alpha i}$ が欠測値であるとき，すべての $x_{\beta j}$, $\beta \geq \alpha$, $j \geq i$ が欠測値になっているとき，単調的であるとよばれる．p 次元正規母集団 $N_p(\mu,\Sigma)$ からの標本であると，$\boldsymbol{x}_\alpha$ の pdf が条件つき pdf をを用いて

$$f(x_\alpha)=f(x_{\alpha 1})f(x_{\alpha 2}|x_{\alpha 1})\cdots f(x_{\alpha p}|x_{\alpha 1},\cdots,x_{\alpha,p-1})$$

と分解され，また，パラメータ $\boldsymbol{\mu},\Sigma$ も一意的に分解される．従って，(欠測値を除いた) 観測値データに基づく最尤推定量が具体的に求められる．欠測値が単調的である場合には，多くの推測問題に関して完全データの場合とほぼ平行的結果が得られている．これ等の文献は，Kariya, Krishnaiah and Rao (1983) に上げられている．彼等は，他の欠測値のタイプについても推測の漸近的性質を調べている．欠測値のタイプが一般な場合の最尤推定量を求める方法として，Dempster, Laird and Rubin (1977) による EM 計算法が知られている．EM 計算法とは，欠測値を Expectation によって推定し，次に仮の完全データを用いて尤度の Maximization を行う反復計算法であって，広く用いられている．他の方法として，尤度方程式を具体的に書き下す研究もなされている (Srivastava (1985))．欠測値が一般の場合には，パラメータの識別可能が保証されない等，厄介な問題が生じる．Anderson and Perlman (1991) は，$N_p(\boldsymbol{\mu},\Sigma)$ からの一般な欠測値をもつ標本に関しても，適当な条件付き独立性条件を導入する事によって単調的な場合と平行的な結果が得られることを示している．

5.4 高次元小標本問題

一般に，多変量推測法においては，標本の大きさが変量の次元より大であることが仮定されている．変量が多いと，この仮定が保証されない場合が生ずる．Dempster (1972) は $N_p(\boldsymbol{\mu}_i, \Sigma)$, $i=1,2$ からの大きさ N_i の標本に関して，$p>N_1+N_2-2$ の場合の検定法を考察している．この場合には，多変量 T^2 検定が利用できなく，直交変換に基づく検定法が提案されている．

Friedman (1989) は，k の母集団Π_i, $i=1,\cdots,k$ の判別問題において高次元小標本および共分散行列の非同等性を考慮した正則化判別法を提案している．第 i 母集団からの大きさ N_i の標本に基づく標本共分散行列を S_i とし，合併不偏推定を S とする．このとき，第 i 母集団の共分散行列 Σ_i を

$$\hat{\Sigma}_i(\lambda, \gamma)=(1-\gamma)\hat{\Sigma}_i(\lambda)+\gamma p^{-1}\operatorname{tr}\hat{\Sigma}(\lambda)I_p$$

で推定している．ただし，$\hat{\Sigma}(\lambda)=\{(1-\lambda)(N_i-1)S_i+\lambda(N-k)S\}/\{(1-\lambda)(N_i-1)+\lambda(N-k)\}$．パラメータ $\lambda, \gamma(0\leq\lambda\leq1, 0\leq\gamma\leq1)$ は誤判別率が小さくなるように決定される．この方法は膨大な計算を用し，計算機を利用することによって実行される統計的方法である．この方法の性質については，あまり知られていない．判別されるべき観測値 $\boldsymbol{x}$ が Π_i に属するときの群内変動行列を W_i とするとき，Saranadasa (1993) は W_i のノルムを利用したいくつかの判別法を提案している．

5.5 多変量回帰の逆推定

回帰分析では説明変量 x が与えられたとき，反応変量 y についての推測が問題となる．これとは逆に，y の値が観測されたとき，x についての推測を行うことを回帰の逆推定といい，古くから研究されている．p 個の反応変量 $y_1,\cdots,y_p$ と q 個の説明変量 $x_1,\cdots x_q$ に関する多変量線形回帰モデルにおける逆推定問題の研究は，Brown (1982) の論文が契機になっている．事前データと $(y_1,\cdots y_p)$ に基づいて，$(x_1,\cdots,x_q)$ を推定する種々の方法が提案されている．1 次元の場合には古典的推定量の平均が有限でないという問題点が生ずるが，多変量の場合にはこの点が解消される．すなわち，$p\geq q+1$ であると，古典的推定量の平均が有限になり，また，$p\geq q+2$ であると，その平均 2 乗誤差が有限になる（Nishii and Krishnaiah (1988)）．一方，$p=q$ の場合には，古典的推定量は最尤推定量であるが，$p>q$ のときは最尤推定量とはならない．さらに，推定量の一致性が保証されなく，また，信頼領域が必ずしも閉領域にならない等，困難な問題点を含んでいる．古典的推定量の他に，逆推定法，ベイズ推定法，リッヂ回帰推定法，等の種々の推定法が提案されているが，これらについては Osborne (1991) のサーベイ論文を参照されたい．

5.6 方向性データの解析

方向性データの解析とは，標本空間が円周または球面である場合の統計解析である．確率モデルとしては，指数分布族のいくつかが考えられているが，その中で次の *pdf* をもつ von Mises-Fisher（または Langevin）分布が代表的である．

$$f(\boldsymbol{x}\,;\boldsymbol{\mu}, \kappa)=a(\kappa)^{-1}\exp\{\kappa\boldsymbol{\mu}'\boldsymbol{x}\}$$

ただし，$\kappa>0, \boldsymbol{\mu}'\boldsymbol{\mu}=1, \boldsymbol{x}'\boldsymbol{x}=1$．この方面のサーベイ論文としては，Mardia (1975), Jupp and Mardia (1988) がある．後者の論文においては，1975 年以降の進展が種々の観点から整理されている．取り上げられている項目は，確率モデル，推測，漸近・近似的結果，分布の特徴，はずれ値・適合性検定，ロバストネス，シュミレーション，ノンパラメトリック・ブートストラップ法，密度推定，曲線適合，非標準的応用，探索的データ解析，その他の標本空間，等多岐にわたっている．最近，Langevin 分布のもとでの検定問題に対して，高次漸近理論が展開され

(Hayakawa (1991), Watamori (1992) 等), また, 局所最適検定に関する研究もなされている (SenGupta and Jammalamadaka (1991)). さらに, Stiefel および Grassmann manifold を標本空間とする問題も研究され, 関連する統計量の漸近展開も導出されている(例えば, Chikuse (1991)). しかし, 標本空間が, 通常のユークリッド空間や離散空間と異なるため, 推測の最適性に関する研究は十分なものとはいえない.

6. さいごに

多変量推測法の発展は, 第1節で述べたように, 例えば, A 1～A 5, B 1～B 13 の如く非常に多岐にわたっており, また, これらの多くについて最近著しい発展がみられる. しかし, 本論文において取り上げられている内容は, これら全般にわたるものではなくかなり限定されたものになっていることをことわっておきたい. なお, 取り上げられていない内容のうちのいくつかは, 別稿および第II部で解説されており, これらを含めると, 多変量推測法のかなりの部分がカバーされることになる. 共分散構造および線形関係モデルについては, 別稿「潜在構造モデル」(浅野長一郎)として統一的に扱われている. 平均ベクトルと共分散行列の推定に関しては, 最近, Stein 推定量の一般化である縮小推定量が構成されている. この発展に関しては別稿「統計的推定論の最近の展開」(久保川達也) で解説されている. なお, 固有値の推定に関しても同様な発展があるが, これについては Muirhead(1987)のサーベイ論文がある. 多変量順序制約下での検定問題に関しては別稿「検定論の最近の展開」(竹村彰通) に解説されている. 判別分析においても多くの進展がみられるが, これらについては McLachlan (1992) による研究書を参照されたい. 最近, 誤判別確率のノンパラメトリック推定法にかなりの関心がよせられている. これは, Efron (1983, 1986) のブートストラップ法に関連するものであって, これについては McLachlan (1987), 別稿「ブートストラップ法とその応用」(小西貞則), 小西・本多 (1992) を参照されたい. 因子分析モデルの最近の発展については, 本稿の第II部および狩野 (1991) のサーベイ論文を参照されたい.

これまでにみてきたように, 多変量推測法は, 最近著しく発展してきている. しかし, これらの大部分は, 多変量正規モデルと線形構造に依存するものである. 従って, 多変量非正規モデル, 非線形構造, および, 実際の多変量データ解析と理論との乖離等に焦点をあてると, 多くの問題点が残されているといえる. また, 計算機の急速な発展と関連して, 計算機を駆使した多変量解析手法, 例えば, 射影追跡法, ブートストラップ法, グラフィカル表現法が発展してきている. これらについてもより厳密な理論展開が必要である.

謝辞：審査委員からは最初の原稿に対して多くの有益なコメントを頂きました. ここに記して謝意を表します.

参考文献

[1] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, Second International Symposium on Information Theory (B. N. Petrov and F. Cráki, eds), Akademiai Kiakó, 267-281.

[2] Anderson, B. M., Anderson, T. W. and Olkin, I. (1986). Maximum likelihood ratio criteria in multivariate components of variance, *Ann. Statist.*, **14**, 405-417.

[3] Anderson, S. A. and Perlman, M. D. (1991). Lattice-ordered conditional independence models for missing data, *Statistics & Probability Letters*, **12**, 465-486.

[4] Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis, John Wiley & Sons.

[5] Anderson, T. W. (1984). An Introduction to Multivariate Statistical Analysis (2nd ed.), John

Wiley & Sons.

[6] Anderson, T. W. (1989). The asymptotic distribution of the likelihood ratio criterion for testing rank in multivariate components of variance, *J. Multivariate Anal.*, **30**, 72-79.

[7] Anderson, T. W. and Amemiya, Y. (1991). Testing dimensionality in the multivariate analysis of variance, *Statistics & Probability Letters*, **12**, 445-463.

[8] Anderson, T. W. and Rubin, H. (1956). Statistical inference in factor analysis, Proc. Third Berkeley Symp. on Math. Statist. & Prob. Vol. V, Univ. of California, 115-150.

[9] Anderson, T. W. and Takemura, A. (1982). A new proof of admissibility of tests in multivariate analysis, *J. Multivariate Anal.*, **12**, 457-468.

[10] Bai, Z. D. and Rao, C. R. (1991). Edgeworth expansion of a function of sample means, *Ann. Statist.* **19**, 1295-1315.

[11] Bai, Z. D., Rao, C. R. and Wu, Y. (1992). M-estimation of multivariate linear regression parameters under a convex discrepancy function, *Statistica Sinica,* **2**, 237-254.

[12] Bhattacharya, R. N. (1985). Some recent results on Cramḗr-Edgeworth expansions with applications, Multivariate Analysis-VI (P. R. Krishnaiah, ed)., Elsevier Science Publishers B. V., 57-75.

[13] Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion, *Ann. Statist.,* **6**, 434-451.

[14] Boik, R. J. (1988). The mixed model for multivariate repeated measures : Validity conditions and an approximate test, *Psychometrika,* **53**, 469-486.

[15] Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria, *Biometrika,* **36**, 317-346.

[16] Brown, P. J. (1982). Multivariate Calibration (with discussion), *J. R. Statist. Soc. B,* **44**, 287-321.

[17] Campbell, N. A. (1980). Robust procedures in multivariate analysis I : Robust covariance estimation, *Appl. Statist.,* **29**, 231-237.

[18] Campbell, N. A. (1982). Robust procedures in multivariate analysis II : Robust canonical analysis, *Appl. Statist.*, **31**, 1-8.

[19] Chandra, T. K. and Ghosh, J. K. (1979). Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square variables, *Sankhayā Ser. A,* **41**, 22-47.

[20] Channra, T. K. and Ghosh, J. K. (1980). Valid asymptotic expansions for the likelihood ratio and other statistic under contiquous alternatives, *Sankhyā Ser. A,* **42**, 170-184.

[21] Chikuse, Y. (1980). Invariant polynomials with matrix arguments and their applications, Multivariate Analysis (R. P. Gupta, ed.), North-Holland, 53-68.

[22] Chikuse, Y. (1991). Asymptotic expansions for distributions of the large sample matrix resultant and related statistics on the Stiefel manifold, *J. Multivariate Anal.,* **39**, 270-283.

[23] Chikuse, Y. and Davis, A. W. (1986). A survey on the invariant polynomials with matrix arguments in relation to econometric distribution theory, *Econometric Theory,* **2**, 232-248.

[24] Chmielewski, M. A. (1981). Elliptically symmetric distributions : A review and bibliography, *Intern. Statist. Rev.,* **49**, 67-74.

[25] Chou, R.-J. and Muirhead, R. J. (1979). On some distribution problems in manova and discriminant analysis, *J. Multivariate Anal.,* **9**, 410-419.

[26] Constantine, A. G. (1963). Some noncentral distribution problems in multivariate analysis, *Ann. Statist. Math.,* **34**, 1270-1285.

[27] Crowder, M. J. and Hand, D. I. (1990). Analysis of Repeated Measures, Chapman and Hall.

[28] Das Gupta, S., Anderson, T. W. and Mudholkar, G. S. (1964). Monotonicity of power functions of some tests of the multivariate linear hypothesis, *Ann. Math. Statist.,* **35**, 200-205.

[29] Davis, A. W. (1979). Invariant polynomials with two matrix arguments extending the zonal polynomials : Applications to multivariate distribution theory, *Ann. Ints. Statist. Math.,* **31A**, 465-485.

[30] Davis, A. W. (1980). Invariant polynomials with matrix arguments, extending the zonal polynomials, Multivariate Analysis-V (P. R. Krishnaiah, ed.), North-Holland, 287-299.

[31] Dempster, A. P. (1960). A significance test for the separation of two highly multivariate small samples, *Biometrics,* **16**, 41-50.

[32] Dempster, A. P. (1971). An overview of multivariate data analysis. *J. Multivariate Anal.,* **1**, 316-

346.
[33] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc., Ser. B* **39**, 29-38.
[34] Dillon, W. R. and Goldstein, M. (1984). Multivariate Analysis: Methods and Applications, John Wiley & Sons.
[35] Efron, B. (1979). Bootstrap method: Another look at the jackknife, *Ann. Statist.*, **7**, 1-26.
[36] Fang, K. T. and Zhang, Y. T. (1990). Generalized Multivariate Analysis, Springer-Verlag.
[37] Fang, K. T. and Anderson, T. W. (eds.) (1990). Statistical Inference in Elliptically Contoured and Related Distributions, Allerton Press Inc.
[38] Fisher, R. A. 1915). Frequency distribution of the values of the correlation coefficient in samples from an infinitely large population, *Biometrika,* **10**, 507-521.
[39] Fisher, R. A. (1928). The general sampling distribution of the multiple correlation coefficient, *Proc. Roy. Soc. London A,* **121**, 654-673.
[40] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems, *Ann. Eugen.,* **7**, 179-188.
[41] Fisher, R. A. (1939). The sampling distribution of some statistics obtained from non-linear equations, *Ann. Eugen.,* **9**, 238-249.
[42] Friedman, J. H. (1989). Regularized discriminant analysis, *J. Amer. Statist. Assoc.,* **84**, 165-175.
[43] Fujikoshi, Y. (1977). Asymptotic expasions for the distributions of some multivariate tests, Multivariate Analysis-IV (P. R. Krishnaiah, ed.), North-Holland, 55-71.
[44] Fujikoshi, Y. (1980). Asymptotic expansions for the distributions of the sample roots under nonnormality, *Biometrika,* **67**, 45-51.
[45] Fujikoshi, Y. (1985a). An error bound for an asymptotic expansion of the distribution function of an estimate in a multivariate linear model, *Ann. Statist.,* **13**, 827-831.
[46] Fujikoshi, Y. (1985b). Selection of variables in discriminant analysis and canonical correlation analysis, Multivariate Analysis-VI (P. R. Krishnaiah, ed.), Elsevier Publishing Publisheres B. V., 219-236.
[47] Fujikoshi, Y. (1988). Comparison of powers of a class of tests for multivariate linear hypothesis and independence, *J. Multivariate Anal.,* **9**, 531-544.
[48] Fujikoshi, Y. (1989). Tests for redundancy of some variables in multivariate analysis, Recent Developments in Statistical Data Analysis and Inference (Y. Dodge, ed.), Elesevier Science Publishers B. V., 141-163.
[49] Fujikoshi, Y. (1993). Error bounds for asymptotic approximations of some distribution functions, to appear in Multivariate Analysis: Future Diections (C. R. Rao, ed.), Elesevier Science Publishers B. V.
[50] Fujikoshi, Y. and Veitch, L. G. (1978). Estimation of dimensionality in canonical correlation analysis, *Biometrika*, **66**, 345-351.
[51] 藤越康祝・清水良一 (1988). ある種の確率分布の漸近展開とその誤差限界—独立確率変数和と尺度混合変数の分布—, 数学, **40**, 28-44.
[52] Gleser, L. J. and Olkin, I. (1970). Linear models in multivariate analysis, Essay in Prob. Statist. (R. C. Bose et al., eds.), Univ. of North Carolina Press, 267-292.
[53] Glynn, W. J. and Muirhead, R. J. (1978). Inference in canonical correlation analysis, *J. Multivariate Anal.,* 8, 468-478.
[54] Hayakawa, T. (1989). Invariant polynomials and related tests, *Amer. J. Math. Management Sci.,* **9**, 69-99.
[55] Hayakawa, T. (1990). On tests for the mean direction of Langevin distribution, *Ann. Inst. Statist. Math.,* **42**, 359-373.
[56] Hotelling, H. (1931). The generalization of Student's ratio, *Ann, Math. Statist.,* **2**, 360-348.
[57] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components, *J. Educational Psychology,* **26**, 139-142.
[58] Hotelling, H. (1936). Relations between two sets of variates, *Biometrika,* **28**, 321-377.
[59] Hsu, P. L. (1939). On the distribution of the roots of certain determinantal equations, *Ann. Eugen.,* **9**, 250-258.

[60] Huber, P. J. (1985). Projection pursuit, *Ann. Statist.*, **13**, 435-475.

[61] Ito, P. K. (1980). Robustness of ANOVA and MANOVA test procedures, Handbook of Statistics -1 (P. R. Krishaiah, ed.), North-Holland Publishing Company, 199-236.

[62] James, A. T. (1960). The distribution of the latent roots of the covariance matrix, *Ann. Math. Statist.*, **31**, 151-158.

[63] James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples, *Ann. Math. Statist.*, **35**, 475-501.

[64] Joe, H. (1993). Parametric families of multivariate distributions with given margins, to appear in *J. Multivariate Anal.*

[65] Jupp, P. E. and Mardia, K. V. (1989). A unified view of the theory of directional statistics, 1975 -1988, *Intern. Statist. Rev.*, **57**, 261-294.

[66] 狩野 裕 (1990). 因子分析における統計的推測：最近の発展, **行動計量学**, **18**, 3-12.

[67] Kariya, T. (1978). The general MANOVA problem, *Ann. Statist.*, **6**, 200-214.

[68] Kariya, T. (1985). Testing in the Multivariate General Linear Model, Kinokunia.

[69] Kariya, T. and Maekawa, K. (1982). A method for approximation to the pdf's of GLSE's and its application to the seemingly unrelated regression model, *Ann. Inst. Statist. Math.*, **34**A, 281-297.

[70] Kariya, T., Krishnaiah, P. R. and Rao, C. R. (1983). Inference on parameters of multivariate normal populations when some data is missing, *Developments in Statistics*-**4** (P. R. Krishnaiah, ed.), Academic Press, 137-184.

[71] Kelker, D. (1970). Distribution theory of spherical distributions and a location scale parameter generalization, *Sankhyā A*, **32**, 419-430.

[72] Khatri, C. G. (1966). A note on a MANOVA model applied to problems in growth curve, *Ann. Inst. Statist. Math.*, **18**, 75-86.

[73] Konishi, S. (1979). Asymptotic expansions for the distributions of statistics based on the sample correlation matrix in principal component analysis, *Hiroshima Math. J.*, **9**, 647-700.

[74] 小西貞則・本多正幸(1992). 判別分析における誤判別率推定とブートストラップ法, **応用統計学**, **21**, 67 -100.

[75] Konishi, S. (1987). Transformations of statistics in multivariate analysis, Advances in Multivariate Statistical Analysis (A. K. Gupta, ed.), *D. Reidel Publishing Company*, 213-231.

[76] Krishnaiah, P. R. (1978). Some recent developments on real multivariate distributions, Development in Statistics-1 (P. R. Krishnaiah, ed.), Academic Press, 35-169.

[77] Krishnaiah, P. R. (1980). Computation of some multivariate distributions, Handbook of Statistics (P. K. Krishnaiah, ed.), North-Holland Publisher Company, 745-991.

[78] Krzanowski, W. J. (1987). Cross-validiation in principal component analysis, *Biometrics*, **43**, 575-584.

[79] Kshirsagar, A. M. (1972). Multivariate Analysis, Marcel Dekker.

[80] Kshirsagar, A. M. and Gupta, R. P. (1990). Direction and collinearity factors of Wilks's Λ—A review, *Commun. Statist.-Theory Meth.*, **19**, 1359-1402.

[81] Kuriki, S. (1993). Likelihood ratio tests for covariance structure in random effects models, to appear in *J. Multivariate Anal*.

[82] Laird, N. M. and Ware, J. H. (1982). Random-effects models for londitudinal data, *Biometrics*, **38**, 963-974.

[83] Lawley, D. N. (1959). Tests of significance in canonical analysis, *Biometrika*, **46**, 59-66.

[84] Lindstrom, M. J. and Bates, D. M. (1988). Newton-Raphson and EM algorithms for linear-effects models for repeated-measures data, *J. Amer. Statist. Assoc.*, **83**, 1014-1022.

[85] Maronna, R. A. (1976). Robust M-estimators of multivariate location and scatter, *Ann. Statist.*, **4**, 51-67.

[86] Mathai, A. M. (1973). A review of the different techniques used for deriving the exact distributions of multivariate test criteria, *Sankhyā A*, **35**, 39-60.

[87] McLachlan, G. J. (1987). Error rate estimation in discriminant analysis: recent advances, Advances in Multivariate Statistical Analysis (A. K. Gupta, ed.), D. Reidel Publishing Company, 233-252.

[88] McLachlan, G. J. (1992). Discriminant Analysis and Statistical Patter Reognition, John Wiley &

Sons.

[89] Muirhead, R. J. (1978). Latent roots of matrix variates: A review of some asymptotic results, *Ann. Statist.,* **6**, 5-33.

[90] Muirhead, R. J. (1982). Aspects of Multivariate Statistical Theory, John Wiley & Sons.

[91] Muirhead, R. J. (1987). Developments in eigenvalue estimation, Advances in Multivariate Statistical Analysis (A. K. Gupta, ed.), D. Reidel Publishing Company, 277-288.

[92] Nagao, H. (1989). Tests of hypotheses for covariance matrices and distributions under multivariate populations, *Amer. J. Math. Management. Sci.,* **9**, 31-49.

[93] Nishii, R. and Krishnaiah, P. R. (1988). On the moments of classical estimates of explanatory variables under a multivariate calibration model, *Sankhyā,* **50**, 137-148.

[94] Osborne, C. (1991). Statistical Calibration: A review, *Intern. Statist. Reviv.,* **49**, 309-336.

[95] Perlman, M. D. and Olkin, I. (1980). Unbiasedness of invariant tests for MANOVA and other multivariate problems, *Ann. Statist.,* 8, 1326-1641.

[96] Pillai, K. C. S. (1976). Distributions of the characteristic roots multivariate analysis: Part I, Null distribution, *Canad. J. Statist. Sec. A & B,* **4**, 157-184.

[97] Pillai, K. C. S. (1977). Distributions of the characteristic roots in multivariate analysis: Part II, Non-null distributions, *Canad. J. Statist. Sec. A & B,* **5**, 1-62.

[98] Pillai, K. C. S. and Jayachandran, K. (1967). Power comparisons of tests of two multivariate hypotheses based on four criteria, *Biometrika,* **54**, 195-210.

[99] Potthoff, R. F. and Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems, *Biometrika,* **51**, 313-326.

[100] Rao, C. R. (1946). Tests with discriminant functions in multivariate analysis, *Sankhyā,* **7**, 407-414.

[101] Rao, C. R. (1948). Tests of significance in multivariate analysis, *Biometrika,* **35**, 58-79.

[102] Rao, C. R. (1959). Some problems involving linear hypotheses in multivariate analysis, *Biometrika,* **46**, 49-58.

[103] Rao, C. R. (1965). Theory of least squares when the parameters are stochastic and its application to the analysis of growth curves, *Biometrika,* **52**, 447-458.

[104] Rao, C. R. (1970). Inference on discriminant function coefficients, Essays in Probability and Statistics (R. C. Bose et al., eds.), Univ. of North Carolina Press, 587-602.

[105] Rao, C. R. (1972). Recent trends of research work in multivariate analysis, *Biometrics,* **28**, 3-22.

[106] Rao, C. R. (1983). Multivariate Analysis: Some reminiscences on its origin and development, *Sankhyā B,* **45**, 284-299.

[107] Rao, C. R. (1987a). Prediction of future observations in growth curve models, *Statist. Sci.,* **2**, 434-471.

[108] Rao, C. R. (1987b). Estimation in linear models with mixed effects: a unified theory, Proc. Second International Tampere Conference in Statistics (T. Pukkila and S. Puntanan, eds.), 73-98, Univ. of Tampere.

[109] Rathie, P. N. (1989). Generalized hypergeometric functions and exact distributions of test statistics, *Amer. J. Math. Management Sci.,* **9**, 155-172.

[110] Reinsel, G. (1982). Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance structure, *J. Amer. Statist. Assoc.,* **77**, 190-195.

[111] Regnsel, G. (1984). Estimation and prediction in a multivariate random effects generalized linear model, *J. Amer. Statist. Assoc.,* **79**, 406-414.

[112] Roy, S. N. (1939). p-statistics or some generalizations in analysis of variance appropriate to multivariate problems, *Sankhyā,* **4**, 381-396.

[113] Schott, J. R. (1984). Optimum bounds for the distributions of some test criteria for tests of dimensionality, *Biometrika,* **71**, 561-567.

[114] Saranadasa, H. (1993). Asymptotic expansion of the misclassification probabilities of D-and A-criteria for discrimination from two high dimensional populations using the theory of large dimensional random matrices, to appear in *J. Multivariate Anal.*

[115] Schervish, J. (1987). A review of multivariate analysis, *Statist. Sci.,* **2**, 396-433.

[116] Schwarz, G. (1978). Estimating the dimension of a model, *Ann. Statist.,* **6**, 461-464.

[117] Sen, P. K. (1986). Contemporary textbooks on multivariate statistical analysis: a panaromic

appraisal and critique, *J. Amer. Statist. Assoc.*, **81**, 560-564.
[118] SenGupta, A. and Jammalamadaka, S. R. (1991). On locally optimal tests for the mean direction of the Langevin distribution, *Statistics & Probability Letters*, **12**, 537-544.
[119] Shimizu, R. (1987). Error bounds for asymptotic expansion of the scale mixtures of the normal distributions, *Ann. Inst. Statist. Math.*, **39**, 611-622.
[120] Sibson, R. (1984). Present position and potential developments: Some personal views multivariate analysis, *J. R. Statist. Soc. A*, **137**, 198-207.
[121] 塩谷 実 (1975). 多変量解析における検定統計量の一般分布にたいする漸近展開—最近の発展にみられる諸方法の概説 I—, 日本統計学会誌, **5**, 87-104.
[122] 塩谷 実 (1976). 多変量解析における検定統計量の一般分布にたいする漸近展開—最近の発展にみられる諸方法の概説 II—, 日本統計学会誌, **6**, 63-87.
[123] 塩谷 実 (1979). 統計的多変量解析の最近の動向, システムと制御, **23**, 203-210.
[124] Siotani, M. (1989). Distributions of Lawley-Hotelling's T_0^2 and related statistics: A review, *Amer. J. Math. Management Sci.*, **9**, 5-30.
[125] Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985). Modern Multivariate Statistical Analysis, American Science Press.
[126] Srivastava, M. S. (1985). Multivariate data with missing observations, *Commun. Statist.-Theor. Meth.*, **14**, 775-792.
[127] Sugiura, N. (1976). Asymptotic expansions of the distributions of the latent roots and the latent vector of the Wishart and Multivariate F matrices, *J. Multivariate Anal.*, **6**, 500-525.
[128] 竹村彰通 (1991). 多変量推測統計の基礎, 共立出版.
[129] 田中 豊 (1992). 多変量解析における感度分析, 行動計量学, **19**, 3-17.
[130] von Rosen, D. (1991). The growth curve model: a review, *Commun. Statist.-Theor. Meth.*, **20**, 2791-2822.
[131] Vonesh, E. F. and Carter, R. L. (1987). Efficient inference for random-coefficient growth curve models with unbalanced data, *Biometrics*, **43**, 617-628.
[132] Watamori, Y. (1992). Tests for a given structure of the mean direction of the Langevin distribution, *Ann. Inst. Statist. Math.*, **44**, 147-156.
[133] Wilks, S. S. (1932). Certain generalizations in the analysis of variance, *Biometrika*, **24**, 471-494.
[134] Wishart, J. (1928). The generalized product moment distribution in samples from a normal multivariate population, *Biometrika*, **20**A, 32-52.
[135] Wold, S. (1978). Cross-validiation estimation of the number of components in factor and principal components models, *Technometrics*, **20**, 397-407.

第 II 部：多変量データ解析法

柳井晴夫

1. はじめに

多変量データ解析 Multivariate Data Analysis とは事象そのもの，またはその事象の背後にあると想定される要因の多元的測定から，（i）事象の簡潔な記述と情報の圧縮（次元の縮小），（ii）事象の背後にある潜在因子の探索，（iii）事象に対する複数な要因の総合化，等を目的とする統計的手法の総称といえよう．なお，分析に用いるデータは必ずしも量的データである必要はなく，この意味で（iv）質的データの数量化に関する手法も多変量データ解析において主要な役割を果たすものである．さらに，同一個体の同一変量に関する測定が時間を追って複数回行われる場合には，それらは3重配列の多変量データとなる．このような，（v）多重配列データの解析法に関する研究が最近急激に増加している．

本稿では多変量データ解析の手法を，次元縮小の方法，質的データの解析法，多重配列デー

タの解析法，に分類し，それぞれに含まれる手法についてできる限り多数の理論的，および応用的研究を取り上げる．国内の雑誌では主に，「日本統計学会誌（J. of Japan Statistical Society）」,「応用統計学」,「行動計量学」,「Behaviormetrika」, 国外の雑誌では多変量データ解析に関する論文が比較的多く掲載されている「Psychometrika」を中心に，その研究動向を探っていくことにしたい．その際，可能なかぎり，我が国の研究者による研究を中心に紹介することにする．なお，本節に示した文献のうち，著書に関するものは第III部の文献を参照して頂きたい．

2. 次元の縮小の方法

本節では，Seber (1984) に従い，主成分分析，多次元尺度法，正準相関分析，判別分析，因子分析を多変量データ解析における次元縮小の方法と捉えて，それぞれの手法の現状と展望を解説する．

2.1 特異値分解

縦に n 個の個体，横に p 変量を割り当てた $n\times p$ 行列，より詳しくいえば，p 個の変量 $x_1,\cdots,x_p$ に対する n 個の個体の測定値を成分とする n 次元ベクトル，$x_1, x_2,\cdots, x_p$ によって記述される $n\times p$ 行列

(1) $$X=[x_1, x_2,\cdots, x_p]$$

を多変量データ行列という。主成分分析は多変量データ解析における最も基本的な次元縮小の方法で，多変量解析の全般的解析書には必ずといってよいほどその記述に多くのページが割かれている．最近になっても，主成分分析に的を絞った著書（Jollife (1986), Jackson (1991)）が出版されており、この意味でも主成分分析は今日においても多変量データ解析における主要な手法といえよう．

(1)式で定義された $n\times p$ 行列 X の列ベクトルで生成される部分空間を $W=S(X)$，さらに，$r=\mathrm{rank}(X)=\dim W$ とするとき，$X'Xg_i=\lambda_i g_i, f_i=Xg_i/\sqrt{\lambda_i}(i=1, r)$ を満たすノルム 1 の p 次元ベクトル g_i と n 次元ベクトル f_i によって X を

(2 a) $$X=\mu_1 f_1 g_1'+\cdots+\mu_r f_r g_r'=F\varDelta G'$$
$$(\text{ただし，} F'F=I_r, G'G=I_r, \mu_i=\sqrt{\lambda_i}, \varDelta=\mathrm{diag}(\mu_1,\ldots, \mu_r))$$

のように分解することを，特異値分解（singular value decomposition)，またはエッカート・ヤング分解（Eckart-Young decomposition）という．μ_j は行列 X の j 番目の特異値で $\mu_j=\mu_j(X)$ と書かれることがある．

なお，特異値分解に関する最適性に関し，$n\times p$ 行列 Y の rank を $s(s\leqq r)$ としたとき，Rao (1979, 1980) により次の性質が導かれた．すなわち，

(2 b) $$\mu_i(X-Y)\geqq\mu_{i+s}(X), i+s\leqq r, \mu_i(X-Y)\geqq 0, i+s>r$$

である．上式における等号成立のための必要十分条件は，X が (2 a) のように分解される場合，Y が，$Y=\mu_1 f_1 g_1'+\cdots+\mu_s f_s g_s'$，となることである．なお，上記のベクトル f_1, f_2，およびベクトル $\mu_1 g_1, \mu_2 g_2$ の各要素からなる $(n+p)$ 個の点を 2 次元上に布置し，これらの各点に対応する事象間の相互関連を容易に把握可能なものにするための手法に Biplot (Gabriel (1971)) がある．

ところで，$n\times p$ 行列 X の他に，n 次，および p 次の正定符号行列A，Bが与えられている場合，$Z=A^{1/2}XB^{1/2}$ の特異値分解は $Z=F\varDelta G'$（ただし，$F'F=I, G'G=I, \varDelta=\mathrm{diag}(\mu_1,\cdots, \mu_r)$, μ_j は Z の特異値）となる．このとき，$U=A^{-1/2}F, V=B^{-1/2}G$ とおくことによって，次の分解が

得られる.

$$X=U\Delta V'(\text{ただし，} U'AU=I,\ V'BV=I) \tag{3}$$

上記の分解は一般化特異値分解（generalized singular value decomposition, Greenacre (1984)))と呼ばれる．特異値分解，一般化特異値分解は，次節以降に示す多変量データ解析の各種手法における数学的な基本原理となるものである．

2.2 主成分分析

主成分分析の数学的原理は多変量データ行列Xの特異値分解に他ならない．より詳しく説明すれば，(2 a) 式におけるf_jは行列Xの第j番目の主成分，$\mu_j g_j$の各成分は行列Xに含まれるp個の変数ベクトルx_iのベクトルf_jへの正射影の長さに相当するもので，主成分負荷量と呼ばれる．ところで，多変量データ行列Xは次のように基準化することができる．

(4 a)　$X_1=Q_nX$　（ただし，$Q_n=I_n-(1/n)1_n1'_n$ は $Q_n^2=Q_n, (Q_n)'=Q_n$, を満たす直交射影行列）

(4 b)　$X_2=XQ_p$　（$Q_p=I_p-(1/p)1_p1'_p, Q_p^2=Q_p$ を満たす）

(4 c)　$X_3=Q_nXQ_p$

通常の主成分分析は（4 a）のように基準化されたデータ行列が用いられる．Gollob (1968) は（4 c）のように基準化された行列の主成分分析を提案した．また，Okamoto (1972) は4種のデータ行列 X_1, X_2, X_3 および，$X_4=X$（素得点行列そのもの）に基づく主成分分析を整理している．なお，実験計画における交互作用要素の処理に関連して，宮川 (1992) は，(4 c) の方法に基づく主成分分析の有用性を指摘している．

主成分分析の対象となる多変量データの各個体について性別，年齢別，出身地域別のような外的基準の情報がある場合を，これらを $n\times m$ 行列Gで表わそう．このとき，直交射影行列 $P_G=G(G'G)^{-1}G', Q_G=I_n-P_G$, の導入により，データ行列$X$，およびその共分散行列 $S=X'X$ は，

$$X=P_GX+Q_GX,\ X'X=X'P_GX+X'Q_GX \tag{5}$$

と分解される．このとき，$X'P_GX$ からは，Gに関連したXの主成分，$X'Q_GX$ からはGの影響を除去した主成分が抽出される．このように、外的基準の情報Gを含めて主成分分析する考え方は，最初 Tukey (1962) によって示唆され，Rao (1964) によって定式化された．これらの適用例には，Yanai (1970)，橋口(1978)がある．ところで，上記の射影行列 P_G および Q_G は，Xの列ベクトルを空間 $S(G)$ とその直交補空間 $S(G)^\perp$ に射影するものであったが，Takane & Shibayama (1990) はXの行ベクトルを $p\times t$ 行列Hによって生成される部分空間 $S(H)$，およびその直交補空間 $S(H)^\perp$ に射影する直交射影行列 P_H, Q_H の導入により，データ行列Xを

$$\begin{aligned} X&=(P_G+Q_G)X(P_H+Q_H)\\ &=P_GXP_H+P_GXQ_H+Q_GXP_H+Q_GXQ_H \end{aligned} \tag{6}$$

と4つの要素に分解しそれぞれの項を多変量データ行列として主成分分析する方法を提案した。高根 (1992) はこれらを含む一般化された主成分分析の方法を制約付き主成分分析（constrained principal component analysis）と呼んでいる．

主成分分析の一般化としては，$b_1x+b_2x^2+b_3y+b_4y^2+b_5xy=c$ といった曲線を2次元平面

上に定めるGnanadesikan (1977) によって提唱された一般化主成分分析 (generalized principal component analysis) があり，この方法の性質がMizuta (1984) によって調べられている．上記の方法はさらに拡張され，多項式以外のより一般的な曲線をあてはめるPrincipal curveとよばれる方法がHastie & Stuetzle (1989) によって提唱されている。

主成分分析の拡張として，多変量データXをそのままのデータとして分析にかけるのでなく，データ行列Xの最適変換と主成分Fの導出といった過程を交互に繰り返すことによって主成分分析を行う非計量主成分分析 (non-metric principal component analysis) 手法であるプリンシパルス (PRINCIPALS) (Takane et. al (1978)) が開発されたが，その後，Saito & Otsu (1988) は異なったアルゴリズムに基づく方法を提唱している．この他，Saito, Kariya & Otsu (1988) により3次以上の高次のモーメントを利用した主成分分析の方法が提唱されている．

主成分分析の応用研究としては，株価の変動といった時系列データへ適用してMTVモデル (Multivariate Variance Component Model ; MTVモデル) を提案した研究 (刈屋 (1987)) などがある．この他，主成分分析の応用としては線形回帰モデル $y=X\beta+\varepsilon$，においてβを推定する場合，計画行列Xの分散共分散行列 ($X'X$) の最大固有値 λ_1 と最小固有値 λ_2 の比 (λ_1/λ_p) の値 (条件数と呼ばれる) に基づく多重共線性の診断 (Belsley et al. (1980)) がある．

2.3 多次元尺度法

主成分分析は，多変量データが与えられている場合の次元縮小の方法であるが，多次元尺度法 (Multi-Dimensional Scaling (MDS)) は複数個の個体間の類似性データが与えられている場合に，個体間の非類似度と空間の距離が何らかの意味でできるかぎり一致するようにそれらの個体をできるかぎり少数の次元を持つ空間に布置する方法である．MDSの方法には類似度が間隔尺度で与えられている場合の手法である計量的MDS (Torgerson (1952))，および類似度が順序尺度，あるいは名義尺度で与えられている場合の非計量MDS (Shepard (1962), Kruskal, (1964), Takane et al. (1977)) がある．計量MDSの最も基本的な解法はn個の個体間の距離の平方を成分とする距離行列 $D_1^{(2)}$ を「ヤングハウスホルダー変換」して得られる行列 $P_1=(-1/2)Q_nD_1^{(2)}Q_n$ の固有値，固有ベクトルを求めるもので，Gower (1966) による主座標分析 (Principal coordinate analysis) と一致する．MDSにおいて分析の対象となる類似度行列は通常は対称行列であるが，n人の被験者のm個の対象に対する好みの評定値の逆数をそれぞれの被験者の理想点と対象点間の距離行列 $D_2^{(2)}$ とみなし，$P_2=(-1/2)Q_nD_2^{(2)}Q_m$ を特異値分解することによって，理想点と対象点を同時に多次元空間に布置する手法に展開法 (unfolding method, Schönemann (1970)) がある．これに対し，非計量MDSは非類似性の順序関係のみに基づいて空間の布置を定めるもので，解を求めるにはかなり複雑な逐次解を必要とする．

これらの手法の原型は1960年代に確立されたが，1970年代に入って複数個の類似度行列からそれらに共通する成分と独自な成分に分解する個人差MDS (Carroll & Chang (1970))，1970年代後半から1980年代にかけての最尤法に基づくMDS (Ramsay (1977), Takane (1978), (1981))，そして1970年代後半から1990年代にかけて我が国の研究者によって非対称類似度行列に基づくMDSの方法 (Okada (1987), Saito & Takeda (1990), Chino (1978, 1990), Saito (1991) が種々開発されている．多次元尺度法の1985年以前の研究に関してのサーヴェイ論文としてHayashi (1985)がある．

2.4 正準相関分析

1970年代に入り，行列の一般逆行列の概念が多変量解析に導入されるようになり，Khatri (1976) は，2変数群，X，Yの標本分散共分散行列，$S_{XX}=X'X$, $S_{YY}=Y'Y$ が正則でない場合に，S_{XX}, S_{YY} の一方の一般逆行列の選び方によらず，正準相関係数が一意に定まることを示した．Rao & Yanai (1979), Rao (1981), Yanai (1981) は正準相関係数の大きさは S_{XX}, S_{YY}

の双方の一般逆行列の選び方によらないことを示した。この結果を説明するためには直交射影行列の利用が不可欠である．X，Y の列ベクトルで生成される部分空間 $S(X), S(Y)$ への直交射影行列を $P_X=X(X'X)^-X'$，および $P_Y=Y(Y'Y)^-Y'$ とすると，正準変数 $F=XA$，および $G=YB$ は

$$(P_XP_Y)XA=XA\varDelta \quad (P_YP_X)YB=YB\varDelta, \tag{7}$$

によって得られる（ただし，A，Bはそれぞれ，$p\times r, q\times r$ 行列で，$r=\mathrm{rank}(X'Y)$，さらに，$\lambda_j=\lambda_j(P_XP_Y)$ とすれば，$\varDelta=\mathrm{diag}(\lambda_1, \lambda_2,\ldots, \lambda_p)$）．したがって，$X$，$Y$ の正準相関係数の平方は $\lambda_j(P_XP_Y)$ となることから，共分散行列 S_{XX}, S_{YY} が正則でない場合にも，正準相関係数は一意に定まることは明らかである．また，X と Y の列ベクトルで生成される部分空間 $S(X\,;\,Y)$ への直交射影行列が，

$$P_{X\cup Y}=P_X+P_{QXY}, \quad \text{または,} \quad =P_Y+P_{QYX} \tag{8}$$

（ただし，P_{QXY}, P_{QYX} はそれぞれ $S(Q_XY), S(Q_YX)$ の上への直交射影行列）と分解されることを利用すると，$F_{XZ}=(X\,;\,Z)$ と $F_{YW}=(Y\,;\,W)$ の正準相関係数の平方和に関して，次の分解が成立する（Yanai (1980)）．

$$\begin{aligned}\mathrm{tr}(P_{X\cup Z}P_{Y\cup W})&=\mathrm{tr}((P_X+P_{QXZ})(P_Y+P_{QYW}))\\&=\mathrm{tr}(P_XP_Y)+\mathrm{tr}(P_{QXZ}P_Y)+\mathrm{tr}(P_XP_{QYW})+\mathrm{tr}(P_{QXZ}P_{QYW})\end{aligned} \tag{9}$$

第2，3項は部分正準相関（part canonical correlation）の平方和，第4項は双偏正準相関（Bipartial Canonical correlation）の平方和である（Timm & Carlson, 1976）．

ところで，藤越康祝は多変量解析における変数間の冗長性について一連の研究（例えば，藤越（1992）を参照せよ）を行っているがそのひとつに正準相関分析における冗長性の研究がある．Fujikoshi (1982) は，X と Y の間の正準相関係数が（$X\,;\,Z$）と（$Y\,;\,W$）の間の正準相関とかわらない，つまり，Z と W が X と Y に関して冗長である (redundant) ための条件を導いているが，この条件は上記の（9）式を用いることにより $\mathrm{tr}(P_{X\cup Z}P_{Y\cup W})=\mathrm{tr}(P_XP_Y)$ によって導かれる．

ところで，正準相関分析における冗長性に関しては，上記の冗長性とは全く異なった文脈で定義されているものがある．（8）式によって得られるA，Bによって定義される構造ベクトルを，

$$V=(v_1, v_2,\ldots, v_p)=(X'X)A, \ W=(Y'Y)B \tag{10}$$

とおけば，Y が与えられた場合の X の冗長性係数（redundancy coefficient）（Levine (1977) を参照）は，次のように展開される．

$$\mathrm{Re}(X/Y)=\sum_{j=1}^{p}(\lambda_j\|v_j\|^2)/p=\mathrm{tr}(X'P_YX)/p \tag{11}$$

上記の冗長性係数は正準相関分析において得られた正準変数の解釈を明確にする為に Stewart & Love (1968) により導入されたものであるが，冗長性係数の概念から派生した冗長性分析（redundancy analysis）（van den Wollenberg, 1977）は上式の右辺の（$X'P_YX$）の主成分分析と等価となる．これらの論文の影響を受けて，Psychometrika 誌上においては，冗長性分析に関連した研究（例えば，Israels (1984)）が発表された．

正準相関分析を記述的な立場からみた場合，大変興味深い性質が存在する．Jewell &

Bloomfield (1983) は下記に示す二組の多変量データ行列X，Yに基づく共分散行列S,

(12)
$$S=\begin{pmatrix} S_{XX} & S_{XY} \\ S_{YX} & S_{YY} \end{pmatrix} \qquad S^{-1}=\begin{pmatrix} S^{XX} & S^{XY} \\ S^{YX} & S^{YY} \end{pmatrix}$$

から計算される正準相関係数の平方 $\lambda_j(S_{XX}^{-1}S_{XY}S_{YY}^{-1}S_{YX})$ の値が，上記に示した逆行列 S^{-1} の要素におきかえて計算される，$\lambda_j((S^{XX})^{-1}S^{XY}(S^{YY})^{-1}S^{YX})$ に等しくなることを導いた.

Khatri (1990)の定式化に従えば，$(A\,;B)=S^{-1}$ を満たす $(p+q)\times p$ 行列 A, $(p+q)\times q$ 行列 Bにより定義される

$$F=(X\,;\,Y)A=Q_YX(X'Q_YX)^{-1},\ G=(X\,;\,Y)B=Q_XY(Y'Q_XY)^{-1}$$

の分散共分散行列が S^{-1} に等しくなることより，$\rho_j(X, Y)$ を二つの変数群X，Yの第j番目の正準相関係数とすれば,

(13)
$$\rho_j(F, G)=\rho_j(Q_YX, Q_XY)=\rho_j(X, Y)$$

が導かれる. Yanai & Puntanen (1993) は上記の結果を3変数群X，Y，Zがある場合に拡張している. この問題に対するもうひとつの拡張の方向はSが特異な場合に上記の関係がどのように変化するかといったものである. Latour et. al (1987), Baksalary et. al (1992) は S^- が対称な反射型一般逆行列の場合の議論を展開し，Khatri (1990) は，S^+ の場合についての議論を展開しているが，Sの任意の一般逆行列についての結果は得られていない.

この他の正準相関分析に関する研究としては,$X=(X_1:X_2)$, $Y=(Y_1\,;\,Y_2)$ とおくとき,$\rho_j(X, Y)\geqq\rho_j(X_1, Y)\geqq\rho_j(X_1, Y_1)$ が成立することを示した研究 (Wang et al. 1987)，2組の多変量データX，Yの相互関連の指標について整理した研究(Ramzay (1984))，3組以上の変数群がある場合の正準相関についての研究 (Kettenring (1971), van de Geer (1984), Campbell & Tomenson (1983))，非線形正準相関の研究 (Burg et. al. (1983)，麻生他 (1987))，次元縮小法としての正準相関分析に関する研究 (Rao (1980))，正準変数に対する重みベクトル a, b, の各成分を1または－1に限定した研究 (DeSarbo (1982))，正準相関分析における解釈法に重点をおいた研究 (Cajo & ter Braak (1990))，正準相関分析で得られる重みベクトル a, b に $Ha=0$, $Gb=0$，といった制約条件をつけた Yanai & Takane (1992) による研究等がある. 正準相関分析は主成分分析と関連づけて用いられることは少なかったが，Gittins (1980) は，植物と動物の全体としての相互関連の強さの計量の必要性を背景にして，生態学における正準相関分析の利用を推奨している.

2.5 判別分析

正準相関分析における二組のデータ行列X，Yのうち，一方のYが個体の所属するq個のグループのいずれかを示すダミー変数行列である場合，S_B を級間分散共分散行列とすれば

(14)
$$(X'P_YX)a=\lambda(X'X)a \Rightarrow S_Ba=\lambda Sa$$

となる. これは，Rao (1952) によって，正準分析 (canonical analysis), Cooley & Rohnes (1962) によって，重判別分析 (multiple discriminant analysis) と呼ばれていたが，最近では正準判別分析 (canonical discriminant analysis) と呼ばれるようになってきている. なお，上記の正の固有値の個数は $m=\mathrm{rank}(S_B)$ で，この値は一般には $\mathrm{Min}(p, q-1)$ と等しくなるが，丘本(1992 a)は，人工データにより，$m<\mathrm{Min}(p, q-1)$ となる例が存在することを指摘している.

一方，p個の変数が多変量正規分布に従うという仮定から最尤法により導かれる判別関数に関しては，これまで多くの研究 (例えば，Krzanowski (1975), Otsu (1975)) がみられるが,

1970年代後半から疫学の分野においては多変量正規性の仮定が必要とされない多重ロジスティック分析が多く用いられるようになってきた（その詳細については，例えば，柳井他（1982）参照）．さらに，70年代から80年代にかけて核関数（kernel function）を用いてデータから確率密度関数を推定するノンパラメトリックな判別分析に関する研究（例えば，van Ness (1979)）が進展した．また，多変量正規分布に基づく判別関数で各群の分散共分散行列 Σ_i が等しいという仮定をおかない場合，２次の判別関数（その適用例については，Kubo (1980) を参照）が導かれるが，このためデータ数が少なくなることにより Σ_i が特異行列となることがある．このような場合に対処する方法（松田他（1990））が提案されている．

なお，判別分析の適用例としては，これまでに適性の診断（Yanai (1973)），病気の自動診断（高橋（1969），古川（1982）），筆跡鑑定（Sugiyama et al. (1986)，大塩他（1991）），形態の類型化（高部（1985））等が挙げられる．

2.6　多変量家族データの分析

身長，体重，胸囲，座高といった多変量の身体計測値の親と同胞間の関連性の程度（級間相関），および同胞間の関連性の程度（級内相関）を多変量的に測定する多変量家族データ（Multivariate familial data）の解析に正準相関分析，および正準判別分析が適用されている（Konishi & Khatri (1990), Konishi et al. (1991), Konishi & Rao (1992)）．また，Rao & Rao (1987) は上記の多変量家族データの分析に関連づけて，同系正準相関（homologous canonical correlation）という概念を提唱している．

2.7　因子分析

因子分析モデル（母数モデル，変量モデル）における基本性質については1940年代から今日まで過去半世紀にわたり多数の文献がある．因子分析における座標軸の回転など記述的側面を重視したものに Harman (1976)，芝（1979），因子分析モデルにおける推測的側面を重視したものに，Lawley & Maxwell (1971)，丘本（1986），Bartholomew (1987)，記述的側面と推測的側面をともに解説したものに柳井・繁桝他（1990）がある．しかし，多変量データの解析手法という観点からいえば，因子分析は主成分分析とほぼ同様の解を生じることがあり，また，不適解に関して統計学的に未知の問題も多く，津村（1971）は「因子分析は有用か」という論文を書いて因子分析の有用性に疑問を投げかけた。しかし，1970年代から80年代にかけて，因子分析に関する研究論文は Psychometrika 誌を中心に増加し，すでに本論の第Ｉ部で述べたように Anderson (1984) においては新しく因子分析の章が付加された．さらに丘本（1987）は，「因子分析は燃えている」という序文とともに因子分析が理論的にも方法論的にも興味ある問題を内包している点を指摘し，最近の因子分析の研究動向をサーヴェイしている．この他の，因子分析に関するサーヴェイ論文としては，Steiger (1979), Mulaik (1986)，市川（1993）がある．本節では，これらを踏まえた上で，因子分析の最近の研究動向のうち，モデルの識別性，共通性の推定値，因子回転の話題に焦点を絞って概観しよう．

因子分析モデル（直交モデル）は p 個の変数 $x_1, x_2, \cdots, x_p$ を列ベクトルの成分とする p 次元確率ベクトル x を

$$x=\Lambda f+e,\quad \text{ただし } E(e)=0,\ Var(f)=I_m,\ Cov(f,e)=O,\ Var(e)=\Psi \tag{15}$$

と分解し，それにより，x の母分散共分散行列を

$$\Sigma=\Lambda\Lambda'+\Psi \tag{16}$$

（ただし，$p\times m$ 行列 Λ は因子負荷行列，$\Psi=\mathrm{diag}(\psi_1, \psi_2, \cdots, \psi_p)$ は独自因子負荷行列）

と分解する．上記（16）式による Σ の分解が存在する場合，回転の不定性を除いて，識別可能

な構造 (Λ, Ψ) を求める問題は古くから研究されており,「$p\times m$ 因子負荷量行列 Λ から, 任意の一行を取り除いたときに二つの正則行列が存在すれば (Λ, Ψ) は識別可能である」という Anderson & Rubin (1956) の条件は良く知られている. この条件は明らかに, $p\geqq 2m+1$ の場合にしか適用できない. $p<2m+1$ の場合の一般的な十分条件は知られていないが, Bekker & de Leeuw (1987) は, $m=p-1$ でしか共通因子分解を持たないための必要十分条件を与えた. $m=3$ の場合については上記の性質で示唆される条件が必要十分条件となることが Tumura & Sato (1980) により証明された. 識別可能性の条件, およびその open problem に関して記述したものに, Shapiro (1985), Sato (1989, 1992) 等があるが, これらの既存の研究の相互関連を踏まえた統一的理論を構築する必要があろう. ところで, 識別可能性の条件が著しく損われたモデルの場合, 不適解 (improper solution) がおこる. 因子分析の最尤解における不適解の問題点とその採るべき処置について, van Driel (1978), 竹内 (1986), Sato (1987) の研究がある.

以下においては, Σ を相関係数行列とみなそう. ここで, 相関係数行列 Σ の逆行列 Σ^{-1} の j 番目の対角要素を r^{jj} とすると, $1-(1/r^{jj})$ は, j 番目の変数を基準変数, 他の $(p-1)$ 個の変数を説明変数とする重相関係数の平方になり, それゆえ, SMC (Squared Multiple Correlation) と呼ばれている. SMC が共通性の下限となることは, 1930 年代に Roff (1936) に示されたが, 最近, 筆者ら (Yanai & Ichikawa (1990)) は, 相関行列 Σ の最小固有値を $\lambda_p(\Sigma)$ とすると $(1-\lambda_p(\Sigma))$ は少なくともひとつの変数に対し, SMC の値を上回るより強い下限になっていることを示した. 数値計算によると, 多くの場合, SMC の値を上まわる下限は 2 つ以上存在するが, 与えられた Σ の関数として, SMC の値を上まわる下限の個数を求めるのは今後の研究課題である。なお, Ihara & Kano (1986) は任意の変数を固定したとき, その変数に関して, 因子負荷行列 Λ が Anderson-Rubin の条件を満たすと仮定して, 独自性 ψ_i を Σ の関数として一意に定め, 反復推定を必要としない Λ の推定量を提案している.

因子分析の推測統計的側面について簡単に触れておこう. まず, 1980 年代に入って因子分析モデルの誤差項に多変量正規分布よりもっと広い範囲の分布を仮定して, 最尤推定, およびその漸近理論の研究が盛んになってきたことである. これらについて, Bentler (1983), Browne (1984), 等によって推進されたのはその一例である. 我が国の研究者による貢献としては, 上記の問題に対する Kano (1991) の研究, および Kano (1983, 1986)) による因子分析における一致性の研究, 最小二乗解に基づく因子推定のアルゴリズムに関する研究 (Okamoto & Ihara (1983, 1984)) 等がある. これら因子分析の推測統計的側面に関する研究の進展についての詳細は, 丘本 (1987), 狩野 (1990) を参照されたい. なお, Ichikawa (1992) は, 誤差項に多変量正規分布を仮定し, 共通性の漸近分散を導いている.

なお, 因子分析のもうひとつの推測統計的問題として, 共通因子数の推定がある. これに関しては, 最尤法因子分析における情報量基準 AIC の導入 (Akaike (1987)), および, Kano (1990) がある.

この他, 因子分析モデルに関連した統計的推測の問題として線形関数関係 (linear functional relationship) および線形構造関係 (linear structural relationship) の解析がある. これらに関する最近の研究には, Anderson (1987), Isogawa (1992) 等がある.

因子分析は心理学の分野で, 知能, 性格および意味の分析などに用いられ, 成果があげられてきた. この理由のひとつに因子の単純構造をめざすバリマックス回転, 与えられた仮説構造をなるべく再現できるようなプロクラステス回転などの因子軸の回転法の導入によるところが大である. プロクラステス回転とは因子分析にとりいれる変量に関する事前の知識を用いて仮説構造行列 G を作り, G の各成分と回転後の因子負荷量行列 $B=AT$ の差の平方和, すなわち

tr[$(G-AT)'(G-AT)$]を最小にする回転行列Tを求めるものである．なお，因子負荷量行列が3つ以上ある場合のプロクラステス回転の研究（Gower (1975), Ten Berge (1977)）も行われている．Gとしてバリマックス回転によって得られる因子負荷量の奇数乗したものを用いるプロクラステス回転がプロマックス回転と呼ばれるもので，近年，バリマックス回転に替えてプロマックス回転を採用した研究が心理学の文献では多く見られるようになっている．その一例として，柳井他（1987）による性格検査の尺度構成がある．

この他の因子分析関連研究に，同一相関係数行列に主成分分析と因子分析を適用し負荷量の大きさの対応関係を調べた研究（Sato (1990)），イメージ因子分析に関する研究（Yanai & Mukherjee (1987)），非計量因子分析に関する研究（Takane et al. (1979)）等がある．この他，共通因子数を固定して，効率よく変数を選択する研究（Yanai (1980), Tanaka (1982)）等がある．

2.8　潜在変数分析一般について

多変量データ解析における潜在変数モデルとしては，因子分析の他に共分散構造分析と項目反応理論，潜在構造分析等がある．共分散構造分析とは共分散行列Σをさまざまな形に分解する仮説モデルを設定して，データである共分散行列Sを用いて，モデルに含まれる種々の母数の推定，およびその検定を行うもので，因子分析モデルもその特別な場合に含まれる．ここでは，共分散構造分析の概要とその適用法に関して Bollen (1989)，豊田 (1992)，小笠原 (1993) を，推測的な面に関しては狩野(1990，1993)，を紹介するにとどめよう．なお，共分散構造分析の理論は，いわば心理学で発展してきた因子分析モデルと経済学の分野で発展した構造方程式モデルに基礎を持つもので，異なった分野の手法を結合することによって新しい手法が生まれた一例といえよう．

なお，共分散構造分析のひとつの手法である LISREL (Linear Structural RELationship) は，因果分析の手法として知られているパス解析（path analysis）をその特殊な場合として含むものである．多変量データに基づく潜在変数モデルとして最近注目されているものに項目反応理論（item response theory, IRT と略記）がある．IRT は計量心理学の分野で開発されたもので，正解，不正解のように採点される一次元的構造を持つ2値反応データから，項目の難易度，識別度，被験者の潜在能力を推定することを目的としたもので，理論的には因子分析の1因子モデルと等価であることが示されている（Takane et al. (1987)）．項目反応理論の理論とその適用例については Hambleton (1985)，芝（1991）を参照されたい．

なお，潜在構造分析に関しては，別稿「潜在構造分析の現状」（浅野長一郎，江島伸興）で解説されている．

上記の潜在変数分析の最近の動向については，柳井他 (1991)，および雑誌数理科学（1992年11月号─93年4月号：リレー連載）を参照されたい．

3.　質的データの構造解析

数量化法とは名義尺度または順序尺度で表される質的変数に対して適当な操作によってある数量を与えることで，広義の意味での尺度化（Scaling）に相当する．このような数量化の方法として代表的なものに，林知己夫による数量化理論がある．数量化理論のうちの1類，2類は質的データを用いた重回帰分析，正準判別分析といってもよいものである．上記の数量化1，2，3，4類からなる数量化理論の方法はあくまで数量化の1つの典型的な方法で，最近では，これら4つの方法を拡張したいくつかの新しい方法が提案されている．この中で，理論的に興味深い方法は数量化3類であろう．数量化3類に類似した方法としては，フィッシャーの交互平均法（Fisher (1938)），ガットマンの尺度解析（Guttman (1941)），西里の双対尺度（Dual

Scaling)（西里 (1982), Nishisato (1980)）がある．さらに，フランスには Benzecri (Benzecri (1973)）を頂点とするフランス学派の対応分析（correspondence analysis)（Lebart et. al (1984), Greenacre (1984)）があり，1981 年より隔年にデータ解析と情報学〈Data Analysis and Informatics〉という国際シンポジウムを開催している．また，それに関連して日仏の交流が促進され，1987 年 3 月には日仏セミナーが東京で開催され，その成果が Hayashi & Diday et al. (1988) によってまとめられている．

ここで，これらの方法の概要を紹介しよう．この方法は，$n\times p$ データ行列 N_{AB}（または，分割表形式でもよい）の n 個の行（要因A）のそれぞれに与える数量ベクトルを a，p 個の列（要因B）のそれぞれに与える数量ベクトルを b, さらに D_A, D_B をそれぞれNの行和，列和を対角要素とする対角行列としたとき，

$$a'N_{AB}b/\sqrt{(a'D_Aa)(b'D_Bb)} \tag{17}$$

を最大にする a，b を求めるもので、その解は次式によって求められる．

$$N_{AB}D_B^{-1}(N_{AB})'a=\lambda D_Aa,\ b=D_B^{-1}(N_{AB})'a/\sqrt{\lambda} \tag{18}$$

このとき，対応分析，または双対尺度法においては，複数個のカテゴリからなる二つの項目へのダミー変数行列を G_A と G_B とおくとき，$N_{AB}=(G_A)'G_B$ と分解されることを強調する．これにより，上記 (18) 式の最大化は数学的には G_A と G_B に基づく正準相関分析と等価になる．この性質を利用することにより，Nishisato (1984) による定比例の原理（principle of constant proportionality), Lebart et al. (1984) による，（i）$N_{AB}=(G_A)'G_B$，（ii）$G=(G_A\,;G_B)$ (iii) $B=G'G$, に基づく対応分析が同一の解をもたらす，と言った興味ある性質が導かれる．さらに，$G_{AC}=(G_A\,;G_C)$ と，$G_{BC}=(G_B\,;G_C)$ に基づく正準相関分析により，要因AとBから要因Cの影響を除去した対応分析，すなわち偏対応分析(partial correspondence analysis (Yanai (1986)) と呼ばれる方法が導かれたが，これに関連して，偏関連係数の定義が Daudin (1980) によって与えられていることを指摘しておこう．

ところで，数量化 3 類の分析対象となるデータ N_{AB} には以下のような自由選択型（2 値データの一方のカテゴリーのみについてのデータを示したもの)，および項目カテゴリ型(2 値データの双方のカテゴリのデータを示したもので，各行の数値の和は項目数（この場合 4）に一致する）といったデータ（Guttman (1950) による）がある．ここで，最近の丘本（1992 c）の研究によると，保有する情報が一見全く同一にみられるこれらの二種類のデータに，数量化 3 類を適用したところ全く異なる解が得られ，固有値の大きさに一貫性がないことが見出された．どのような付加条件により，自由選択型に比べ項目カテゴリ型の固有値が大きくなるかについては今後の課題であろう．

(自由選択型)

$N_{AB}=$

	b_1	b_2	b_3	b_4
a_1	0	0	0	0
a_2	1	0	0	0
a_3	1	1	0	0
a_4	1	1	1	0
a_5	1	1	1	1

(項目カテゴリ型)

$N_A=$

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
c_1	0	1	0	1	0	1	0	1
c_2	1	0	0	1	0	1	0	1
c_3	1	0	1	0	0	1	0	1
c_4	1	0	1	0	1	0	0	1
c_5	1	0	1	0	1	0	1	0

数量化 3 類は通常，個人データ x 項目といった 2 元データに適用されるが，対応分析は分割表形式のデータに直接適用することも可能で，その意味で汎用性があり，その幾何学的性質

(Greenacre et al. (1987)),制約付き解法 (Giula & Haberman (1988), Bockenholt & Bockenholt, 1990),対数線形モデルとの融合 (Choulukian (1988)),人工データによる1次元および2次元構造の研究(丘本 (1992 b))など多数の論文が生まれている.また,ter Braak (1986) による正準対応分析 (canonical correspondence analysis) は,制約付き対応分析と等価であることが示されている (Takane, Yanai, Mayekawa, 1991).

上記の対応分析は,基本的には分割表の特異値分解で,各カテゴリで得られた数量の大きさに関する統計的推論は行われていないが,Takane et al. (1987) において導入された理想点判別分析 (ideal point discriminant analysis) によって,最尤推定に基づく,母数,および標準誤差の推定が可能となった.この他,尤度原理に基づき分割表の各カテゴリに数量化を与える統計モデルに,Goodman (1981) による連関モデル (association model) があり,これに関し順序カテゴリカルデータに関するいくつかの研究(例えば,Tsujitani (1988, 1992))がある.この他の数量化理論の性質,およびその拡張に関する研究には佐藤 (1985) の研究がある.また,質的データに基づく因子分析の動向については繁桝 (1990) を参照されたい.

数量化理論の応用研究は多数あるが,ここでは,国際データの比較研究に数量化理論を適用した Hayashi et al. (1984) および Hayashi et al. (1992) を挙げておく.

4. 多重配列データの分析法

同一変量,同一被験者による複数回繰り返し測定データのような3次元の多変量データに関する研究としては,Tucker (1966) による三相因子分析法,先に示した多次元的尺度法における個人差 MDS, Takane & Young (1977) による ALSCAL がある.

タッカーの方法から派生したものにはさまざまなものがあるがそのひとつに Kroonenberg & deLeeuw (1980) による Tucker 2,村上 (Murakami (1983),村上 (1990)),ten Berge et al. (1987), Kroonenberg & ten Berge (1989) による三相の主成分分析の方法がある.なお,同一変数を複数の集団に適用して得られる多変量データの主成分分析とその関連手法についての記述が成書 (Flury (1988)) として出版されている.

また,質的データの解析にも高次元のデータが利用される.3つの項目の3元分割表データが与えられている場合,それぞれの項目のカテゴリ間の関連が強くなるような数量化の方法に関しては岩坪 (1975),多重配列データの特異値分解に関しては吉沢 (1975, 1976) らの先駆的研究がある.1970年代の後半から1980年代にかけて,この種の研究は多数みられるようになった.Lastovicka (1981) はタッカーの3相因子分析を拡張して4相の主成分分析を定式化した。この研究を拡張して,Kapteyn et al. (1986) は multimode component analysis を定式化した.この研究において表わされる n 重配列データ $y_{j_1j_2-j_n}$ に関してクロネッカー積による表現を与えており,多重配列データの表現,およびその解析法の導入にあたりクロネッカー積が有効となることが示唆された.また,同種の議論は Kiers (1991) においても展開されている.クロネッカー積を多変量データ解析に導入した他の研究としては,数人の医師が複数の患者に対して同一評定項目の評定を行って得られた多特性多方法相関係数行列にクロネッカー積によってモデル化された仮説検証的因子分析を適用した研究 (Browne (1984)) がある.なお,クロネッカー積についての数学的詳細については,Magunus & Neudecker (1988) を参照されたい.

多重配列データに関するもうひとつの研究方向に,高次元データの rank をどのように定義するかという興味深い問題がある.Kruskal (1977, 1989) は $n_1 \times n_2 \times n_3$ 次からなる3元データ行列 $Y=(y_{ijk})$ の rank を

(19)
$$y_{ijk}=\sum_{j=1}^{R}a_{ir}b_{ir}c_{kr}$$

が成立する最小のRによって定義した．この定義は二次元の場合の特異値分解の定義を内包するもので，多重配列データのrankとして自然なものであろう(この点の記述に関しては，文献［170］に掲載されている今井・大西他論文を参考にした)．付言すれば，上記の3次元データ分解モデルはHarshman (1970) によるParafacモデルに対応するものである．なお，Parafacモデルの適用例にはHayashi & Hayashi (1982)，最尤解に基づくParafacモデルの解法についてはMayekawa (1987) がある．

以上のようにいくつかの研究はみられるが多重配列データを解析する研究はまだ端緒についたばかりである．Multiway dataに関する国際集会が1988年3月にローマで開催され，プロシーディングスが出版されている (Coppi & Bolasoco (1989))．また，我が国においては，吉沢正(筑波大)を代表とする文部省科研費による研究班が1989年4月から3年間にわたって精力的な研究を続け，その研究成果の集大成となる報告書 (吉沢 (1992)) をまとめている．

5. その他

多変量データの解析にあたっては，分析結果を大きく左右する少数の外れ値の摘出，特に多変量外れ値 (multivariate outlier) の摘出が不可欠である．多変量外れ値の存在を見出す方法のひとつに，感度分析(sensitivity analysis)と呼ばれるものがある．我が国においては田中豊による一連の研究 (例えばTanaka & Tarumi (1986), Tanaka (1989)，詳しくは田中 (1992) を参照せよ) がある．また，書物としてはChatterjee (1988) が興味深い．この本では，重回帰分析において各種の診断法としての各種偏回帰プロットの方法が紹介されている．このような手法に基づいて検証される外れ値や，クラスター分析におけるデンドログラムの表示に関しては，近年のコンピュータによるグラフィカル接近法，すなわち，統計的グラフィックス (例えば，Cleveland (1987)，後藤他 (1988) を参照) の利用が薦められている．統計的グラフィックスの著しい発展によってもたらされた多変量データ解析のもうひとつの手法に射影追跡 (projection pursuit) がある．

射影追跡とは多変量データ行列Xによって特徴づけられるn個の個体の散らばりの興味深い構造を最もよく表わす部分空間への射影を求めるもので，その結果をグラフィック端末などに表示することにより，分析を行う者が実際にその構造を目で見ることが可能となる．これまでに示したところの多変量データ解析の全ての方法をその特殊の場合として含むものであるが，通常の多変量データ解析の手法では把握できない構造を探ることが可能となるような様々な射影指標が考案されている．射影追跡の原理，その応用については，Friedman & Tukey (1974), Huber (1985)，岩崎 (1989, 1990) 等を参照されたい．なお，本稿では触れなかったクラスター分析については，統計学辞典 (竹内他 (1989)) のIII—6において最近の動向の的確な解説がなされている．なお，クラスター分析を含む，分類(classification)に関する種々の方法の理論的および応用的研究は多くの分野に跨がって行われている．1987年7月には第1回の国際分類学会の会合が開かれ，そのプロシーディングス (Bock (1988)) が発行されている．また，アメリカでは1984年に，J. of Classificationという雑誌が刊行された．

6. 多変量データ解析の応用およびソフトウエアについて

これまでの各節で概観した多変量データ解析の各種手法は，1970年代後半から1980年代前半にかけてのデイスプレイ装置を備えたパーソナルコンピューターの普及および後述する各種

多変量解析プログラムの普及により，法律学，政治学，経済学、経営学，心理学，教育学，社会学，工学、農学，人類学，医学，生物学，看護学などの行動科学とよばれる幅広い分野に於けるデータ解析の手法として脚光を浴びるようになってきた．一例を挙げよう．図は1970年から1985年までの「日本公衆衛生誌」と「American J. of Epidemiology」誌における多変量解析手法の使用頻度を示したものであるが，1970年代から80年にかけて加速度的に増加していることがみてとれる．同様な傾向は因子分析が最も使用される頻度の高い心理学研究，教育心理学研究にもみられる（柳井・市川（1985））．多変量解析に関する研究会も各地で開かれ，その成果も多数出版されている．大学入試センター研究開発部で1986年11月から毎月開催されている多変量解析研究会は，1993年7月で第80回を迎えたが，第36回までの講演内容の一部が柳井・岩坪・石塚（1990）として出版された．また，過去20年にわたって実施されている日本科学技術連盟主催の多変量解析研究会の成果も吉沢・芳賀（1992）として出版されている．この著にはTQC（Total Quality Control，総合的品質管理）の分野における多変量解析の適用事例が多数紹介されている．

多変量データ解析に関する各種手法のソフトウエアとして世界的にみて最も定評があるのはSPSS, BMDP, SASといったアメリカで開発されたプログラムであろう．SASによる多変量解析プログラムの利用に関しては，大橋・市川（1987）が参考となる．この他に，我が国の研究者によって作成されたものに，田中・垂水・脇本（1984）に基づくプログラム，芳賀（1984）によるCDA，柳井・高木（1986），高木（1992）に基づくプログラム（HALBAU），日本科学技術研修所によるJUSE-MA 1，およびJUSE/MDSAなどがあり，それぞれ，多くの研究領域で利用されている．

また，多変量解析および多変量データ解析に関する研究発表は，日本統計学会，日本行動計量学会，応用統計学会はもちろんのこと，法律，経済，心理，教育，工学，医学等の多数の領域で行われており，その全貌を明らかにすることは困難であるが，今後もますます増加の一途をたどるものと期待される．しかし，研究者が自分の保有するデータに多変量データ解析を適用するにあたって，数々の疑問を持っていることが木下（1992）によって明らかにされた．したがって，多変量データ解析の利用者が自己の領域に固有のデータに多変量データ解析の各種手法を適切に利用できるようさまざまな啓蒙的研究活動がなされていく必要があろう．そのひとつの方向としては，Gale（1986）に示唆されるような人工知能的アプローチを多変量データ解析のプログラムに採り入れることであろう．中野他（1991）は知識ベース型重回帰分析支援プログラムを開発しているが，他の多変量データ解析の手法にも人工知能的アプローチが採り入れられていくことが望まれる．

なお，第1部の多変量推測論と，第2部の多変量データ解析は独立に開発されたものではなく，相互に関連を持ちながら発展してきたものである．この点に関し，Rao（1983）と竹村(1987)の総説は有益であることを指摘しておく．

謝辞：複数の審査員および丘本正氏（追手門学院大学教授），市川雅教氏（東京外国語大学助教授）からは最初の原稿に対して有益な多数のコメントを頂きました．ここに記して謝意を表します．

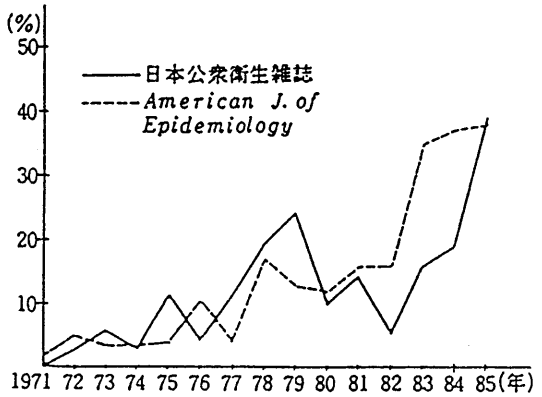

図 「日本公衆衛生雑誌」と「American J. of Epidemiology」における多変量データ解析の技法を用いた論文の割合
(雑誌「Basic 数学」現代数学社 (1987年3月号，柳井・高木論文) より転載)

参 考 文 献

[1] Akaike, H. (1987) Factor analysis and AIC, *Psychometrika*, **52**, 317-332.

[2] Anderson, T. W. & Rubin, H. (1956) Statistical inference in factor analysis, *Proc. Third Berkley Symp. on Math. Statist. Prob.*, **5**, 111-150, Univ. of California Press, Berkeley.

[3] Anderson, T. W. (1987) Multivariate linear relations, Pukkila & Puntanen Ed. *Proc. Second International Tampere Conference in Statistics*, 9-36.

[4] 麻生英樹・栗田多喜夫・大津展之 (1987) 正準相関分析および判別分析の非線形の定式化による解釈について，行動計量学，**14**, 1-9.

[5] Baksalary, J. K., Puntanen, S. & Yanai, H. (1992) Canonical correlation associated with symmetric reflexive generalized inverses of the dispersion matrix, *Linear Algebra and its Applications*, **176**, 61-74.

[6] Bekker, P. A. and de Leeuw, J. (1987) The rank of reduced dispersion matrices, *Psychometrika*, **52**, 125-135.

[7] Bentler, P. M. (1983) Some contribution to efficient statistics in structural models: specification and estimation of moment structures, *Psychometrika*, **48**, 493-517.

[8] Bockenholt, U. & Bockenholt, I. (1990) Canonical analysis of contingency tables with linear constraints. *Psychometrika*, **55**, 633-639.

[9] Browne, N. W. (1984) Asymptotic distribution-free methods for the analysis of covariance structure, *British J. of Mathematical and Statistical Psychology*, **37**, 62-83.

[10] Browne M. W. (1984) The decomposition of multitrait-multimethod matrices, *British J. of Math. and Statist. Psychology*, **37**, 1-21.

[11] Burg, E. van der, and de Leeuw, J. (1983) Non-linear canonical correlation. *British J. of Math. Statist. Psychology*, **36**, 54-80.

[12] Cajo, J. F. & ter Braak (1990) Interpreting canonical correlation analysis through biplots of structure correlations and weights, *Psychometrika,*, **55**, 519-532.

[13] Caroll, J. D. & Chang, J. J. (1970) Analysis of individuall differences in multidimensional scaling by an N-way Eckart-Young decomposition, *Psychometrika*, **35**, 282-319.

[14] Campbell, N. A. & Tomenson, J. A. (1983) Canonical variable analysis for several sets of data, *Biometrics*, **39**, 425-435.

[15] Chino, N. (1978) A graphical technique for representing the asymmetric relationship between N objects, *Behaviormetrika*, **5**, 23-40.

[16] Chino, N. (1990) A generalized inner product model for the analysis of asymmetry, *Behaviormetrika*, **27**, 25-46.

[17] Choulukian, V. (1988) Exploratory analysis of contingency table by loglinear formulation and

generalization of correspondence analysis, *Psychometrika*, **53**, 235-250.

[18] Cleveland, W. S. (1987) Research in statistical graphics. *J. Amer. Statist. Assoc.* **82**, 419-423.

[19] Daudin, J. J. (1980) Partial association measure and an application to qualitative regression, *Biometrika*, **67**, 581-590.

[20] DeSarbo, W. S et al. (1982) Constrained canonical correlation, *Psychometrika*, **47**, 489-516.

[21] Fisher ,R. A. (1938) The statistical utilization of of multiple measurements, *Annals of Eugenics*, 8, 376-386.

[22] Fujikoshi, Y. (1982) A test for additional information in canonical correlation analysis, *Ann. Inst. Statist. Math.* **34**, 137-144.

[23] 藤越康祝 (1992) 多変量解析における変量の冗長性, 行動計量学 **19**, 1, 18-28.

[24] Friedman, J. H. & Tukey, J. W. (1974) A projection pursuit algorithm for exploratory data analysis, *IEEE Trans. Computers* C-**23**, 881-890.

[25] Gabriel, K. R. (1971) The biplot graphic display of matrices with application to principal component analysis, *Biometrika*, **58**, 453-467.

[26] Giula, Z & Haberman, S. J. (1988) The analysis of contingency tables by restricted canonical and restricted association models. *J. Amer. Statist. Assco.* **83**, 760-771.

[27] Gollob, H. F. (1968) A statistical models which combines features of factor analysis and analysis of variance, *Psychometrika*, **33**, 73-115.

[28] Goodman, L. A. (1981) Association models and canonical correlation in the analysis of cross classification having ordered categories, *J. Amer. Statist. Assoc.*, **76**, 320-334.

[29] 後藤昌司, 松原義弘, 脇本和昌 (1988) グラフィカル接近法の最近の発展, 行動計量学, **29**, 45-70.

[30] Gower J. C. (1966) Some distance properties of latent roots and vector methods used in multivariate analysis, *Biometrika*, **53**, 325-338.

[31] Gower, J. C. (1975) Generalized Procrustes analysis, *Psychometrika*, **40**, 33-51.

[32] Greenacre, M. & Hastie, T. (1987) The geometric interpretation of correspondence analysis, *J. Amer. Statist. Assco.*, **82**, 437-447.

[33] Guttman, L. (1941) The quantification of a class of attributes: a theory and method of scale construction. In the Committee on Social Adjustment (ed.), The *Prediction of Personal adjustment*, Social Science Council, Wiley, New York, 312-361.

[34] Guttman, L. (1950) The principal components of scale analysis. In Stouffer, S. A. (ed.) *Measurement and Prediction*, Wiley New York, 312-361.

[35] 芳賀敏郎 (1984) 対話型データ解析システム, 応用統計学, 13-3, 125.

[36] Harshman, R. A. (1970) Foundation of the PARAFAC procedure: Models and conditions for an explanatory multimode factor analysis, Unpublished thesis, Los Angeles, Univ. of California.

[37] 橋口渉子 (1978) 多変量データの解析における直交関数の利用について, 応用統計学, 7, 111-124.

[38] Hastie, T. & Stuetzle, W. (1989) Principal curves, *J. Amer. Statist. Assoc.* **84**, 502-516.

[39] Hayashi, C. Suzuki, T. &, Hayashi, F. (1984) Comparative study of lifestyle and quality of life: Japan and France, *Behaviormetrika*, **15**, 1-18.

[40] Hayashi, C. (1985) Recent theoretical and methodological developments in multidimensional scaling and its related methods in Japan, *Behaviormetrika*, **18**, 67-79.

[41] Hayashi, C. & Hayashi, F. (1982) A new algorithm to solve PARAFAC model, *Behaviormetrika*, **11**, 49-60.

[42] Huber, P. J. (1985) Projection pursuit (with discussion), *Ann. Statist.* **13**, 435-525.

[43] Ichikawa, M. (1992) Asymptotic distribution of the estimators of communalities in factor analysis, *Psychometrika*, **57**, 399-404.

[44] 市川雅教 (1993) 因子分析モデル—多変量解析の最近の動向 (2), 数理科学, 12, 66-71.

[45] Ihara, M. & Kano, Y. (1986) A new estimator of the uniqueness in factor analys is, *Psychometrika*, **51**, 563-566.

[46] Israels, A. Z. (1984) Redundancy analysis for qualitative variables, *Psychometrika*, **49**, 331-346.

[47] Isogawa, Y. (1992) Asymptotic distributions of a test statistic in multivariate linear relationships, *J. of Japan Statist. Soc.*, **22**, 201-210.

[48] 岩崎 学・福永真美 (1989) 多項式指標による射影追跡, 応用統計学 18, 103-128.

[49] 岩崎 学 (1990) 射影追跡と多変量データ解析, 柳井・岩坪・石塚編 (1990) 人間行動の計量分析 東

京大学出版会.
[50] 岩坪秀一 (1975) 3-way 離散データを分類する二つの技法—相関比と 3 次相関係数による分類 行動計量学, 2, 1, 54-65.
[51] Jewell, N. P & Bloomfield, (1983) Canonical correlations of past and future for time series: definitions and theory, *Ann. of Statist*. **11**, 837-847.
[52] Kano, Y. (1983) Consistency of estimators in factor analysis, *J. Japan Statist. Soc*, **13**, 137-144.
[53] Kano, Y. (1986) Conditions on consistency of estimators in covariance structure model, *J. Japan Statist. Soc.*, **16**, 75-80.
[54] Kano, Y. (1990) Noniterative estimation and the choice of factors in exploratory factor analysis, *Psychometrika*, **55**, 277-291.
[55] 狩野 裕 (1990) 因子分析における統計的推測：最近の発展, 行動計量学, **18**, 1, 3-12.
[56] Kano, Y. (1991) The asymptotic distribution of a non-iterative estimator in exploratory factor analysis, *Ann. Statist*., **19**, 272-282.
[57] 狩野 裕 (1993) 共分散構造分析モデルと統計的推測, 数理科学, **355**, 77-84.
[58] Kapteyn A., Neudecker H. & Wansbeek T. (1986) An approach to n-mode component analysis, *Psychometrika*, **51**, 269-275.
[59] 刈屋武昭 (1987) 第 5 章 MTV (多変量時系列変動要因モデル) と株価の予測, 鈴木雪夫・竹内 啓編, 社会科学の計量分析—多変量解析の理論と応用：東大出版会.
[60] Kettenring, J. R. (1971) Canonical analysis of several sets of variables, *Biometrika*, **58**, 433-451.
[61] Khatri, C. G. (1976) A note on multiple and canonical correlations for a singular covariance matrix, *Psychometrika*, **41**, 465-470.
[62] Khatri, C. G. (1990) Some properties of BLUE in a linear model and canonical correlations associated with linear transformations, *J. Multivar. Anal.* **34**, 211-226.
[63] Kiers, H. A. L. (1991) Hierarchical relations among the three-way method, *Psychometrika*, **56**, 449-470.
[64] 木下冨雄 (1992) 多変量解析に対するユーザーのニーズ, 行動計量学, 19, 1, 40-48.
[65] Konishi, S. and Khatri, C.G. (1990) Inferences on interclass and intraclass correlations in multivariate familial data, *Ann. Inst. Statisti. Math.*, **42**, 561-580.
[66] Konishi, S., Khatri, C. G.and Rao, C. R. (1991) Inferences on multivariate measures of interclass and intraclass correlations in familial data, *J. Royal Statist. Soc. Series B*. **53**, 649-659.
[67] Konishi, S. and Rao, C. R. (1992) Principal component analysis for multivariate familial data, *Biometrika*, **79**, 631-41.
[68] Kruskal, J. B. (1964) Multidimensional scaling by optimizing a goodness of fit to a nonmetric hypothesis, *Psychometrika*, **29**, 1-29.
[69] Kruskal, J. B. (1977) Three-way arrays: Rank and uniqueness of trilinear decomposition, with application to arithmetic complexity and statistics, *Linear Algebra and its Applications*, **18**, 95-138.
[70] Kruskal, J. B. (1989) Rank, decomposition, and uniqueness for 3-way and N-way arrays, *(Coppi, & Bolasco, Ed.) Multiway Data Analysis*, 7-18, North-Holland.
[71] Kroonenberg, P. M. & de Leeuw, J. (1980) Principal component analysis of three-mode data by means of alternating least squares algorithm, *Psychometrika*, **45**, 69-97.
[72] Kroonenberg, P. M. & ten Berge ,J. M. F. (1989) Three mode principal component analysis and perfect congruence analysis for sets of covariance matrices. *Psychometerika*, **42**, 63-80.
[73] Krzanowski, W. J. (1975) Discrimination and classification using both binary and continuous variables, *J. Amer. Statist. Assoc.* **70**, 782-790.
[74] Kubo, T. (1980) Quadratic discriminant analysis of cephalo-pelvic disproportion, *Behaviormetrika*, 8, 1-14.
[75] Lastovicka, J. L. (1981) The extension of component analysis to four-mode matrices, *Psychometrika*, **45**, 47-57.
[76] Latour, D., S. Puntanen & Styan P. H. (1987) Equalities and inequalities for the canonical correlation associated with some partitioned generalized inverses of a covariance matrix, *Proc. Second International Tampere Conference in Statistics*, 541-553.
[77] 松田眞一, 藤本 隆, 吉村 功, (1990) 分散行列の縮小推定量を用いたロバストな 2 次判別関数, 応用統計学, **19**, 33-52.

[78] Mayekawa, S. (1987) Maximum liklifood solution to the PARAFAC model, *Behaviormetrika*, **21**, 45-63.

[79] 宮川雅巳，(1992) 交互作用要素に基づく多様性実験データの要因解析，応用統計学，**21**, 27-36.

[80] Mizuta, M. (1984) Generalized principal component analysis invariant under rotations of a coordinate system, *J. Japan Statist. Soc*. **14**, 1-9.

[81] Mulaik, S. A. (1986) Factor analysis and Psychometrika : major development, *Psychometrika*, **51**, 23-33.

[82] Murakami, T. (1983) Quasi three-mode principal component analysis, A method assesing the factor change, *Behaviormetrika*, **14**, 27-48.

[83] 村上 隆 (1990) 3相データの階層的主成分分析，柳井晴夫他編 (1990) 人間行動の計量分析 第10章，東大出版会.

[84] 中野純司，山本由和，岡田雅史，(1991) 知識ベース重回帰分析支援システム，応用統計学，**20**, 11-24.

[85] Nishisato, S. (1984) Forced clasification: A simple application of a quantification method, *Psychometrika*, **49**, 25-36.

[86] 小笠原春彦 (1993) 共分散構造分析の諸モデル，数理科学，**356**, 57-65.

[87] Okada A. (1987) Nonmetric multidimensional scaling of asymmetric proximities, *Behaviormetrika*, **21**, 81-96.

[88] Okamoto, M. (1972) Four techniques of principal component analysis, *J. Japan Statist. Soc*. **2**, 63-69.

[89] Okamoto, M. & Ihara, M. (1983) A new algorithm for the least-squares solution in factor analysis, *Psychometrika*, **48**, 597-605.

[90] Okamoto, M. & Ihara, M. (1984) Partial Gauss Newton algorithm for least-squares and maximum likelifood methods in factor analysis, *J. Japan Statist. Soc.*, **14**, 137-144.

[91] 丘本 正 (1987) 因子分析法の最近の発展，鈴木雪夫・竹内 啓編，社会科学の計量分析，東京大学出版会，7-26.

[92] 丘本 正 (1992 a) 数量化法第2類の人工的データ，日本統計学会誌 **22**, 95-102.

[93] 丘本 正 (1992 b) 数量化法第3類の人工的データ，行動計量学 **19**, 1, 75-82.

[94] 丘本 正 (1992 c) 数量化法第3類の諸問題，日本統計学会誌 **22**, 229-239.

[95] 大塩達一郎・長谷部文夫 (1991) 矩形診断法による日本文筆跡の計量と識別―第1部：筆跡の計量と筆記能力の発達：第2部：筆跡の鑑定，行動計量学 **18**, 2, 9-24, 25-34.

[96] Otsu, N. (1975) Nonlinear discriminant analysis as a natural extension of the linear case, *Behaviormetrika*, **2**, 45-59.

[97] Ramsay, J. O. (1977) Maximum likelihood estimation in multidimensional scaling, *Psychometrika*, **42**, 241-266.

[98] Ramsay, J. O., Berge, J. T. & Styan, G. P. H. (1984) Matrix correlation, *Psychometrika*, **49**, 403-425.

[99] Rao, C. R. (1964) The use and interpretation of principal component analysis, *Sankhya A*, **26**, 329-358.

[100] Rao, C. R. (1979) Separation theorem for singular values of matrices and their applications in multivariate analysis, *J. Multivar. Anal*. **9**, 362-377.

[101] Rao, C. R. & Yanai, H. (1979) General definition of a projector, its decom-position and application to statistical problems, *J. Statist. Planning and Inference*, 3, 1, 1-17.

[102] Rao, C. R. (1980) Matrix approximation and reduction of dimensionality in multivariate statistical analysis. *Multivariate Analysis V* (P. R. Krishnaiar, ed.) North-Holland. Amsterdam, 3-22.

[103] Rao, C. R. (1981) A lemma on g-inverse of a matrix and computation of correlation coefficient in the singular case, *Commun. Statist. A, Theory Methods* **10**, 1-10.

[104] Rao, C. R. (1983) Multivariate analysis. Some reminiscences on its origin and development. *Sankhya, B*, **45**, 284-299 (柳井晴夫，竹内 啓訳 (1983) 多変量解析―その起源と発展に関する回想―応用統計学 **12**, 69-78).

[105] Rao, C. R. & Rao, C. V. (1987) Stationary values of the product of two Releigh quatients: Homologous canonical correlation, *Sankhya, B* **2**, 113-125.

[106] Roff, M. (1936) Some properties of the communalities in multiple factor theory, *Psychometrika*, **1**, 1-6

[107] Saito, T., & Otsu (1988) A method of optimal scaling for multivariate ordinal data and its extensions, *Psychometrika*, **53**, 5-25.

[108] Saito, T., Kariya, T. & Otsu, T. (1988) A generalization of the principal component analysis, *J. of Japan Statist. Soc.,* **18**, 187-193.

[109] Saito, T. & Takeda, S. (1990) Multidimensional scaling of asymmetric proximity model and method, *Behaviormetrika*, **28**, 49-80.

[110] Saito, T. (1991) Analysis of asymmetric proximity matrix by a model of distance and additive terms, *Behaviormetrika*, **29**, 45-60.

[111] 佐藤俊哉(1985) 離散変数を用いた判別分析—RaoのV統計量による尺度化の試み，**行動計量学**，13, 1, 1-7.

[112] Sato, M. (1987) Pragmatic treatment of improper solutions in factor analysis, *Ann. of Inst. Statist. Math.,* **39**, 443-455.

[113] Sato, M. (1990) Some remarks on principal component analysis as a substitute for factor analysis in monofactor cases, *J. Japan Statist.Soc.,* **20**, 23-32.

[114] Sato, M. (1989) Some comment's on Shapiro's paper : Identifiablity of factor analysis, *Technical Report 249, Statistical Research Group,* Hiroshima Univ.

[115] Sato, M. (1992) A study of an identification problem and substitute use of principal component analysis in factor analysis, *Hiroshima Math. J.,* **22**, 479-524.

[116] Schönemann, P. H. (1970) On metric multidimensional unfolding, *Psychometrika*, **35**, 349-366.

[117] Shapiro, A. (1985) Identifiability of factor analysis :some results and open problems, *Linear Algebra and its Application*, **70**, 1-7.

[118] Shepard, R. N. (1962) The analysis of proximities : Multidimensional scaling with an unknown distance function, I& II, *Psychometrika,* **27**, 125-140, 219-246.

[119] 繁桝算男 (1990) カテゴリカルデータの因子分析 **行動計量学**，**18**, 41-51.

[120] Steiger, J. H. (1979) Factor indeterminacy in the 1930's and the 1970's : some interesting pararells, *Psychometrika*, **44**, 157-167.

[121] Stewart, D. & Love, W. (1968) A general canonical correlation index, *Psychometrika*, **70**, 160-163.

[122] Sugiyama, T.& Kurauchi, H. (1986) Identification of handwriting in Chinese characters using discriminant analysis, *Behaviormetrika*, **19**, 55-72.

[123] 高部啓子(1985) 着衣基体としての人体の形態類型化に関する研究(第2報)—判別分析による人体の形態類型化—**応用統計学**，**14**, 113-130.

[124] 高木広文 (1992) 統計パッケージ HALBAU，**計算機統計学**，**5**, 165-170.

[125] Takane , Y., Young, F. W. et al. (1977) Nonmetric individual differences multidimensional scaling with optimal scaling feature, *Psychometrika*, **42**, 7-67.

[126] Takane, Y., Young, F. W. et al. (1978) The principal component of mixed measurement level multivariate data : An alternating least squares method with optimal scaling feature, *Psychometrika*, **43**, 279-281.

[127] Takane, Y. (1978) A maximum likelihood method for nonmetric multidimensional scaliongg :1. The case in which all empirical pairwise orderings are independent-theory, *Jap. Psycholo. Res.*, 20 -1, 7-17.

[128] Takane, Y., Young, F. W. & de Leeuw , J. (1979) Nonmetric common factor analysis : An alternating least squares method with optimal scaling, *Behaviormetrika*, **6**, 45-56.

[129] Takane, Y. (1981) Multidimensional successive categories scaling : A maximum likelihood method, *Psychometrika*, **46**, 389-405.

[130] Takane, Y. & de Leeuw, J. (1987) On the relationship between item response theory and factor analysis, *Psychometrika*, **52**, 393-408.

[131] Takane, Y., Bozdogan, H. & Shibayama, T. (1987) Ideal point discriminant analysis, *Psychometrika*, **52**, 371-392.

[132] Takane, Y. & Shibayama, T. (1990) Principal component analysis with external criteria on both subjects and variables, *Psychometrika*, **56**, 97-120.

[133] Takane, Y. Yanai, H. & Mayekawa, S. (1991) Relationship among several methods of linearly constrained correspondence analysis, *Psychometrika*, **56**, 667-684.

[134] 高根芳雄 (1992) 制約付き主成分分析法について：**行動計量学**，19, 1, 29-39.

[135] 竹村彰通（1987） 多変量記述統計と標本分布論，鈴木雪夫・竹内 啓編 社会科学の計量分析 東京大学出版会，245－263.

[136] 竹内 啓（1986） 因子分析モデルにおける最尤推定量の構造について，応用統計学，**15**, 29-45.

[137] Tanaka, Y & Kodake, K. (1981) A method of variable selection in factor analysis and its numerical investigation, *Behaviormetrika*, **10**, 49-62.

[138] Tanaka, Y. & Tarumi, T (1986) Sensitivity analysis in Hayashi's second method of Quantification, *J. Japan Ststisti. Soc.* **16**, 44-60.

[139] Tanaka, Y. & Odaka, Y (1989) Influential observations in principal factor analysis, *Psychometrika*, **54**, 475-85.

[140] 田中 豊（1992） 多変量解析における感度分析，行動計量学，19, 1, 3-17.

[141] Ten Berge, J. M. F. (1977) Orthogonal Procrustes rotation for two or more matrices, *Psychometrika*, **42**, 267-276.

[142] Ten Berge, J. M. F., deLeeuw, J. and Kroonenberg, P. M. (1987) Some additional results on principal component analysis of three-mode data by means of alternating least squares algorithm. *Psychometrika*, **32**, 183-191.

[143] ter Braak, C. J. F. (1986) Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis, *Ecology*, **67**, 1167-1179. 143).

[144] Timm, N. H. & Carlson, J. E. (1976) Part and bipartial canonical correlation analysis, *Psychometrika*, **41**, 159-176.

[145] Tsujitani, M. (1988) Optimal scaling for association models when category scores have a natural ordering, *Statist. & Prob. Letters* **6**, 175-180.

[146] Tsujitani, T. (1992) Diagnostic for association models for the analysis of cross-classifications having ordered categories, *J. Japan Statist. Soc.*, 22, 1, 19-32.

[147] Tucker, L. R. (1966) Some mathematical notes on three-mode factor analysis. *Psychometrika*, **31**, 279-311.

[148] Tukey, J. W. (1962) The future of data analysis, *Ann. Math.Statist.* **33**, 1-67.

[149] 津村善郎（1971） 因子分析は有用か，科学，**41**, 437-441.

[150] Tumura, Y. & Sato, M. (1980) On the identification in factor analysis, *TRU Math*, **16**, 121-131.

[151] van de Geer, J. P. (1984) Linear relations among k sets of variables, *Psychometrika*, **49**, 79-94.

[152] van den Wollenberg, A. L.(1977) Redundancy Analysis :An alternative for canonical correlation analysis, *Psychometrika*, **42**, 207-219.

[153] van Driel, O. P. (1978) On various causes of improper solutions in maximum likelihood factor analysis, *Psychometrika*, **43**, 225-243.

[154] van Ness, J. W. (1979) On the effect of dimension in discriminant analysis for unqual covariance population, *Technometrics*, **21**, 119-127.

[155] Yanai, H. (1970) Factor analysis with external criteria. *J. of Jap. Psycho. Res.*, **12**, 143-153.

[156] Yanai, H. (1973) Aptitude diagnosis for various courses of university, *Jap. J. of Psycholo. Res.*, **14**, 190-204.

[157] Yanai, H. (1980) A proposition of generalized method for foward selection of variables, *Behaviormetrika*, **7**, 95-107.

[158] Yanai, H. (1981) Explicit expressions of projectors on canonical variables and distances between centroids of groups; *J. Japan Statist. Soc.* **11**, 43-53.

[159] 柳井晴夫他（1982） 多変量解析法（その3）公衆衛生学・疫学への適用例をめぐって一日本公衆衛生誌，**47**, 744-751.

[160] 柳井晴夫・市川雅教（1985） 多変量解析とパソコンプログラム，心理学評論，**28**, 392-403.

[161] Yanai, H. (1986) Some generalization of correspondence analysis in terms of projection operators, *(Diday Ed). Data Analysis and Informatics*, IV North-Holland, 193-207.

[162] Yanai, H. & Mukherjee, B. N. (1987) A generalized method of image analysis from an intercorrelation matrix which may be singular, *Psychometrika*, **52**, 555-564.

[163] 柳井晴夫・柏木繁男・国生理枝子（1987） プロマックス回転法における新性格検査の作成について，心理学研究，**58**, 158-165.

[164] Yanai, H. & Ichikawa, M. (1990) New lower and upper bounds for communality in factor analysis, *Psychometrika*, **55**, 405-410.

[165] 柳井晴夫他 (1991) 多変量解析における潜在変数モデルの理論と応用 統計数理, 39-1, 97-124.
[166] Yanai, H. & Takane, Y. (1992) Canonical correlation with linear constraints, *Linear Algebra & its Applications* **176**, 75-89.
[167] Yanai, H. & Puntanen, S. (1993) Partial canonical correlation associated with symmetric reflexive generalized inverses of the dispersion matrix, *Matusita, K, et al.(Ed) Statistical Sciences and Data Analysis ; Proceedings of the Third Pacific Area Statist. Conference, VS, International Sciences Publishers, Zeist (Netherlands)* (in press).
[168] 吉沢 正 (1975) 分割表における数量化モデル—その理論的検討, 行動計量学, 3, 1, 1-11.
[169] 吉沢 正 (1976) 交互作用概念の一般化と多重配列の特異値分解, 行動計量学, 4, 1, 32-43.
[170] 吉沢正他 (1992) 高度な相互関連性を持つ多重配列データの新しい解析システムの開発 平成1, 2, 3年度科学研究費用補助金研究成果報告書.
[171] Wang Song-Gui & Chow Shein-Chung (1987) Some results on canonical correlations and measures of multivariate association. *Commun. Statist, Theory and Methods,* **16**, 339-351.

第III部 多変量解析に関する専門書とプロシーディングス

藤越康祝・柳井晴夫

最近, 多変量解析の理論と応用は著しく発展している. これに伴って, この方面の書物の発刊が増加し, とくに, 1970年代には加速度的な増加があり, また, 1980年以降も引き続き多くの書物が発刊されている. この増加は多変量解析の個々分野において, それぞれ独立した専門書が発刊されるようになったことによるものである. この第III部においては, 多変量解析に関する専門書とプロシーディングスをリストしている. 専門書は全般的なものと分野別のものとに分け, 後者は分野毎に分類している. なお, プロシーディングスの中で, 本の色彩が強いものは専門書として扱っている.

専門書:

(1) 全 般

塩谷 実, 浅野長一郎 (1966). 多変量解析論, 共立出版.
伊藤孝一 (1969). 多変量解析の理論, 培風館.
竹内 啓, 柳井晴夫 (1972). 多変量解析の基礎, 東洋経済新報社.
河口至商 (1973). 多変量解析入門I, 森北出版.
後藤昌司 (1973). 多変量データの解析法, 科学情報社.
芝 祐順 (1975). 行動科学における相関分析法 (第2版), 東大出版会. (初版1967).
奥野忠一他 (1976). 続多変量解析法, 日科技連出版.
河口至商 (1978). 多変量解析入門II, 森北出版.
奥野忠一他 (1981). 多変量解析法 (第2版), 日科技連出版. (初版1971).
田中 豊, 脇本和昌 (1983). 多変量統計解析法, 現代数学社.
田中 豊, 垂水共之, 脇本和昌 (1984). パソコン統計解析ハンドブックII—多変量解析編, 共立出版.
柳井晴夫, 高根芳雄 (1985). 新版多変量解析法, 朝倉書店. (初版1976).
柳井晴夫, 高木広文編 (1986). 多変量解析ハンドブック, 現代数学社.
鈴木雪夫, 竹内 啓編 (1987). 社会科学の計量分析—多変量解析の理論と応用—, 東大出版会.
鷲尾泰俊, 大橋靖雄 (1989). 多次元データの解析, 岩波書店.
柳井晴夫, 岩坪秀一, 石塚智一編 (1990). 人間行動の計量分析—多変量解析の理論と応用—, 東大出版会.
塩谷 実 (1990). 多変量解析概論, 朝倉書店.
竹村彰通 (1991). 多変量推測統計の基礎, 共立出版.
Rao, C. R. (1952). Advanced Statistical Methods in Biometric Research, John Wiley & Sons.
Roy. S. N. (1957). Some Aspects of Multivariate Analysis, John Wiley & Sons.
Cooley, W. W. and Lohnes, P. R. (1962). Multivariate Procedure for the Behavioral Sciences, John Wiley & Sons.
Seal, H. (1964). Multivariate Statistical Analysis for Biologist, Methuen (塩谷 実訳 (1970). 多変量解

析入門—生物学を題材にして—,共立出版).
Kendall, M. G. (1968). A Course in Multivariate Analysis, Charles Griffin (浦 昭二,竹並輝之訳 (1972). 多変量解析の基礎,サイエンス社).
Dempster, A. P. (1969). Elements of Continuous Multivariate Analysis, Addison Wesley.
Cooley, W. W. and Lohnes, P. R. (1971). Multivariate Data Analysis (2nd ed.), John Wiley & Sons (井口晴弘,藤沢武久,守谷栄一訳 (1973). 行動科学のための多変量解析,鹿島出版会).
Roy, S. N., Gnanadesikan, R., and Srivastava, J. N. (1971). Analysis and Design of Certain Quantitative Multi-Response Experiments, Pergamon.
Kshirsagar, A. M. (1972). Multivariate Analysis, Marcel Dekker.
Rao, C. R. (1973). Linear Statistical Inference and Its Applications, John Wiley & Sons. (1st ed., 1965).
Finn, J. D. (1974). A General Model for Multivariate Analysis. New York : Holt, Rinehart and Winston.
Bock, R. D. (1975). Multivariate Statistical Methods in Behavioral Science, McGraw-Hill.
Kendall, M. (1975). Multivariate Analysis (2nd ed.), Charles Griffin (奥野忠一,大橋靖雄訳 (1981). 多変量解析,培風館). (1st ed., 1957).
Harris, R. J. (1975). A Primer of Multivariate Statistics, Academic Press.
Timm, N. H. (1975). Multivariate Analysis with Applications in Education and Psychology, Brooks/Cole Publishing Company.
Gnanadesikan, R. (1977). Methods for Statistical Data Analysis, John Wiley & Sons (丘本 正,磯貝恭史訳 (1979). 統計的多変量データ解析,日科技連).
Giri, N. C. (1977). Multivariate Statistical Inference, Academic Press.
Maxwell, A. E. (1977). Multivariate Analysis in Behavioral Research, Halsted Press.
Srivastava, M. S., and Khatri, C. G. (1979). An Introduction to Multivariate Analysis, North-Holland.
Mardia, K. B., Kent, J. T. and Bibby, J. M. (1979). Multivariate Analysis, Academic Press.
Chaffield, C. and Collins, A. J. (1980). Introduction to Multivariate Analysis, Chapman and Hall (福場庸,大沢 豊,田畑吉雄訳 (1986). 多変量解析入門,培風館).
Arnold, S. F. (1981). Theory of Linear Models and Multivariate Analysis, John Wiley & Sons.
Barnet, V. (1981). Interpreting Multivariate Data, John Wiley & Sons.
Fornell, C. (1982). A Second Generation of Multivariate Data Analysis, Praeger Publisher.
Hawkins, D. M. (ed.) (1982). Topics in Applied Multivariate Analysis, Cambridge Univ. Press (医学統計研究会訳 (1988). 多変量解析の理論と実際,MPC).
Muirhead, R. J. (1982). Aspects of Multivariate Statistical Theory, John Wiley & Sons.
Press, S. J. (1982). Applied Multivariate Analysis, R. E. Krieger.
Takeuchi K., Yanai H. and Mukherjee, B. N. (1982). The Foundation of Multivariate Analysis, Wiley Eastern (Halested Press).
Eaton, M. L. (1983). Multivariate Statistics: A Vector Space Approach, John Wiley & Sons.
Srivastava, M. S. and Carter, E. M. (1983). An Introduction to Applied Multivariate Statistics, North-Holland.
Anderson, T. W. (1984). An Introduction to Multivariate Statistical Analysis (2nd ed.), John Wiley & Sons. (1st ed., 1958).
Chatfield, C. and Colins, A. (1984). Introduction to Multivariate Analysis, Chapman and Hall (福場 庸他訳 (1986). 多変量解析入門,培風館).
Dillon, W. R. and Goldstein, M. (1984). Multivariate Analysis: Methods and Applications, John Wiley & Sons.
Seber, F. G. A. (1984). Multivariate Observations, John Wiley & Sons.
Siotani, M., Hayakawa, T. and Fujikoshi, Y. (1985). Modern Multivariate Statistical Analysis, American Science Press.
Manly, B. F. J. (1986). Multivariate Statistical Methods: A Primer, Chapman and Hall (村上正康,田栗正章 (1992). 多変量解析の基礎,培風館).
Krzanowski, W. J. (1988). Principle of Multivariate Analysis: A User's Perspectives, Oxford Science Publications.
Flury, B., and Riedwyl, H. (1988). Multivariate Statistics: A Practical Approach, Chapman and Hall.
Tatsuoka, M. M. (1988). Multivariate Data Analysis: Techniques for Educational and Psychological Research (2nd ed.), John Wiley & Sons. (1st ed., 1971).

Gifi, A. (1989). Monlinear Multivariate Analysis, John Wiley & Sons.
Morrison, D. F. (1990). Multivariate Statistical Methods (3rd ed.), McGraw-Hill.
Jambu, M. (1991). Exploratory and Multivariate Data Analysis, Academic Press.
Fang, K. T., and Zhang, Y. T. (1992). Generalized Multivariate Analysis, Springer-Verlag.
（2） 多変量線形モデル
刈屋武昭 (1979). 回帰分析の理論，岩波書店.
Finn, J. D. (1972). Multivariance: Univariate and Multivariate Analysis of Variance, Covariance and Regression, National Education Resources.
Belsley, D. A., Kuh, E. and Welsh, R. E. (1980). Regression Diagnostic: Identifying Influential Data and Sources of Collinearity, Jhon Wiley & Sons.
Krishnaiah, P. R. (ed.) (1980). Handbook of Statistics, Vol. 1: Analysis of Variance, North-Halland.
Kariya, T. (1985). Testing in the Multivariate General Linear Model, Kinokuniya.
Hand, D. J., and Taylor, C. C. (1987). Multivariate Analysis of Variance and Repeated Measures, Chapman and Hall.
Crower, M. J., and Hand, D. J. (1990). Analysis of Repeated Measures, Chapman and Hall.
（3） 潜在変数分析（因子分析・主成分分析・共分散構造分析・項目反応理論）
浅野長一郎 (1971). 因子分析法通論，共立出版.
芝 祐順 (1979). 因子分析法（改訂版），東大出版会.
芳賀敏郎，橋本茂司 (1980). 回帰分析と主成分分析，日科技連出版.
丘本 正 (1986). 因子分析の基礎，日科技連出版.
柳井晴夫，繁桝算男，前川眞一，市川雅教 (1990). 因子分析—その理論と方法—，東大出版会.
芝 祐順 (1991). 項目反応理論—基礎と応用—，東大出版会.
豊田秀樹 (1992). SAS による共分散構造分析，東大出版会.
Thurstone, L. L. (1947). Multiple Factor Analysis, University of Chicago Press.
Harman, H. H. (1976). Modern Factor Analysis (3rd ed.), University of Chicago Press.
Horst, P. (1965). Factor Analysis of Data Matrices, Halt Linehart Winston（柏木繁男他訳 (1978). コンピュータによる因子分析法，科学技術出版社).
Lazasfeld, P. F. and Henry, N. W. (1967). Latent Structure Analysis, Houghton Mifflin Company.
Rummel, R. J. (1970). Applied Factor Analysis, Northwestern University Press.
Lawley, D. N. and Maxell, A. E. (1971). Factor Analysis as a Statistical Method (2nd ed.), Butterworth (1st ed., 1963;丘本 正監訳 (1970). 因子分析法，日科技連出版).
Mulaik, S. A. (1972). The Foundation of Factor Analysis, McGraw-Hill.
Comrey, A. L. (1973). A First course in Factor Analysis, Academic Press（芝 祐順訳 (1979). 因子分析入門，サイエンス社).
Jöreskog, K. G. and Sörbom, D. (1979). Advances in Factor Analysis and Structure Equation Models, Abt Books.
Everitt, B. S. (1984). An Introduction to Latent Variable Models, Chapman and Hall.
Hambleton, R. K. and Swaminathan, H. (1985). Item Response Theory-Principle and Applications, Kluwer Nijhoff Publishing.
MacDonald, R. P. (1985). Factor Analysis and Related Methods, Lawrence Erlbaum Associates.
Jolliffe, I. T. (1986). Principal Component Analysis, Springer-Verlag.
Fuller, W. A. (1987). Measurement Error Models, John Wiley & Sons.
Bartholomew D. J. (1987). Latent Variable Models and Factor Analysis, Charles Griffin.
Flury, B. (1988). Common Principal Component and Related Multivariate Models, John Wiley & Sons.
Bollen, K. A. (1989). Structural Equations with Latent Variables, John Wiley & Sons.
Lohmoler, J. B. (1989). Latent Variable Path Modelling with Partial Least Squares, Physica-Verlag.
Jackson, J.E. (1991). A User's Guide to Principal Component Analysis, Wiley Interscience.
（4） 数量化関連（質的データの構造解析）
林知己夫，樋口伊佐夫，駒沢 勉 (1970). 情報処理と統計数理，産業図書.
林知己夫 (1974). 数量化の方法，東洋経済新報社.
林知己夫 (1977). データ解析の方法，東洋経済新報社.
西里静彦 (1982). 質的データの数量化，朝倉書店.
駒沢 勉 (1982). 数量化理論とデータ処理，朝倉書店.

坂元慶行（1985）．カテゴリカルデータのモデル分析，共立出版．
岩坪秀一（1987）．数量化の基礎，朝倉書店．
駒沢　勉他（1990）．パソコン数量化分析，朝倉書店．
Benzecri, J. P. (1973). L'Analyse des Donnes, Tom 2, Dunod.
Cailliez, F. and Pages, J. P. (1976). Introduction a l'Analyse des Donnees, SMASH.
Nishisato, S. (1980). Analysis of Categorical Data: Dual Scaling and Its Application, University of Toront Press.
Agresti, A. (1984). Theory and Application of Correspondence Analysis, Academic Press.
Greenacre, M. J. (1984). Theory and Application of Correspondence Analysis, Academic Press.
Lebart, L., Morineau, A. K. and Warwick, M. (1984). Multivariate Descriptive Statistical Analysis, John Wiley & Sons.
van Rijckevorsel and de Leeuw, J. (1988). Component and Correspondence Analysis, John Wiley & Sons.
（5）離散データの多変量解析
柳川　堯（1986）．離散多変量データ解析，共立出版．
松田紀之（1987）．質的情報の多変量解析，朝倉書店．
Cox, D. R. and Snell, E. J. (1989). Analysis of Binary Data, (2nd ed.) Chapmann and Hall（1st ed., 1970：後藤昌司他訳（1980），2値データの解析，朝倉書店）．
Bishop, Y. M. M., Fienberg, S. E. and Holland, P. (1975). Discrete Multivariate Analysis: Theory and Practice, The MIT Press.
Haberman, S. J. (1978). Analysis of Qualitative Data, Academic Press.
Read, T. R. C., and Cressie, N. A. C. (1988). Goodness-of-Fit Statistics for DIscrete Multivariate Data, Springer-Verlag.
（6）多次元尺度法
林知己男，飽戸　弘編（1976）．多次元尺度法，サイエンス社．
高根芳雄（1980）．多次元尺度法，東大出版会．
斎藤堯幸（1980）．多次元尺度構成法，朝倉書店．
Tongerson, W. S. (1958). Theory and Methods of Scaling, John Wiley & Sons.
Kruskal, J. B. and Wish, M. (1978). Multidimensional Scaling, Saga Publications（高根芳雄訳（1980）．多次元尺度法，朝倉書店）．
Davison, M. L. (1983). Multidimensional Scaling, John Wiely & Sons.
Borg, I. and Lingoes, J. (1987). Multidimensional Similarity Structure Analysis, Springer-Verlag.
Arabie, P. Carroll and D. W. S. (1987). Three-way Scaling and Clustering, Sage Publications（岡太彬訓，今泉　忠訳（1990）．3元データの分析，共立出版会）．
（7）判別分析・クラスター分析
大隅　昇（1979）．データ解析と管理技法，朝倉書店．
Sokal, P. R. and Sneath, P. H. (1963). Principle of Numerical Taxonomy, Freeman.
Everitt, B. (1974). Cluster Analysis, Heinemann Educational Books.
Hartigan, J. A. (1975). Clustering Algorithms, John Wiley & Sons（西田春彦，吉田光雄，平松　闊，田中邦夫訳（1978）．クラスター分析，日本コンピューター協会）．
Lachenbruch, P, A. (1975). Discriminant Analysis, Hafner Press（鈴木義一郎，三宅章彦訳（1979）．判別分析，現代数学社）．
Goldstein, M., and Dillon, W. R. (1978). Discrete Discriminant Analysis, John Wiley & Sons.
Bezdek, J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press.
Hand, D. J. (1981). Discrimination and Classification, John Wiley & Sons.
Krishnaiah, R. P., and Kanal, L. N. (eds.) (1982). Handbook of Statistics (Vol. 2): Classification, Pattern Recognition and Reduction of Dimensionality, North-Holland.
McLachlan, G. J. (1992). Discriminant Analysis and Statistical Pattern Recognition, John Wiley & Sons.
（8）正準分析
Gittins, R. (1980). Canonical Analysis: A Review with Applications in Ecology. Spring-verlag.
Levine, M. S. (1977). Canonical Analysis and Factor Comparison, SAGA publication（柳井晴夫，新田祐史訳（1984）．多変量相関分析の方法，朝倉書店）．
（9）ロバストネス
Hampel, F. R., Ronchelti, E. M., Rousseeuw, P. J., and Stahel, W. A. (1986). Robust Statistics: The

Approach based on Influence Function, John Wiley & Sons.
Kariya, T., and Sinha, B. K. (1989). Robustness of Statistical Tests, Academic Press.

(10) 多変量順序制約

Barlow, R. E., Bartholomew, D. J., Bremner, J. M., and Burnk, H. D. (1972). The Theory and Application of Isotonic Regression, John Wiley & Sons.
Robertson, T., Wright, F. T., and Dykstra, R. L. (1988). Order Restricted Statistical Inference, John Wiley & Sons.

(11) 多変量解析の数学理論

Farrel, R. H. (1976). Techniques of Multivariate Calculation, Lecture Notes in Mathematics, Springer-Verlag.
Farrel, R. H. (1985). Multivariate Calculation : Use of the Continuous Groups, Springer-Verlag.
Magnus, J. R. and Neudecker, H. (1988). Matrix Differential Calculus with Applications in Statistics and Economics, John Wiley & Sons.

(12) 多変量分布

Miller, K. S. (1964). Multidimensional Gaussian Distributions, John Wiley & Sons.
Johnson, N. L., and Kotz, S. (1972). Distributions in Statistics : Continuous Multivariate Distributions, John Wiley & Sons.
Tong, Y. L. (1990). The Multivariate Normal Distribution, Springer-Verlag.
Fang, K.-T., Kotz, S., and Kai, W. N. (1992), Symmetric Multivariate and Related Distributions, Chapman and Hall.

(13) 多変量ノンパラメトリック法

Puri, M. L., and Sen, P. K. (1971). Nonparametric Methods in Multivariate Analysis, John Wiley & Sons.
Puri, M. L., and Sen, P. K. (1985). Nonparametric Methods in General Linear Model, John Wiley & Sons.
Scott, D. W. (1992). Multivariate Density Estimation, John Wiley & Sons.

(14) 方向性データ解析

Mardia, K. V. (1972). Statistics of Directional Data, Academic Press.
Watson, G. S. (1983). Statistics on Spheres, John Wiley & Sons.

(15) 多変量グラフ解析

脇本和昌，後藤昌司，松原義弘 (1983). 多変量グラフ解析，朝倉書店.
Everitt, B. (1978). Graphical Techniques for Multivariate Data; Hainemann Educational Books LTD (医学統計研究会全訳 (1982). 多変量グラフィカル表現法，MPC).
Whittaker, J. (1990). Graphical Models in Applied Mathematical Multivariate Statistics, John Willey & Sons.

(16) 多変量解析の各分野への応用

高橋晄正編 (1969). 計量診断学，東大出版会.
本田正久，島田一明 (1977). 経営者のための多変量解析法，産業能率大学出版.
奥野忠一，山田文道 (1978). 情報化時代の経営分析，東大出版会.
安田三郎，海野道郎 (1979). 社会統計学 (改訂２版)，丸善出版.
古川俊之 (1982). コンピュータ診断，共立出版.
林知己夫編 (1984). 健康管理の計量化，共立出版.
奥野忠一他 (1986). 工業における多変量データの解析，日科技連.
林知己夫，鈴木達三 (1986). 社会調査と数量化一国際比較におけるデータ解析一，岩波書店.
大橋靖雄，市川伸一 (1987). SAS によるデータ解析入門，東大出版会.
林知己夫，水野欽司編 (1990). 計量生物学・行動計量学，放送大学教育振興会.
浜島信之 (1990). 多変量解析による臨床研究，名古屋大学出版会.
吉沢　正，芳賀敏郎 (1992). 多変量解析事例集，日本科学技術連盟.
渡部　洋編 (1992). 心理教育のための多変量解析一事例編，福村出版.
Hubert, L. J. (1987). Multivariate Interpretation of Clinical Laboratory Data, Marcel Dekker.
Van Vark, G. N. and Howells, W.W. (1984). Multivariate Statistical Methods in Physical Anthropology, Reidel.
Hayashi, C., Suzuki, T. and Sasaki, M. (1992). Data Analysis for Comparative Social Research : International Perspectives, North-Holland Publishing Company.

(17) その他

竹内 啓他編 (1989). 統計学辞典, 東洋経済新報社.

Gale W. A. (1986). Artificial Intelligence and Statistics, Addison-Wesley Rublishing Company.

Johnson, M. E. (1987). Multivariate Statistical Simulation, John Wiley & Sons.

Stone, M. (1987). Cordinate-Free Multivariate Statistics: An Illustrated Geometric Progression from Halmos to Gauss and Bayes, Oxford University Press.

Chatterjee, S. and Hadi, A. S. (1987). Sensitivity Analysis in Linear Regression, Wiley Interscience.

Magnus, J. R. and Neudecker, H. (1988). Matrix Differential Calculas with Applications in Statistics and Econometrics, John Wiley & Sons.

プロシーディングス:

Krishnaiah, P. R. (ed.) (1966). Multivariate Analysis, Academic Press.

Krishnaiah, P. R. (ed.) (1969). Multivariate Analysis-II, Academic Press.

Krishnaiah, P. R. (ed.) (1973). Multivariate Analysis-III, Academic Press.

Kabe, D. G. and Gupta, R. P. (eds.) (1973). Multivariate Statistical Inference, North-Holland Publishing Company.

Krishnaiah, P. R. (ed.) (1977). Multivariate Analysis-IV, North-Holland Publishing Company.

Wang, P. C. C. (ed.) (1978). Graphical Representation of Multivariate Data, Academic Press.

Diday, E. et al. (ed.) (1979). Data Analysis and Informatics II, North-Holland.

Krishnaiah, P. R. (ed.) (1980). Multivariate Analysis-V, North-Holland Publishing Company.

Gupta, R. P. (ed.) (1980). Multivariate Statistical Analysis, North-Holland Publishing Company.

Diday, E. et al. (ed.) (1983). Data Analysis and Informatics III, Elsevier Science Publisher B. V.

Krishnaiah, P. R. (ed.) (1985). Multivariate Analysis-VI, Elsevier Science Publishers B. V.

Diday, E. et al. (ed.) (1986). Data Analysis and Informatics IV, Elsevier Science Publisher B. V.

Gupta, A. K. (ed.) (1987). Advances in Multivariate Statistical Analysis, Dordrecht Reidel Publishing Company.

Bozdogan, H. and Gupta, A. K. (eds.) (1987). Multivariate Statistical Modeling and Data Analysis, Dordrecht Reidel Publishing Company.

Bock, H. H. (ed.) (1988). Classification and Related Methods of Data Analysis, North-Holland Publishing Company.

Das Gupta, S. and Ghosh, J. K. (1988). Advances in Multivariate Statistical Analysis, Indian Statistical Institute.

Diday, E. et al. (ed.) (1988). Data Analysis and Informatics V, Elsevier Science Publisher B. V.

Hayashi, C., Diday, E., Jambu, M. and Ohsumi, N. (1988). Recent Development in Clustering and Data Analysis, Academic Press.

Coppi, R. and Bolasco, (ed.) (1989). Multiway Data Analysis, North-Holland Publishing Company.

Rao, C. R. and Rao, M. M. (eds.) (1989). Multivariate Statistics and Probability, Academic Press.

Fang, K.-T. and Anderson, T. W. (eds.) (1990). Statistical Inference in Elliptical Contoured and Related Distributions, Allerton Press Inc.