

多重配列データと母数空間の構造

1977年9月16日

吉澤 正

山梨大学工学部計算機科学科

Tadashi Yoshizawa

Structure of Multiarray Data and Population Spaces

Ph.D. Thesis

Tōkyō University

gen-abstract

は し か き

本論は、多重配列データと母数空間の構造に関する研究をまとめ、東京大学学位審査のために提出するものである。

本研究は、一応、統計学の範疇に属するものであるが、統計学の各種の立場や応用分野の見られる中で、統計情報工学ともいべき枠組の中で進められているものである。

序論の1節「研究の背景」で述べるように、情報工学や行動諸科学といった学際的領域の発展と歩調を合わせて、各種の現象解析あるいは情報解析の統計的接近法の新らかな開発が必要になっている。例えば、従来の量的データ中心の解析手法に対して質的データの解析、連続量に対して離散データの解析、実験データや調査データに対し無作為化などの統計的管理の欠けた状況で得られる観察データの解析などがあげられよう。そこでは解析すべきデータの性格や構造の多様化がみられ、従来の統計学の再検討がその基盤からせまられている。

一般的にいて、データ解析のもっとも重要な目標の一つは、その現象と解明する確率モデルを構成することであろう。その目標のためには、誤差やバラツキの発生源に及ぼす確率モデルの妥当性の検討や、モデル構成の基本となる母数空間の構造の追求が必要となろう。同

時に、大量かつ多様な性格や構造を持つデータに対しては、その収集・検査・蓄積から各種の解析に至る幅広いデータ処理としての側面、すなわち、電子計算機の利用を前提としてのデータ処理システムや使用する算法の効率の研究も無視できない。

そのような立場から、数値計算における丸め誤差の確率モデル、粒子の混合問題における確率モデルとその応用、多重分割表での記述的数量化手法に対する確率モデルによる基礎付け、母数空間の構造の幾何学接近法による基礎的研究、多重分割表などの多重配列構造を持つデータと母数空間の研究、さらには、電子計算機による統計計算やデータ解析システムに関わる研究等を行ってきた。

ここでは、多重配列の分解、とくに特異値分解の理論と多重分割表などの解析への応用、ならびに母数空間の構造への幾何学的接近法による研究について一応の成果が得られたと考え、「多重配列データと母数空間の構造」と題して報告する次第である。

1977年9月

吉澤 正

目 次

はしがき

1. 序論

- 1.1 研究の背景
- 1.2 研究の概要
- 1.3 従来の研究及び文献
- 2. 多重配列構造を持つ線形母数空間
 - 2.1 効果・交互作用
 - 2.2 階層的モデル
 - 2.3 内積の一般化
- 3. 多重配列の特異値分解
 - 3.1 算法
 - 3.2 特異値分解の性質
- 4. 多重分割表への応用 (I)
 - 4.1 加法モデルと対数線形モデル
 - 4.2 平方根モデル
- 5. 多重分割表への応用 (II)
 - 5.1 2重分割表における数量化
 - 5.2 3重分割表における数量化とモデル
 - 5.3 多重分割表における数量化
 - 5.4 平方根分解
 - 5.5 数値例
- 6. 多重配列の混合型分解
 - 6.1 混合型分解
 - 6.2 数値例
- 7. あとがき

謝辞

参考文献

付録

1. 序 論

1.1 研究の背景

何らかの現象を統計的現象として解析することの重要な目標の一つは、その現象の特性を一つのあるいは複数の確率変数として捉え、その確率変数の分布を把握することであろう。その目標は追求対象である現象を統計的あるいは確率的現象として説明する確率モデルを構成することともいえよう。

統計学の歴史から見れば、19世紀初頭以来、現象解析の統計的接近法の根本思想は大量データの観察によって現象の統計的規則性を見出そうという点にある。しかしながら、一つの現象を解析するため一般的には多数の特性を同時にかつ多次元的に考慮する必要がある、これらの特性の多次元分布を定めることは限られた大きさの標本をもってしては困難であり、場合によっては不可能に近い。

そこで、20世紀以後の統計学では、確率分布や標本分布論から始まり、実験計画法や標本調査法の進歩を伴って、いわゆる推測統計学、さらにはベイズ流推測等の統計的決定理論の発展をみた。そこでは、有限の大きさの標本にもとづく統計的推測の基礎を近代的確率論に求めるために、実験計画においてはいわゆる

Fisher の三原則: 反復 (replication); 無作為化 (randomization); 局所管理 (local control) を, 標本調査法においては何らかの無作為抽出 (random sampling) をデータ収集の基盤とする。さらに, 解析の対象とする母集団の特性に対してある枠内の分布族を想定し, 上で, その分布族に対応する母数 (parameters) についてのモデルあるいは仮説についての統計的推測を行うという形式を採用する。

一方, 従来から因子分析法や多次元尺度構成法など独自の手法を發展させてきた心理学, 社会学, 疫学等の行動諸科学をはじめ, 現象解析への統計的接近の必要な分野が広まるにつれ, 上述の無作為化や無作為抽出の原則に従うとはみなせないデータ, しばしば実験データや調査データに対して観察 (observational) データと呼ばれるものの解析, あるいは量的データに対して質的カテゴリカルなデータの解析, 量的であっても離散的データの解析が必要になっている。

これらのデータ解析の場合では, 今後解析あるいは發展させべきものとして, 以下に述べるような諸問題が生じていると考えられる。

観察データの解析に確率論を応用する根拠をどこに求めるか。観察データを不完全な無作為抽出によるデータとみるなら, 完全な無作為抽出の

場合と異なり、ある種の偏りを持つデータとなろう。その偏りをいかにとらえるかも重要な問題であろう。

調査データを含めて、質的データの典型的場合である多重分割表の解析とそのモデル構成をいかに進めるか。分散分析における母数モデルからの類推のような形で対数線形モデルや加減モデルの研究はある程度の成果をあげている。さらにデータの性格やデータの抽出過程の現実的多様性に応じた手法やその頑健性の研究が望まれる。分散分析の変量モデルに対応する多重分割表のモデルとその解析もその一つである。筆者等は多種粒子の混合問題での不完全混合状態での確率モデルの構成から出発して、多重分割表の変量モデルを考えている。そのモデルは品質管理での不完全管理状態、例えばロット内でも製品の不良率にムラがあるような状態の取り扱いにも通じるものがある。

上記の多重分割表をはじめ、一般にコンジョイント・メジャメント (conjoint measurement) データなどと呼ばれる多重配列構造を持つデータの解析手法をいかに発展させるか。これに対して筆者は多重配列の特異値分解を考え、多重分割表の解析に応用している。行列の特異値分解が多変量解析において重要な役割を果たしている

のに対して、多重配列データの解析にその特異値分解が理論的基礎としても有効に利用されることが期待される。

主成分分析や回帰分析など従来の多変量解析手法では、いわゆるプロフィール (profile) データ、すなわち、多数の特性についての観測対象集合に関する行列構造のデータを抱えるものであったが、上述の多重配列構造をはじめとして、各種の構造や性格を持つデータが出現している。例えば、枝分かれあるいは木構造を持つデータや、対象間の類似性も表わす近接性 (proximity) データ、対象間の優劣関係も表現する優越 (dominance) データなど。このようなデータも対応とする数量化手法、最適尺度法、一対比較法などを含めて広い意味での多次元尺度構成法の現状は満足できるものではない。それらの手法は全般的にデータの記述的分析の段階にとどまっており、それらの有効性や適用の限界も、それぞれの分野の実質科学的な経験の累積によって明らかにするとともに、実質科学に即した確率モデルの構成を通じて統計的推測をも可能にすることが必要とされていくことが不可欠であろう。

特性の分布型を追求すること。従来の統計手法や理論は連続量なら正規分布、離散量ならポアソン分布や二項分布というように、特

性の分布を比較的簡単なものに仮定した上で作られているものが多く、それらが形式的に利用されるという弊害も生じている。ある確率モデルが考えられるとき、そのモデルでの仮定から導かれる結果としての統計量がその現象を表面的にはよく説明するところがあるだろう。しかし、それではそのモデルの妥当性が保証されるわけではなく、別のモデルでも同じ現象を説明できるかもしれない。確率モデルについては、モデルの根本的仮定に立ち戻って徹底的に検討し追求する必要がある。つまり、誤差やバラツキの発生源にできるだけさかのぼってモデルの妥当性を、諸特性の分布を追求することによって、検討することが望まれる。物理現象や筆者なども研究している数値計算での丸め誤差の現象などでは、そのような検討が比較的容易に行えるところがあるが、それぞれの現象でもモデルの妥当性の追求を避けるわけにはいかない。

最後に、データ処理としての側面に触れておこう。大量かつ構造的にも複雑化したデータを扱うには、電子計算機を用いた良質のデータ処理、あるいはデータ解析のシステムが不可欠である。通常のデータ処理におけるデータの検査、データファイルの更新、分類(sorting)、報告書作成などのほかに、解析に用いられる算法の効率と異常データの統計

的検出が重要課題として指摘されよう。例えば多重配列データの解析や多次元尺度構成法などでの算法はプログラムの良し悪しによって著しく処理時間が異なる。異常データへの関心はその存在が以後の解析に影響も及ぼすためだけではない。従来の統計学が何々を捨象して集団の統計的性格の分析を主目的としたのに対し、集団の中での何の相対的立場をみる必要が高まっていることによろう。集団の中で変わって異なるものとしての異常データの検出とその適切な取り扱いはその意味で重要である。本論では、確率モデルからのずれ、あるいはモデルの不適合の要因となる何の追求という立場から異常データの検出を考える。

1.2 研究の概要

本研究は次の二部に分けることができる：

- (I) 多重配列構造を持つ母数空間と多重配列の特異値分解の研究；
- (II) 多重分割表データの解析等への応用。

以上のほか、(I)に関連する Riemann 構造を導入した母数空間の研究を附録として含む。

その概要を述べたにあたって、多重配列構造を持つデータと母数の出所を示しておくことは本研究の応用の場を考える上で意義がありう。説明を簡潔にするために、寸法 $r \times s \times t$ の 3 重配列をとりあげ、それを $A = (A_{ijk})$, $\forall i \in \bar{r}, \forall j \in \bar{s}, \forall k \in \bar{t}$ と書く。ここで、 $\bar{r}, \bar{s}, \bar{t}$ は 3 個の指標 i, j, k のそれぞれがとる名前あるいは番号の添字集合 (index sets) である。現実の応用の際には、多重配列の各要素とその添字集合はさまざまな物理的意味を持つが、3 重配列の出所と添字集合の異同によって 3 種に分類しておく。

- (1) 添字集合 $\bar{r}, \bar{s}, \bar{t}$ がすべて異なる場合。その代表的な出所として、実験計画法での要因実験に対する多変配置型データと調査データのクロス集計としての多重分割表が挙げられよう。とくに、後者は本論における主要な応用分野として取り扱われるものである。要因実験においては、 $\bar{r}, \bar{s}, \bar{t}$ はそれぞれ別個の一つの因子に対する水準

の集合を示し、配列要素 A_{ijk} は水準の組 (i, j, k) での実験の観測値として、あるいはモデルでのその観測値に対する期待値として現われる。要因実験で扱われる通常のモデルは期待値として A_{ijk} に関する線形モデルである。多重分割表の場合、 $\bar{r}$, $\bar{s}$, 及び $\bar{t}$ はしばしば項目 (item) と呼ばれる属性に対応し、 i, j, k はそれぞれの項目を構成する排他的なカテゴリ (category) を意味する。添字要素の組合せのひとつひとつをセル (cell) といい、 (i, j, k) セルなどとも書く。多重分割表の要素 A_{ijk} は (i, j, k) セルに該当する観測対象の度数 (以下セル度数ともいう)、何らかの確率分布を想定したモデルでのセル度数の期待値や各セルでの確率 (セル確率) としての意味を持つ。多重分割におけるモデルとしては、セル度数の期待値に関する線形モデル (多重分割表の加法モデル) やその期待値の対数をとった母数に関する線形モデル (多重分割表の対数線形モデル) などが一般的なものとして研究されている。本論ではセル度数の平方根をとった多重配列の分解も論じる。多重配列構造のデータで、その添字集合が要因の水準や項目のカテゴリに対応して各セルでの値が観測されるようなものを多次元尺度構成法などでは conjoint measurement データと称することもある。

(ロ) 添字集合 $\bar{r}, \bar{s}, \bar{t}$ の二つが同一である場合. 例えば, 大学等の入学者選抜試験をとりあげよう. $\bar{r}$ と $\bar{s}$ はともに入試の教科の集合, 例えば {国語, 社会, 外国語, 数学, 理科} と, $\bar{t}$ は数年における年度を表わすものとする. そのとき A_{ijk} は k 年度における教科 i と j との相関係数などとして出現する. そゝでは数年度における相関係数行列の構造の変化を分析することが問題となる. 一般に, $\bar{r}$ と $\bar{s}$ がともに何らかの物の集合で, その集合に属する任意の二つの物の間の関係が直接あるいは間接に測定され, その測定が複数回の異質な集団で行われるような場合, A_{ijk} は k 集団で測定された i と j の物の間の関係を表わす. その関係は類似性の度合や優劣の関係などいろいろな場合がある.

(ハ) 添字集合 $\bar{r}, \bar{s}, \bar{t}$ がすべて同一である場合. $\bar{r}, \bar{s}$, 及び $\bar{t}$ がいずれも同一の人間の集合を表わし, 3人のあらゆる組について何らかの関係, 例えば 3人組による作業の能率が測定されるとき, その測定値の並びが 3重配列を表現する. また, マルコフ連鎖の拡張として, $\bar{r}, \bar{s}, \bar{t}$ も同一の状態の集合と考え, 状態 i から j を経て k に移る遷移確率を A_{ijk} とすることもできる.

多重配列構造を持つデータは上記のほかにも多様な出所があり,

それぞれが固有の性格を有している。その解析およびモデル構成は、その出所である実質科学上の性格に即して行われるべきものであり、一般的な解析手法やモデルがあらわけてはいない。しかし、線形性を持つモデルや本論で示す多重配列のある種の分解定理が多重配列構造を持つデータの解析やそのモデル構成の重要な基礎になると考えられる。

さて、次項においてこの分野における従来の研究とその文献をあげて概観するのについで、2章では多重配列構造を持つ母数に対して各指標ごとのベクトル空間のテンソル積を考え、後章の基礎となる諸概念の定義や性質を本論の立場から記述する。3章では行列、すなわち二重配列の特異値分解の拡張に相当する多重配列の特異値分解を論じ、4章及び5章で多重分割表を直接の対象とする応用を行う。6章では、多重配列の添字集合のいくつかが同じ物理的意味を持つ場合についてその分解と応用を論じる。附録では母数空間に基本計量テンソルを導入して母数空間をRiemann空間とみなし、其分散安定化変換が存在するための必要十分条件をはじめ、二三の応用を論じる。以上が本論の基本的枠組であるが、以下に各章内の概要を述べる。

第2章では、 m 重配列の構造を持つ母数空間を、各指標ごとのベクトル空間 $\Omega^{(i)}$, $i = 1, 2, \dots, m$ のテンソル積 Ω と考え、以後の章で必要とな

る基礎的概念を導入する。母数空間 Ω の部分空間としての効果空間と母数 A の因子効果 あるいは交互作用が各ベクトル空間 $\Omega^{(i)}$ で指定されるベクトル $u^{(i)}$, $i=1, 2, \dots, m$ に相対的な形式で定義される。2.1 節では、その一般的定義にもとづく多重配列母数 A の因子効果の表現, A の分解表現, A の平方和分解などに関する定理が与えられる。内容的には完全要因計画における分散分析とそのモデルからの類推によって理解されよう。2.2 節では汎用性の高いモデルとして階層的モデルや分解型モデルを論じる。それらは多重分割表のモデルの選択を行うときに重要となる。2.3 節ではベクトル空間における内積の定義を一般化したときの問題点も注意する。後述の加法モデルや多重分割表における数値化手法の理論展開に必要となる。

本章では、多重配列データを多重一次形式の係数と考え、その多重一次形式の最大化問題を取りあげ、それより多重配列の特異値分解を導く。

多重一次形式は

$$\rho = \sum_{\lambda_1, \dots, \lambda_m} A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots u_{\lambda_m}^{(m)}$$

で与えられる。ここで、 $A_{\lambda_1 \lambda_2 \dots \lambda_m}$ は m 重配列の要素, $u_{\lambda_i}^{(i)}$ はベクトル $u^{(i)} \in \Omega^{(i)}$ の要素である。この多重一次形式は、 m 次の相関関数、あるいは適当な条件の下で m 次の相関係数としての統計的意味を

持つ。3.1節では、まず上述の最大化問題の解の存在と決定方程式を示す。次にその決定方程式を再帰的に用いる算法を定義し、その算法の有限性、その算法によって得られるベクトルの各ベクトル空間 $\Omega^{(i)}$ での正規直交性を示す。また、最大化問題の各種の拡張とその算法との関係を論じる。3.2節では、上述の算法から多重配列の特異値分解ともいふべき分解表現が得られることを示し、その分解による多重配列の因子効果成分や平方和の分解に関する特異値分解の諸性質を示す。

才4章1節では、従来からの多重分割表に対する代表的モデルとして研究されている加法モデルと対数線形モデルを本論の立場から検討する。とくに、加法モデルと特異値分解との密接な関係を指摘する。4.2節では多重分割表のセル度数あるいはセル確率の平方根をもとに多重配列の線形モデルを論じる。

才5章では、特異値分解の直接的応用として、多重分割表における数量化手法と平方根分解を扱う。5.1節で従来の2重分割表における手法の基本的性質を整理し、5.2節において3重分割表における数量化をもくわしく扱う。従来提案されている手法が加法モデルによって理論的に扱えることを示し、新たな数量化手法の提案を行う。

5.3 節では一般の多重分割表における数量化について、数量化以前の
問題として多重分割表の構造をさぐるモデルの選択が必要であり、分割
表の周辺和から直接に推定される構造を持つ分解型モデルを中心と
する数量化モデルを考えることが有効であることを指摘する。5.4
節では多重分割表のセル度数の平方根の配列についての特異値分解
を平方根分解と呼び、多重分割表におけるモデルの選択に有効に
利用できることを述べる。5.5 節では平方根分解を中心として、数
値例によってこれまでに述べた各種のモデルの比較を行う。とくに、
平方根分解による各因子効果の平方和の表を用いるモデルの選択
が有効であることを示す。

第6章では、ある対称性を持つ多重配列の要素を一次形式と二次
形式の混合形式の係数とみなして、その混合形式の最大化問題と
その応用に触れる。6.1 節ではその問題に対する決定方程式を示
し、第3章の特異値分解と併行して議論が可能であることを指摘
する。6.2 節では数値例を示して応用に触れる。

附録においては、母数空間の構造に関する本研究に属する
ものとして、Riemann 構造の導入される母数空間を扱う。その研究
は、多重配列の特異値分解や多重分割表の平方根分解の研究の動機

を与え、かつ本研究全体の思想的背景でもある。附録1では、漸近的に正規分布に従う一致推定量を持つ多次元母数について、その漸近的分散共分散行列の逆行列を基本計量テンソルとすることによってその母数空間を Riemann 空間と考える。次に、重要な幾何学的不変量である Riemann-Christoffel 曲率テンソルの統計的意味付けが与えられる。すなわち、その曲率テンソルが消えることが、共分散安定化変換 (covariance stabilizing transformation) の存在のための必要十分条件となる。以下、いくつかの母数空間が定曲率空間となることが示され、測地距離の応用が論じられる。例えば、多項分布での確率の母数空間は正の定曲率空間となり、その測地距離はオ4章とオ5章で用いる Bhattacharya の距離となる。附録2では、一般に位置及び尺度母数を持つ確率分布に対する二次元母数空間は負の定曲率空間となり、測地距離が比較的簡単に表現されることが示される。

1.3 従来の研究及び文献

この節では本論に関連する従来の研究を文献によって概観する。本研究の直接的な基盤となる分野は、多重分割表に対するパラメトリックなモデルによるデータ解析と多重分割表を含むカテゴリカルデータに対する数量化あるいは最適尺度構成の研究分野である。

多重分割表に対するパラメトリックなモデルとそれによるデータ解析の研究はこの15年間に急速に進み、さらに発展を続けている。最近の主要な研究は次の三種類のモデルに沿って分類できよう。すなわち、対数線形モデル、線形モデル、混合モデルである。対数線形モデルについては、G. U. Yule, R. A. Fisher 以来、M. S. Bartlett, J. A. Roy, M. A. Kastenbaum を経て、I. J. Good, S. Kullback, M. W. Birch, L. A. Goodman, Y. M. Bishop, S. M. Fienberg, S. J. Haberman など多くの研究者があり、分割表の研究の主流ともいえよう。線形モデルや混合モデルは、K. Pearson, R. A. Fisher から H. O. Lancaster を中心とする流れがあり、とくに線形モデルには J. E. Grizzle, C. F. Starmer, G. G. Koch 等の研究がある。その線形モデルは分散分析のモデルからの単純な類推であり、Lancaster による加法 (additive) モデルとは異なる。

さて、分割表の研究は1900年当初 K. Pearson と G. U. Yule に

その端緒を見ることが出来る。Pearson (1900 a, b) は有名なカイ二乗適合度検定を提案し, 2×2 分割表に対する tetrachoric correlation coefficient を発展させている。Pearson は分割表を多次元正規分布などの連続分布の離散化に対応する標本と見る立場をとり, その背景としての正規分布についての相関係数を推定することに関心があった。また, Pearson (1904) は, 二つの確率測度の相違を測る尺度として, Pearson,あるいは後に Hellinger の名によって呼ばれる汎関数 ϕ^2 を定義した。その ϕ^2 は

$$\phi^2 + 1 = \int (d\nu)^2 / d\mu = \int (d\nu/d\mu)^2 d\mu = \int \Lambda^2 d\mu$$

と定義される。ここで, μ と ν は二つの確率測度, $\Lambda = d\nu/d\mu$ は μ に関する ν の Radon-Nikodym 導関数である。 ϕ^2 は Λ の分散と解釈される。この汎関数がときに Hellinger の名で呼ばれるのは [例えば, Cramér (1946)], Hellinger が一般の 2 次元分布関数 $F(x, y)$ について

$$\phi^2 + 1 = \iint_{-\infty}^{\infty} [dF(x, y)]^2 / [dG(x) dH(y)]$$

と置いて研究したことから。ここで, $G(x)$ と $H(y)$ は $F(x, y)$ の周辺分布である。2 次元離散分布を, 二つの確率変数 X と Y について

$$\Pr\{X = x_i, Y = y_j\} = p_{ij}$$

と書くなら, Hellinger の ϕ^2 は

$$\phi^2 = \sum_{i,j} p_{ij}^2 / (p_{i+} p_{+j}) - 1$$

となる. ここで, $p_{i+} = \sum_j p_{ij}$, $p_{+j} = \sum_i p_{ij}$ である. もし, この p_{ij} を 2重分割表データの各セルでの標本確率とみなせば, 上式の ϕ^2 は, いわゆる 2重分割表の独立性についての Pearson の χ^2 を標本の大きさを N で割ったものにほかならない. その後この方向の研究は Lancaster などによって受け継がれる.

一方, Yule (1900, 1912) は 2×2 分割表における二項目間の連関の強さを測る尺度 (measure of association) として, しばしば Yule の Q 及び Y (あるいは w) と記される二つの量を提案した. Yule は, Pearson と異なり, 分割表を本来的に離散的データとして考えているようである. また, 上記の尺度が対数線形モデルの基本ともなり, 2×2 分割表での cross product ratio の関数になるというある種の不変性を持つ. このようなことから, 対数線形モデル派ともいえる研究者群は歴史的に Yule をこの分野の研究の出発点とするようである [例えば, Fienberg (1975)].

次に, Fisher (1922, 1924) は, データから母数が推定されるべきの適合度検定の χ^2 統計量が漸近的に χ^2 分布に従うことの初の

証明を与えたことをはじめ、2重分割表についての現在の理論がその多くを彼の研究に依るほどの寄与を残している。Fisherの貢献について、本論の立場では、次の二点を強調しておくべきと思われる。オ一は、3重分割表における高次の交互作用の基準を示唆して Bartlett (1935)の重要な研究を導いたことである。3重分割表での高次の交互作用の概念をどのように定式化するという問題こそが、多重分割表におけるパラメトリックなモデルの発展を促したともいえる。オニは、Maung (1942)に引用されているように、Hotelling (1936)による正準相関分析の考えを2重分割表に適用したことである。これについては、同じ考えが以前にも発表されているという説もあるが[例えば西里(1977)参照]、実質的影響はFisherによる。この手法は分割表における数量化あるいは最適尺度法といわれる分野の基礎となる(オ5章参照)。

前述したBartlett (1935)は $2 \times 2 \times 2$ 分割表で2次交互作用が存在しないという概念の一つの定式化を示した。その後20年近い空白を経て、1951年に入って Simpson (1951)はBartlettの定式化が cross product ratio に関係し必然的に乗法的モデルにつながることを示した。この立場は対数線形モデルへつながる。

一方, Lancasterは3重分割表について, Pearsonの χ^2 , らういはそれと同値な χ^2 の分解から, 2次の交互作用非存在の別の定式化として加法モデルを提案した. その加法モデルは, 3重分割表のセル確率を p_{ijk} , $i=1, \dots, r$, $j=1, \dots, s$, $k=1, \dots, t$ とするとき, セル確率が

$$p_{ijk} = p_{ij+} p_{+jk} + p_{i+k} p_{+j+} + p_{+jk} p_{i++} - 2 p_{i++} p_{+j+} p_{+jk}$$

という関係を満たすものとして表現された. 加法モデル及びそれに伴う χ^2 の分解については4.1節で本論の立場から検討を加える. さらに, Lancaster (1960)は交互作用概念の多次元分布への拡張を行っている. また, Bahadur (1961)とLazarsfeld (1961)は 2^m 分割表についての分解表現を与えているが, 両者は表現こそ違うが, 加法モデルと同値なものになっている.

再び, 対数線形モデルの流れに目を向けると, Bartlettによる定式化がRoyとKastenbaum (1956)に至って, 一般の $r \times s \times t$ 分割表に拡張されたことが特筆される. その後, Good (1963)によるエントロピー最大化による分割表データの解析, Kullback (1959)による情報理論的接近が顕著なものとしてあげられる. そして, Birch (1963, 1964, 1965)による最尤推定法による接近に至り, それ

までの比較簡単な仮説の検定を主体とする研究から、対数線形モデルによるモデル構成の方向の研究への門戸が開かれた。対数線形モデルは、セル確率 p_{ijk} の対数について

$$\log p_{ijk} = \mu + \mu_i^{(1)} + \mu_j^{(2)} + \mu_k^{(3)} + \mu_{ij}^{(12)} + \mu_{ik}^{(13)} + \mu_{jk}^{(23)} + \mu_{ijk}^{(123)}$$

をその飽和型とし、右辺の項を種々に取捨選択した一連のモデルとして表現される。上式ではセル確率のかめりにセル度数の期待値を用いることも多い。例えば、 $\mu_{ijk}^{(123)}$ をゼロとした場合は Roy と Kastenbaum による二次交互作用（本論では 3 因子効果という用語を使うが）の非存在の定式化と同値である。対数線形モデルは、その後 Goodman (1964, 1969, 1970, 1971a,b), Bishop (1967, 1969) のほか、多くの研究者によって発展させられている。これらの成果は、Plackett (1964), McKinlay (1975) による総合報告として、また Plackett (1974), Bishop, Fienberg & Holland (1975), Haberman (1974) によって成書としてまとめられている。対数線形モデルに関連する研究は数理統計学的基本理論と適用範囲の拡大への努力に分けられよう。前者については、Birch 以来分割表の標本抽出過程をポアソンモデルや多項分布モデルに仮定した上で、十分統計量や最尤推定法との関連における対数線形モデルの良い性質の追

求が中心である。標本抽出過程をより複雑なものにすれば、この種の研究は一段と難しくなり、簡単な前提での手法の頑健性の研究を含めて、今後の発展を待たなければならない。適用範囲あるいは応用性の拡大という後者の側面にも種々の研究があるが、本論に関係するモデルの選択の問題に触れておこう。 m 重の分割表での対数線形モデルを考えると、分散分析モデルからの類推でめめるように、多因子効果を含めて 2^m 個の効果・交互作用が考えられる。分割表の解析では事後的なモデル構成もしばしば必要となり、 2^m 個の因子効果の中から有意なものを選択しないという問題が生じる。形式的には 2^m 個の因子効果の組み合わせによって $2^{(2^m)}$ 個のモデルが可能である。そこで、2.2 節で述べるように、階層的モデルや分解型 (closed form) モデルの概念が必要となる。Bishop (1967) は階層的モデルでは最尤推定が、周辺和も、観測値とその期待値の推定値について、一致させるという反復計算で得られることを示した。また、Goodman (1970) 及び Bishop (1971) は 4 重分割表以下では分解型モデルでのセル確率の最尤推定値が反復計算を要せずに陽の形に求まることを示し、さらに Goodman (1971b) は任意のモデルが分解型であるかどうかを判定する構成的算法を与えた。Haberman

(1974)はそれを再帰的算法に書き直している。

モデル選択の現実的な算法としては、Goodman (1971a)が重回帰分析における段階的変数選択の手法を応用した手順を提案している。また、最近我国では赤池(1973)による情報量規準による選択の有向性が検証されつつある。情報量規準の一つの解釈としては竹内(1976)、多重分割表への応用例としては吉澤・藤本(1976)がある。第5章では平方根分解によるモデル選択が言及される。

以上のほか、線形モデルの系統の研究には Grizzle, Starmer 及び Koch (1969)を中心とするものがあるが、単純な線形モデルであって、本論でのモデル構成の立場と多少異なる方向にもあるのでその詳細には触れない。なお、多重分割表での高次交互作用の非存在についての加法モデルと対数線形モデル(この場合は乗法モデルと言ってもよい)との比較は Darroch (1974)にくわしい。分割表に関する文献リストが Killian と Zahn (1976)によって与えられている。

次に分割表における数量化手法の関係について述べる。ここでいう数量化手法は、多重分割表における名義尺度あるいは順序尺度の

特性である項目について、項目を構成するカテゴリを分類したり、その相互の関係も把握することとを目的に、何らかの規準に従って各カテゴリに数値を付与しようとするものである。一般のカテゴリカルなデータに対する数量化手法を含めて最適尺度法といわれることもある。この分野の文献目録が西田(1977)によって作られている。その多くは、0-1データの行列が与えられるときに、行のカテゴリと列のカテゴリに同時に数値を付与して比較の尺度を作ろうとするものである。本論で興味のある多重配列データも扱う数量化には岩坪(1971, 1974)がある。そこでは3次の相関係数最大化の規準による数量化と相関比による数量化が提案されている。吉澤(1975)ではそれらの手法が分割表の加法モデルと密接な内容を持つことが指摘され、従来記述的手法として考えられていた分野へ確率論的モデルを持ち込むことによって推測的扱いをする試みが行われている。

多重分割表と実験計画法での要因実験以外の分野での多重配列データの研究はあまり多くないが、正準相関分析の拡張としての Kettenring (1971)の研究、3相因子分析についての Tucker (1964)があげられる。多重分割表でも潜在構造分析とされる Lazarsfeld & Henry (1968)の研究と前述の3相因子分析は、

本論における特異値分解と形式的類似性がある.

2. 多重配列構造を持つ線形母数空間

本章では、実数値 m 重配列の構造を持つ母数空間を配列の各指標に対応するベクトル空間のテンソル積と考え、効果・交互作用の概念とその諸性質、ならびに階層的モデルや分解型モデルなどの母数空間における典型的モデルについて論じる。

2.1 効果・交互作用

実数値 m 重配列の構造を持つ母数を A とし、その要素を

$$A_{\lambda_1 \lambda_2 \dots \lambda_m} \quad (\forall \lambda_i \in \overline{r_i}, \forall i \in \overline{m})$$

と書く。ここで、 $\overline{r_i}$ は配列の i 指標の添字集合 $\{1, 2, \dots, r_i\}$ 、 $\overline{m}$ は 1 から m までの指標の番号の集合 $\{1, 2, \dots, m\}$ である。一般に正の整数 k に対し、 $\overline{k}$ は k 以下の正整数の集合を意味するものとする。配列 A の全要素数を n と書く。明らかに

$$n = r_1 r_2 \dots r_m$$

である。この配列 A を寸法 $r_1 \times r_2 \times \dots \times r_m$ の m 重配列と呼び、例えば $2 \times 3 \times 5$ の 3 重配列、あるいは略して $2 \times 3 \times 5$ 配列などを書くこともある。

配列の各指標に対して以下に述べる内積の定義される r_i 次元のベクトル空間 (内積空間) $\Omega^{(i)}$ を考える。 $\Omega^{(i)}$ に属するベクトルを

$$\underline{u}^{(i)} = (u_{\lambda_i}^{(i)}), \quad v^{(i)} = (v_{\lambda_i}^{(i)})$$

あるいは,

$$\underline{u}^{(i)} = (u_{\lambda_i}^{(i)}), \quad \underline{u}^{(i)} = (u_{\lambda_i}^{(i)}), \dots$$

などと書く。さしあたって、各ベクトル空間 $\Omega^{(i)} (\forall i \in \bar{m})$ における内積は

$$(\underline{u}^{(i)}, \underline{v}^{(i)}) = \sum_{\lambda_i} u_{\lambda_i}^{(i)} v_{\lambda_i}^{(i)} \quad (2.1)$$

によって定義する[†]。

さて、実数値をとり、寸法 $r_1 \times r_2 \times \dots \times r_m$ の m 重配列の全体から成る n 次元実数空間 R^n を、Halmos (1958) の定義によって、 m 個のベクトル空間 $\Omega^{(i)} (\forall i \in \bar{m})$ のテンソル積 (空間) と考える。そのテンソル積を

$$\Omega = \Omega^{(1)} \otimes \Omega^{(2)} \otimes \dots \otimes \Omega^{(m)} \quad (2.2)$$

と書く。上式の右辺については、まとめて

$$\bigotimes_{i=1}^m \Omega^{(i)} \quad (2.3)$$

と記すことがある。記号 $\otimes$ はベクトルの直積や行列の Kronecker 積にも用いるが、その区別は文脈から明らかであろう。Halmos によれば、テンソル積 Ω は $\underline{u}^{(i)} \in \Omega^{(i)} (\forall i \in \bar{m})$ に関するあらゆる多重一次形式によって構成されるベクトル空間の双対空間として定義される。

[†] $\sum_{\lambda_i}$ は $\forall \lambda_i \in \bar{r}_i$ について和をとることを意味する。

m 個のベクトル $u^{(i)} \in \Omega^{(i)} \ (\forall i \in \bar{m})$ の直積を, n とおりの積

$$u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \cdots u_{\lambda_m}^{(m)} \quad (2.4)$$

を要素とする n 次元ベクトルで定義し, それを

$$\underline{u} = \bigotimes_{i=1}^m \underline{u}^{(i)} \quad (2.5)$$

と書く. (2.4) の量は本来 m 個の添字を持つが, これを n 次元ベクトルとみなすとき $\sim$ の記号を用いる. m 重配列 A などについて全体の要素をベクトルとして扱うときは必要に応じて $\underline{A}$ と書く. 添字の組 $(\lambda_1, \lambda_2, \dots, \lambda_m)$ の一次元的並べ方は, 一定の規則に従いさえすればよいが, ここでは列ごと (column-wise) の順を原則として採用する. また, ベクトルは, とくに断らなければ列ベクトルとし, ベクトルと行列の転置には ' を用いる.

任意の寸法の行列

$$U^{(i)} = (U_{\kappa_i \lambda_i}^{(i)}), \quad (\forall i \in \bar{m})$$

の Kronecker 積を

$$U_{\kappa_1 \kappa_2 \cdots \kappa_m \lambda_1 \lambda_2 \cdots \lambda_m} = U_{\kappa_1 \lambda_1}^{(1)} U_{\kappa_2 \lambda_2}^{(2)} \cdots U_{\kappa_m \lambda_m}^{(m)} \quad (2.6)$$

を要素とし, 添字 $(\kappa_1, \kappa_2, \dots, \kappa_m)$ と $(\lambda_1, \lambda_2, \dots, \lambda_m)$ を行及び列の指標とする行列 U によって定義し,

$$U = U^{(1)} \otimes U^{(2)} \otimes \cdots \otimes U^{(m)} = \bigotimes_{i=1}^m U^{(i)} \quad (2.7)$$

と書く. 添字の組 $(\kappa_1, \kappa_2, \dots, \kappa_m)$ と $(\lambda_1, \lambda_2, \dots, \lambda_m)$ は $\mathbb{U}$ や $\mathbb{A}$ の場合と同様に列ごとに並べるものとする.

以下に, テンソル積空間や Kronecker 積の性質として本論で用いるものを補題として挙げておく. Halmos (1958) あるいは Haberman (1974) などの記述から容易に証明されるので, それらの補題の証明は省略する. なお, $\bar{m}$ の部分集合を $\alpha, \alpha', \dots$ などと書き, 空集合を ϕ と記す. 集合については包含記号 $\subset$ や同値記号 $\equiv$ を必要に応じて用いる. また, 指示関数 (indicator) $\chi_\alpha(i)$ を, α に i が含まれれば 1, そうでなければゼロの値をとる関数, すなわち

$$\chi_\alpha(i) = \begin{cases} 1, & i \in \alpha \text{ のとき,} \\ 0, & i \notin \alpha \text{ のとき} \end{cases} \quad (2.8)$$

と定義する.

補題 2.1. $\Omega^{(i)}$ の一つの直和分解を

$$\Omega^{(i)} = \omega_0^{(i)} \oplus \omega_1^{(i)} \quad (\forall i \in \bar{m}) \quad (2.9)$$

とする. そのとき, $\Omega^{(i)}$ のテンソル積 Ω は次のように書ける:

$$\begin{aligned} \Omega &= (\omega_0^{(1)} \oplus \omega_1^{(1)}) \otimes (\omega_0^{(2)} \oplus \omega_1^{(2)}) \otimes \cdots \otimes (\omega_0^{(m)} \oplus \omega_1^{(m)}) \\ &= (\omega_0^{(1)} \otimes \omega_0^{(2)} \otimes \cdots \otimes \omega_0^{(m)}) \oplus (\omega_1^{(1)} \otimes \omega_0^{(2)} \otimes \cdots \otimes \omega_0^{(m)}) \\ &\quad \oplus \cdots \oplus (\omega_1^{(1)} \otimes \omega_1^{(2)} \otimes \cdots \otimes \omega_1^{(m)}) \end{aligned}$$

あるいは

$$\begin{aligned}\Omega &= \bigotimes_{i=1}^m (\omega_0^{(i)} \oplus \omega_1^{(i)}) \\ &= \bigoplus_{\alpha} \bigotimes_{i=1}^m \omega_{\chi_{\alpha}(i)}^{(i)}.\end{aligned}\quad (2.10)$$

すなわち、演算 $\bigotimes$ と $\bigoplus$ について分配律が成り立つ。ここで、 α は $\bar{m}$ の任意の部分集合であり、 $\bigoplus_{\alpha}$ は $\bar{m}$ の 2^m 個のあらゆる部分集合についての直和を意味するものとする。

補題 2.2. 補題 2.1 において $\omega_0^{(i)}$ と $\omega_1^{(i)}$ が直交するなら ($\forall i \in \bar{m}$), (2.10) での各項

$$\bigotimes_{i=1}^m \omega_{\chi_{\alpha}(i)}^{(i)}$$

は異なる $\alpha \subset \bar{m}$ について互いに直交する。

補題 2.3. 任意の $i \in \bar{m}$ について、 $\Omega^{(i)}$ の任意の部分空間を $\omega^{(i)}$, その次元 $\dim(\omega^{(i)})$ を q_i , かつ $\omega^{(i)}$ のある基底を

$$\left\{ \underset{\sim}{u}^{k_i(i)}, \forall k_i \in \bar{q}_i \right\}$$

とする。そのとき、基底向のあらゆる組のベクトルの直積の集合

$$S = \left\{ \bigotimes_{i=1}^m \underset{\sim}{u}^{k_i(i)}, \forall k_i \in \bar{q}_i, \forall i \in \bar{m} \right\} \quad (2.11)$$

を作れば、 S はテンソル積

$$\omega = \bigotimes_{i=1}^m \omega^{(i)}$$

の基底となる。また、 ω の次元は

$$\dim(\omega) = \prod_{i=1}^m \dim(\omega^{(i)}) = \prod_{i=1}^m q_i \quad (2.12)$$

で与えられる.

注意. この補題で, $q_i = r_i$ ($\forall i \in \overline{m}$), すなわち $\omega^{(i)} \in \Omega^{(i)}$ のものとすれば, $\Omega^{(i)}$ の基底 $\{\underline{u}^{(i)}_{k_i}, \forall k_i \in \overline{r_i}\}$ を与えて, テンソル積 Ω の基底

$$S = \left\{ \bigotimes_{i=1}^m \underline{u}^{(i)}_{k_i}, \forall k_i \in \overline{r_i}, \forall i \in \overline{m} \right\}$$

を作ることができる. 逆にテンソル積 Ω も, S を基底とするベクトル空間として, 定義することも可能である.

補題 2.4. Kronecker 積についても, 補題 2.1 と同様に, 分配律が成り立つ. すなわち

$$\begin{aligned} U &= \bigotimes_{i=1}^m (U_0^{(i)} + U_1^{(i)}) \\ &= \sum_{\alpha} \bigotimes_{i=1}^m U_{\alpha^{(i)}}^{(i)} \end{aligned}$$

ここで, $U_0^{(i)}$ と $U_1^{(i)}$ は各 i ごとに同じ寸法の行列であり, $\sum_{\alpha}$ は $\overline{m}$ のあらゆる部分集合についての和を意味する.

補題 2.5. Kronecker 積 $U = \bigotimes_{i=1}^m U^{(i)}$ の階数 (rank) は

$$\text{rank } U = \prod_{i=1}^m \text{rank } U^{(i)} \quad (2.13)$$

で与えられる. 二つの Kronecker 積

$$U = \bigotimes_{i=1}^m U^{(i)}, \quad V = \bigotimes_{i=1}^m V^{(i)}$$

について、普通の行列積 $U^{(i)} V^{(i)}$ が可能なら ($\forall i \in \bar{m}$)

$$U V = \bigotimes_{i=1}^m (U^{(i)} V^{(i)}) \quad (2.14)$$

が成り立つ。 U の転置について、 $U' = \bigotimes_{i=1}^m (U^{(i)'})$ が成り立つ。

以上の準備にもとづいて、本論で用いる‘効果’の概念を定義する。

定義 2.1. 効果空間. あるベクトル $\underline{u}^{(i)} \in \Omega^{(i)}$ ($\forall i \in \bar{m}$) が与えられ、その $\underline{u}^{(i)}$ が張る $\Omega^{(i)}$ の部分空間を $\text{span}(\underline{u}^{(i)})$ と書いて、それを $\omega_0^{(i)}$ とする。

$$\omega_0^{(i)} = \text{span}(\underline{u}^{(i)}) \quad (2.15)$$

とする。また、 $\omega_0^{(i)}$ の $\Omega^{(i)}$ に関する直交補空間を $\omega_1^{(i)}$ とする。このとき、 $\bar{m}$ の任意の部分集合 α に対し、

$$\Omega(\alpha) = \bigotimes_{i=1}^m \omega_{\chi_\alpha(i)}^{(i)} \quad (2.16)$$

によって与えられる Ω の部分空間 $\Omega(\alpha)$ を α -効果空間と定義する。

注意. α -効果空間は部分集合 α のほかに、 $\underline{u}^{(i)}$ ($\forall i \in \bar{m}$) の与え方に依存し、 $\Omega(\alpha | \underline{u}^{(i)})$ とでも書くべきところであるが、本章では $\underline{u}^{(i)}$ が与えられたものとして $\Omega(\alpha)$ とだけ書く。ただし、 $\underline{u}^{(i)}$ ($\forall i \in \bar{m}$) はゼロベクトルではないものとする。

定理 2.1. すべての $\alpha \subset \bar{m}$ に対する α -効果空間は Ω の直交直和分解を与える。また、 α -効果空間 $\Omega(\alpha)$ の次元は

$$\dim(\Omega(\alpha)) = \prod_{i \in \alpha} (r_i - 1) \quad (2.17)$$

で与えられる. $\prod_{i \in \alpha}$ は α に含まれる i について積をとり, α が空集合中なら 1 とする.

証明. 前者は補題 2.1 と 2.2 から明らかである. (2.15) で与えられる $\omega_0^{(i)}$ の次元は 1 であるから, 補題 2.3 の (2.12) によって

$$\begin{aligned} \dim(\Omega(\alpha)) &= \prod_{i=1}^m \dim(\omega_{\chi_\alpha(i)}^{(i)}) \\ &= \prod_{i \in \alpha} (r_i - 1) \end{aligned}$$

となり, 後者が証明される. |

定義 2.2. α -因子効果. 多重配列構造を持つ母数 A をテンソル積 Ω に属するベクトルとすると, A の α -効果空間 $\Omega(\alpha)$ の成分を (ベクトル $\underline{u}^{(i)}$ ($\forall i \in \bar{m}$) に相対的な) A の α -因子効果と定義する. これを $B(\alpha)$ と書く. ここで, 成分とは直交射影成分を意味する.

注意. $B(\alpha)$ は A と同一の多重配列構造を持つ. $B(\alpha)$ の一つの表現は次の定理で与えられる.

定理 2.2. 与えられるベクトル $\underline{u}^{(i)} \in \Omega^{(i)}$ ($\forall i \in \bar{m}$) について

$$U_0^{(i)} = \frac{1}{\underline{u}^{(i)'} \underline{u}^{(i)}} \underline{u}^{(i)} \underline{u}^{(i)'} \quad (2.18)$$

$$U_1^{(i)} = I_{r_i} - U_0^{(i)} \quad (2.19)$$

によって, $r_i \times r_i$ の行列 $U_0^{(i)}$ と $U_1^{(i)}$ を作る. ここで I_{r_i} は $r_i \times$

r_i の単位行列である。そのとき、任意の $\alpha \in \bar{m}$ について A の α -因子効果 $B(\alpha)$ は次式で与えられる:

$$\underline{B}(\alpha) = \left(\bigotimes_{i=1}^m U_{x_\alpha(i)}^{(i)} \right) \underline{A}. \quad (2.20)$$

証明. 任意の $i \in \bar{m}$ について, $U_0^{(i)}$ と $U_1^{(i)}$ は次の性質を持つ:

$$\left. \begin{aligned} U_k^{(i)} &= U_k^{(i)'} \\ U_k^{(i)} U_k^{(i)} &= U_k^{(i)} \end{aligned} \right\}, \quad k=0, 1 \quad (2.21)$$

$$U_0^{(i)} U_1^{(i)} = 0. \quad (2.22)$$

これらは $U_0^{(i)}$ と $U_1^{(i)}$ の定義式 (2.18) と (2.19) から, 互いに証明される。従って, $U_0^{(i)}$ は $\Omega^{(i)}$ における $\underline{u}^{(i)}$ の張る部分空間 $\omega_0^{(i)}$ への射影変換行列 (射影子, projector) に, $U_1^{(i)}$ は $\omega_0^{(i)}$ の直交補空間 $\omega_1^{(i)}$ への射影子になっている [例えば, Halmos (1958, p.73)]。そこで, (2.20) の右辺の左の因子を $U(\alpha)$, すなわち

$$U(\alpha) = \bigotimes_{i=1}^m U_{x_\alpha(i)}^{(i)} \quad (2.23)$$

と置くと, $U(\alpha)$ は対称かつ中等な行列となる。すなわち,

$$\left. \begin{aligned} U(\alpha) &= U(\alpha)', \\ U(\alpha) U(\alpha) &= U(\alpha). \end{aligned} \right\} \quad (2.24)$$

(2.24) の二つの式は補題 2.5 から, (2.21) を用いて, 互いに導かれる。

従って, $U(\alpha)$ が Ω における一つの射影子であり, それが Ω から α -効

果空間への射影子であることも示される(補題2.3の基底を用いてもよい).

よって, (2.20)を得る. |

定理2.3. 多重配列 A はそのすべての因子効果の和として表わされる.
すなわち,

$$A = \sum_{\alpha} B(\alpha). \quad (2.25)$$

また, 相異なる α と $\alpha' \in \bar{m}$ に対して $B(\alpha)$ と $B(\alpha')$ は直交する.

すなわち,

$$\underline{B}(\alpha)' \underline{B}(\alpha') = 0. \quad (2.26)$$

証明. 定理2.1により, すべての α -効果空間は Ω の直交直和分解を与えらる. α -因子効果の定義2.2から, (2.25) は明らかである. また, (2.23) と補題2.5を用いて

$$\begin{aligned} U(\alpha)' U(\alpha') &= \left(\bigotimes_{i=1}^m U_{\chi_{\alpha}(i)}^{(i)} \right)' \left(\bigotimes_{i=1}^m U_{\chi_{\alpha'}(i)}^{(i)} \right) \\ &= \bigotimes_{i=1}^m \left(U_{\chi_{\alpha}(i)}^{(i)} U_{\chi_{\alpha'}(i)}^{(i)} \right) \end{aligned}$$

と書ける. $\alpha \neq \alpha'$ のとき, $\chi_{\alpha}(i) \neq \chi_{\alpha'}(i)$ となる i が存在し, その i に対しては (2.22) によって 上式の因子 $U_{\chi_{\alpha}(i)}^{(i)} U_{\chi_{\alpha'}(i)}^{(i)}$ がゼロとなる. よって, $U(\alpha)' U(\alpha') = 0$ であり

$$\underline{B}(\alpha)' \underline{B}(\alpha') = \underline{A}' U(\alpha)' U(\alpha') \underline{A} = 0$$

すなわち, (2.26) を得る. |

定理 2.4. 多重配列 A の要素の平方和は, それぞれの因子効果の平方和の総和に等しい. すなわち,

$$\underline{A}'\underline{A} = \sum_{\alpha} \underline{B}(\alpha)' \underline{B}(\alpha) \quad (2.27)$$

証明. 定理 2.3 の (2.25) の表現と $B(\alpha)$ 間の直交性 (2.26) からそれぞれ (2.27) が導かれる. |

以上の議論は完全要因計画における分散分析の母数型モデルにおける周知の概念を, ベクトル $\underline{u}^{(i)} \in \Omega^{(i)}$ の取り方を任意にするという点で, 一般化して構成したものである. 本節におけるベクトル $\underline{u}^{(i)}$ を定数ベクトル, 例えば,

$$u_{\lambda_i}^{(i)} = 1, \quad \forall \lambda_i \in \bar{\gamma}_i, \quad \forall i \in \bar{m}, \quad (2.28)$$

とすれば, 本節の定理 2.2 ~ 4 はそのまま分散分析モデルの定理となる.

なお, $\bar{m}$ の部分集合 α を具体的に与えたとき, $\{1, 2\}$ -因子効果, $\{1, 2, 4\}$ -因子効果などと書く. α に含まれる要素数 εk とするとき, 総称的に k 因子効果と呼ぶことがある. 分散分析モデルでは, いま, $k=0, 1, 2, \dots$ に応じて 総平均, 主効果, 1次交互作用, $\dots$, あるいは 最高次交互作用 という用語を使うので, 本論でもときに流用することがある.

2.2 階層的モデル

多重分割表や完全要因計画のモデルとして、多数の α -効果空間のうち適当なものを選び、それらの直和の部分空間に母数が含まれるという形式をとることが多い。多重配列の母数空間の α -効果空間を規定する $\overline{m}$ の部分集合 α の取り方は空集合中や $\overline{m}$ 自身を含めて $2^{\overline{m}}$ 個の場合がある。その全体は集合の包含関係を半順序とみなすことによって m 次元ブール束(lattice)としての代数的構造を持つことはよく知られている。その構造を m が2, 3, 4の場合について図2.1に示しておこう。図2.1では 点に添えてある数字列が $\overline{m}$ の部分集合を表わし、直線が2つの部分集合の包含関係(下のものが上のものを含む)を示す。ここでは、 α の適当な集合をクラスと呼ぶ。クラスを例示する場合 $\{\{12\}, \{34\}\}$, あるいは中の $\{\}$ を省略して $\{12, 34\}$ などと書く。

$\overline{m}$ の部分集合のあるクラス $\mathcal{K}$ について、すべての $\alpha \in \mathcal{K}$ に対する α -効果空間の直和を $\Omega[\mathcal{K}]$ と書く。すなわち

$$\Omega[\mathcal{K}] = \bigoplus_{\alpha \in \mathcal{K}} \Omega(\alpha) \quad (2.29)$$

である。ここでは、母数 A に関するモデルを

$$A \in \Omega[\mathcal{K}] \quad (2.30)$$

という形式で表現して、特殊なクラスとモデルとの関係を本論で必

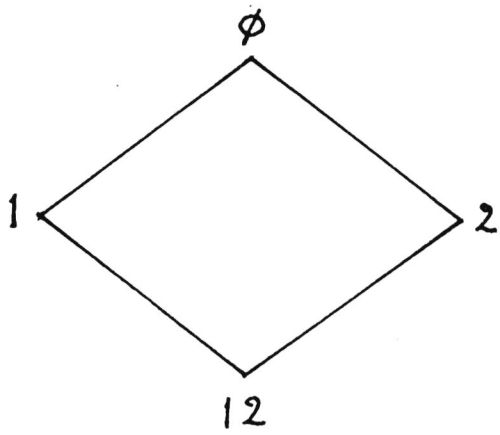
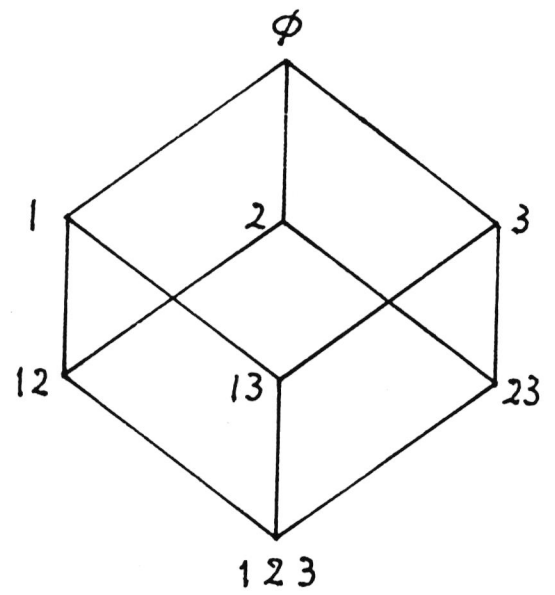
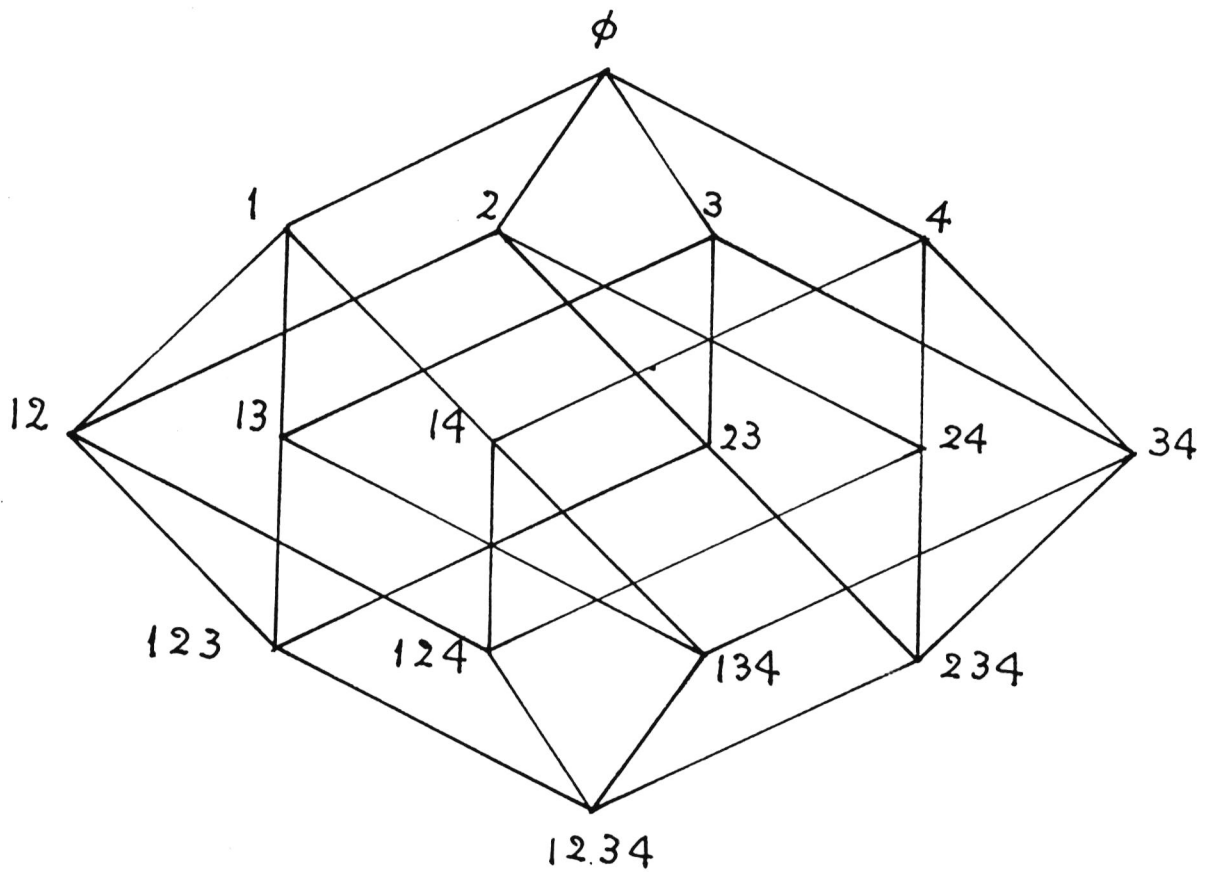
(a) $m=2$ (b) $m=3$ (c) $m=4$

Fig 2.1

要とする範囲で論じる。

定義 2.3. 生成クラス. M の部分集合のクラス $\mathcal{L}$ が次の条件を満たすとき、その $\mathcal{L}$ を生成クラス (generating class) と呼ぶ: すなわち、 $\alpha, \alpha' \in \mathcal{L}$ のとき、もし $\alpha \subset \alpha'$ なら $\alpha \equiv \alpha'$ が成り立つ。

例 2.1. 以下の例では図 2.1 を参照されたい。

$$\mathcal{L} = \{34, 123, 124\}$$

は生成クラスである。 $\{23, 123, 124\}$ は生成クラスでない。

定義 2.4. 階層的クラス. 生成クラス $\mathcal{L}$ によって生成されるクラス、すなわち $\mathcal{L}$ の任意の要素のすべての部分集合からなるクラスを階層的クラスという。それを $H(\mathcal{L})$ と書く。

例 2.2. 生成クラス $\mathcal{L}$

$$\mathcal{L} = \{12, 23\}$$

とすると、

$$H(\mathcal{L}) = \{\emptyset, 1, 2, 3, 12, 23\}$$

であり、 $\{\emptyset, 1, 3, 12, 23\}$ は階層的でない。

なお、階層的クラス $H(\mathcal{L})$ とその生成クラスによって、上の例では $H(\mathcal{L}) = \{12/23\}$ のように、生成クラスの要素を斜線で区切って書く。

定義2.5. 階層的モデル. ある階層的クラスに対応するモデル:

$$A \in \Omega[H(L)] \quad (2.31)$$

を階層的モデルという.

階層的モデルは非階層的モデルの場合より簡単な表現が可能である.

(2.31) のモデル表現を

$$\left. \begin{aligned} A &= \sum_{\forall \alpha \in H(L)} B(\alpha) \\ B(\alpha) &\in \Omega(\alpha) \end{aligned} \right\} \quad (2.32)$$

と書くこともできる. これに対し, 生成クラス L の各要素に対し

$$\omega(\alpha) = \bigoplus_{\forall \alpha' \subset \alpha} \Omega(\alpha') \quad (2.33)$$

すなわち, $\omega(\alpha)$ を α に含まれるすべての部分集合 α' に関する α' -効果空間の直和とし, (2.32) にかえて

$$\left. \begin{aligned} A &= \sum_{\forall \alpha \in L} b(\alpha) \\ b(\alpha) &\in \omega(\alpha) \end{aligned} \right\} \quad (2.34)$$

と表現することもできる. 一般に $H(L)$ の要素数は L の要素にくらべてはるかに多くなるので (2.34) の表現は (2.32) より簡潔である.

例2.3. 生成クラスを例2.2の $\{12, 23\}$ とする. また, $r \times s \times t$ の3重配列をとりあげ, A の要素を $A_{\kappa\lambda\mu}$, $B(\alpha)$ と $b(\alpha)$ について $B_{\kappa\lambda\mu}^{(\alpha)}$ と $b_{\kappa\lambda\mu}^{(\alpha)}$ と書く. このとき, (2.32) の第1式は

$$A_{\kappa\lambda\mu} = B_{\kappa\lambda\mu}^{(0)} + B_{\kappa\lambda\mu}^{(1)} + B_{\kappa\lambda\mu}^{(2)} + B_{\kappa\lambda\mu}^{(3)} \\ + B_{\kappa\lambda\mu}^{(12)} + B_{\kappa\lambda\mu}^{(23)} \quad (\forall \kappa \in \bar{\tau}, \forall \lambda \in \bar{\sigma}, \forall \mu \in \bar{\xi})$$

(2.34) の第1式は

$$A_{\kappa\lambda\mu} = b_{\kappa\lambda\mu}^{(12)} + b_{\kappa\lambda\mu}^{(23)} \quad (\forall \kappa \in \bar{\tau}, \forall \lambda \in \bar{\sigma}, \forall \mu \in \bar{\xi})$$

となる.

(2.33) の $\omega(\alpha)$ への Ω からの射影影子は, Ω から $\Omega(\alpha)$ への射影影子 $U(\alpha)$ [(2.23) 式] の和 Σ となることによって求められる. それを $V(\alpha)$ とすると,

$$V(\alpha) = \sum_{\forall \alpha' \in H(\alpha)} U(\alpha') \\ = \bigotimes_{i=1}^m V_{\chi_{\alpha}(i)}^{(i)} \quad (2.35)$$

を得る. ここで, $V_0^{(i)} = U_0^{(i)}$, $V_1^{(i)} = I_{r_i}$, $U_0^{(i)}$ は (2.18) で定義され $\Omega^{(i)}$ から $\text{span}(U_{\alpha}^{(i)})$ への射影影子である. ところで, A の $\omega(\alpha)$ への射影成分は

$$\underline{b}(\alpha | V(\alpha)) = V(\alpha) A \quad (2.36)$$

によって求められる. ただし, $V(\alpha)$ と $V(\alpha')$ ($\alpha \neq \alpha'$) は直交せず, (2.34) の $b(\alpha)$ と (2.36) の $\underline{b}(\alpha | V(\alpha))$ によって一意に決まるものとみなすことはできない. (2.32) の表現では $B(\alpha)$ を (2.20) で与えることができる. (2.32) の第2式は (2.20) と同値である.

階層的モデルを採用する利点は、モデル表現の単純化のほかに、モデルの部分空間 $\Omega[\mathcal{M}(\mathcal{L})]$ を張る基底を簡単に求めらんとができること、多重分割表のセル確率をそのモデルのもとで推定する場合の都合のよさなどがある [例えば, Habermann (1974) 参照]。

次に、生成クラス $\mathcal{L}$ の分解可能性の概念を定義する。

定義2.6. 生成クラス $\mathcal{L}$ が n 個の集合要素から成るが、あるいは、 $\mathcal{L}$ が共通要素を持たない (disjoint な) n 個の分解可能なクラス $\mathcal{L}_1$ と $\mathcal{L}_2$ に分け、かつ $\alpha_1 \in \mathcal{L}_1$, $\alpha_2 \in \mathcal{L}_2$ が存在して

$$[\cup(\mathcal{L}_1)] \cap [\cup(\mathcal{L}_2)] = \alpha_1 \cap \alpha_2 \quad (2.37)$$

となるとき、生成クラス $\mathcal{L}$ は分解可能であるという。

例2.4 $\{12\}$ は分解可能である (一つの集合要素からなる)。

$$\mathcal{L}_1 = \{12, 23\}$$

は、分解可能な $\{12\}$ と $\{23\}$ からなり、(2.37) の条件を満たすので分解可能な生成クラスである。

$$\mathcal{L}_2 = \{12, 13, 23\}$$

は分解可能でない。

$$\mathcal{L}_3 = \{34, 123, 124\}$$

は分解可能でない。

定義 2.7. 分解型モデル. 分解可能な生成クラス $\mathcal{L}$ に対する階層的モデルを分解型 (あるいは, closed form) モデルと呼ぶ.

定義 2.8. 生成クラス $\mathcal{L}$ が空でない二つのクラス $\mathcal{L}_1$ と $\mathcal{L}_2$ に分解され

$$(\cup \mathcal{L}_1) \cap (\cup \mathcal{L}_2) = \phi \quad (2.38)$$

となるとき, その $\mathcal{L}$ に対するモデルを部分独立モデルという. とくに, $\mathcal{L} = \{1, 2, \dots, m\}$ のときは, それに対するモデルを相互独立モデルという.

定義 2.9. 生成クラス $\mathcal{L}$ のすべての要素の共通部分が空でないとき, その共通部分を α_0 とする. すなわち

$$\alpha_0 = \cap \mathcal{L} (\neq \phi)$$

そのとき, $\mathcal{L}$ の各要素から α_0 を除いたクラスに対するモデルが部分独立モデルであれば, $\mathcal{L}$ に対するモデルを α_0 について条件独立モデルという.

例 2.5. $m=3$ の場合について整理しておこう. $\mathcal{L}$ が $\{1, 2, 3\}$ であるとき 相互独立モデル, $\mathcal{L}$ が $\{1, 23\}, \{2, 13\}, \{3, 12\}$ のいずれかの場合 部分独立モデル, $\mathcal{L}$ が $\{12, 13\}, \{12, 23\}, \{13, 23\}$ のいずれかの場合は条件独立モデルが作られる. $\mathcal{L}$ が $\{12, 13, 23\}$ のときは上記いずれのモデルでもなく 連関モデルと呼んでおく. $\mathcal{L}$ が $\{123\}$ のときは, しばしば飽和モデルといわれる.

2.3 内積の一般化

多重分割表における加法モデルや Lancaster (1967) による多次元分布の標準表現に関連するベクトル空間を論じるには, 2.1 節で用いた内積の定義 (2.1) を一般化しておく必要がある. そこで, 多重配列の i 指標に対応するベクトル空間 $\Omega^{(i)}$ における内積を次のように定義する: 二つのベクトル $\underline{u}^{(i)}, \underline{v}^{(i)} \in \Omega^{(i)}$ に対して, その内積を

$$\begin{aligned} (\underline{u}^{(i)}, \Sigma^{(i)} \underline{v}^{(i)}) &= \underline{u}^{(i)'} \Sigma^{(i)} \underline{v}^{(i)} \\ &= \sum_{\lambda_i, \lambda'_i} \sigma_{\lambda_i \lambda'_i}^{(i)} u_{\lambda_i}^{(i)} v_{\lambda'_i}^{(i)} \quad (\forall i \in \bar{m}) \end{aligned} \quad (2.40)$$

とする. ここで, $\Sigma^{(i)} = (\sigma_{\lambda_i \lambda'_i}^{(i)})$ は $r_i \times r_i$ の正值対称行列で $\Omega^{(i)}$ ごとに与えられるものとする. 現実的応用では $\Sigma^{(i)}$ を対角行列にとることも多い.

2.1 節にならって,

$$\underline{u} = \bigotimes_{i=1}^m \underline{u}^{(i)}, \quad \underline{v} = \bigotimes_{i=1}^m \underline{v}^{(i)}$$

とし, $\Sigma^{(i)} (\forall i \in \bar{m})$ の Kronecker 積を Σ , となゆ

$$\Sigma = \bigotimes_{i=1}^m \Sigma^{(i)} \quad (2.41)$$

と置く. 次の補題は補題 2.5 から容易に導かれる.

補題 2.6. 次式が成り立つ:

$$\underline{u}' \Sigma \underline{v} = \prod_{i=1}^m (\underline{u}^{(i)'} \Sigma^{(i)} \underline{v}^{(i)}). \quad (2.42)$$

補題 2.7. Σ は正值対称行列である.

以上の補題から, $\Omega^{(i)} (\forall i \in \bar{m})$ のテンソル積空間 Ω では, $A, B \in \Omega$ に対し, (2.41) の Σ を用いて

$$\underline{A}' \Sigma \underline{B} \quad (2.43)$$

によって, その内積を定義することができる. そして, $\underline{A}' \Sigma \underline{B} = 0$ となるとき, $\underline{A}$ と $\underline{B}$ は Σ -直交であるという. 同様に, 二つのベクトル $\underline{u}^{(i)}, \underline{v}^{(i)} \in \Omega^{(i)}$ が $\underline{u}^{(i)'} \Sigma^{(i)} \underline{v}^{(i)} = 0$ となるとき, $\underline{u}^{(i)}$ と $\underline{v}^{(i)}$ は $\Sigma^{(i)}$ -直交であるという. 次に $\Sigma^{(i)\frac{1}{2}}$ を次式を満たす $r_i \times r_i$ の正則行列とする:

$$(\Sigma^{(i)\frac{1}{2}})' (\Sigma^{(i)\frac{1}{2}}) = \Sigma^{(i)} \quad (\forall i \in \bar{m}) \quad (2.44)$$

その存在は $\Sigma^{(i)}$ の正值対称性の仮定から保証される.

$$\Sigma^{\frac{1}{2}} = \otimes_{i=1}^m (\Sigma^{(i)\frac{1}{2}}) \quad (2.45)$$

と置くと,

$$(\Sigma^{\frac{1}{2}})' (\Sigma^{\frac{1}{2}}) = \Sigma$$

となることは, 補題 2.5 から明らかである. 同様に, 逆行列についても

$$\left. \begin{aligned} (\Sigma^{\frac{1}{2}})^{-1} &= \Sigma^{-\frac{1}{2}} = \otimes_{i=1}^m (\Sigma^{(i)-\frac{1}{2}}) \\ (\Sigma^{\frac{1}{2}})' (\Sigma^{-\frac{1}{2}}) &= \Sigma^{-1} \end{aligned} \right\} \quad (2.46)$$

が成り立つ.

さて, テンソル積 Ω に属する多重配列 A を

$$\underline{\bar{A}} = \Sigma^{\frac{1}{2}} \underline{A} \quad (2.47)$$

と変換し, $\bar{A}$ の空間を $\bar{\Omega}$ とする. これを形式的に

$$\bar{\Omega} = \Sigma^{\frac{1}{2}} \Omega \quad (2.48)$$

と書き, 同様の意味で

$$\bar{\Omega}^{(i)} = \Sigma^{(i)\frac{1}{2}} \Omega^{(i)} \quad (\forall i \in \bar{m}) \quad (2.49)$$

とすると, $\bar{\Omega}$ は $\bar{\Omega}^{(i)}$ のテンソル積である. $\bar{\Omega}^{(i)}$ での内積は, 任意の $\underline{w}^{(i)}, \underline{x}^{(i)} \in \bar{\Omega}^{(i)}$ に対し, (2.49) の変換によって それぞれに対応する

$$\underline{u}^{(i)} = \Sigma^{-\frac{1}{2}} \underline{w}^{(i)} \quad (2.50)$$

$$\underline{v}^{(i)} = \Sigma^{-\frac{1}{2}} \underline{x}^{(i)}$$

から

$$(\underline{w}^{(i)}, \underline{x}^{(i)}) = \underline{w}^{(i)'} \underline{x}^{(i)} = \underline{u}^{(i)} \Sigma^{(i)} \underline{v}^{(i)} \quad (2.51)$$

としよう. そこで, $\bar{\Omega}$ と $\bar{A}$ について; $w^{(i)} \in \bar{\Omega}^{(i)} (\forall i \in \bar{m})$ を与えたとする $\bar{\Omega}$ の α -効果空間, $\bar{A}$ の α -因子効果を 2.1 節の議論に従って求めることができる.

いま, $\underline{w}^{(i)} \in \bar{\Omega} (\forall i \in \bar{m})$ が与えられたとする. 簡単のため

$$\underline{w}^{(i)'} \underline{w}^{(i)} = 1 \quad (\forall i \in \bar{m}) \quad (2.52)$$

としておく. この $\underline{w}^{(i)}$ から 射影子

$$\left. \begin{aligned} W_0^{(i)} &= w^{(i)} w^{(i)'} \\ W_1^{(i)} &= I_{r_i} - W_0^{(i)} \end{aligned} \right\} \quad (2.53)$$

を定義しておく. 定理 2.2 と 2.3 から

$$\bar{A} = \sum_{\alpha} \left(\otimes_{i=1}^m W_{\chi_{\alpha}(i)}^{(i)} \right) \bar{A}$$

を得る. 上式 $\bar{A}$ に $\Sigma^{\frac{1}{2}} A$ [(2.47) 式] を代入し, 両辺に $\Sigma^{-\frac{1}{2}}$ をかけると,

$$A = \Sigma^{-\frac{1}{2}} \sum_{\alpha} \left(\otimes_{i=1}^m W_{\chi_{\alpha}(i)}^{(i)} \right) \Sigma^{\frac{1}{2}} A$$

を得る. ここで (2.45), (2.46), (2.50) を用い, さらに $\bar{U}_0^{(i)}$ と $\bar{U}_1^{(i)}$ を

$$\left. \begin{aligned} \bar{U}_0^{(i)} &= (\Sigma^{(i)})^{-\frac{1}{2}} \underline{W}^{(i)} \underline{W}^{(i)'} (\Sigma^{(i)})^{\frac{1}{2}} \\ &= \underline{U}^{(i)} \underline{U}^{(i)'} \Sigma^{(i)} \\ \bar{U}_1^{(i)} &= I_{r_i} - \bar{U}_0^{(i)} \end{aligned} \right\} \quad (2.54)$$

としておく,

$$\bar{A} = \sum_{\alpha} \left(\otimes_{i=1}^m \bar{U}_{\chi_{\alpha}(i)}^{(i)} \right) A \quad (2.55)$$

を得る. 上式の各項を $B(\alpha; \Sigma)$, 3 はわら

$$\bar{B}(\alpha; \Sigma) = \otimes_{i=1}^m \bar{U}_{\chi_{\alpha}(i)}^{(i)} A \quad (2.56)$$

と置くと, $B(\alpha; \Sigma)$ は r_i 個に Σ -直交であることがわかる. 従って,

定理 2.4 の平方和分解は

$$A' \Sigma A = \sum_{\alpha} B(\alpha; \Sigma)' \Sigma B(\alpha; \Sigma) \quad (2.57)$$

に帰着する.

本章以後では, 多重配列 $A \in \underline{U}^{(i)} \in \Omega^{(i)} (\forall i \in \bar{m})$ についての多重-次

形式

$$\sum_{\lambda_1 \dots \lambda_m} A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots u_{\lambda_m}^{(m)} = \underline{A}' \underline{u} \quad (2.58)$$

の係数と考えろ。そのとき、 A はテンソル積空間の双対(dual)空間の要素とみなされる。従って、 $\Omega^{(i)}$ での内積を(2.40)で定義したときは、 Ω の双対空間では Σ の逆行列 Σ^{-1} を用いて内積を定義することになる。そこで、変換

$$\underline{\hat{A}} = (\Sigma^{-1/2})' \underline{A} \quad (2.59)$$

を(2.47)のかたちに用いればこれまでの議論が併行して適用される。なお、 $\underline{u}^{(i)}$ と $\underline{w}^{(i)}$ についてはこれまでどおりとする[(2.50), (2.52)参照]。

そこで、(2.54)の $\underline{\bar{U}}_0^{(i)}$ と $\underline{\bar{U}}_1^{(i)}$ に対応して、 $\underline{\hat{U}}_0^{(i)}$ と $\underline{\hat{U}}_1^{(i)}$ を

$$\left. \begin{aligned} \underline{\hat{U}}_0^{(i)} &= \Sigma^{(i)} \underline{u}^{(i)} \underline{u}^{(i)'} \\ \underline{\hat{U}}_1^{(i)} &= I_{r_i} - \underline{\hat{U}}_0^{(i)} \end{aligned} \right\} \quad (\forall i \in \bar{m}) \quad (2.60)$$

とすると、(2.55)に対応して

$$\underline{A} = \sum_{\alpha} \left(\bigotimes_{i=1}^m \underline{\hat{U}}_{\chi_{\alpha}(i)}^{(i)} \right) \underline{A} \quad (2.61)$$

を得る。上式の各項を

$$\underline{B}(\alpha; \Sigma^{-1}) = \left(\bigotimes_{i=1}^m \underline{\hat{U}}_{\chi_{\alpha}(i)}^{(i)} \right) \underline{A} \quad (2.62)$$

と置くと、 $\underline{B}(\alpha; \Sigma^{-1})$ は異なる α について互いに Σ^{-1} -直交である。

従って、平方和の分解は

$$\underline{A}' \Sigma^{-1} \underline{A} = \sum_{\alpha} \underline{B}(\alpha; \Sigma^{-1})' \Sigma^{-1} \underline{B}(\alpha; \Sigma^{-1}) \quad (2.63)$$

という形式で行われる。なお、(2.50)の多重一次形式の値は、(2.59)の変換と

$$\begin{aligned}\underline{w} &= \bigotimes_{i=1}^m \underline{w}^{(i)} = \bigotimes_{i=1}^m \Sigma^{(i)\frac{1}{2}} \underline{u}^{(i)} \\ &= \Sigma^{\frac{1}{2}} \underline{u}\end{aligned}$$

の変換によって、不変である。すなわち

$$\hat{A}' \underline{w} = (\Sigma^{\frac{1}{2}'} A)' \Sigma^{\frac{1}{2}} \underline{u} = A' \underline{u}$$

となる。

最後に、(2.54)の $\overline{U}_0^{(i)}$ と (2.60)の $\hat{U}_0^{(i)}$ の意味について注意しておく。

$\overline{U}_0^{(i)}$ は内積が (2.40) で定義された空間 $\Omega^{(i)}$ でベクトル $\underline{u}^{(i)}$ の張る空間 $\text{span}(\underline{u}^{(i)})$ への $\Sigma^{(i)}$ -直交の方向への射影子になっている。一方の $\hat{U}_0^{(i)}$ は、 $\overline{U}_0^{(i)}$ を転置したもので、 $\Omega^{(i)}$ の双対空間において $\text{span}(\underline{u}^{(i)})$ に双対な部分空間に沿った射影子になっている。すなわち、 $\overline{U}_0^{(i)}$ と $\hat{U}_0^{(i)}$ は双対な射影子である【例えば、Halmos (1958)】。

3. 多重配列の特異値分解

この章では、二重の配列、すなわち行列の特異値分解の多重配列への拡張を試みる。与えられた多重配列の要素を係数とする多重一次形式の最大化問題から始めて、その問題の決定方程式を反復的に適用する算法を記述し、多重配列の分解表現を導く。その分解表現に対応する部分空間の次元や平方和の分解に触れ、特異値分解の諸性質を調べる。

3.1 算法

前章で論じた母数空間としての Ω における効果空間や多重配列構造を持つ母数 A の因子効果の定義は、ベクトル $\underline{u}^{(i)} \in \Omega^{(i)}$ の選択に依存する形式で一般的に与えられた。そこで、統計的に意味のあるベクトル $\underline{u}^{(i)}$ の抽出法の一つとして、次の問題を考える。

問題 1. 寸法 $r_1 \times r_2 \times \cdots \times r_m$ の m 重配列 A が与えられたとき、条件

$$\underline{u}^{(i)'} \underline{u}^{(i)} = \sum_{\lambda_i} \{ u_{\lambda_i}^{(i)} \}^2 = 1 \quad (\forall i \in \bar{m}) \quad (3.1)$$

のもとで

$$\begin{aligned} p &= A' \underline{u} \\ &= \sum A_{\lambda_1 \lambda_2 \cdots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \cdots u_{\lambda_m}^{(m)} \end{aligned} \quad (3.2)$$

を最大にするベクトル $\underline{u}^{(i)} \in \Omega^{(i)} \quad (\forall i \in \bar{m})$ を求めること[†]。

[†] $\sum$ 記号で和をとる添字を省略したときは A と $\underline{u}$ の因子で一致する添字について和をとるものとする。なお、 $\underline{u} = \otimes_{i=1}^m \underline{u}^{(i)}$ である。

ここでは、条件(3.1)を正規化条件、 ρ を最高次相関関数と呼ぶ。まず、 ρ の上限を与える次の定理を示す。

定理3.1. 最高次相関関数 ρ について、次の不等式が成立立つ：

$$\rho^2 \leq \underline{A}' \underline{A}. \quad (3.3)$$

等号が成立するのは、 A が $\underline{u}$ に比例するとき、そのときに限る。

証明. Schwarzの不等式により、(3.2)から

$$\rho^2 \leq (\underline{A}' \underline{A}) (\underline{u}' \underline{u})$$

が成立立つ。正規化条件(3.1)により、補題2.6で $\Sigma^{(i)} = I_{r_i} \ (\forall i \in \bar{m})$, $\underline{u}' \underline{u} = 1$ とすれば

$$\underline{u}' \underline{u} = 1 \quad (3.4)$$

となるので、 ρ に対して(3.3)が導かれる。等号をとるための条件もSchwarzの不等式から明らかである。 |

定理3.2. 問題1の解は存在する。また、その解は次の決定方程式の最大根に対応するベクトルを条件(3.1)で正規化したものである：

$$\left. \begin{aligned} \sum A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_2}^{(2)} u_{\lambda_3}^{(3)} \dots u_{\lambda_m}^{(m)} &= \rho u_{\lambda_1}^{(1)}, \\ \sum A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_3}^{(3)} \dots u_{\lambda_m}^{(m)} &= \rho u_{\lambda_2}^{(2)}, \\ &\dots \dots \dots \\ \sum A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots u_{\lambda_{m-1}}^{(m-1)} &= \rho u_{\lambda_m}^{(m)}. \end{aligned} \right\} \quad (3.5)$$

証明. 問題1の m 個の正規化条件(3.1)のJacobianの階数は m であり, ベクトル $u^{(i)}$ ($\forall i \in \bar{m}$) のすべての要素が有界閉領域を動くことから問題1の解が存在し, かつ, その解がLagrange未定乗数法によって得られる決定方程式(制約条件(3.1)を含んで)の解で与えられることは既存の定理[例えば, 竹内(1972, p.37)参照]から容易に証明される. ここで決定方程式(3.5)を導出しておく. すると,

$$\rho^* = \sum A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots u_{\lambda_m}^{(m)} - \sum_{i=1}^m \frac{\mu_i}{2} (\sum \{u_{\lambda_i}^{(i)}\}^2 - 1)$$
と置く. ここで, μ_i ($\forall i \in \bar{m}$) はLagrange乗数である. ρ^* を変数 $u_{\lambda_i}^{(i)}$ ($\forall \lambda_i \in \bar{\gamma}_i, \forall i \in \bar{m}$) と乗数 μ_i ($\forall i \in \bar{m}$) で偏微分して

$$\frac{\partial \rho^*}{\partial u_{\lambda_i}^{(i)}} = \sum A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots \cancel{u_{\lambda_i}^{(i)}} \dots u_{\lambda_m}^{(m)} - \mu_i u_{\lambda_i}^{(i)} = 0$$

$$\frac{\partial \rho^*}{\partial \mu_i} = \sum \{u_{\lambda_i}^{(i)}\}^2 - 1 = 0$$

を得る. ここで, μ_i は u の*i*因子 $u_{\lambda_i}^{(i)}$ を除くことを意味する. すると

$$\sum \frac{\partial \rho^*}{\partial u_{\lambda_i}^{(i)}} u_{\lambda_i}^{(i)}$$

を評価すれば,

$$\mu_i = \sum A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots u_{\lambda_m}^{(m)} = \rho \quad (\forall i \in \bar{m}) \quad (3.6)$$

が得られ, 決定方程式(3.5)が導かれる. なお, (3.5)の*i*式は $\forall \lambda_i \in \bar{\gamma}_i$ について考えられることに注意しておく. |

決定方程式 (3.5) の最大根を ρ_1 , ρ_1 に対応するベクトルを $\underline{u}^{(1)} (i \in \bar{m})$ と書く. このベクトル $\underline{u}^{(1)} (i \in \bar{m})$ が得られたとき, それに関する多重配列 A の最高次交互作用, すなわち $\bar{m}$ -因子効果を求め, それを $\tilde{A}^2$ とする. この $\tilde{A}^2$ を新しい m 重配列として問題 1 を解く. この操作を反復する過程を次のように定義する.

算法 1. m 重配列 A が与えられたとする.

- $\tilde{A}^1 = A$ とする.
- $k = 1, 2, \dots$ について以下を反復する.
 - $\tilde{A}^k$ のすべての要素がゼロなら反復を終了する.
 - $\tilde{A}^k$ について決定方程式 (3.5) を解く. その最大根を ρ_k , 対応するベクトルを $\underline{u}^{(k)} (i \in \bar{m})$ とする. ベクトル $\underline{u}^{(k)} (i \in \bar{m})$ は正規化条件を満たすものとする.
 - $\underline{u}^{(k)} (i \in \bar{m})$ に関する $\tilde{A}^k$ の最高次交互作用を $\tilde{A}^{k+1}$ とする.

注意. $\tilde{A}^{k+1}$ は定理 2.2 の (2.20) を適用して

$$\tilde{A}^{k+1} = \left(\bigotimes_{i=1}^m \underline{u}_i^{(k)} \right) \tilde{A}^k \quad (3.7)$$

で求められる. これを $\tilde{B}^k(\bar{m})$ と書くこともある. $\underline{u}_i^{(k)}$ は (2.19) に対応する射影子で, $\underline{u}_0^{(k)}$ とともに, 条件 (3.1) により

$$\left. \begin{aligned} \underline{U}_0^{(i)} &= \underline{u}^{(i)} \underline{u}^{(i)'} \\ \underline{U}_1^{(i)} &= I_{r_i} - \underline{u}^{(i)} \underline{u}^{(i)'} \end{aligned} \right\} (\forall i \in \bar{m}, k=1, 2, \dots) \quad (3.8)$$

で与えられる。なお、後の引用のために、 $\overset{k+1}{A}$ の添字ジとの定義式と $\overset{k}{A}$ についての決定方程式を記しておく。

$$\begin{aligned} \overset{k+1}{A}_{\lambda_1 \lambda_2 \dots \lambda_m} &= \sum \left[(\delta_{\lambda_1 \lambda'_1} - \underline{u}_{\lambda_1}^{(1)} \underline{u}_{\lambda'_1}^{(1)}) (\delta_{\lambda_2 \lambda'_2} - \underline{u}_{\lambda_2}^{(2)} \underline{u}_{\lambda'_2}^{(2)}) \dots \right. \\ &\quad \left. (\delta_{\lambda_m \lambda'_m} - \underline{u}_{\lambda_m}^{(m)} \underline{u}_{\lambda'_m}^{(m)}) \right] \overset{k}{A}_{\lambda'_1 \lambda'_2 \dots \lambda'_m} \end{aligned} \quad (3.9)$$

ここで、 $\delta_{\lambda_i \lambda'_i}$ は Kronecker のデルタで $\lambda_i = \lambda'_i$ のとき 1 でその他のときはゼロの値をとる。 $\sum$ は $\lambda'_1, \dots, \lambda'_m$ についてとる。第 k 段での決定方程式は

$$\begin{aligned} \sum \overset{k}{A}_{\lambda_1 \lambda_2 \dots \lambda_m} \underline{u}_{\lambda_1}^{(1)} \underline{u}_{\lambda_2}^{(2)} \dots \underline{u}_{\lambda_m}^{(m)} &= \rho_k \underline{u}_{\lambda_i}^{(i)} \\ (\forall \lambda_i \in \bar{\gamma}_i, \forall i \in \bar{m}) \end{aligned} \quad (3.10)$$

となる。

次に、算法 1 の有限性を示す。

定理 3.3. 任意の実数値配列 A に対し、算法 1 は k について高々 r 回の反復で終了する。すなわち

$$\overset{r+1}{A}_{\lambda_1 \lambda_2 \dots \lambda_m} = 0 \quad (\forall \lambda_i \in \bar{\gamma}_i, \forall i \in \bar{m}) \quad (3.11)$$

となる。ここで、

$$r = \min(r_1, r_2, \dots, r_m) \quad (3.12)$$

である.

この定理の証明は以下の補題を示してから行う.

補題3.1. 任意の $\underline{u}^{(i)} \in \Omega^{(i)}$ ($\forall i \in \bar{m}$) に関する A の α -因子効果 $B(\alpha)$

について, $i \in \alpha$ なら次式が成り立つ:

$$\sum_{\lambda_i} B_{\lambda_1 \lambda_2 \dots \lambda_m}(\alpha) \cdot u_{\lambda_i}^{(i)} = 0 \quad (\forall i \in \alpha) \quad (3.13)$$

注意. (3.13) は λ_i を除く任意の指標の値で成立する. (2.32) のようなモデル表現において (3.13) を 因子効果のパラメータに対する制約とすることもできる.

証明. $B(\alpha)$ を与える (2.20) の右辺の射影因子

$$\otimes_{i=1}^m U_{\chi_\alpha(i)}^{(i)}$$

には, $i \in \alpha$ なら $U_1^{(i)}$ が含まれる. $U_1^{(i)}$ の定義 (2.19) から

$$U_1^{(i)} \underline{u}^{(i)} = 0$$

となり, (3.13) が得られる. |

補題3.2. 算法1で求められる $\hat{A}^k, \hat{\underline{u}}^{(i)}, k=1, 2, \dots$ について

$$\sum_{\lambda_i} \hat{A}_{\lambda_1 \lambda_2 \dots \lambda_m}^k \hat{u}_{\lambda_i}^{(i)} = 0 \quad (j < k) \quad (3.14)$$

及び

$$\sum_{\lambda_i} \hat{u}_{\lambda_i}^{(i)} \hat{u}_{\lambda_i}^{(i)} = 0 \quad (j \neq k) \quad (3.15)$$

が 任意の $i \in \bar{m}$ について成り立つ.

注意. (3.15) は 算法1によって $\Omega^{(i)}$ における正規直交系 $\{\hat{\underline{u}}^{(i)}, \hat{\underline{u}}^{(i)}, \dots,$

$\underline{u}^{(i)}\}$ が得られることを示す.

証明. (3.15) は j と k について対称であるから, (3.14) と同様に $j < k$ のときについて証明すれば十分である. 証明は帰納法による. まず, j を固定する. $k = j+1$ のとき, $\overset{j+1}{A}$ は $\overset{j}{A}$ に対する $\underline{u}^{(i)}$ ($\forall i \in \bar{m}$) を用いたときの $\bar{m}$ -因子効果であるから, 補題 3.1 によって (3.14) が成り立つ. 才 k 段での反復の決定方程式 (3.10) の両辺に $\underline{u}_{\lambda_i}^{(i)}$ をかけて λ_i について和をとる. すなわち,

$$\sum \left(\sum_{\lambda_i} \overset{k}{A}_{\lambda_1 \lambda_2 \dots \lambda_m} \underline{u}_{\lambda_i}^{(i)} \right) \overset{k(1)}{u}_{\lambda_1} \overset{k(2)}{u}_{\lambda_2} \dots \overset{k(m)}{u}_{\lambda_m} = \sum_{\lambda_i} \overset{k}{p}_k \overset{k(i)}{u}_{\lambda_i} \overset{j(i)}{u}_{\lambda_i}$$

しなわち, (3.14) が成立すれば上式の左辺がゼロになり, その右辺がゼロ, すなわち, (3.15) を得る. よって, まず $k = j+1$ のときは (3.14) と (3.15) が成立する. 次に, $k = j+1, \dots, k'$ まで (3.14) と (3.15) が成立すると仮定する. そのとき, (3.9) の $k \in k'$ に変え, その両辺に $\underline{u}_{\lambda_i}^{(i)}$ を掛けて λ_i について和をとる. その左辺は

$$\sum_{\lambda_i} \overset{k'+1}{A}_{\lambda_1 \lambda_2 \dots \lambda_m} \underline{u}_{\lambda_i}^{(i)}$$

である. 右辺の因子で λ_i と λ'_i を添字に持つものをとり出して処理すると,

$$\begin{aligned} & \sum_{\lambda_i, \lambda'_i} \left[(\delta_{\lambda_i \lambda'_i} - \overset{k'(i)}{u}_{\lambda_i} \overset{k'(i)}{u}_{\lambda'_i}) \overset{k'}{A}_{\lambda_1 \lambda_2 \dots \lambda_m} \underline{u}_{\lambda_i}^{(i)} \right] \\ &= \sum_{\lambda'_i} \overset{k'}{A}_{\lambda_1 \lambda'_2 \dots \lambda'_m} \underline{u}_{\lambda'_i}^{(i)} - \left(\sum_{\lambda_i} \overset{k'(i)}{u}_{\lambda_i} \underline{u}_{\lambda_i}^{(i)} \right) \left(\sum_{\lambda'_i} \overset{k'(i)}{u}_{\lambda'_i} \overset{k'}{A}_{\lambda_1 \lambda'_2 \dots \lambda'_m} \right) \end{aligned}$$

となる. 仮定により, 上式はゼロとなる. よって, j と $k = k'+1$ の組に対し,

(3.14) が成り立ち, したがって (3.15) も成立する. 以上の議論が特定の j に依存しないことは明らかである. よって, 補題が証明される. |

定理 3.3 の証明. ベクトル空間 $\Omega^{(i)}$ の次え r_i の最小値である r を越えて, $r+1$ の要素でゼロでないものが存在すると仮定する. そこで, $r+1$ について決定方程式 (3.10) を解けば, ゼロベクトルでない解 $\underline{u}^{(i)} \in \Omega^{(i)}$ ($\forall i \in \bar{m}$) が得られる. なぜなら, 定理 3.2 によりその解が存在し, ゼロベクトルでは正規化条件 (3.1) を満たさないからである. 補題 3.2 により, $\underline{u}^{(1)}, \dots, \underline{u}^{(i)}, \underline{u}^{(i+1)}$ は互いに直交する. ところが, r_i が r に等しいベクトル空間 $\Omega^{(i)}$ では $r+1$ 個の互いに直交するベクトルは存在しない. よって, 上記の仮定は矛盾を生じ, 定理が証明される. |

問題 1 はベクトル空間 $\Omega^{(i)}$ ($\forall i \in \bar{m}$) のテンソル積 Ω において

$$\underline{u} = \bigotimes_{i=1}^m \underline{u}^{(i)}$$

の形式のベクトルで, 与えられた多重配列 A に相関関数の意味でもっとも近いものを求めている. そこで, 問題 1 は次の問題と同値であることを示される.

問題 2. 問題 1 と同じ条件 (3.1) のもとで

$$Y = (\underline{A} - \rho \underline{u})' (\underline{A} - \rho \underline{u}) \quad (3.16)$$

を最小にすること. ここで, ρ を正としておく.

定理 3.4. 問題 1 と問題 2 は同値である。すなわち、両者の解 $\underline{u}^{(i)}$ ($\forall i \in \bar{m}$) は一致し、問題 1 の (3.2) の最大値 ρ_1 は問題 2 の解の ρ に一致する。ただし、両者の解が一意でなければ $\underline{u}^{(i)}$ の一貫性は向わない。

証明. 任意の $\underline{u}^{(i)} \in \Omega^{(i)}$ ($\forall i \in \bar{m}$) に対し

$$\rho_1 \geq \underline{A}' \underline{u}.$$

ここで、 $\underline{u} = \otimes_{i=1}^m \underline{u}^{(i)}$ である。また、 $(\rho_1 - \rho)^2 \geq 0$ より、 $\rho_1^2 \geq 2\rho\rho_1 - \rho^2$ である。以上の不等式を用いて (3.16) を変形し

$$\begin{aligned} \gamma &= \underline{A}' \underline{A} - 2\rho \underline{A}' \underline{u} + \rho^2 \\ &\geq \underline{A}' \underline{A} - 2\rho \rho_1 + \rho^2 \\ &\geq \underline{A}' \underline{A} - \rho_1^2 \end{aligned}$$

を得る。等号の成り立つのは、 $\rho_1 = \underline{A}' \underline{u}$ 、かつ $\rho = \rho_1$ のときに限る。よって、定理が証明された。 |

注意. 問題 1 の決定方程式 (3.5) の形から、 $\rho_1, \underline{u}^{(i)}$ ($\forall i \in \bar{m}$) が (3.5) の解であるとき、 $\rho_1 \rightarrow -\rho_1$ に、奇数個の i について $\underline{u}^{(i)} \rightarrow -\underline{u}^{(i)}$ に変えたものも (3.5) の解であることは明らかである。問題 2 の (3.16) でも $\rho \underline{u}$ の形から同様のことがいえる。問題 1 の ρ を最大にする要求から ρ_1 は正とし、問題 2 でもそれに応じて ρ を正として扱う。

算法 1 では補題 3.2 の (3.15) が示すように互いに直交するベクトル群

$\underline{u}^{(i)}$, $k=1, 2, \dots, r$ ($\forall i \in \bar{m}$) が得られる. 逆に, すでに与えられた直交ベクトル群にさらに直交するという条件で $\underline{A}'\underline{u}$ を最大にする問題を考えよう.

問題3. 各ベクトル空間 $\Omega^{(i)}$ で互いに正規直交する $k-1$ ($< r$) 個のベクトル

$$\underline{u}_1^{(i)}, \underline{u}_2^{(i)}, \dots, \underline{u}_{k-1}^{(i)} \quad (\forall i \in \bar{m})$$

が与えられたとき, 条件

$$\sum_{\lambda_i} \underline{u}_{\lambda_i}^{(i)} \underline{u}_{\lambda_i}^{(i)} = 0, \quad l=1, 2, \dots, k-1 \quad (\forall i \in \bar{m}) \quad (3.17)$$

及び, 条件 (3.1) のもとで, 最高次相関関数 $\underline{A}'\underline{u}$ (3.2) を最大にするベクトル $\underline{u}^{(i)} \in \Omega^{(i)}$ ($\forall i \in \bar{m}$) を求めること.

定理 3.5. 問題3の解は,

$$\underline{\underline{A}}^k = \left(\bigotimes_{i=1}^m \underline{V}_1^{k-1(i)} \right) \underline{A} \quad (3.18)$$

で与えられる $\underline{\underline{A}}^k$ に対して問題1を解くこと, すなわち決定方程式 (3.10) を解くことによって与えられる. ここで, $\underline{V}_1^{k-1(i)}$ は

$$\underline{V}_1^{k-1(i)} = I_{r_i} - \sum_{l=1}^{k-1} \underline{u}_l^{(i)} \underline{u}_l^{(i)'} \quad (\forall i \in \bar{m}) \quad (3.19)$$

で与えられる.

注意. (3.18) の $\underline{\underline{A}}^k$ は, 算法1において決定方程式 (3.10) を解く部分省略し, その才段の反覆で問題3での $\underline{u}^{(i)}$ ($\forall i \in \bar{m}$) が与えられる

とし, $\underline{u}^{(i)}$ ($\forall i \in \bar{m}$) に関する最高次交互作用を求めることも $k-1$ 回反復したときに得られるものである. 問題 3 を $k=1, 2, \dots$ について再帰的に扱えば, 算法 1 が直接にその解を与えろことになる. もっとも, 問題 3 で与えられるベクトル $\underline{u}^{(i)}$, $i=1, \dots, k-1$ は算法 1 の解である必要はない.

定理の証明に先立って二つの補題を示しておく.

補題 3.3. (3.18) の $\underline{A}^k$ は, 漸化式

$$\left. \begin{aligned} \underline{A}^1 &= A \\ \underline{A}^{l+1} &= \left(\bigotimes_{i=1}^m \underline{U}_1^{(i)} \right)^l \underline{A}^1, \quad l=1, 2, \dots, k-1 \end{aligned} \right\} \quad (3.20)$$

を解くことで得られる. ここで, $\underline{U}_1^{(i)}$ は $\underline{u}^{(i)}$ を用いた射影子 (3.8) である ($\forall i \in \bar{m}$). また, 次の 2 式が成り立つ:

$$\bigotimes_{i=1}^m \underline{V}_1^{(i)} = \left(\bigotimes_{i=1}^m \underline{U}_1^{(i)} \right) \left(\bigotimes_{i=1}^m \underline{V}_1^{(i)} \right) \dots \left(\bigotimes_{i=1}^m \underline{U}_1^{(i)} \right) \quad (3.21)$$

$$\underline{V}_1^{(i)} \underline{u}^{(i)} = \underline{u}^{(i)} \quad (\forall i \in \bar{m}) \quad (3.22)$$

ここで, $\underline{u}^{(i)}$ は条件 (3.17) を満たすベクトルである.

証明. 漸化式を解いて

$$\underline{A}^k = \left(\bigotimes_{i=1}^m \underline{U}_1^{(i)} \right)^{k-1} \left(\bigotimes_{i=1}^m \underline{U}_1^{(i)} \right)^{k-2} \dots \left(\bigotimes_{i=1}^m \underline{U}_1^{(i)} \right)^1 A$$

を得る. 上式右辺の $\underline{A}$ にかかっている因子が (3.19) の $\underline{V}_1^{(i)}$ の積に等しいことはその因子を展開すればただちにわかる. (3.22) は (3.19) の右辺に $\underline{u}^{(i)}$

をかけた条件(3.17)を用いれば明らかである。 |

注意. $\overset{k}{A} \in (3.20)$ の漸化式で求めるにしても, 与えられたベクトル $\underline{u}^{(i)}$, $i=1, \dots, k-1$ の順序に依存しないとは (3.19) の $\overset{k-1}{V}_1^{(i)}$ から明らかである. また, $\overset{k-1}{V}_1^{(i)}$ は $\Omega^{(i)}$ における $\{\underline{u}^{(i)}, i=1, \dots, k-1\}$ を基底とする部分空間の直交補空間への射影子になっている.

補題3.4. 条件(3.17)を満たす $\underline{u}^{(i)} (i \in \bar{m})$ と (3.18) の $\overset{k}{A}$ について, 次式が成り立つ:

$$\begin{aligned} & \sum \overset{k}{A}_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots \underset{i}{\wedge} \dots u_{\lambda_m}^{(m)} \\ &= \sum A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots \underset{i}{\wedge} \dots u_{\lambda_m}^{(m)} - \sum_{i=1}^{k-1} v_{i\ell} \underline{u}_{\lambda_i}^{(i)} \quad (i \in \bar{m}) \end{aligned} \quad (3.23)$$

ここで,

$$v_{i\ell} = \sum A_{\lambda_1 \lambda_2 \dots \lambda_m} \underline{u}_{\lambda_i}^{(i)} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots \underset{i}{\wedge} \dots u_{\lambda_m}^{(m)} \quad (3.24)$$

であり.

証明. (3.23) の左辺は

$$(\underline{u}^{(1)} \otimes \underline{u}^{(2)} \otimes \dots \otimes I_{r_i} \otimes \dots \otimes \underline{u}^{(m)})' \overset{k}{A}$$

と書ける. この式の $\overset{k}{A}$ に (3.18) の右辺を代入して補題2.5を用いると

$$\begin{aligned} & (\underline{u}^{(1)} \otimes \underline{u}^{(2)} \otimes \dots \otimes I_{r_i} \otimes \dots \otimes \underline{u}^{(m)})' (\otimes_{i=1}^m \overset{k-1}{V}_1^{(i)}) \overset{k}{A} \\ &= ((\underline{u}^{(1)})' \overset{k-1}{V}_1^{(1)}) \otimes \dots \otimes \overset{k-1}{V}_1^{(i)} \otimes \dots \otimes (\underline{u}^{(m)}' \overset{k-1}{V}_1^{(m)}) \overset{k}{A} \end{aligned}$$

となる. そして, (3.22) を用いて展開すれば (3.23) を得る. |

定理 3.5 の証明. 問題 3 の解が存在し, その解が Lagrange の未定乗数法を適用して得られることは問題 1 の場合と同じである. そこで,

$$\rho^* = \sum A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots u_{\lambda_m}^{(m)} - \sum_{i=1}^m \sum_{\ell=1}^{k-1} \left(\nu_{i\ell} \sum_{\lambda_i} u_{\lambda_i}^{(\ell)} u_{\lambda_i}^{(i)} \right) - \frac{\mu_i}{2} \left(\sum \{ u_{\lambda_i}^{(i)} \}^2 - 1 \right)$$

と置く. ここで, $\nu_{i\ell}$, μ_i は Lagrange 乗数である. $\rho^* \in u_{\lambda_i}^{(i)}$ で偏微分して決定方程式を求めると, 定理 3.2 の場合と同じように μ_i は i に依らないことがわかり,

$$\sum A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots u_{\lambda_i}^{(i)} u_{\lambda_m}^{(m)} - \sum_{\ell=1}^{k-1} \nu_{i\ell} u_{\lambda_i}^{(\ell)} = \rho u_{\lambda_i}^{(i)} \quad (\forall \lambda_i \in \bar{\gamma}_i, \forall i \in \bar{m}) \quad (3.25)$$

を得る. ここで, $\nu_{i\ell}$ は (3.24) で与えられるものとなる. 補題 3.4 によって,

(3.25) の左辺は (3.23) の左辺によっておきかえられ, 決定方程式

$$\sum_k A_{\lambda_1 \lambda_2 \dots \lambda_m} u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \dots u_{\lambda_i}^{(i)} u_{\lambda_m}^{(m)} = \rho u_{\lambda_i}^{(i)}$$

を得る. 上式は (3.10) にほかならない. よって, 定理が証明される. |

問題 3 をさらに一般化して次の問題を考えようともできる.

問題 4. $\Omega^{(i)}$ の部分空間 $\omega^{(i)}$ ($\forall i \in \bar{m}$) が指定されるとき, $\underline{u}^{(i)}$ がその部分空間 $\omega^{(i)}$ に直交するという条件, 及び $\underline{u}^{(i)}$ の正規化条件のもとで, 最高次相関関数 $A' \underline{u}$ を最大にすること.

定理 3.6. $\omega^{(i)}$ への直交射影子 $V^{(i)}$ が与えられれば, 問題 4 の解は

次式で与えられる $\tilde{A}^*$ についての決定方程式 (3.10) の解で与えられる:

$$\tilde{A}^* = \left(\bigotimes_{i=1}^m (I_{r_i} - V^{(i)}) \right) \tilde{A} \quad (3.26)$$

ただし, $V^{(i)}$ の階数は r_i より小さいものとする.

注意. 部分空間 $\omega^{(i)}$ の正規直交基底 $\{\underline{u}^{(i)}_\ell, \ell=1, \dots, k_i-1\}$ が与えられるれば, 問題 4 は 問題 3 と k_i が i に依存する点のみが異なる. 実は, 定理 3.5 は 全ての $k \in k_i$ として i によって与えられるベクトルの個数を変えても成立する. 部分空間の指定が $\underline{u}^{(i)}$ の制約条件

$$C^{(i)'} \underline{u}^{(i)} = 0 \quad (\forall i \in \overline{m}) \quad (3.27)$$

という形で与えられることも多い. ここで, $C^{(i)}$ は $r_i \times k_i$ の行列である. そのときは射影子 $V^{(i)}$ を

$$V^{(i)} = C^{(i)} (C^{(i)'} C^{(i)})^- C^{(i)'} \quad (3.28)$$

で与えればよい. ここで, X^- は X の一般化行列を意味する.

定理 3.6 の証明は定理 3.5 の場合と並行した議論で行えるので省略する.

3.2 特異値分解の性質

算法1は、 m が2の場合には行列の特異値分解を漸化的に表現したものにほかならず、一般には多重配列の分解を与えるものである。多重配列に対する因子効果の概念を用いてその事情を明らかにする。

まず、算法1で得られる多重配列 $\overset{k}{A}$ に対し、ベクトル $\underline{\overset{k}{u}}^{(i)}$ ($\forall i \in \bar{m}$)に相対的な α -因子効果を $\underline{\overset{k}{B}}(\alpha)$ と書き、(2.20)にならって、それを

$$\underline{\overset{k}{B}}(\alpha) = \left(\otimes_{i=1}^m \underline{\overset{k}{U}}_{\chi_{\alpha}(i)}^{(i)} \right) \underline{\overset{k}{A}} \quad (\forall \alpha \subset \bar{m}) \quad (3.29)$$

で定義する。ここで、

$$\left. \begin{aligned} \underline{\overset{k}{U}}_0^{(i)} &= \underline{\overset{k}{u}}^{(i)} \underline{\overset{k}{u}}^{(i)'} \\ \underline{\overset{k}{U}}_i^{(i)} &= I_{r_i} - \underline{\overset{k}{U}}_0^{(i)} \end{aligned} \right\} \quad (3.30)$$

である。2.1節の最後に触れのように、 $\bar{m}$ の部分集合 α の要素数によって、 $\underline{\overset{k}{B}}(\alpha)$ についても1因子効果、2因子効果などとまとめることがある。また、

$$\overset{k+1}{A} = \overset{k}{B}(\bar{m})$$

であることに注意しておく。

算法1で得られる直交系 $\underline{\overset{k}{u}}^{(i)}$, $k=1, 2, \dots, r$ とそれによって定義される $\overset{k}{A}$ の因子効果の特徴は次の定理が成り立つのである。

定理 3.7. $\overset{k}{A}$ の1因子効果が消える。すなわち、 α が $\bar{m}$ の1つの要素から成る部分集合 $\{j\}$ であるとき、

$$\overset{k}{B}(\{j\}) = 0 \quad (\forall j \in \overline{m}) \quad (3.31)$$

となる。

証明, $\alpha = \{j\}$ のとき, (3.29) で定義した $\overset{k}{B}(\{j\})$ は

$$\begin{aligned} \overset{k}{B}(\{j\}) &= (\overset{k}{U}_0^{(1)} \otimes \cdots \otimes I_{r_j} \otimes \cdots \otimes \overset{k}{U}_0^{(m)}) \overset{k}{A} \\ &\quad - (\overset{k}{U}_0^{(1)} \otimes \cdots \otimes \overset{k}{U}_0^{(j)} \otimes \cdots \otimes \overset{k}{U}_0^{(m)}) \overset{k}{A} \end{aligned}$$

と書ける。これを要素ごとに書いてまとめると

$$\begin{aligned} \overset{k}{B}_{\lambda_1 \lambda_2 \cdots \lambda_m}(\{j\}) \\ = \overset{k}{u}_{\lambda_1}^{(1)} \cdots \overset{k}{u}_{\lambda_m}^{(m)} \left(\sum \overset{k}{A}_{\lambda_1 \lambda_2 \cdots \lambda_m} \overset{k}{u}_{\lambda_1}^{(1)} \cdots \overset{k}{u}_{\lambda_m}^{(m)} - \rho_k \overset{k}{u}_{\lambda_j}^{(j)} \right) \end{aligned}$$

となり, この式の右辺の括弧内は決定方程式 (3.10) からゼロとなり, よって, 定理を得る。 |

次に, 与えられた多重配列 A の分解表現を $\overset{k}{A}$ の因子効果を用いて示す。前述の定理 3.6 によって $\overset{k}{A}$ の 1 因子効果が消えるので, 部分集合 α をこれに含まれる要素数で分類し, 要素数 s 個の α のクラスを $\mathcal{S}_s$ としておく。

定理 3.8 与えられた m 重配列 A を算法 1 によって, 次のように分解することができる。これを特異値分解と呼ぶ。

$$\left. \begin{aligned} \overset{1}{A} &= A, \\ \overset{k}{A} &= \overset{k}{B}(\phi) + \sum_{s=2}^{m-1} \sum_{\alpha \in \mathcal{S}_s} \overset{k}{B}(\alpha) + \overset{k+1}{A}, \quad k=1, 2, \dots, r. \end{aligned} \right\} \quad (3.32)$$

証明. 算法1の定義から A は $B^k(\bar{m})$ であり, 定理3.6で1因子効果が消える. そんで, A に定理2.3を適用すれば (3.32) の表現を得る. |

注意. A の主効果に相当する $B^k(\phi)$ の要素は

$$B_{\lambda_1 \lambda_2 \dots \lambda_m}^k(\phi) = \rho_k u_{\lambda_1}^{k(1)} u_{\lambda_2}^{k(2)} \dots u_{\lambda_m}^{k(m)} \quad (3.33)$$

と書ける. $m=2$ のとき, (3.32) は

$$\begin{aligned} A &= \sum_{k=1}^r B^k(\phi) \\ &= \sum_{k=1}^r \rho_k u_{\lambda_1}^{k(1)} u_{\lambda_2}^{k(2)} \end{aligned} \quad (3.34)$$

となり, 行列 A の特異値分解に行かからない (例えば, Forsythe & Moler (1967) 参照). これを3重配列の場合に拡張すると, 一見

$$A = \sum_{k=1}^r \rho_k u_{\lambda_1}^{k(1)} u_{\lambda_2}^{k(2)} u_{\lambda_3}^{k(3)}$$

となりそうであるが, 一般には不可能である. 3重配列の場合は2因子効果が消えない. 補題2.3のあとの注意として述べたように, 一般に $\Omega^{(i)}$ の基底 $\{u^{k_i(i)}, \forall k_i \in \bar{\gamma}_i\}$ について, $S = \{\otimes_{i=1}^m u^{k_i(i)}, \forall k_i \in \bar{\gamma}_i, \forall i \in \bar{m}\}$ の n 個のベクトルがテンソル Ω の基底となり, m 重配列 A はその一次結合として表現できる. そのことを要素ごとに書けば

$$A_{\lambda_1 \lambda_2 \dots \lambda_m} = \sum_{k_1} \dots \sum_{k_m} \rho_{k_1 \dots k_m} u_{\lambda_1}^{k_1(1)} \dots u_{\lambda_m}^{k_m(m)} \quad (3.35)$$

となる. A の α -因子効果は上式右辺の n 個を適当に分類してまとめ

ものになっている。算法1で求めたベクトル $\underline{\underline{U}}^{(i)}$, $k=1, \dots, r \in \Omega^{(i)}$ の基底の一部とすると, $k_1 = \dots = k_m = k$ の項が $\underline{\underline{B}}^k(\phi)$ に一対一対応し, $\underline{\underline{B}}^1(\{1\})$ は $k_2 = \dots = k_m = 1$ で k_1 の変わる r_1 項から $\underline{\underline{B}}(\phi)$ と除いた $r_1 - 1$ 項が, $\underline{\underline{B}}^1(\bar{m})$ には $k_1, k_2, \dots, k_m$ がいずれも2以上の項 $(r_1 - 1)(r_2 - 1) \dots (r_m - 1)$ 項が対応する。このような対応は (3.29) の $\underline{\underline{B}}^k(\alpha)$ の定義から行うことができる。さらには, 後述の定理3.10を参照されたい。

次に, (3.22)の分解表現に対応する平方和分解の定理を示す。

定理3.9. 与えられた m 重配列 A の平方和を, 算法1の結果を用いて, 次のように段階的に分解することができる:

$$\underline{\underline{A}}^{k, k} = \rho_k^2 + \sum_{s=2}^{m-1} \sum_{\alpha \in \mathcal{B}_s} \underline{\underline{B}}^k(\alpha)' \underline{\underline{B}}^k(\alpha) + \underline{\underline{A}}^{k+1, k+1} \quad (3.36)$$

$k = 1, 2, \dots, r$

ここで, $\underline{\underline{A}}^1 = \underline{\underline{A}}$ である。

証明. 定理2.3から異なる α と α' に対し

$$\underline{\underline{B}}^k(\alpha)' \underline{\underline{B}}^k(\alpha') = 0 \quad (3.37)$$

が成り立つ ($k \neq 1$ のときも定理2.3のときと同じように証明される)。

よって, (3.32)の右辺の内積を整理し, $\underline{\underline{B}}^k(\phi)' \underline{\underline{B}}^k(\phi)$ については (3.33) を $\underline{\underline{U}}^{(i)}$ の正規化条件で変形して, (3.36)を得る。 |

この定理から 6.2.1 に次の系を得る.

系 3.1 ρ_k^2 の上限は次の不等式で与えられる:

$$\rho_k^2 \leq \tilde{A}^k \tilde{A}^k, \quad k=1, 2, \dots, r. \quad (3.38)$$

等号が成立するのは, $\tilde{A}^k$ があるベクトル $\underline{u} = \bigotimes_{i=1}^m \underline{u}^{(i)}$ に比例するとき, そのときに限る.

注意. この系は定理 3.1 に対応する. 証明は (3.36) の各項が非負であることから明らかであるので省略する.

本節の以上の議論は算法 1 で得られる正規直交系 $\{\underline{u}^{(i)}\}^k$ にもとづくものである. それに対して, それぞれのベクトル空間 $\Omega^{(i)}$ で配列 A と無関係に正規直交系を与えた場合は, (3.32) と (3.36) で 1 因子効果の項を加えることによって, 配列 A 及びその平方和の段階的な分解が与えられる. こゝでは, 任意に与えられた正規直交系 $\{\underline{u}^{(i)}\}^k$ に対して (3.29) の $\tilde{B}^k(\alpha)$ の定義に対応する Ω の部分空間 $\tilde{\omega}^k(\alpha)$ を定義してその部分空間の次元を求めておこう. 前節の問題 3 のときと同様に考え, $A \in \tilde{A}^k$ に射影する射影子 E (3.18) によって

$$\bigotimes_{i=1}^m \frac{\tilde{V}^{(i)}(i)}{V_1^{(i)}} \quad (3.39)$$

とし, $\tilde{A}^k$ を $\tilde{B}^k(\alpha)$ に射影する射影子は (3.29) によって

$$\bigotimes_{i=1}^m \frac{\tilde{U}^k(i)}{X_{\alpha}(i)} \quad (3.40)$$

とする. ここで, $\overset{k-1}{V}_1^{(i)}$ は (3.19) で, $\overset{k}{U}_0^{(i)}$ と $\overset{k}{U}_1^{(i)}$ は (3.30) で与えられて
 いるものである. よって, $A \in \overset{k}{B}(\alpha)$ に射影する射影子は (3.39) と (3.40)
 の積, すなわち,

$$\left(\bigotimes_{i=1}^m \overset{k}{U}_{\chi_\alpha(i)}^{(i)} \right) \left(\bigotimes_{i=1}^m \overset{k-1}{V}_1^{(i)} \right) = \bigotimes_{i=1}^m \left(\overset{k}{U}_{\chi_\alpha(i)}^{(i)} \overset{k-1}{V}_1^{(i)} \right) \quad (3.41)$$

となる. そで, $\overset{k}{W}(\alpha) \subset \Omega$ から (3.41) の射影子によって射影される
 Ω の部分空間と定義する. $\overset{k}{W}(\alpha)$ の次元について次の定理が導かれる.

定理 3.10. それぞれのベクトル空間 $\Omega^{(i)}$ で任意に与えられる正規直
 交系 $\{\underline{u}^{(i)}, \dots, \underline{u}^{(i)}\}$ ($\forall i \in \bar{m}$) について 射影子 (3.41) で規定される
 Ω の部分空間 $\overset{k}{W}(\alpha)$ の次元は, 任意の $\alpha \in \bar{m}$ に関して

$$\dim(\overset{k}{W}(\alpha)) = \prod_{i \in \alpha} (r_i - k) \quad (3.42)$$

で与えられる.

証明. $\overset{k}{W}(\alpha)$ の定義から, その次元は射影子 (3.41) の階数で与え
 られる. 補題 2.5 から

$$\text{rank} \left[\bigotimes_{i=1}^m \left(\overset{k}{U}_{\chi_\alpha(i)}^{(i)} \overset{k-1}{V}_1^{(i)} \right) \right] = \prod_{i=1}^m \text{rank} \left(\overset{k}{U}_{\chi_\alpha(i)}^{(i)} \overset{k-1}{V}_1^{(i)} \right).$$

一方, $\overset{k}{U}_{\chi_\alpha(i)}^{(i)} \overset{k-1}{V}_1^{(i)}$ はそれぞれの定義 (3.30), (3.19) より

$$\overset{k}{U}_{\chi_\alpha(i)}^{(i)} \overset{k-1}{V}_1^{(i)} = \begin{cases} \overset{k}{V}_1^{(i)} = I_{r_i} - \sum_{i=1}^k \underline{u}^{(i)} \underline{u}^{(i)'} & i \in \alpha \text{ のとき,} \\ \overset{k}{U}_0^{(i)} = \underline{u}^{(i)} \underline{u}^{(i)'} & i \notin \alpha \text{ のとき} \end{cases} \quad (3.43)$$

となり, その階数は

$$\text{rank} \left(\bigcup_{\alpha(i)}^{(k)} \bigvee_1^{(k-1)} \right) = \begin{cases} r_i - k, & i \in \alpha \text{ のとき,} \\ 1, & i \notin \alpha \text{ のとき} \end{cases} \quad (3.44)$$

で与えられる。よって, (3.42) を得る。

注意. 要因計画のモデルでは $\omega(\alpha)$ の次元を平方和 $\tilde{B}(\alpha)$ $\tilde{B}(\alpha)$ の自由度 (degrees of freedom) と呼ぶ。算法 1 を用いる場合は総平均に相当する ρ_k を最大にし, 主効果に相当する $\tilde{B}(\{j\})$ ($\forall j \in \bar{m}$) をゼロにするようにベクトル $\tilde{u}^{(k(i))}$ を求めている。その場合, ρ_k^2 に対する自由度は $\tilde{u}^{(k(i))}$ の数学的に独立なパラメータの個数を加えて

$$\sum_{i=1}^m r_i - mk + 1 \quad (3.45)$$

とすべきであろう。これは 総平均 と 1 因子効果の平方和に対する自由度とまとめたものともいえる。

最後に, 多重配列の一般的分解表現である (3.35) の右辺の n 項のうち, 特異値分解 (3.32) によって消える $\tilde{B}(\{j\})$, $\forall j \in m$ に相当する項数を示しておこう。それは $\alpha = \{j\}$, $\forall j \in m$ に対する $\omega(\alpha)$ の次元数に等しい。そのことは射影子 (3.41) と (3.35) の右辺にかけることによりわかる。 $k=1, 2, \dots, r$ についての $\dim(\omega(\{j\}))$ ($\forall j \in \bar{m}$) の和は簡単な計算で求まる。ここでは $r_1 = \dots = r_m = r$ としたときの結果のみを示しておく。求める項数は

$$m \frac{r(r-1)}{2}$$

となる。これに対する全項数は r^m であり, $m=2$ のときは r 項のみが残り, (3.34) の行列の特異値分解の項数に一致する。 m と r のいくつかの値について, 上記の数も表 3.1 に示しておく。

表 3.1

m	r	r^m	$mr(r-1)/2$	残り
2	r	r^2	r^2-r	r
3	2	8	3	5
3	5	125	30	95
3	10	1000	135	865
4	2	16	4	12
4	5	625	40	585

4. 多重分割表への応用(I)

前章までに論じた多重配列の分解と多重分割表データの解析のためのモデルとの関係を図べる。まず、多重分割表における従来の主要なモデルである加法モデルと対数線形モデルをとりあげ、次に平方根モデルを論じる。

4.1. 加法モデルと対数線形モデル

多重分割表に関する研究では、その項目間の高次の交互作用をどういうモデルで表現するか、1.3節で述べたように、長らく問題にされている。その問題から代表的なモデルとして、対数線形モデル(log-linear model)や加法モデル(additive model)が生み出されている。ここでは、 $r \times s \times t$ 分割表をとりあげて、それらのモデルと前章までに論じた多重配列構造を持つ母数空間との関係を調べる。

$r \times s \times t$ 分割表のセル度数、対応するセル確率、及びセル度数の期待値をそれぞれ

$$x_{ijk}, p_{ijk}, m_{ijk},$$

$$(\forall i \in \mathcal{R}, \forall j \in \mathcal{S}, \forall k \in \mathcal{T})$$

と書く。これらの量の周辺和を示すのに

$$x_{ij+} = \sum_k x_{ijk}, \quad x_{i++} = \sum_j x_{ij+}, \quad x_{+++} = \sum_i x_{i++}$$

のように和をとる添字の位置に + をつける. セル度数の総和 α_{+++} とくに N とする. 小文字の n はセルの総数 $r \cdot s \cdot t$ に用いる.

本節では, セル確率 p_{ijk} ($\forall i \in \bar{r}, \forall j \in \bar{s}, \forall k \in \bar{t}$) を基本的な母数として以下のモデルの記述に用いる. セル確率は非負で,

$$p_{+++} = 1 \quad (4.1)$$

すなわち, 総和が 1 である.

さて, 本節で論じる加法モデルは多重配列母数 A を

$$A = (p_{ijk}) \quad (\forall i \in \bar{r}, \forall j \in \bar{s}, \forall k \in \bar{t}) \quad (4.2)$$

としたときの, 対数線形モデルは A を

$$A = (\log p_{ijk}) \quad (\forall i \in \bar{r}, \forall j \in \bar{s}, \forall k \in \bar{t}) \quad (4.3)$$

とした場合の線形モデルにはかならない. 次節ではセル確率の平方根を母数とする平方根モデル, すなわち

$$A = (\sqrt{p_{ijk}}) \quad (\forall i \in \bar{r}, \forall j \in \bar{s}, \forall k \in \bar{t}) \quad (4.4)$$

の場合を扱う.

a) 加法モデル 3種の添字に対応するベクトル空間 $\Omega^{(1)}, \Omega^{(2)}$, および $\Omega^{(3)}$ における内積を次のような対角行列を重みとして (2.40) で定義する.

$$\left. \begin{aligned} \Sigma^{(1)} &= \text{diag}(p_{i++}) \\ \Sigma^{(2)} &= \text{diag}(p_{+j+}) \end{aligned} \right\} \quad (4.5)$$

$$\Sigma^{(3)} = \text{diag}(p_{++k})$$

以下, 2.3節の後半の議論に従って, まず, (2.59)の変換を用い,

$$\hat{A}_{ijk} = \frac{p_{ijk}}{\sqrt{p_{i++}} \sqrt{p_{+j+}} \sqrt{p_{++k}}} \quad (\forall i \in \bar{r}, \forall j \in \bar{s}, \forall k \in \bar{t}) \quad (4.6)$$

と置く. この $\hat{A}$ に対する因子効果を定義するベクトル $\underline{w}^{(1)}, \underline{w}^{(2)}, \underline{w}^{(3)}$ を

$$\left. \begin{aligned} w_i^{(1)} &= \sqrt{p_{i++}} & (\forall i \in \bar{r}) \\ w_j^{(2)} &= \sqrt{p_{+j+}} & (\forall j \in \bar{s}) \\ w_k^{(3)} &= \sqrt{p_{++k}} & (\forall k \in \bar{t}) \end{aligned} \right\} \quad (4.7)$$

とする. この $\underline{w}^{(1)}, \underline{w}^{(2)}, \underline{w}^{(3)}$ の変換 (2.50) によるベクトル $\underline{u}^{(1)}, \underline{u}^{(2)}, \underline{u}^{(3)}$ が

$$u_i^{(1)} = 1 \quad (\forall i \in \bar{r}), \quad u_j^{(2)} = 1 \quad (\forall j \in \bar{s}), \quad u_k^{(3)} = 1 \quad (\forall k \in \bar{t}) \quad (4.8)$$

と, すべての要素が1になることが特徴である.

このように定義される $\{\hat{A}, \underline{w}^{(1)}, \underline{w}^{(2)}, \underline{w}^{(3)}\}$ あるいは $\{A = (p_{ijk}), \underline{u}^{(1)}, \underline{u}^{(2)}, \underline{u}^{(3)}, \Sigma^{(1)}, \Sigma^{(2)}, \Sigma^{(3)}\}$ によって構成される一連のモデル (2.2節, (2.30)参照) を $r \times s \times t$ 分割表の加法モデルと呼ぶ. そこで, $\hat{A}$ の α -因子効果を $\hat{B}(\alpha)$, $\hat{B}(\alpha)$ を (4.6) によって p_{ijk} の分解に変換したものを $B(\alpha)$ と書いて, それらの表現を示しておく.

ここで、記述の簡単のために、

$$g_{ijk} = p_{i+} p_{+j+} p_{++k} \quad (4.9)$$

とおき、 $\alpha = \phi, \{1\}, \{12\}, \{123\}$ などに対して略式に $\hat{B}(\alpha)$, $\hat{B}(1)$, $\hat{B}(12)$, $\hat{B}(123)$ と書く. α -因子効果の定義(2.20)を適用して次式を得る:

$$\hat{B}(\phi) = (\sqrt{g_{ijk}}) \quad (4.10)$$

$$\hat{B}(1) = \hat{B}(2) = \hat{B}(3) = 0 \quad (4.11)$$

$$\left. \begin{aligned} \hat{B}(12) &= (p_{ij+} p_{++k} / \sqrt{g_{ijk}}) - \hat{B}(\phi) \\ \hat{B}(13) &= (p_{i+k} p_{+j+} / \sqrt{g_{ijk}}) - \hat{B}(\phi) \\ \hat{B}(23) &= (p_{+jk} p_{i++} / \sqrt{g_{ijk}}) - \hat{B}(\phi) \end{aligned} \right\} \quad (4.12)$$

$$\hat{B}(123) = \hat{A} - \hat{B}(12) - \hat{B}(13) - \hat{B}(23) - \hat{B}(\phi) \quad (4.13)$$

ここで、 $(\sqrt{g_{ijk}})$ などには 3 重配列であることを示す. (4.11) で 1-因子効果が消えることは、 $\hat{A}$ についての決定方程式(3.5)が(4.7)のベクトル $\underline{w}^{(1)}$, $\underline{w}^{(2)}$, $\underline{w}^{(3)}$ で満たされることを示す. そのとき

$$\rho = \sum_{i,j,k} \hat{A}_{ijk} w_i^{(1)} w_j^{(2)} w_k^{(3)} = 1 \quad (4.14)$$

となる. これは(4.10)の導出に使われている. 次に、 $\hat{B}(\alpha)$ を $B(\alpha)$ に(2.62)を用いて変換する. 実は $\hat{B}(\alpha)$ の各要素に $\sqrt{g_{ijk}}$ をかければよい. (2.62)によれば、 $B(\alpha)$ は $B(\alpha; \Sigma^{-1})$ と書くべきであ

るが、ここでは簡単に $B(\alpha)$ としておく。次式を得る:

$$B(\phi) = (q_{ijk}) = (p_{i++} p_{+j+} p_{++k}), \quad (4.15)$$

$$\left. \begin{aligned} B(12) &= ((p_{ij+} - p_{i++} p_{+j+}) p_{++k}), \\ B(13) &= ((p_{i+k} - p_{i++} p_{++k}) p_{+j+}), \\ B(23) &= ((p_{+jk} - p_{+j+} p_{++k}) p_{i++}), \end{aligned} \right\} \quad (4.16)$$

$$\begin{aligned} B(123) &= (p_{ijk} - p_{ij+} p_{++k} - p_{i+k} p_{+j+} \\ &\quad - p_{+jk} p_{i++} + 2 q_{ijk}). \end{aligned} \quad (4.17)$$

以上の式から

$$\hat{A} = \sum_{\alpha} \hat{B}(\alpha), \quad (4.18)$$

あるいは,

$$A = (p_{ijk}) = \sum_{\alpha} B(\alpha) \quad (4.19)$$

を確かめることは容易であろう。

そもそも加法モデルは 3 重分割表での 3 因子効果 (2 次交互作用) が存在しないということを表現するために考えられたもので [Lancaster (1951)], (4.17) の $B(123)$ の効果をゼロと置いて得られるモデルである。このモデルを H_a と書くと,

$$H_a: p_{ijk} = p_{ij+} p_{++k} + p_{i+k} p_{+j+} + p_{+jk} p_{i++} - 2 q_{ijk}. \quad (4.20)$$

を得る。

平方和 $\hat{A}'\hat{A}$ あるいは $A'\Sigma^{-1}A$ の分解は (2.63) から

$$\hat{A}'\hat{A} = A'\Sigma^{-1}A = \sum_{\alpha} B(\alpha)'\Sigma^{-1}B(\alpha) \quad (4.21)$$

で与えられる。ここで、 Σ は $n \times n$ の行列で

$$\Sigma = \Sigma^{(1)} \otimes \Sigma^{(2)} \otimes \Sigma^{(3)} = \text{diag}(g_{ijk}) \quad (4.22)$$

である。(4.21)は次のように書ける:

$$\begin{aligned} \sum_{i,j,k} \frac{p_{ijk}^2}{p_{i++}p_{+j+}p_{++k}} &= 1 + \sum_{i,j} \frac{(p_{ij+} - p_{i++}p_{+j+})^2}{p_{i++}p_{+j+}} \\ &\quad + \sum_{i,k} \frac{(p_{i+k} - p_{i++}p_{++k})^2}{p_{i++}p_{++k}} \\ &\quad + \sum_{j,k} \frac{(p_{+jk} - p_{+j+}p_{++k})^2}{p_{+j+}p_{++k}} \\ &\quad + \sum_{i,j,k} \frac{(p_{ijk} - ap_{ijk})^2}{p_{i++}p_{+j+}p_{++k}} \end{aligned} \quad (4.23)$$

ここで、 ap_{ijk} は

$$ap_{ijk} = p_{ij+}p_{++k} + p_{i+k}p_{+j+} + p_{+jk}p_{i++} - 2g_{ijk} \quad (4.24)$$

としたものである。(4.23)の分解は、よく知られている Lancaster (1951) による χ^2 の分割にほかならない。実際、(4.23)の左辺から右辺の1を引いて $N (= \chi_{+++})$ 倍したものは、その p_{ijk} を推定値

$$\hat{p}_{ijk} = \chi_{ijk} / N$$

で置きかえれば、いわゆる相互独立性の検定における χ^2 にほかならな

い. その χ^2 は

$$\begin{aligned}\chi^2 &= N \left(\sum_{i,j,k} \frac{\hat{p}_{ijk}^2}{\hat{p}_{i++} \hat{p}_{+j+} \hat{p}_{++k}} - 1 \right) \\ &= \sum_{i,j,k} \frac{(x_{ijk} - \hat{m}_{ijk})^2}{\hat{m}_{ijk}}\end{aligned}\quad (4.25)$$

と書かれる. ここで, $\hat{m}_{ijk} = N \hat{p}_{i++} \hat{p}_{+j+} \hat{p}_{++k}$ である.

モデル H_a のほかに各種の階層的モデルを考えることができる. 2.2節の記法に従って, 階層的クラス $\{1/2/3\}, \{12\}, \{12/23\}$ とすると, それぞれに応じたモデル

$$H_1: p_{ijk} = p_{i++} p_{+j+} p_{++k} \quad (4.26)$$

$$H_2: p_{ijk} = p_{ij+} p_{++k} \quad (4.27)$$

$$H_{3a}: p_{ijk} = p_{ij+} p_{++k} + p_{+jk} p_{i++} - p_{i++} p_{+j+} p_{++k} \quad (4.28)$$

などが得られる. H_1 を分割表の相互独立モデル, H_2 を部分独立モデルという. H_2 は H_2 と同様に添字の入れかえによって 三つおりのモデルがある. H_3 は, いって呼ぶなら, 加法モデルでの条件独立モデルともいえようが, 普通, 分割表での条件独立モデルは

$$H_{3c}: p_{ijk} = \frac{p_{ij+} p_{i+k}}{p_{i++}} \quad (4.29)$$

の形式のものをいう. 非階層的モデルとして, $B(13)=0$ とするもの

を考えると, (4.16) から

$$p_{i+k} = p_{i+t} p_{t+k} \quad (4.30)$$

を得る. このモデルを周辺独立モデルということがある. なお, 以上の議論はより多重の分割表に拡張されることに注意しておく.

b) 対数線形モデル 対数線形モデルは, 多重配列 A を

$$A = (\log p_{ijk}), \quad (4.31)$$

ベクトル $u^{(1)}, u^{(2)}, u^{(3)}$ を (4.8) のようにすべての要素に 1 を与えるものとして, 2.1, 2.2 節の記述に従って定義されるものである. その飽和型は

$$\log p_{ijk} = \mu + \mu_i^{(1)} + \mu_j^{(2)} + \mu_k^{(3)} + \mu_{ij}^{(12)} + \mu_{ik}^{(13)} + \mu_{jk}^{(23)} + \mu_{ijk}^{(123)} \quad (4.32)$$

と表現される. このモデルは実験計画法における要因実験での母数モデルと同形式であるが, セル確率の総和が 1 という制約 (4.1) が附加されている. 従って, 総平均に相当する μ はその制約を満たすように決められる.

加法モデルでの H_a に対応する 3 因子効果非存在のモデルは (4.32) から $\mu_{ijk}^{(123)}$ を除いたものであり, 階層的モデルであるから, その別表現として

$$\log p_{ijk} = \alpha_{ij} + \beta_{ik} + \gamma_{jk} \quad (4.33)$$

と書いてもよい [(2.34) 参照]. また, 乗法的表現として

$$H_m: p_{ijk} = \theta_{ij} \phi_{ik} \psi_{jk} \quad (4.34)$$

としても同値である [Darroch (1962, 1974)]。これをモデル H_m とする。(4.32)の右辺の因子効果を適当に取捨選択して各種のモデルが作られ、これは 2.2 節で述べたとおりである (例 2.5 参照)。階層的クラス $\{1/2/3\}$, $\{12\}$, $\{12/23\}$ に応じて前述の H_1 , H_2 , H_{3c} のモデルが得られる。加法モデルでは $\{12/13\}$ に対して H_{3a} (4.28) が得られることが異なっている。

加法モデルの H_a と対数線形モデルの H_m との比較は Darroch (1974) によってなされている。モデルの一般的性質としては両者それぞれ長短があるが、近似的には H_a が成り立てば H_m が満たされることが示されている。

以上のモデルの実際のデータにもとづく検定や母数の推定問題の精密な議論は、分割表データの標本抽出過程に即して確率モデルを構成することによって行われる。各セルの度数がそれぞれ独立に母数 m_{ijk} のポアソン分布に従うというポアソンモデル、大きさ N の標本の無作為抽出と考える多項分布モデル (あるいは部分的な表組との多項分布モデル)、及びある条件付のポアソンモデルが取り扱いの比較的簡単な一般的確率モデルとして考えられる。これらの確率モデルのもとでは、対数

線形モデルが十分統計量, 最尤推定法, 尤度比検定などとの関係で秀れに性質を持っている [Birch (1963), Habermann (1974)].
 ここでその詳細は論じないが, 対数線形モデルの階層的モデルではその生成クラスに対応するセル度数の周辺和, 例えば, 生成クラスが $\{12, 13, 23\}$ なら, $x_{ij+}, x_{i+k}, x_{+jk}$ の組が最小完備十分統計量となり, 期待値 m_{ijk} の推定値 $\hat{m}_{ijk}$ の周辺和をそれらに一致させるという条件でそのモデルの最尤推定が行われる. 尤度比検定からは, 検定統計量

$$G^2 = 2 \sum x_{ijk} \log \frac{x_{ijk}}{\hat{m}_{ijk}} \quad (4.35)$$

が得られ, 前述の確率モデルのもとで漸近的に χ^2 分布に従う.

ここで, $\hat{m}_{ijk}$ は各モデルのもとでの推定値であり, χ^2 の自由度はそのモデルの部分空間の次元に等しい. 分解型モデル (定義 2.7) に対しては対応する周辺和から陽にセル確率の推定値が得られる.

例えば, H_{3C} に対して

$$\begin{aligned} \hat{p}_{ijk} &= \frac{(x_{ij+}/N) \cdot (x_{i+k}/N)}{(x_{i++}/N)} \\ &= \frac{x_{ij+} x_{i+k}}{N x_{i++}} \end{aligned} \quad (4.36)$$

が最尤推定値となる。分解型でないモデルに対してはそれぞれ何らかの反復計算が必要である。

現実には、標本抽出過程は複雑多様であり、才1章で述べたような観察データを解析しにくいことが多い。標準的な手法の各種の標本抽出過程での頑健性の研究は今後の重要な課題であろう。

なお、加法モデルの利用や別の観点からの問題点については才5章で述べる。

4.2 平方根モデル

多重分割表に対し、そのセル確率あるいはセル度数の期待値の平方根を母数とする母数空間における才2章の記述に従う線形モデルを平方根モデル、あるいは平方根線形モデルと呼ぶ。寸法 $\gamma_1 \times \gamma_2 \times \cdots \times \gamma_m$ の m 重分割表のセル度数、対応するセル確率、及びセル度数の期待値をそれぞれ、 $x_{\lambda_1 \lambda_2 \cdots \lambda_m}$, $p_{\lambda_1 \lambda_2 \cdots \lambda_m}$, $m_{\lambda_1 \lambda_2 \cdots \lambda_m}$ ($\forall i \in \bar{\gamma}_i, \forall i \in \bar{m}$) とする。ここではセル確率を用い、多重配列の母数 A を

$$A = \left(\sqrt{p_{\lambda_1 \lambda_2 \cdots \lambda_m}} \right) \quad (4.37)$$

とおく。各指標に対するベクトル空間 $\Omega^{(i)}$ におけるベクトル $u^{(i)}$ で

$$u_{\lambda_i}^{(i)} = 1 \quad (\forall i \in \bar{\gamma}_i, \forall i \in \bar{m}) \quad (4.38)$$

であるものを採用し、それによって Ω での α -効果空間 $\omega(\alpha)$, A の α -因子効果 $B(\alpha)$, 及び各種のモデル

$$A \in \Omega[\mathcal{H}] \quad (4.39)$$

を規定する。ここで $\Omega[\mathcal{H}]$ は 2.2 節の記述に従い、 $\bar{m}$ の部分集合のクラス $\mathcal{H}$ について

$$\Omega[\mathcal{H}] = \bigoplus_{\forall \alpha \in \mathcal{H}} \omega(\alpha) \quad (4.40)$$

で与えられる [(2.29), (2.30) 参照]. (4.38) のようにベクトル $u^{(i)}$

をあらかじめ与えずに、(4.37) の多重配列に特異値分解を適用するこ

とについては次章で検討する。

ある種の平方根変換は分散安定化変換 (variance stabilizing transformation) としてしばしば論じられている [例えば, Rao (1973)] が, ここで平方根モデルを導入した理由の一つは分散安定化変換にある。セル度数もポアソン分布に従う確率変数とみなすポアソンモデルではセル度数の平方根が漸近的に分散が $1/4$ の正規分布に従うとみなされるので, 分散分析での母数モデルからの類推で多重分割表を解析することが可能である。多重分割表のデータもセル確率もパラメータとする多項分布によるものとする多項モデルでは次のように考えられる。二つの多項分布間の相違を測る量として Hellinger distance [例えば, Rao (1973, p.352)] がある。多重分割表の場合, モデルとしての分布 $\{p_{\lambda_1\lambda_2\cdots\lambda_m}\}$ と経験分布 $\{x_{\lambda_1\lambda_2\cdots\lambda_m}/N\}$ と考えると, その量は

$$\cos^{-1} \sum \sqrt{\frac{x_{\lambda_1\lambda_2\cdots\lambda_m}}{N} \cdot p_{\lambda_1\lambda_2\cdots\lambda_m}} \quad (4.41)$$

で与えられる。これは, より一般に二つの分布間で定義される

Bhattacharya divergence [Bhattacharya (1946)] の一

例になっており, (4.41) と Bhattacharya の距離値ということも

ある。この距離値の意味づけは付録でも与えられるが, 二項分布の

逆正弦変換の一般化になっている。あるモデルのもとで, (4.41) を最小にする母数の推定は, Pearson の χ^2 , Neyman の修正 χ^2 [Neyman (1949)], Kullback-Leibler 判別情報量 [Kullback (1959)] などによる推定とともに, 適当な正規条件の下で一緻かつ有効推定量を与える[†]。いま, 多重配列 Y を

$$Y = (\sqrt{x_{\lambda_1 \lambda_2 \dots \lambda_m} / N}) \quad (4.42)$$

とおくと, (4.41) は (4.37) と (4.42) によって

$$\cos^{-1}(Y' \hat{A}) \quad (4.43)$$

と書け, これは最小にするとは, 逆余弦関数の性質から, 明らかに $Y' \hat{A}$ を最大にすることになる。この事実によって, 平方根モデルでの分析が以下に述べるように簡単になる。

まず, Y と A についてはセル確率の総和が 1 であることから (4.1),

$$Y' Y = A' A = 1 \quad (4.44)$$

であり, A の推定値 $\hat{A}$ についても

$$\hat{A}' \hat{A} = 1 \quad (4.45)$$

[†] Rao (1961, 1962, 1963) の意味で一次の有効性を持つ。Rao の定義による二次の有効性についてはわずかな差があるが, 現実的应用で問題にされるほどではないと思われる。

と制約を置く。そのとき

$$(\underline{Y} - \hat{\underline{A}})'(\underline{Y} - \hat{\underline{A}}) = 2 - 2\underline{Y}'\hat{\underline{A}} \quad (4.46)$$

が成り立つので, $\underline{Y}'\hat{\underline{A}}$ を最大にするかわりに, 上式の右辺を最小, すなわち最小二乗法として考えられる。

さて, モデル (4.39) と制約 (4.45) のもとで (4.41) を最小にする解は, 簡単な計算から

$$\hat{\underline{A}} = \frac{1}{\lambda} \sum_{\forall \alpha \in H} \hat{\underline{B}}(\alpha) \quad (4.47)$$

で与えられる。ここで, $\hat{\underline{B}}(\alpha)$ は

$$\hat{\underline{B}}(\alpha) = \left(\bigotimes_{i=1}^m U_{\chi_\alpha(i)}^{(i)} \right) \underline{Y} \quad (4.48)$$

すなわち, (2.20) で定義される $\underline{Y}$ の α -因子効果であり, λ は

$$\lambda = \left(\sum_{\forall \alpha \in H} \hat{\underline{B}}(\alpha)' \hat{\underline{B}}(\alpha) \right)^{1/2} \quad (4.49)$$

で与えられる。このときの Hellinger distance (4.41) は $\cos^{-1} \lambda$

となる。一方, 多項分布モデルの前提のもとでは

$$\chi^2 = 4N(\cos^{-1} \lambda)^2 \quad (4.50)$$

が漸近的に χ^2 分布に従うことが知られている [例えば, Rao (1973)] ので, モデル (4.39) の検定が可能である。その χ^2 の自由度は定理 2.1 を用いて次式で与えられる:

$$\text{d.f.} = n - \sum_{\forall \alpha \in H} \left(\prod_{\forall i \in \alpha} (r_i - 1) \right). \quad (4.51)$$

ここで、平方根モデルの特徴を述べておこう。第一に、飽和モデル
 における各因子効果への分解が容易なことである。対数線形モデル
 ではゼロの対数が無限大になるという問題が生じ、多重分割表で
 はセル度数の一部がゼロ (zero-cell) になることも稀でない。
 第二に、対数線形モデルでは、モデルごとにセル確率を反復計算に
 よって推定しかつ検定統計量を計算しなからので、モデルの選択、
 あるいは有意な因子効果の抽出が面倒であるのに対し、平方根
 モデルでは、飽和モデルで因子効果ごとの平方和の表を一度作っ
 ておけば、一連の検定を行うことができる。第三は、これは欠点であるが
 対数線形モデルにおける分解型 (closed form) モデルのように、
 独立、条件独立といったモデルの確率論的解釈が困難な点で
 ある。この点に関しては平方根モデルを使ってある程度の見込みをつけ
 てから そんで選択されたモデルに近い分解型の乗法モデルを検
 定すればよからう。多重分割表データの解析の重要な目標の一つは
 与えられたデータから適切なモデルを構成することであるが、モデルの
 選択はきわめて事後的な分析であることに注意しなければならない。
 しなからって、一つのモデルや手法にこだわることなく、幅広い観点か
 らデータの分析にあたることが肝心であろう。

5. 多重分割表への応用(II)

才3章で述べた多重配列の特異値分解を多重分割表データの解析に応用する。まず、2重分割表に対する数量化手法から始め、3重分割表に対する数量化手法の型を整理して、それらと加法モデルとの関連を論じる。次に、一般の多重分割表に対する数量化手法として分解型モデルからのずれも表現することなども考え、最後に平方根分解の応用を論じる。

5.1 2重分割表における数量化

多重分割表に対する数量化手法は、1.3節で述べたように、実質的には Fisher が正準相関分析の考え方を2重分割表に適用したことと端を発する。ここでいう分割表の数量化手法は、optimal scaling method (最適尺度法)の一つと考えられることもある。この分野の文献は西田(1977)に詳しい。

2重分割表の数量化問題は次のように述べられる。

問題5. $r \times c$ 分割表のセル確率 p_{ij} ($\forall i \in \bar{r}, \forall j \in \bar{c}$) が与えられるとき、

条件

$$\sum_i p_{i+} x_i = 0; \quad \sum_j p_{+j} y_j = 0; \quad (5.1)$$

$$\sum_i p_{i+} x_i^2 = 1; \quad \sum_j p_{+j} y_j^2 = 1 \quad (5.2)$$

の下で、

$$\rho = \sum_{i,j} p_{ij} x_i y_j \quad (5.3)$$

を最大にする x_i ($\forall i \in \bar{r}$) と y_j ($\forall j \in \bar{c}$) を求めること。

この種の数量化の目的は、本来名義尺度の項目の各カテゴリに x_i ($\forall i \in \bar{r}$) と y_j ($\forall j \in \bar{c}$) という数値を付与して、カテゴリの分類あるいは比較を行うことにある。同様の目的のためには広津(1973, 1976)による多重比較の観測からの手法があるが、そこでは項目ごとにカテゴリ間の類似性の比較が行われるのに対し、問題5では二項目に対して同時に数量化する点が大きく異なっている。

問題5の条件(5.1)は、各項目を x_i ($\forall i \in \bar{r}$) と y_j ($\forall j \in \bar{c}$) という値をとる離散的確率変数 X と Y とみなすとき、 X と Y の期待値がゼロであるという制約を表わす。条件(5.2)は X^2 と Y^2 の期待値が1、条件(5.1)と合わせて X と Y の分散が1と解釈される。したがって、(5.3)の ρ は X と Y の相関係数も意味することになる。

この問題に対して次のような事実はずでによく知られている[例えば、Kendall & Stuart (1961) 参照]。まず、決定方程式は

$$\left. \begin{aligned} \sum p_{ij} y_j &= \rho p_{i+} x_i & (\forall i \in \bar{r}), \\ \sum p_{ij} x_i &= \rho p_{+j} y_j & (\forall j \in \bar{c}) \end{aligned} \right\} \quad (5.4)$$

で与えられる。(5.4)の最大特異値 ρ_1 は1で、対応するベクトルは単位

ベクトルに比例するものである。そのベクトルは条件(5.1)を満たさず、 ρ_2 に大きい特異値と対応するベクトルが問題5の解となる。さらに、 ρ_3 以後の特異値も求めていき、それらを ρ_k , 対応するベクトルを $x_i^{(k)}, y_j^{(k)}, k=2, 3, \dots$ とする。そのとき、いわゆる Fisher の恒等式

$$p_{ij} = p_{i+} p_{+j} \left(1 + \sum_{k=2}^{\min(r,c)} \rho_k x_i^{(k)} y_j^{(k)} \right) \quad (5.5)$$

$(\forall i \in \bar{r}, \forall j \in \bar{c})$

及び、平方和分解

$$\sum_{i,j} \frac{p_{ij}^2}{p_{i+} p_{+j}} = 1 + \sum_{k=2}^{\min(r,c)} \rho_k^2 \quad (5.6)$$

を得る。(5.6)の右辺の1を左辺に移項して両辺を N 倍すると、いわゆる

る2重分割表における独立性の検定統計量 χ^2 の分解が得られる:

$$\chi^2 = N \left(\sum_{i,j} \frac{p_{ij}^2}{p_{i+} p_{+j}} - 1 \right) = N \sum_{k=2}^{\min(r,c)} \rho_k^2 \quad (5.7)$$

ここで、 N は 分割表 データ (x_{ij}) が与えられたときのセル度数の総和である。

以上の事実を、行列

$$A = \left(\frac{p_{ij}}{\sqrt{p_{i+} p_{+j}}} \right) \quad (5.8)$$

に対する特異値分解によって導かれるものにほかならない。このことは前章の加法モデルの項で述べたことから容易に理解されよう。(5.8)の特異値分解で上記の $x_i^{(k)}, y_j^{(k)}$ に対して、 $\sqrt{p_{i+}} x_i^{(k)}, \sqrt{p_{+j}} y_j^{(k)}$ が計算されることになる。

5.2 3重分割表における数量化とモデル

本節では、従来提案されている3重分割表の数量化手法の理論的基礎付けを与え、それらの手法の拡張を検討する[†]。

a) 数量化の形式 3重分割表に対する記法は第4章と同じとする。ここでは、 $r \times s \times t$ 分割表における数量化として、次の三つの場合を扱う：

$$(A) \quad i \leftarrow u_i \quad (\forall i \in \bar{r}), \quad j \leftarrow v_j \quad (\forall j \in \bar{s}), \quad k \leftarrow w_k \quad (\forall k \in \bar{t});$$

$$(B) \quad (i, j) \leftarrow u_{ij} \quad (\forall i \in \bar{r}, \forall j \in \bar{s}), \quad k \leftarrow w_k \quad (\forall k \in \bar{t});$$

$$(C) \quad [(i, j) \leftarrow v_{ij} \quad (\forall j \in \bar{s}), (i, k) \leftarrow w_{ik} \quad (\forall k \in \bar{t})] \quad (\forall i \in \bar{r}).$$

これを上記の順に A 型、B 型、および C 型数量化と呼ぶことにする。後二者については項目の組合せによってそれぞれ三つおりの変形がありうる。A 型はそれぞれの項目ごとにそれを構成するカテゴリにある実数値を付与するものである。B 型は、第1項目と第2項目を複合して rs 個のカテゴリ (i, j) , $(\forall i \in \bar{r}, \forall j \in \bar{s})$ から成るひとつの項目とみなし、1の項目と第3項目との2重分割表における数量化にほかならない。C 型数量化は第1項の各カテゴリごとに $s \times t$ 分割表があるとみて、カテゴリ $i \in \bar{r}$ ごとに2重分割表における数量化を行おうというものである。

岩坪(1974)は A 型に対して3次の相関係数と定義した量 ρ の最大化

[†] 本節の内容は吉澤(1975)に報告されている。

を規準とする数量化を、およびB型に対しては z_k に関する相関比の最大化を規準とする数量化を提案している。この種の数量化手法を含めて一般に最適尺度法や多次元尺度構成法といわれる分野では各手法の根底に確率論的モデルを想定してモデルのデータへの適合性を議論することが困難である。ここでは、3重分割表の数量化手法と前述の多重配列の特異値分解及び加法モデルとの関連を指摘して、数量化手法の基礎的理論を構成する。

b) 加法モデルの拡張と3次相関係数による数量化 離散的確率変数 X, Y , および Z を想定し、分割表のセル確率をこれらの変数の同時分布と与えるものと解釈する。すなわち、

$$p_{ijk} = \Pr(X=u_i, Y=v_j, Z=w_k) \quad (\forall i \in \bar{r}, \forall j \in \bar{s}, \forall k \in \bar{t}).$$

岩坪(1974)による3次の相関係数の最大化を規準とするA型の数量化は、制約条件、

$$\sum_i p_{i++} u_i = 0, \quad \sum_j p_{+j+} v_j = 0, \quad \sum_k p_{++k} w_k = 0. \quad (5.9)$$

$$\sum_i p_{i++} u_i^2 = 1, \quad \sum_j p_{+j+} v_j^2 = 1, \quad \sum_k p_{++k} w_k^2 = 1 \quad (5.10)$$

の下で、確率変数 X, Y , および Z の3次の相関係数を最大にする、すなわち、

$$\rho = \sum_{i,j,k} p_{ijk} u_i v_j w_k \rightarrow \max \quad (5.11)$$

とする $\{u_i\}$, $\{v_j\}$, および $\{w_k\}$ を求めるものである。(5.9)と(5.10)は確率変数 X, Y , および Z のそれぞれの期待値がゼロ、分散が1という統

計的意味を持つ。岩坪(1974)はこの問題の決定方程式を導出し、 ρ^2 が有界であること、および不等式

$$\rho^2 < \sum_{i,j,k} \frac{p_{ijk}^2}{p_{i++} p_{+j+} p_{++k}} \quad (5.12)$$

が成り立つことを示している。

一方、この問題は第3章における問題3や定理3.5、および第4章で述べた加法モデルの性質から次の定理として述べるように解釈される。

定理5.1. 条件(5.9)と(5.10)の下で、3次の相関係数 ρ (5.11)を最大にする問題は、同じ条件の下で

$$\rho = \sum_{i,j,k} (p_{ijk} - ap_{ijk}) u_i v_j w_k \quad (5.13)$$

を最大にする $\{u_i\}$, $\{v_j\}$, $\{w_k\}$ を求める問題と同値である。ここで、 ap_{ijk} は(4.24)で与えられるものである。

注意. (5.13)の右辺の因子 $p_{ijk} - ap_{ijk}$ は4.1節で述べたように3重分割表での加法モデルにおける最高次交互作用(3因子効果)の成分にほかならない。

証明. $\Sigma^{(1)}$, $\Sigma^{(2)}$, $\Sigma^{(3)}$ は(4.5)と同様に与えれば、(5.9)は各ベクトル $\underline{u} = (u_i)$, $\underline{v} = (v_j)$, $\underline{w} = (w_k)$ がそれぞれ単位ベクトルと重みつきで直交していることを意味する。そして2.3節で述べた内積の一般化と3.1節

の問題3に適用すれば定理3.5からただちに本定理を得る。 |

この定理5.1から定理3.1及び定理3.4に対応するものとして次の定理が得られる。証明は定理5.1と同様に内積の定義に注意して定理3.1及び定理3.4の証明とそれを組み合わせればよいので省略する。

定理5.2. 3次の相関係数について、次の不等式が成り立つ:

$$\rho^2 \leq \sum_{i,j,k} \frac{(p_{ijk} - ap_{ijk})^2}{p_{i++} p_{+j+} p_{++k}} \quad (5.14)$$

等号が成り立つのは、ある $\{u_i\}$, $\{v_j\}$, $\{w_k\}$ に対して、 $p_{ijk} - ap_{ijk}$ が $p_{i++} p_{+j+} p_{++k} u_i v_j w_k$ に比例するとき、そのときに限る。

注意. 不等式(5.14)が(5.12)より、 ρ^2 の精密な評価を与えていることに注意されたい。(4.23)からそのことは明らかである。

定理5.3. 制約条件(5.9)と(5.10)の下で、3次の相関係数 ρ を最大にするとは、同じ条件の下で、

$$Q = \sum_{i,j,k} \frac{(p_{ijk} - ap_{ijk} - \rho g_{ijk} u_i v_j w_k)^2}{g_{ijk}} \quad (5.15)$$

を最小にする問題と同値である。ここで、

$$g_{ijk} = p_{i++} p_{+j+} p_{++k} \quad (5.16)$$

とする。

注意. 以下で式が長くなるときは g_{ijk} の記法を用いる. なお, 定理 3.4 のあとの注意に述べたように ρ, u_i, v_j, w_k の符号の任意性を除く,

定理 5.1 からわかるように, 3 次相関係数の最大化と規準とする数量化は, 加法モデルにおける 3 因子効果だけが問題になる. そのことを次の定理として述べておく.

定理 5.4. 3 次の相関係数 ρ (5.11) が 条件 (5.9) と (5.10) のもとで恒等的にゼロになるための必要十分条件は 加法モデルでの 3 因子効果が存在しないことである.

注意. 4.1 節で述べた階層的モデル H_1, H_2, H_{3a}, H_4 は皆 ρ がゼロになるための十分条件になっている.

証明. (4.17) によって $p_{ijk} - \alpha p_{ijk}$ は 加法モデルでの 3 因子効果の要素である. よって, (5.14) の不等式から, 仮定から定理を得る. |

以上の議論によって, この A 型数量化が, 加法モデルで定義される 3 因子効果 (4.17) の一部を抽出するものであることが明らかとなった. しからうて, その 3 因子効果を 4.1 節にならって $B(123)$ と書くと, この数量化が意味のあるものであるかどうかは, $B(123)$ が存在して, それ自身が量的にも意味があり, かつ,

$$B(123) = \rho \sum_{ijk} q_{ijk} u_i v_j w_k \quad (5.17)$$

というモデルがほぼ成り立つかどうかということに依存しよう。加法モデルの各種のモデルに (5.17) の右辺の項を加えた形式のモデルの検定については後述 (9 項) する。

なお、不等式 (5.14) の等号の成り立つ分割表の例を示しておこう。

例 5.1. $2 \times 2 \times 2$ 分割表. $r=s=t=2$ の場合は 2 つの条件 (5.9)

と (5.10) から $\{u_i\}, \{v_j\}, \{w_k\}$ は付号を除いて一意に決まってしまう。すなわち、

$$u_1 = \sqrt{\frac{p_{2++}}{p_{1++}}}, \quad v_1 = \sqrt{\frac{p_{+2+}}{p_{+1+}}}, \quad w_1 = \sqrt{\frac{p_{++2}}{p_{++1}}}$$

と正に決まれば、

$$u_2 = -\sqrt{\frac{p_{1++}}{p_{2++}}}, \quad v_2 = -\sqrt{\frac{p_{+1+}}{p_{+2+}}}, \quad w_2 = -\sqrt{\frac{p_{++1}}{p_{++2}}}$$

となる。このとき、3 因子効果が (5.17) を満たすことも直接な計算によっても、3 因子効果の次元 $(r-1)(s-1)(t-1)$ がこの場合 1 であることから示され、不等式 (5.14) で等号が成り立つ場合であることがわかる。

例 5.2. 1 要素集中型分割表. 一つのセルを除いて残りのセルの確率が一定であるような特殊な分割表を 1 要素集中型と呼ぼう。そのセル確率を

$$p_{ijk} = \frac{1}{N} (1 + d \delta_{ia} \delta_{jb} \delta_{kc}) \quad (5.18)$$

と置く。ここで、 $N = rst + d$ 、 d は定数、 δ_{ia} などはKroneckerのデルタとする。観測されるセル度数で見れば、ある一つのセル (a, b, c) だけで $1 + d$ となり、残りのセルは1の値をとる分割表である。この場合、

$$p_{ijk} = a p_{ijk} + \frac{d(N-2d)}{N^3} (r \delta_{ia} - 1)(s \delta_{jb} - 1)(t \delta_{kc} - 1)$$

と表現できるとが確かめられる。

C) 相関比による数量化に対するモデル 岩坪(1974)は一つの項目について相関比最大化の規準による数量化を提案している。又三項目をとりあげ、数量化も記号的に

$$k \leftarrow w_k \quad (\forall k \in T)$$

とする。相関比はある確率変数 Z について Z の何らかの条件つき期待値の分散と Z 自身の分散との比を意味する。そこで、条件

$$\sum_k p_{++k} w_k = 0 \quad (5.19)$$

の下で、

$$\eta^2 = \frac{\sum_i \sum_j p_{ij+} \left(\sum_k \frac{p_{ijk}}{p_{ij+}} w_k \right)^2}{\sum_k p_{++k} w_k^2} \quad (5.20)$$

を最大にする $\{w_k\}$ を求めろ。この考え方も相関比最大化の規準による数量化という。

この数量化は、実は、単純なB型数量化にすぎないことがわかる。オ
 ー項目とオ二項目を複合して、 rs 個のカテゴリ (i, j) ($\forall i \in \bar{r}, \forall j \in \bar{s}$)
 からなる新しい項目を作る。与えられた3重分割表をその複合項目とオ
 三項目による $(rs) \times t$ の2重分割表と考える。そこで、5.1節の2重分割
 表に対する数量化もB型数量化として用いることができる。その数量化
 は次のようである：条件

$$\sum_{i,j} p_{ij+} u_{ij} = 0; \quad \sum_k p_{++k} w_k = 0; \quad (5.21)$$

$$\sum_{i,j} p_{ij+} u_{ij}^2 = 1; \quad \sum_k p_{++k} w_k^2 = 1 \quad (5.22)$$

の下で、

$$\rho = \sum_{i,j,k} p_{ijk} u_{ij} w_k \quad (5.23)$$

を最大にする $\{u_{ij}\}$ と $\{w_k\}$ を求めること。これに対する決定方程式
 は(5.4)と本質的に同じで

$$\left. \begin{aligned} \sum_k p_{ijk} w_k &= \rho p_{ij+} u_{ij}, \\ \sum_{i,j} p_{ijk} u_{ij} &= \rho p_{++k} w_k \end{aligned} \right\} \quad (5.24)$$

と与えられる。(5.24)から u_{ij} を消去すると

$$\sum_{k'=1}^t \left(\sum_{i,j} \frac{p_{ijk}}{p_{++k}} \cdot \frac{p_{ijk'}}{p_{ij+}} \right) w_{k'} = \rho^2 w_k \quad (5.25)$$

という固有方程式が得られる。これが相関比 η^2 (5.20) の最大化に対する決定方程式にほかならないことは容易に示される。

したがって、相関比最大化も基準とする数量化の諸性質は 5.1 節で述べた 2 重分割表の性質から導かれる。ここでは次の定理だけを示しておく。証明は省略する。

定理 5.5. (5.20) の相関比 η^2 について次の不等式が成り立つ:

$$\begin{aligned}\eta^2 &\leq \sum_{i,j,k} \frac{(p_{ijk} - p_{ij+} p_{++k})^2}{p_{ij+} p_{++k}} \\ &= \sum_{i,j,k} \frac{p_{ijk}^2}{p_{ij+} p_{++k}} - 1.\end{aligned}\quad (5.26)$$

以上の考察から、相関比最大化の規準による数量化が意味を持つのは、セル確率について、モデル

$$H_r: p_{ijk} = p_{ij+} p_{++k} (1 + \rho u_{ij} w_k) \quad (5.27)$$

が成り立つようなときであるといえよう。なお、B 型数量化としては、添字の入れかえによって

$$(B') \quad (i, k) \leftarrow w_{ik} \quad (\forall i \in \bar{r}, \forall k \in \bar{t}), \quad j \leftarrow v_j \quad (\forall j \in \bar{s})$$

$$(B'') \quad (j, k) \leftarrow v_{jk} \quad (\forall j \in \bar{s}, \forall k \in \bar{t}), \quad i \leftarrow u_i \quad (\forall i \in \bar{r})$$

という形式が考えられ、それぞれに対応するモデル

$$H_{r'}: p_{ijk} = p_{i+k} p_{j+} (1 + \rho w_{ik} v_j) \quad (5.28)$$

$$H_{r''}: p_{ijk} = p_{+jk} p_{i++} (1 + \rho v_{jk} u_i) \quad (5.29)$$

を考えるとすることができる。これは部分独立モデルからの式も抽出している。

d) 条件付き独立性からの数量化とそのモデル 一つの変数Xの値が与えられた条件のもとでYとZが独立である場合、すなわち、条件付独立モデル

$$H_{3c}: p_{ijk} = \frac{p_{ij+} p_{i+k}}{p_{i++}} \quad (5.30)$$

を中心にいくつかの数量化手法も考えることができる。まず、このモデルが $i \in \bar{r}$ のいくつかのカテゴリについて成り立たないとき、そのカテゴリの部分集合を $\bar{r}'$ とする。各々の $i \in \bar{r}'$ について $\{p_{ijk}\}$ を j と k について $s \times t$ の2重分割表と考え、C型の数量化を行うことができる。すなわち、 $\forall i \in \bar{r}'$ に対し、条件

$$\sum_j p_{ij+} v_{ij} = 0; \quad \sum_k p_{i+k} w_{ik} = 0; \quad (5.31)$$

$$\sum_j p_{ij+} v_{ij}^2 = 1; \quad \sum_k p_{i+k} w_{ik}^2 = 1 \quad (5.32)$$

の下で、

$$f_i = \sum \frac{p_{ijk}}{p_{i++}} v_{ij} w_{ik} \longrightarrow \max \quad (5.33)$$

とする $\{v_{ij}\}$, $\{w_{ik}\}$ を求める。これに対するモデルを

$$H_d: p_{ijk} = \frac{p_{ijr} p_{i+k}}{p_{i+r}} (1 + p_i v_{ij} w_{ik}) \quad (v_i \in \bar{r}) \quad (5.34)$$

としておこう.

e) 加法モデルの一般化. これまで述べた数量化手法では (5.9), (5.24), (5.31) のように数値を付与して想定された確率変数の期待値がゼロであるという制約が置かれていた. それらの制約は, それぞれの添字に対応するベクトル空間に付与すべき数値ベクトルに対して重みつきで直交する単位ベクトル $(1, 1, \dots, 1)'$ をあらかじめ想定していること, しかるしてその単位ベクトルに相対的な最高次の交互作用に対する成分のみが数量化の対象となるという点で重要な意味を持つ. いま, これらの制約は位置に関する制約, あるいは位置条件といわれる. 一方, 分散が 1 であるという (5.10) のような条件は尺度に関する制約, あるいは尺度条件といわれる. 尺度条件は, ベクトル空間 (2.3 節の意味で) の内積の定義に即していれば, 右辺の 1 は本質的な制約ではない.

ここでは, これまでのモデルを一般化して, 位置の制約の意味を別の観点から論じておこう. モデルを

$$H_g: p_{ijk} = p_{ijk}^0 + \rho g_{ijk} u_i v_j w_k$$

とする. ここで, $\{p_{ijk}^0\}$ は与えられるセル確率 $\{p_{ijk}\}$ とは無関係に設定しても, あるいは加法モデルでの $a p_{ijk}$ のようにセル確率

$\{p_{ijk}\}$ から作られるものとしてもよい。二つの離散分布 $\{p_{ijk}\}$ と $\{p_{ijk}^0\}$ の相違を $\rho g_{ijk} u_i v_j w_k$ で表わすと考えてもよい。このモデルのパラメータ $\{u_i\}, \{v_j\}, \{w_k\}$ については以下に条件を付加する。

まず、尺度条件

$$(C1) \quad \sum_i p_{i++} u_i^2 = \sum_j p_{+j+} v_j^2 = \sum_k p_{++k} w_k^2 = 1 \quad (5.35)$$

については前述のようにあまり議論のないところであろう。Hg における $u_i v_j w_k$ という積の形から何らかの尺度条件が必要であり、 $\{u_i\}, \{v_j\}, \{w_k\}$ について対称的な条件を置くかわりに ρ がパラメータとして追加されているともいえる。

次に、セル確率の総和が1であるという要求から、 $\{p_{ijk}^0\}$ についても

$$(C2) \quad \sum_{i,j,k} p_{ijk}^0 = 1 \quad (5.36)$$

とする。これまでの位置条件を

$$(C3) \quad \sum_i p_{i++} u_i = \sum_j p_{+j+} v_j = \sum_k p_{++k} w_k = 0 \quad (5.37)$$

としておく。セル確率の総和が1ということから、モデル Hg では

$$\sum_{i,j,k} \rho g_{ijk} u_i v_j w_k = 0$$

であればよく、例えば (C3) のかわりに

$$(C4) \quad \sum_i p_{i++} u_i = 0 \quad (5.38)$$

という条件だけを置くことも考えられる。

以上の制約条件を一部の組合せで含め三通りのモデルを検討しよう。

$$Hg_1: p_{ijk} = p_{ijk}^0 + \rho g_{ijk} u_i v_j w_k,$$

制約条件 C1, C2, C3;

$$Hg_2: p_{ijk} = p_{ijk}^0 + \rho g_{ijk} u_i v_j w_k,$$

制約条件 C1, C2, C4;

$$Hg_3: p_{ijk} = p_{ijk}^0 + \rho g_{ijk} u_i v_j w_k,$$

制約条件 C1, C2.

各モデルの構造を決定する規準として、いずれのモデルでも、それぞれの制約条件のもとで

$$Q = \sum_{i,j,k} \frac{(p_{ijk} - p_{ijk}^0 - \rho g_{ijk} u_i v_j w_k)^2}{g_{ijk}} \rightarrow \min \quad (5.39)$$

という規準を採用する。この規準は

$$\rho = \sum_{i,j,k} (p_{ijk} - p_{ijk}^0) u_i v_j w_k \rightarrow \max \quad (5.40)$$

と同値である。それぞれのモデルに対する決定方程式は次のようになる。

Hg_1 の場合:

$$\left. \begin{aligned} \sum_{j,k} \{ (p_{ijk} - p_{ijk}^0) - p_{i++} (p_{+jk} - p_{+jk}^0) \} v_j w_k &= \rho p_{i++} u_i \\ \sum_{i,k} \{ (p_{ijk} - p_{ijk}^0) - p_{+j+} (p_{i+k} - p_{i+k}^0) \} u_i w_k &= \rho p_{+j+} v_j \\ \sum_{i,j} \{ (p_{ijk} - p_{ijk}^0) - p_{++k} (p_{ij+} - p_{ij+}^0) \} u_i v_j &= \rho p_{++k} w_k \end{aligned} \right\} \quad (5.41)$$

H_{g_2} の場合:

$$\left. \begin{aligned} \sum_{j,k} \{ (p_{ijk} - p_{ijk}^0) - p_{i++} (p_{+jk} - p_{+jk}^0) \} v_j w_k &= \rho p_{i++} u_i \\ \sum_{i,k} (p_{ijk} - p_{ijk}^0) u_i w_k &= \rho p_{+j+} v_j \\ \sum_{i,j} (p_{ijk} - p_{ijk}^0) u_i v_j &= \rho p_{++k} w_k \end{aligned} \right\} (5.42)$$

H_{g_3} の場合:

$$\left. \begin{aligned} \sum_{j,k} (p_{ijk} - p_{ijk}^0) v_j w_k &= \rho p_{i++} u_i \\ \sum_{i,k} (p_{ijk} - p_{ijk}^0) u_i w_k &= \rho p_{+j+} v_j \\ \sum_{i,j} (p_{ijk} - p_{ijk}^0) u_i v_j &= \rho p_{++k} w_k \end{aligned} \right\} (5.43)$$

これらの決定方程式を比較するとによって, $\{u_i\}$, $\{v_j\}$, $\{w_k\}$ の位置条件の違いが決定方程式のどの部分へ影響するかが明らかであろう. 逆に $\{p_{ijk}\}$ と $\{p_{ijk}^0\}$ とが適当な関係にあれば, (5.42) や (5.43) の解が条件 C3 を満たすことがある. 例えば, 2変数周辺和が互いに一致する場合, すなわち

$$p_{ij+} = p_{ij+}^0, \quad p_{i+k} = p_{i+k}^0, \quad p_{+jk} = p_{+jk}^0 \quad (5.44)$$

となるとき, (5.41) と (5.42) はそれぞれ (5.43) に帰着する. したがって, (5.44) が満たされるならば, (5.43) の解は H_{g_3} に含まれていながら制約 C3 を満たすことになる. 次に

$$p_{ijk}^0 = p_{i++} p_{+jk}^0, \quad p_{ijk}^0 = p_{+j+} p_{i+k}^0, \quad p_{ijk}^0 = p_{++k} p_{ij+}^0 \quad (5.45)$$

を満足するとは, (5.41) が次のように変形される:

$$\left. \begin{aligned} \sum_{j,k} (p_{ijk} - p_{i+r} p_{jk}) v_j w_k &= \rho p_{i+r} u_i, \\ \sum_{i,k} (p_{ijk} - p_{+j+} p_{i+k}) u_i w_k &= \rho p_{+j+} v_j, \\ \sum_{i,j} (p_{ijk} - p_{++k} p_{ij+}) u_i v_j &= \rho p_{++k} w_k. \end{aligned} \right\} \quad (5.46)$$

これは, 岩坪 (1974, p.61) が導出した3次の相関係数と規準とする数量化の決定方程式にはかならない. 加法モデルで3因子効果と除いたモデルを p_{ijk}^0 とすれば (5.44) を満たす. (5.45) は満たされるとはかぎらないが, (5.41) の p_{ijk} へ代入し条件C3を用いれば, (5.41) は (5.46) の形に変形される.

なお, 条件独立モデルを使って

$$p_{ijk}^0 = \frac{p_{ij+} p_{i+k}}{p_{i+r}} \quad (5.47)$$

とすれば, (5.44) のオ1とオ2式を満足する. したがって, そのときは (5.32) を解けば, その解は条件C3を満たす. $\{p_{ijk}^0\}$ の候補としては, 多くのものが考えられるが加法モデルや乗法モデルに由来するものを表5.1と表5.2に示す. 表5.1と5.2に示される次元や自由度は g 項で説明する.

士) 乗法モデル. 3次の相関係数の最大化を規準とする数量化は3項目間の同時関連性, すなわち3因子効果と $u_i v_j w_k$ のような積の形で

抽出する手法であるといつてよい。a) 項では加法モデルによって定義される3因子効果 $p_{ijk} - \alpha_{ijk}$ とその数量化手法との関連が指摘されている。

一方、分割表の代表的モデルである乗法モデルあるいは対数線形モデルによって定義される3因子効果を利用することも考えられる。4.1節で述べたように、対数線形モデルの階層的モデルは標本抽出過程のあり様の仮定の下ではその最大推定量がいくつかの周辺和を合わせるという方針で比較的容易に求められる。例えば、3因子効果非存在のモデル $\mu_{ijk}^{(123)} = 0$ [(4.32) 参照]でのセル確率の推定値は、 $\hat{m}_{ijk}$ と書くと

$$p_{i+j} = \hat{m}_{i+j}, \quad p_{i+k} = \hat{m}_{i+k}, \quad p_{+jk} = \hat{m}_{+jk} \quad (5.48)$$

の関係にある。そこで、モデル H_g における p_{ijk} と $\hat{m}_{ijk}$ にとると、上式(5.48)は(5.44)に相当し、(5.29)の規準のもとで決定方程式(5.43)を解けば条件C3も満足する $\{u_i\}, \{v_j\}, \{w_k\}$ が得られる。

乗法的なモデルを拡張していけば、 H_g のかわりに

$$H_n: p_{ijk} = \rho p_{ijk}^0 u_i v_j w_k \quad (5.49)$$

という形式のモデルも考えることが出来る。 $\{p_{ijk}^0\}$ は $\{p_{ijk}\}$ と無関係に設定されるものであってよい。母数 $\rho, \{u_i\}, \{v_j\}, \{w_k\}$ は尺度条件C1と次式を満足するものとする:

$$\sum_{i,j,k} \rho p_{ijk}^0 u_i v_j w_k = 1. \quad (5.50)$$

一般的にいて、乗法的なモデルでは最小=乗法的な規準より、Kullback の情報量 [Kullback (1959)] を利用したほうが便利である。

そこで、モデル H_n に対し、規準

$$\sum_{i,j,k} p_{ijk} \log \frac{p_{ijk}}{\rho p_{ijk}^0 u_i v_j w_k} \rightarrow \min \quad (5.51)$$

によって H_n の構造を決定することにしよう。その決定方程式は

$$\left. \begin{aligned} \rho \sum_{j,k} (p_{ijk}^0 v_j w_k) u_i &= p_{i++} \\ \rho \sum_{i,k} (p_{ijk}^0 u_i w_k) v_j &= p_{+j+} \\ \rho \sum_{i,j} (p_{ijk}^0 u_i v_j) w_k &= p_{++k} \end{aligned} \right\} \quad (5.52)$$

となる。この解に対し、

$$n p_{ijk} = \rho p_{ijk}^0 u_i v_j w_k$$

とみると、(5.52) は

$$n p_{i++} = p_{i++}, \quad n p_{+j+} = p_{+j+}, \quad n p_{++k} = p_{++k} \quad (5.53)$$

と書ける。対数線形モデルではセル確率の推定値 $\hat{p}_{ijk}$ が

$$\hat{p}_{i++} = p_{i++}, \quad \hat{p}_{+j+} = p_{+j+}, \quad \hat{p}_{++k} = p_{++k}$$

の関係を保持してしまうので、 $\rho = u_i = v_j = w_k = 1$ とし、 p_{ijk}^0 を

その $\hat{p}_{ijk}$ にとれば (5.52) が満たされる。その解が (5.51) を最小

にするかどうかは保留しても、 H_n は普通の乗法モデルに帰着してしまう。

p_{ijk}^0 を周辺和の関数にとり場合でも

$$p_{ijk}^0 = \frac{p_{+jk} p_{i+k} p_{ij+}}{p_{i++} p_{+j+} p_{++k}} \quad (5.54)$$

とすれば, p_{ijk}^0 は乗法モデル H_m での一つの推定値 $\hat{m}p_{ijk}$ とは異なり, ある程度の意味が残されている。Darroch (1962, 1974) や Lancaster (1971) で議論されているように, 'Perfect two-variable marginal probabilities' を持つ分割表の場合, すなわち, 条件

$$\sum_i \frac{p_{ij+} p_{++k}}{p_{i++}} = p_{+j+} p_{++k} \quad (5.55)$$

が, 添字を入れかえに三とありの場合について, 成り立つとき, (5.64) は $\hat{m}p_{ijk}$ に一致する。そこで, (5.54) の p_{ijk}^0 を用いるモデル H_n は 'imperfection' の程度, すなわち (5.65) の成り立たない程度を示す数量化のモデルといってもよいであろう。

g) モデルの検定および選択. 一般に, 統計的モデルのあいはいは仮説の検定は, 'より広いモデル G の中で限定されたモデル H が成立するといふことができるか' という形式で行われる。Good (1973) は, モデルに含まれる数学的に独立な母数の個数をそのモデルの次元 $\dim(\cdot)$ とし, モデル G の中でのモデル H の検定の自由度を

$$df(H|G) = \dim(G) - \dim(H) \quad (5.56)$$

と定義すると提案している. 2.2 節で述べた一般モデル (2.30)

の次元は

$$\dim(H) = \sum_{\forall \alpha \in \mathcal{H}} \dim(\Omega(\alpha)) \quad (5.57)$$

で与えられる. ここで, $\Omega(\alpha)$ は母数空間 Ω での α -効果空間,

$\mathcal{H}$ は $\overline{m}$ の部分集合の任意のクラスであり, そのモデルが

$$H: A \in \bigoplus_{\forall \alpha \in \mathcal{H}} \Omega(\alpha),$$

すなわち, $\Omega(\alpha)$ のある直和空間に母数が含まれるという形式のもの

であったことに注意しておく. そして, より広いモデルを母数空間

Ω 全体に与えれば

$$G: A \in \Omega$$

かつ

$$\dim(G) = n$$

と書ける. そしてその G の中で H の検定の自由度を (5.56) で求める

ものとすることができよう.

表5.1と表5.2では本節で扱ったモデルを一般的にモデル H_g

$$H_g: p_{ijk} = p_{ijk}^0 + \rho g_{ijk} u_i v_j w_k$$

とし, p_{ijk}^0 の各種の場合と H_g の次元, および

$$G: \sum_{i,j,k} p_{ijk} = 1$$

というもっとも広いモデルの中での H_g の検定の自由度を表にまとめて

ある。このとき, $\dim(G)$ は $rst-1$ である。

実際の検定は, つまびきに述べた標本抽出過程のポアソンモデルや多項モデルの仮定のもとでは,

$$\chi^2 = \sum_{i,j,k} \frac{(x_{ijk} - \hat{m}_{ijk})^2}{\hat{m}_{ijk}}, \quad (5.58)$$

$$G^2 = 2 \sum_{i,j,k} x_{ijk} \log \frac{x_{ijk}}{\hat{m}_{ijk}}, \quad (5.59)$$

あるいは Hellinger distance (4.41), Freeman-Tukey (1950) による統計量などを用いることができよう。ここで $\{x_{ijk}\}$ は与えられた重分割表のセル度数, $\hat{m}_{ijk}$ は各モデルの下での x_{ijk} の期待値の推定値である。対数線形モデルでは上記の検定統計量が漸近的に χ^2 分布に従う。しかし, 加法モデルや数量化のために $\rho u_i v_j w_k$ のような項を追加したモデルでは, (4.5) での $\Sigma^{(1)}$, $\Sigma^{(2)}$, $\Sigma^{(3)}$ の定義からわかるようにセル確率自体を用いて内積を定義して因子効果を抽出するという事実から, χ^2 や G^2 が漸近的に χ^2 分布に従うとすることはできない。近似的に χ^2 分布を用いた検定を行っても経験的にほとんど正しいかはわからないが, 注意を要するところである。

モデルの選択については, 1.3 節で述べたように, 対数線形モデルについての Goodman (1971) の方式や 情報量規準 (AIC) の利用〔

吉沢・藤本(1976) 参照] などによって行うことができよう。従来の
数量化手法は、何らかの規準を満たすように、与えらるべき数値を
作り出すだけであつた。そこで与えられたデータに対する適合性がよい
のか悪いのかを判断することができない。近似的にしろ、これまで
に示したモデルの適合性を検定することによって数量化の妥当性に
対する一つの判断基準が与えられるという点に本節の理論の意味
があると思われる。

5.3 多重分割表に対する数量化手法

前節では 3重分割表に対する数量化手法とそのモデルも詳しく扱った。その中の A 型, B 型, C 型という数量化の形式は一般の多重分割表ではさらに複雑になる。しかし, B 型や C 型の数量化が 2重分割表の A 型数量化に帰着しるように, 一般の多重分割表においても複合項目を考え方によって より少ない指標の因子列の問題に帰着させられることもある。

一方, 多重分割表に対する数量化手法として 最高次相関係数の最大化も規準として各項目ごとに数量化するものはあまり有効とは思えない。その規準は, 3重分割表のときと同様に, 最高次交互作用の一部も抽出する意味を持ち, それもノイズも含めや高い高次の交互作用の不安定さの影響を受けてしまう。もっとも, 高次の交互作用も異常なセル度数がみかけ上大きくなっていうようなときは, 数量化としての意味より異常データの検出という目的に使用できるときがある。

そこで, 一般の多重分割表に対しては, 確率論的に解釈しやすい相互独立モデル, 部分独立モデル, あるいは条件独立モデルを中心として, それらの一つが適合するならそのモデル空間の中での数量化を行う, 逆に
あるいは, 不適合なモデルの空間からのずれを数量化し, ときには異常値

の検出手段として利用するという方向が望ましいと思われる。

4.1 節では加法モデルと多重配列の特異値分解との関連を述べた。前節の加法モデルと数量化の関係は、得られたベクトル $\underline{u} = (u_i)$, $\underline{v} = (v_j)$, $\underline{w} = (w_k)$ にさらに直交するベクトルを特異値分解によって得られることを示している。一般的には定理 3.8 によって多重配列が分解され、高次の交互作用がその次元とともに大きく残っているときには意味があるかもしれない。しかし、数量化の観点からの実用性には疑問が残る。

5.4 平方根分解

本節では、与えられた m 重配列のセル度数を $x_{\lambda_1 \lambda_2 \dots \lambda_m}$ とするとき、その平方根を要素とする m 重配列

$$A = (\sqrt{x_{\lambda_1 \lambda_2 \dots \lambda_m}}) \quad (5.60)$$

の特異値分解に簡単に触れ、次節での数値例による検討の準備とする。この分解を平方根分解と呼ぶ [吉澤 (1974)]。4.2 節で述べた平方根線形モデルはポアソンモデルなどでの分散安定化変換の効果もぬらいとし、分散分析の母数モデルと同様の線形モデルであった。次節での数値例で示るように、多重分割表データの平方根分解によるモデルの選択は対数線形モデルや加法モデルの場合とほとんど同じ結論を導く。その点に関しては、平方根線形モデルは事情が異なり、 α -因子効果の定義の相違が他のモデルとの間で相当に現われてくる。

平方根分解は、上述のように、(5.60)の特異値分解であり、第3章の諸定理が直接に適用される。定理3.6によって分解の各段での1因子効果が消えることを除くと、モデルの検定や選択には4.2節で平方根線形モデルについて述べた手法が利用できる。平方根線形モデルではセル確率の平方根を要素とする配列を扱った。ここでは次節の数値例や実用上の観点から(5.60)のようにセル度数の平方根を用いるので、(4.50)の修正を示しておこう。

特異値分解の第1段で得られるベクトル $\underline{u}^{(i)}$ ($\forall i \in \bar{m}$) についての (5.60) の A の α -因子効果を $B(\alpha)$ とする。階層的 (非階層的でもよいが) モデルに 対する $\bar{m}$ の部分集合 α のクラスを $\mathcal{A}$ とする。そのとき,

$$\lambda = \left(\frac{1}{N} \sum_{\alpha \in \mathcal{A}} \underline{B}(\alpha)' \underline{B}(\alpha) \right)^{1/2} \quad (5.61)$$

とし,

$$D^2 = 4N (\cos^{-1} \lambda)^2 \quad (5.62)$$

を求める。この D^2 が近似的に χ^2 分布に従うことをモデルの検定に用いる。

(5.61) と (4.49) の相違は $1/N$ の因子の有無である。 D^2 の自由度は,

1 因子効果が消えることから $\underline{B}(\phi)' \underline{B}(\phi)$ に対して

$$\sum_{i=1}^m r_i - m + 1 \quad (5.63)$$

を割り当て、 α が空でない平方和には

$$\prod_{\alpha \in \mathcal{A}} (r_i - 1) \quad (5.64)$$

として、 $\mathcal{A}$ に含まれる α についての総和を n から引いて求める。 n はセル総数 $r_1 \times r_2 \times \dots \times r_m$ である。

4.2 節の終りに平方根線形モデルの特徴をあげたが、その $\alpha=1$ と $\alpha=2$ の特徴は平方根分解にも同様である。対数線形モデルの検定で (5.59) の G^2 や (5.58) の χ^2 を用いるには $\hat{m}_{ijk}$ の計算がモデルごと必要であり、その都度 G^2 などとその式によって計算しなおさなければならな

い。平方根分解では次節で示すように一回の特異値分解によって平方和の表を作っておけば、回帰分析における変数選択 [例えば, 奥野他 (1969) 参照] のような考え方で因子効果の選択が可能となり, それぞれのモデルに対する D^2 の計算もきわめて簡単である。一方, 加法モデルでの平方和分解は, 3重分割表の場合について (4.23) 式で示したように, 例えば 3 因子効果が存在しないというモデル H_a に対する平方和

$$\sum_{i,j,k} \frac{(p_{ijk} - ap_{ijk})^2}{p_{i++} p_{+j+} p_{++k}}$$

は, いわゆる χ^2 検定統計量 (5.58) の

$$\sum_{i,j,k} \frac{(p_{ijk} - ap_{ijk})^2}{ap_{ijk}}$$

には一致しない。これに対し (5.60) の D^2 は (5.58) の χ^2 や (5.59) の G^2 と漸近的には同一の分布に従うものになっている。また, 多重-次形式 (3.2) の $\underline{A}'\underline{u}$ の最大化は,

$$u_{\lambda_1}^{(1)} u_{\lambda_2}^{(2)} \cdots u_{\lambda_m}^{(m)} = \sqrt{p_{\lambda_1 \lambda_2 \cdots \lambda_m}}$$

とおくことによって, 相互独立モデルの方向への Hellinger distance の最小化とみる事ができる。相互独立モデルの方向への何らかの距離の最小化は Darroch (1974) が指摘している 加法モデルと

対数線形モデルがともに持つ性質である。次節に示すように、平方根分解、加法モデル、及び対数線形モデルが類似性を持つのは、その性質によろう。3重分割表における3因子効果非存在の加法モデルは近似的にその対数線形モデルを満たすことは Darroch (1974) に示されている。

5.5 数値例

平方根分解による多重分割表モデルの選択を中心に、これまで述べた加法モデルや対数線形モデルとの比較検討を二三の現実のデータを用いて行う。数値計算は山梨大学工学部計算機科学科のFACOM 230-45Sにより、FORTRAN プログラムの倍精度計算を使用した。第3章の算法1を中心として加法モデル、平方根線形モデル、及び平方根モデルが統一的に扱える機能を持つプログラムを作製したほか、対数線形モデル用には別途にプログラムを用意した。

多重配列データやテンソル積のプログラム上の取扱いは、プログラムの巧拙によって計算所要時間が著しく異なってくるので注意を要する。我々のプログラムでは (3.29) による因子効果

$$\tilde{B}(\alpha) = \left(\bigotimes_{i=1}^m \tilde{U}_{x_{\alpha(i)}}^{(i)} \right)^k \tilde{A} \quad (5.65)$$

をすべての反復 $k=1, 2, \dots$ とすべての $\alpha \in \mathcal{M}$ について求めるようにしたのでこの計算にもっとも気を付けなければならない。決定方程式 (3.5) の解法がもう一つの要諦であるが、これについては単純な反復計算によってもセル数が1000程度までは経験的に実用上の問題はあまりない。(5.65) の右辺の括弧内は $n \times n$ の行列、 $\tilde{A}$ は n 元ベクトルとして普通の行列とベクトルの内積をとると、プログラミングは簡単であるが、 n^2 に比例する時間がかかる。

その括弧内の行列はそもそもテンソル積の構造を持っているので、高速フーリエ変換 (FFT) の算法のように、多重配列の指標ごとに処理していくべきである。そうすれば $n \log n$ に比例する時間で済む。さらに $U_0^{(i)}$ や $U_1^{(i)}$ の構造を利用すると n に比例する時間に短縮される。(5.65)の計算のプログラム全体に占める時間はきわめて大きく、もっともまづいプログラムで $8 \times 9 \times 10$ の配列について 40分程度かかるものが、上述の工夫によって 40秒に短縮された例がある。

a) 入試データ. ある大学の工学部の入学者選抜試験は5教科について行われ、うち1教科(社会)は5科目のうち1科目を選択して受験することになっている。その5科目を A, B, C, D, E とする。選択科目については問題の難易等によるある意味の不公平が生じることがある。この問題について吉澤・藤本(1976)による対数線形モデルを用いた解析が報告されているが、ここではそんで使用された二種の $5 \times 4 \times 2$ の3重分割表を用いて各種のモデルの適合性を調べる。表5.3と表5.4にその分割表のデータを示す。項目は、両者とも、科目(5カテゴリ)、年度(4)、合否(合格と不合格の2カテゴリ)である。表5.3は実際の合格判定による科目×年度×合否の分割表である。一方、表5.4は選択のある教科を除いた残り4教科の総点による合格判定を行ったときの比較用の分割表である。

まず、表5.3と表5.4に対する平方根分解、平方根線形モデル、加法モデル(1)、及び加法モデル(2)による要因別の平方和分解の様子を、それぞれ表5.5 a~d と表5.6 a~d に示す。加法モデル(2)は4.1節の記述に従ってセル確率(相対度数)を用いたもので、加法モデル(1)は χ^2 の分解になるようにしたものである。セル度数を x_{ijk} とすると、加法モデル(1)は

$$\frac{N x_{ijk}}{\sqrt{x_{i++} x_{+j+} x_{++k}}}$$

を要素と列配列について特異値分解を行うことに相当する。表5.5と表5.6の上の見出しは左から、FACTOR(要因 α)、DIM(α -効果空間の次元)、S.S. (α -因子効果に対する平方和)、以上が特異値分解の第1段、以下は特異値分解の第2段以後について繰り返される。まず、平方根線形モデルを除いて1因子効果は理論どおりに消えていること、平方根分解と平方根線形モデルの第1段での平方和の合計は5761でデータの総和に等しいこと、加法モデルでの μ -効果(総平均)に対する平方和はモデル(1)で5761、モデル(2)で1と予想どおりであることが観測される。

次に、3重分割表に対する8通りの階層的モデルに対する検定統計量の値を平方根分解(D^2)、加法モデル[(4.23)の分解による χ^2_a]、対数線形モデル(χ^2 と G^2)について、表5.5と表5.6に対し表5.7と表5.8に示す。表5.7と表5.8の最後の三欄は、モデルの検定の自由度、1の自由

度と対数線形モデルでの χ^2 による下側確率, 情報量規準(AIC)の値である。赤池(1973)による情報量規準は

$$AIC = G^2 - 2 \times \text{自由度}$$

で計算され, その値が小さいモデルは相対的に妥当なモデルと判断しようという規準である。加法モデルではモデルから除かれる要因に対する平方和を加えることによってその χ^2_a が計算される。例えば, 階層的モデル 12/13 に対しては要因として {23}- と {123}-効果を除かれ表5.7では

$$31.00 + 5.09 = 36.09$$

が χ^2_a の値となる。平方根分解の D^2 は(5.62), 対数線形モデルの χ^2 と G^2 はそれぞれ(5.58)と(5.59)による。

表5.7と表5.8で, 対数線形モデルに対する χ^2 と G^2 がほぼ等しいのは当然としても, D^2 と χ^2_a ともかなり近い値を示しているのは注目に値しよう。なお, 階層的モデルは2.2節で述べたようにその生成クラスで記号的に示されている。1/2/3 は相互独立, 1/23, 2/13, 3/12 は部分独立, 12/13, 12/23, 13/23 は条件独立, 12/13/23 は連関モデルといわれる。結論としては表5.7 ではいずれの階層的モデルも採択されず, 表5.8 では 12/13, すなわち 科目ごとに年度と合否は独立というモデルが一応採用される。この結論だけで, 選択科目間の不公平さを云々

するのは早急にすぎよう。しかし、選択のあり教科を除いたときの分割表(表5.4)にみられる構造化性、すなわち、科目ごとには年度と合否が独立という構造化性があり、その教科を加えることによってみだされていることは確かであろう。

以上、平方根分解、加法モデル、及び対数線形モデルの三者が似たような傾向を持つのに対し、平方根線形モデルはそれらとは異なる様相を示す。平方根モデルの定義から1因子効果が消えない。階層的モデルの選択を行えば、形式的には前三者のときと同じモデルが選択されるが、その解釈は条件独立モデルというような確率論的に正確なものにはならない。一方で、平方根線形モデルではオ2項目の主効果(121-因子効果)が小さいということとは、{12}-因子効果は存在するにもかかわらず、主効果としては無視できる、すなわち、非階層的なモデルを考えてもよいことを意味する。現実には、非階層的モデルは利用しづらいが、階層的モデルでなおかつ主効果が無視できるなら、例えば、部分独立モデル

$$p_{ijk} = p_{i++} p_{+jk}$$

に対し、 p_{i++} が i に依存しないというこのようなモデルを導くので、その意味は大きい。実際、表5.3や表5.4のデータではオ2項目(年度)の周辺和はそれほど変動しないという事実が平方根モデルによって

- 尺、強く現われたと解釈される。

b) モデルの段階的選択. Goodman (1971) は、前述したように、対数線形モデルによるモデル構成のための段階的選択手続の一例を検討している。ここでは、Goodman が同じ使い込データに平方根分解を適用し、その平方和の分解表をもとにしてモデルの選択を試みる。

データを表 5.9 に示す。 $2 \times 2 \times 2 \times 3$ の 4 重分割表である。このデータは洗剤の消費者テストによるものであるが、ここではその詳細に触れず、モデルの選択のみに限って分析しよう。なお、Goodman の使用した項目記号は $T \times U \times P \times S$ であり、例えば、同じ要因 134 と書くことは $\{TPS\}$ に対応する。

平方根分解による平方和の分解の結果を表 5.10 にまとめ、最後の欄はそれぞれの平方和に対応する因子効果の次元（いわゆる平方和の自由度）で割ったもので、1 次元あたりの平方和と呼ぶ。ここでは、表 5.10 をもとに、階層的モデルの選択を行う。表 5.11 に一部の階層的モデルに対する D^2 、Goodman の計算による G^2 [Goodman の記号は X^2]、自由度、及び

$$AICR = D^2 - 2 \times \text{自由度}$$

による情報量規準の値を示す。ここでも D^2 と G^2 はよく類似している。

前述に示すように、あらゆる階層的モデルについて D^2 、 G^2 あるいは情報

量規準を計算するのは実用的でない。とくに、対数線形モデルではモデルごとにセル度数の期待値も反復計算によって推定しなければならない。そこで、Goodman (1971) は重回帰分析における変数選択の手続と利用して二通りの手続を試みる。オ1は、すべての2因子効果を含むモデルから出発して変数増減法〔例えば、奥野他 (1971) 参照〕の考えを適用するもの、オ2は、すべての3因子効果を含むモデルから出発して変数減少法を適用するものである。あるモデル H_1 のもとでの G^2 を $G^2(H_1)$ と書くと、 H_1 の中でのそれに含まれるモデル H_2 の検定は

$$G^2(H_2|H_1) = G^2(H_2) - G^2(H_1)$$

により χ^2 検定を用いる。そこで有意となれば先に進む。これが Goodman の方法である。この方法は χ^2 検定の有意水準の取り方に依存する。

この手続に対し、平方根分解を用いれば、まずオ1に、 G^2 のかわりに D^2 を用いてめとは完全に Goodman と同じ手続をいふことが考えられる。この場合は同一の結論が得られ、モデルごとの反復計算が不要である分だけ簡単になる。オ2に二つのモデルの比較に AICR を用いる案を変更する。これは二つのモデルの間に階層性がない場合でも利用できるが、ここでは階層的モデルの比較であり、上述の $G^2(H_2|H_1)$ を用いるのとあまり違いはない。実際、有意水準を 5% 程度にとったときのオ1の手続と

同じ結論が得られる。一般のデータについて常にオ1とオ2の手順が同一の結論を与えらるわけではないが、一つだけのモデルを選択してそれを絶対的によいモデルと考えるわけではないので、実用上の差はないと思われらる。オ3は手続の出発点となりモデルの決定方針を回帰分析のときのようにゼロまたはすべての因子効果も含むモデルとすることと考えられらる。表5.10を見れば、{ ϕ }-因子効果のつぎにはそれぞれに{23}-因子効果も選択されて、{ ϕ }も出発点にうつらつていこうである。しかし、一般には、Goodman の出発点のとり方で十分であらう。

平方根分解の良さは次の点にもある。いま、2因子効果として{23}だけも含むモデル 1/23/4 に対し次にとりいれらるべき因子効果を選ぶとすらる。そのとき、 D^2 の算出式から D^2 の減少については平方和最大(残りの因子で)、AICRについては次えづとに平方和を比較するところによってそれぞれに選ぶべき因子効果が決定されらる。この場合、AICRに関しては1次元の中からは{13}も、2次元の中からは{14}も選ぶことになる。そしてモデル 13/23/4, と 14/23 の AICR -13.83 と -14.09 を比較し 14/23 のほうがゆづかによいと考えられらる。

この方針も{ ϕ }-因子効果(独立モデル)から出発して徹底的に使用すると図5.1に示るようにな{ ϕ , 23, 14, 234, 13, 123}の順に選択され、そ

のときの AICR の値は -17.66 となる。非階層的モデルとしてはこれが採用されよう。もし、階層的モデルを選択したいならば、 $\{\phi, 23, 14, 234, 23, 123\}$ の生成クラス $\{14/234/123\}$ を考え、このクラスの部分集合について AICR を計算して選ぶことができる。もう一度、増加法または減少法の考えに従って選択してもよい。その結果、 $\{14/123\}$ が選択され、そのときの AICR は -15.53 となる。表 5.11 の中からその他の場合の AICR をさがせよう。表 5.11 の G^2 の欄の *印は Goodman の論文に与えられていないものである。その *印に相当する値、及び上述の手続きで選択されたモデル $\{14/123\}$ でのセル期待値などの対数線形モデルでの分析を付録 3 に示しておく。

上述の算法を一般化することは容易であり、その効率性も期待される。実際の応用にあたっては、一つの規準や算法による最良のモデルを選ぶのではなく、それに近い良さを示す相当数のモデルを抽出して解析者に提示する必要がある。

表5.1 モデル H_g における p_{ijk}^0 の候補 (1)

	p_{ijk}^0	次元	自由度
I	q_{ijk}	$2(r+s+t)-8$	$rst-2(r+s+t)+7$
II	$p_{i++}p_{+jk}$	$st+2r+s+t-7$	$(r-1)(st-1)-(r+s+t-5)$
	$p_{+j+}p_{i+k}$	$rt+r+2s+t-7$	$(s-1)(rt-1)-(r+s+t-5)$
	$p_{++k}p_{ij+}$	$rs+r+s+2t-7$	$(t-1)(rs-1)-(r+s+t-5)$
III	$p_{+j+}p_{i+k}+p_{++k}p_{ij+}-q_{ijk}$	$rs+rt+s+t-6$	$r(s-1)(t-1)-(r+s+t-5)$
	$p_{i++}p_{+jk}+p_{++k}p_{ij+}-q_{ijk}$	$rs+st+r+t-6$	$s(r-1)(t-1)-(r+s+t-5)$
	$p_{i++}p_{+jk}+p_{+j+}p_{i+k}-q_{ijk}$	$st+rt+r+s-6$	$t(r-1)(s-1)-(r+s+t-5)$
IV _a	$p_{i++}p_{+jk}+p_{+j+}p_{i+k}+p_{++k}p_{ij+}-2q_{ijk}$	$rs+st+rt-5$	$(r-1)(s-1)(t-1)-(r+s+t-5)$

表5.2 モデル H_g における p_{ijk}^0 の候補 (2)

	p_{ijk}^0		
III	$p_{ij+}p_{i+k}/p_{i++}$	$rs+rt+s+t-6$	$r(s-1)(t-1)-(r+s+t-5)$
	$p_{ij+}p_{+jk}/p_{+j+}$	$rs+st+r+t-6$	$s(r-1)(t-1)-(r+s+t-5)$
	$p_{+jk}p_{i+k}/p_{++k}$	$st+rt+r+s-6$	$t(r-1)(s-1)-(r+s+t-5)$
IV _m	$m p_{ijk}$	$rs+st+rt-5$	$(r-1)(s-1)(t-1)-(r+s+t-5)$

表5.3 入試データ(5教科)

合格者

年度 科目	'73	'74	'75	'76	+
A	3	3	4	2	12
B	29	28	31	31	119
C	200	196	125	130	651
D	153	162	195	224	734
E	65	61	95	63	284
+	450	450	450	450	1800

不合格者

年度 科目	'73	'74	'75	'76	+
A	12	16	13	5	46
B	104	101	100	86	391
C	371	394	315	287	1367
D	363	364	405	360	1492
E	172	190	146	157	665
+	1022	1065	979	895	3961

表5.4 入試データ(4教科)

合格者

年度 科目	'73	'74	'75	'76	+
A	3	2	4	2	11
B	28	25	32	31	116
C	180	189	130	135	634
D	176	178	207	219	780
E	63	56	77	63	259
+	450	450	450	450	1800

不合格者

年度 科目	'73	'74	'75	'76	+
A	12	17	13	5	47
B	105	104	99	86	394
C	391	401	310	282	1384
D	340	348	393	365	1446
E	174	195	164	157	690
+	1022	1065	979	895	3961

表5.5 平方和分解表 (表5.3のデータ)

a) 平方根分解

FACTOR	DIM	S.S.	DIM	S.S.
0	1	5733.129	1	4.871
1	4	0.0	3	0.0
2	3	0.0	2	0.0
3	1	0.0	0	0.0
12	12	13.533	6	2.443
13	4	5.834	0	0.0
23	3	1.190	0	0.0
123	12	7.314	0	0.0

b) 平方根線形モデル

FACTOR	DIM	S.S.	DIM	S.S.
0	1	4459.875	1	3.541
1	4	1057.721	3	0.0
2	3	2.799	2	0.0
3	1	195.229	0	0.0
12	12	14.807	6	1.332
13	4	23.288	0	0.0
23	3	2.408	0	0.0
123	12	4.873	0	0.0

c) 加法モデル(1)

FACTOR	DIM	S.S.	DIM	S.S.
0	1	5761.	1	19.732
1	4	0.0	3	0.0
2	3	0.0	2	0.0
3	1	0.0	0	0.0
12	12	54.427	6	11.266
13	4	22.701	0	0.0
23	3	5.093	0	0.0
123	12	30.999	0	0.0

d) 加法モデル(2)

FACTOR	DIM	S.S.	DIM	S.S.
0	1	1.	1	0.003
1	4	0.0	3	0.0
2	3	0.0	2	0.0
3	1	0.0	0	0.0
12	12	0.0094475	6	0.002
13	4	0.0039405	0	0.0
23	3	0.0008840	0	0.0
123	12	0.0053809	0	0.0

表5.6 平方和分解表 (表5.4のデータ)

a) 平方根分解

FACTOR	DIM	S.S.	DIM	S.S.
0	1	5732.644	1	1.951
1	4	0.0	3	0.0
2	3	0.0	2	0.0
3	1	0.0	0	0.0
12	12	13.625	6	0.371
13	4	11.210	0	0.0
23	3	1.199	0	0.0
123	12	2.322	0	0.0

b) 平方根線形モデル

FACTOR	DIM	S.S.	DIM	S.S.
0	1	4452.071	1	1.074
1	4	1066.851	3	0.0
2	3	2.572	2	0.0
3	1	200.801	0	0.0
12	12	13.278	6	0.122
13	4	21.325	0	0.0
23	3	2.907	0	0.0
123	12	1.196	0	0.0

c) 加法モデル(1)

FACTOR	DIM	S.S.	DIM	S.S.
0	1	5761.	1	6.926
1	4	0.0	3	0.0
2	3	0.0	2	0.0
3	1	0.0	0	0.0
12	12	54.427	6	1.792
13	4	43.081	0	0.0
23	3	5.093	0	0.0
123	12	8.718	0	0.0

d) 加法モデル(2)

FACTOR	DIM	S.S.	DIM	S.S.
0	1	1.	1	0.001
1	4	0.0	3	0.0
2	3	0.0	2	0.0
3	1	0.0	0	0.0
12	12	0.009475	6	0.000
13	4	0.0074781	0	0.0
23	3	0.0088401	0	0.0
123	12	0.0015132	0	0.0

表5.7 モデルの選択 (表5.3のデータ)

モデル	D^2	χ^2_a	χ^2	G^2	DF	PROB.	AIC
1/2/3	111.67	113.22	113.22	111.94	31	100.0	49.94
1/23	106.89	108.13	107.40	106.87	28	100.0	50.87
2/13	88.26	90.52	88.63	88.21	27	100.0	34.21
3/12	57.40	58.79	56.00	56.84	19	100.0	18.84
12/13	34.04	36.09	33.56	33.10	15	99.6	3.10
12/23	52.64	53.70	50.92	51.77	16	100.0	19.77
13/23	83.49	85.42	83.09	83.13	24	100.0	35.13
12/13/23	29.27	31.00	28.68	28.35	12	99.6	4.35

表5.8 モデルの選択 (表5.4のデータ)

モデル	D^2	χ^2_a	χ^2	G^2	DF	PROB.	AIC
1/2/3	113.61	111.32	111.32	112.52	31	100.0	50.52
1/23	108.80	106.23	105.48	107.45	28	100.0	51.45
2/13	68.65	68.24	67.51	68.14	27	100.0	14.14
3/12	58.98	56.89	54.84	57.43	19	100.0	19.43
12/13	14.09	13.81	12.93	13.05	15	39.2	-16.95
12/23	54.17	51.80	50.14	52.35	16	100.0	20.35
13/23	63.85	63.14	62.53	63.07	24	100.0	15.07
12/13/23	9.29	8.72	8.94	9.01	12	29.2	-14.99

表 5.9 $2 \times 2 \times 2 \times 3$ 分割表P=1, S=1

U	X=	1	2
1		19	57
2		29	63

P=2, S=1

U	X=	1	2
1		29	49
2		27	53

P=1, S=2

U	X=	1	2
1		23	47
2		33	66

P=2, S=2

U	X=	1	2
1		47	55
2		23	50

P=1, S=3

U	X=	1	2
1		24	37
2		42	68

P=2, S=3

U	X=	1	2
1		43	53
2		30	42

表 5.10 平方根分解による平方和

FACTOR	DIM	S.S.	S.S./DIM
0	1	997.3548	
1	1	---	
2	1	---	
3	1	---	
4	2	---	
12	1	0.2267	0.2267
13	1	0.9734	0.9734
14	2	1.5349	0.7675
23	1	5.1372	5.1372
24	2	0.2556	0.1278
34	2	0.0595	0.0298
123	1	0.6570	0.6570
124	2	0.3758	0.1879
134	2	0.0284	0.0142
234	2	1.2594	0.6297
1234	2	0.1375	0.0688
TOTAL	24	1008.	

表 5.11 モデルの検定

MODEL	D ²	G ²	DF	AICR
1/2/3/4	42.73	42.93	18	6.73
すべての2因子効果	9.84	9.85	9	-8.16
すべての3因子効果	0.55	0.74	2	-3.45
1/4/23	22.07	22.35	17	-11.93
4/13/23	18.17	17.99	16	-13.83
14/23	15.91	16.25	15	-14.09
2/3/14	36.55	*	16	4.55
1/234	15.76	*	11	-6.24
4/123	14.62	*	14	-13.38
14/234	9.60	*	9	-8.40
14/123	8.47	*	12	-15.53
234/123	8.31	*	8	-7.69
14/234/123	2.17	*	6	-9.83

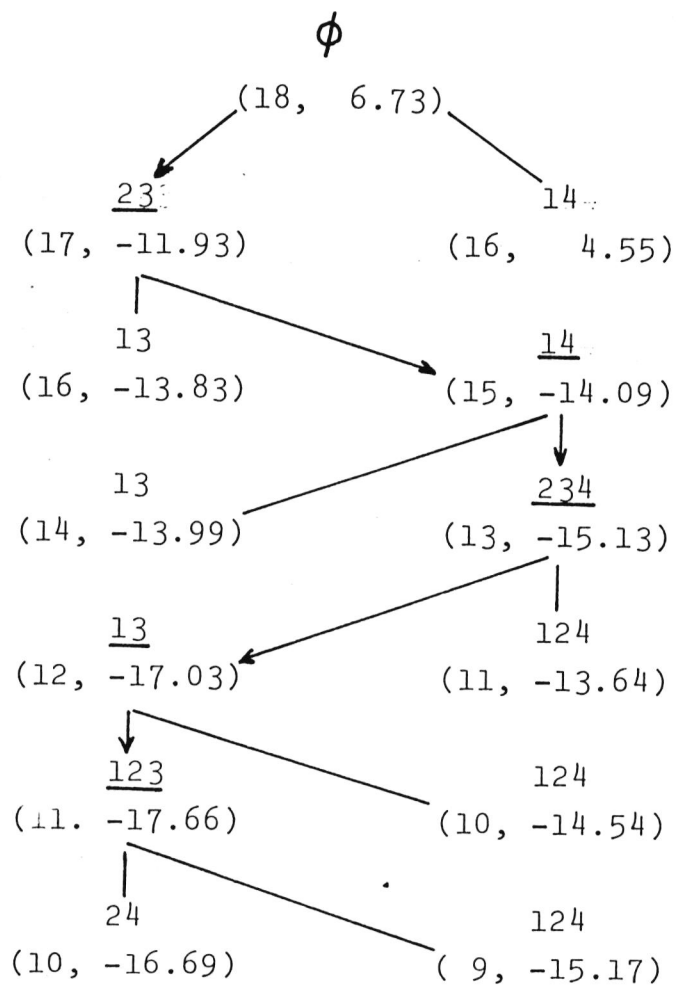


図5.1 非階層的モデルの選択

φ は 独立モデル 1/2/3/4 に相当する。
 括弧の上 が 要因α, アンダラインのありα-
 効果がとり入れられ, 矢印の方向に進む。
 括弧内は それぞれに とり入れら 因子効果に
 よる モデルの 検定の自由度 と AICR と示す。

6. 多重配列の混合型分解

オ2章からオ5章までの議論は m とおりの異なる項目を背景にして得られる m 重配列を対象とするものであった。オ1章で多重配列の出所として述べた二つの項目が同一のものである場合、例えば同一の変数群に対する相関係数行列や分散共分散行列が複数组ある場合など、その多重配列の要素を一次形式と二次形式の混合した形式の係数とみなすことによって、ここで混合型分解と称する多重配列の分解が可能となる。本章では混合型分解と特異値分解との関係、その算法、および簡単な数値例を示す。

6.1 混合型分解

ここでは、寸法 $r \times s \times s$ の3重配列をとりあげ、その要素を

$$A_{ijk} \quad (\forall i \in \bar{r}, \forall j \in \bar{s}, \forall k \in \bar{s}) \quad (6.1)$$

とする。添字 j と k は同一の項目のカテゴリ、あるいは同一の変数群に属する変数に対応するものとする。この3重配列 A について オ2 と オ3 の添字に関する対称性、すなわち

$$A_{ijk} = A_{ikj} \quad (\forall i \in \bar{r}, \forall j, k \in \bar{s}) \quad (6.2)$$

を仮定する。そこで、次の問題を考える。

問題6. 制約条件

$$\left. \begin{aligned} \sum_i u_i^2 &= 1 & (\forall i \in \bar{r}); \\ \sum_j v_j^2 &= 1 & (\forall j \in \bar{s}) \end{aligned} \right\} \quad (6.3)$$

の下で, 混合形式

$$\rho = \sum_{i,j,k} A_{ijk} u_i v_j v_k \quad (6.4)$$

を最大にする u_i ($\forall i \in \bar{r}$) と v_j ($\forall j \in \bar{s}$) を求めること.

第3章において問題1から定理3.1と3.2を導き, さらに算法1を提示してそれに関する諸定理を証明したのと, まったく並行な議論によって問題6から3重配列 A の分解表現をうつことができた. まず, 次の二つの定理を示す. 証明は定理3.1と3.2の場合と同様であるので省略する.

定理6.1. (6.4)の混合形式 ρ について, 次の不等式が成り立つ:

$$\rho^2 \leq \sum A_{ijk}^2. \quad (6.5)$$

等号が成立するのは, A_{ijk} が添字によらず $u_i v_j v_k$ に比例するとき, そのときに限る.

定理6.2. 問題6の解は存在する. また, その解は次の決定方程式の ρ の最大根に対応するベクトル $\underline{u} = (u_i)$, $\underline{v} = (v_j)$ を (6.3) で正規化したものである:

$$\left. \begin{aligned} \sum_{j,k} A_{ijk} v_j v_k &= \rho u_i & (\forall i \in \bar{r}); \\ \sum_{i,j} A_{ijk} u_i v_j &= \rho v_k & (\forall k \in \bar{s}). \end{aligned} \right\} \quad (6.6)$$

(6.6) の決定方程式の数値解法はいろいろ考えられるが、問題1に対する決定方程式 (3.5) を用いることも可能である。この場合、(3.5) を

$$\left. \begin{aligned} \sum_{j,k} A_{ijk} v_j w_k &= \rho u_i \quad (\forall i \in \bar{r}), \\ \sum_{i,k} A_{ijk} u_i w_k &= \rho v_j \quad (\forall j \in \bar{s}), \\ \sum_{i,j} A_{ijk} u_i v_j &= \rho w_k \quad (\forall k \in \bar{s}) \end{aligned} \right\} \quad (6.7)$$

と書く。 A_{ijk} に関する対称性 (6.2) から (6.7) の解では

$$v_j = w_j \quad (\forall j \in \bar{s})$$

となり、(6.5) の解と一致するようになる。そこで、算法1に対して次の算法を定義する。

算法2. 対称性 (6.2) を持つ3重配列 (A_{ijk}) が与えられたとする。

- $\hat{A}_{ijk} = A_{ijk}$ とする。
- $l = 1, 2, \dots$ について以下を反復する。
 - $\hat{A}_{ijk}$ がすべてゼロなら反復を終了する。
 - $\hat{A}_{ijk}$ について決定方程式 (6.6) を解く。その最大根を ρ_l , 対応するベクトルを $\underline{\hat{u}}^l = (\hat{u}_i^l)$, $\underline{\hat{v}}^l = (\hat{v}_j^l)$ とする。ベクトルは (6.3) を満たすように正規化する。
 - $\hat{A}_{ijk}^{l+1} = \hat{A}_{ijk}^l - \hat{u}_i^l \sum_i \hat{A}_{ijk}^l \hat{u}_i^l - \hat{v}_j^l \sum_j \hat{A}_{ijk}^l \hat{v}_j^l - \hat{v}_k^l \sum_k \hat{A}_{ijk}^l \hat{v}_k^l + 2\rho_l \hat{u}_i^l \hat{v}_j^l \hat{v}_k^l.$ (6.8)

注意. (6.8) で計算される A_{ijk}^{l+1} は A_{ijk}^l の 3 因子効果に相当する。

A_{ijk}^l の α -因子効果を $B_{ijk}^l(\alpha)$ と書き, (3.29) にならって, 次のように定義できよう:

$$\left. \begin{aligned} B_{ijk}^l(\phi) &= \rho_l \bar{u}_i^l \bar{v}_j^l \bar{v}_k^l; \\ B_{ijk}^l(12) &= \bar{v}_k^l \sum_k A_{ijk}^l \bar{v}_k^l - \rho_l \bar{u}_i^l \bar{v}_j^l \bar{v}_k^l; \\ B_{ijk}^l(13) &= \bar{v}_j^l \sum_j A_{ijk}^l \bar{v}_j^l - \rho_l \bar{u}_i^l \bar{v}_j^l \bar{v}_k^l; \\ B_{ijk}^l(23) &= \bar{u}_i^l \sum_i A_{ijk}^l \bar{u}_i^l - \rho_l \bar{u}_i^l \bar{v}_j^l \bar{v}_k^l. \end{aligned} \right\} \quad (6.9)$$

そのとき

$$A_{ijk}^{l+1} = A_{ijk}^l - B_{ijk}^l(12) - B_{ijk}^l(13) - B_{ijk}^l(23) - B_{ijk}^l(\phi) \quad (6.10)$$

と書け, 3 因子効果に相当する形になっている. 1 因子効果に相当する項が消えるのも算法 1 と同じで, (6.9) でその分がキャンセルしているといえる. ここで,

$$B_{ijk}^l(12) = B_{ikj}^l(13) \quad (6.11)$$

となるが, $B_{ijk}^l(12)$ や $B_{ijk}^l(13)$ それぞれでは j と k についての対称性はない. 第 2 と第 3 指標とは同等のものでありながら, $\{12\}$ -因子効果と $\{13\}$ -因子効果を区別する本質的な理由はなからうが, 特異値分解との対応のために一応分離しておく. (6.11) から $\{12\}$ -と $\{13\}$ -因子効果のそれぞれの平方和が一致することは明らかである.

算法 2 は 対称性 (6.2) を持つ 3 重配列に対し 算法 1 と同じことであるが

ら、第3章における定理3.3以下の定理や補題に述べられていることが、わずかな字句の修正で算法2にも適用される。おもな性質だけを示しておこう。

定理6.3. 算法2は $\min(r, s)$ 回の反復で終了する。すなわち

$$A_{ijk}^{(\min(r,s)+1)} = 0 \quad (\forall i \in \bar{r}, \forall j, k \in \bar{s}) \quad (6.12)$$

となる。

補題3.1に対して、例えば

$$\sum_i^l B_{ijk}^{(12)} \cdot \tilde{u}_i = \sum_j^l B_{ijk}^{(12)} \cdot \tilde{v}_j = 0$$

となるように、 $B_{ijk}^{(\alpha)}$ について $1 \in \alpha$ なら $\tilde{u}_i$ 、 $2 \in \alpha$ なら $\tilde{v}_j$ 、 $3 \in \alpha$ なら $\tilde{v}_k$ との積和がゼロになることがわかる。補題3.2の(3.15)に対応する次の性質は重要である。

定理6.4. 算法2で得られるベクトル $\{\tilde{u}_1, \tilde{u}_2, \dots\}$, $\{\tilde{v}_1, \tilde{v}_2, \dots\}$

はそれぞれ直交系となる。すなわち、

$$\sum_i^l \tilde{u}_i \tilde{u}_i^{l'} = 0, \quad \sum_j^l \tilde{v}_j \tilde{v}_j^{l'} = 0 \quad (l \neq l'). \quad (6.13)$$

この定理から次の問題が簡単に処理される。

問題7. r 次元、 s 次元ベクトル空間 $\Omega^{(1)}$ と $\Omega^{(2)}$ でそれぞれ正規直交する $l-1$ 個ずつのベクトル $\{\tilde{u}_1, \dots, \tilde{u}_{l-1}\}$ と $\{\tilde{v}_1, \dots, \tilde{v}_{l-1}\}$ が与えられるとき、条件

$$\sum_i^l \tilde{u}_i \tilde{u}_i^{l'} = \sum_j^l \tilde{v}_j \tilde{v}_j^{l'} = 0, \quad l' = 1, 2, \dots, l-1, \quad (6.14)$$

及び、条件 (6.3) のもとで、混合形式

$$\rho = \sum_{i,j,k} A_{ijk} u_i v_j v_k$$

を最大にするベクトル $\underline{u} = (u_i)$ と $\underline{v} = (v_j)$ を求めると.

定理 3.5 に対応して次の定理を得る:

定理 6.5. 問題 7 の解は

$$\overset{\ell}{A}_{ijk} = \sum_{i',j',k'} \overset{\ell}{V}_{ii'}^{(1)} \overset{\ell}{V}_{jj'}^{(2)} \overset{\ell}{V}_{kk'}^{(2)} A_{i'j'k'} \quad (6.15)$$

で与えられる $\overset{\ell}{A}_{ijk}$ について問題 6 を解くこと、すなわち決定方程式 (6.6) を解くことによって与えられる. ここで,

$$\left. \begin{aligned} \overset{\ell}{V}_{ii'}^{(1)} &= \delta_{ii'} - \sum_{\ell'=1}^{\ell-1} \overset{\ell'}{u}_i \overset{\ell'}{u}_{i'} \quad (\forall i, i' \in \bar{r}), \\ \overset{\ell}{V}_{jj'}^{(2)} &= \delta_{jj'} - \sum_{\ell'=1}^{\ell-1} \overset{\ell'}{v}_j \overset{\ell'}{v}_{j'} \quad (\forall j, j' \in \bar{s}) \end{aligned} \right\} \quad (6.16)$$

で与えられる.

注意 (6.15) の $\overset{\ell}{A}_{ijk}$ は次の漸化式でも求められる:

$$\begin{aligned} \overset{1}{A}_{ijk} &= A_{ijk}; \\ \overset{\ell'+1}{A}_{ijk} &= \overset{\ell'}{A}_{ijk} - \overset{\ell'}{B}_{ijk}(12) - \overset{\ell'}{B}_{ijk}(13) - \overset{\ell'}{B}_{ijk}(23) - \overset{\ell'}{B}_{ijk}(11) - \overset{\ell'}{B}_{ijk}(22) \\ &\quad - \overset{\ell'}{B}_{ijk}(33) - \overset{\ell'}{B}_{ijk}(\phi), \quad \ell' = 1, 2, \dots, \ell-1. \end{aligned} \quad (6.17)$$

(6.17) には 1 因子効果に相当する $\overset{\ell'}{B}_{ijk}(1)$, $\overset{\ell'}{B}_{ijk}(2)$, $\overset{\ell'}{B}_{ijk}(3)$ の項が入っているが (6.10) と (6.17) の相違である. 問題 7 で与えられるベクトルはかならずしも決定方程式 (6.6) を満足していない. したがって α -因子

効果は, $\rho_{\ell'} \in$

$$\text{と} \quad \rho_{\ell'} = \sum \overset{\ell'}{A}_{ijk} \overset{\ell'}{u}_i \overset{\ell'}{v}_j \overset{\ell'}{v}_k \quad (6.18)$$

$$\overset{\ell'}{B}_{ijk}(\phi) = \rho_{\ell'} \overset{\ell'}{u}_i \overset{\ell'}{v}_j \overset{\ell'}{v}_k, \quad (6.19)$$

$$\left. \begin{aligned} \overset{\ell'}{B}_{ijk}(1) &= \overset{\ell'}{v}_j \overset{\ell'}{v}_k \sum_{i,k} \overset{\ell'}{A}_{ijk} \overset{\ell'}{v}_j \overset{\ell'}{v}_k - \rho \overset{\ell'}{u}_i \overset{\ell'}{v}_j \overset{\ell'}{v}_k, \\ \overset{\ell'}{B}_{ijk}(2) &= \overset{\ell'}{u}_i \overset{\ell'}{v}_k \sum_{i,k} \overset{\ell'}{A}_{ijk} \overset{\ell'}{u}_i \overset{\ell'}{v}_k - \rho \overset{\ell'}{u}_i \overset{\ell'}{v}_j \overset{\ell'}{v}_k, \\ \overset{\ell'}{B}_{ijk}(3) &= \overset{\ell'}{u}_i \overset{\ell'}{v}_j \sum_{i,j} \overset{\ell'}{A}_{ijk} \overset{\ell'}{u}_i \overset{\ell'}{v}_j - \rho \overset{\ell'}{u}_i \overset{\ell'}{v}_j \overset{\ell'}{v}_k, \end{aligned} \right\} \quad (6.20)$$

及び

$$\left. \begin{aligned} \overset{\ell'}{B}_{ijk}(12) &= \overset{\ell'}{v}_k \sum_k \overset{\ell'}{A}_{ijk} \overset{\ell'}{v}_k - \overset{\ell'}{B}_{ijk}(1) - \overset{\ell'}{B}_{ijk}(2) - \overset{\ell'}{B}_{ijk}(\phi), \\ \overset{\ell'}{B}_{ijk}(13) &= \overset{\ell'}{v}_j \sum_j \overset{\ell'}{A}_{ijk} \overset{\ell'}{v}_j - \overset{\ell'}{B}_{ijk}(1) - \overset{\ell'}{B}_{ijk}(3) - \overset{\ell'}{B}_{ijk}(\phi), \\ \overset{\ell'}{B}_{ijk}(23) &= \overset{\ell'}{u}_i \sum_i \overset{\ell'}{A}_{ijk} \overset{\ell'}{u}_i - \overset{\ell'}{B}_{ijk}(2) - \overset{\ell'}{B}_{ijk}(3) - \overset{\ell'}{B}_{ijk}(\phi) \end{aligned} \right\} \quad (6.21)$$

で求められる. (6.9) と $\overset{\ell'}{B}_{ijk}(\phi)$ にだけ変らない.

以上のほか, 算法1に対する定理3.6, 3.7, 及び3.8, さらに係数3.1, 定理3.9が, 算法2に適用される.

定理6.6. 算法2で得られる $\overset{\ell}{A}_{ijk}$ に対する (6.20) の1因子効果はゼロである.

定理6.7. 対称性 (6.2) を持つ $r \times s \times s$ の3重配列 A と 算法2 により, 次のように分解できる. これを混合型分解と呼ぶ.

$$\overset{\ell}{A}_{ijk} = \overset{\ell}{B}_{ijk}(\phi) + \overset{\ell}{B}_{ijk}(12) + \overset{\ell}{B}_{ijk}(13) + \overset{\ell}{B}_{ijk}(23) + \overset{\ell+1}{A}_{ijk}, \quad (6.22)$$

$\ell = 1, 2, \dots, \min(r, s).$

そこで, $\hat{A}_{ijk} = A_{ijk}$ とする. $\hat{B}_{ijk}(\alpha)$ は (6.9) で与えられる.

定理 6.8. 前定理の分解表現 (6.22) に対し, 平方和分解

$$\begin{aligned} \sum (\hat{A}_{ijk})^2 &= \rho_\ell^2 + \sum (\hat{B}_{ijk}(12))^2 + \sum (\hat{B}_{ijk}(13))^2 \\ &\quad + \sum (\hat{B}_{ijk}(23))^2 + \sum (\hat{A}_{ijk}^{\ell+1})^2, \\ \ell &= 1, 2, \dots, \min(r, s) \end{aligned} \quad (6.23)$$

が得られる.

注意. $\hat{A}_{ijk}$ の α -因子効果 $\hat{A}_{ijk}$ が互いに直交するから. また,

$$\sum_{i,j,k} (\hat{B}_{ijk}(12))^2 = \sum_{i,j} \left(\sum_k \hat{A}_{ijk} v_k - \rho_\ell u_i v_j \right)^2$$

と変形することもできる. $\hat{B}_{ijk}(13)$ と $\hat{B}_{ijk}(23)$ についても同様である.

系 6.1. 次の不等式が成り立つ:

$$\rho_\ell^2 \leq \sum_{i,j,k} (\hat{A}_{ijk})^2. \quad (6.24)$$

等号が成り立つのは $\hat{A}_{ijk}$ が $u_i v_j v_k$ に比例する形のとて, そのときに限る.

平方和の自由度, あるいは $\hat{A}_{ijk}$ に対する α 因子効果 $\hat{B}_{ijk}(\alpha)$ における数学的に独立なパラメータの個数については, (6.2) の対称性から, 次のように考えられる. A_{ijk} 全体では (6.2) の対称性から

$$r s (s+1) / 2 \text{ 個}$$

が独立である. 以下, $\rho, u_i v_j v_k$ に対しては条件 (6.3) から

$$1 + r - 1 + s - 1 = r + s - 1,$$

以下, $\dot{B}_{ijk}(12)$ と $\dot{B}_{ijk}(13)$ は (6.11) から $-\bar{\sigma}$ のみと見做せばよく,
独立なものは

$$(r-1)(s-1) \text{ 個},$$

$\dot{B}_{ijk}(23)$ は j, k についての対称性もあり

$$(s-1)s/2 \text{ 個}$$

が独立である. $\dot{B}_{ijk}(123)$, すなわち $\dot{A}_{ijk}$ については

$$(r-1)(s-1)s/2 \text{ 個}$$

が独立である. ここで, 恒等式

$$\frac{rs(s+1)}{2} = (r+s-1) + (r-1)(s-1) + \frac{s(s-1)}{2} + \frac{(r-1)(s-1)s}{2} \quad (6.25)$$

に注意しておこう. 以上の考えを $\dot{A}_{ijk}$ に進めて次の結果を得る.

‘各平方和 p_l^2 , $\{\sum \dot{B}_{ijk}^l(12)^2 + \sum \dot{B}_{ijk}^l(13)^2\}$, $\sum \dot{B}_{ijk}^l(23)^2$,
 $\sum \dot{B}_{ijk}^l(123)^2$ に対し, その自由度を $r+s-2l+1$, $(r-l)(s-l)$,
 $(s-l)(s-l+1)/2$, $(r-l)(s-l)(s-l+1)/2$ と考えられること
ができる’.

6.2 数値例

本節では、混合型分解の簡単な数値例を示す。前節で述べたように、多重配列の混合型分解は、算法2を用いて行われる。決定方程式(6.6)を解く反復解法は何とおりが考えられ、また、オ3章で述べた算法1の決定方程式を解くことによって可能である。実際に二・三の数値算法を試みれば、ここではその詳細は省略する。実用上、特異値分解のプログラム(5.5節参照)を用いれば十分である。

表6.1に $5 \times 5 \times 4$ の3重配列のデータを示す。このデータは入試の5教科についての年度ごとの相関係数行列(5×5)を4年分並べたものである。対称性を表示するために、6.1節の記号と異ってオ1とオ2項目の添字が一致するデータになっている。このデータでは、教科の一次結合でその分散(標準化されたデータでの)を年度ごとに求め、年度についての重み付きの分散和を最大にするようにその一次結合の係数と年度の重みを同時に求める、というのが混合形式9の最大化問題(問題6)の意味である。ある意味で主成分分析[例えば、奥野他(1971)参照]の拡張になっている。

表6.2に、混合型分解も特異値分解とに行なったときの平方和の分解の結果を示す。実際は、4段の分解が可能であるがオ2段まで

を示した。なお、 $\rho_1 = 4.61$, $\rho_2 = 0.22$, $\rho_3 = 0.12$, $\rho_4 = 0.06$ であった。 ρ_1 に対するベクトルは 教科に対して $(.32, .42, .48, .48, .51)$, 年度に対して $(.51, .52, .45, .51)$ である。教科に対する上記のベクトルは主成分分析での第1固有ベクトルに相当し、年度ごとの相関行列の第1固有ベクトルの一種の平均になっている。年度に対するベクトルを見ると第3年度だけがその重みがいさく、その年度の特異性を示している。

表6.2では121の効果、すなわち教科間の交互作用に相当する平方和が大きくなっていることが重要である。このデータでは第2段以後の分解は決して意味を持たない。主成分分析の観点からは、教科間の交互作用の項は、相関行列から第1主成分の影響を除いたものであり、したがって、その交互作用項をさらに分解して第2主成分に相当するものを抽出してゆくことが意味があろう。このような分析を我が国における大企業の多年にわたる財務指標間相関行列についても試みているが、年度間の変動はそれほど大きくはなく、この例と同じような事情にある。これ以上の分析は主成分分析と同じようになるのでその詳細は省略しよう。

表 6.1 4年度分相関係数行列

Z=1

Y	X= 1	2	3	4	5
1	1.0	0.230	0.283	0.170	0.218
2	0.230	1.0	0.367	0.248	0.396
3	0.283	0.367	1.0	0.409	0.481
4	0.170	0.248	0.409	1.0	0.532
5	0.218	0.396	0.481	0.532	1.0

Z=2

Y	X= 1	2	3	4	5
1	1.0	0.290	0.324	0.195	0.240
2	0.290	1.0	0.301	0.275	0.369
3	0.324	0.301	1.0	0.441	0.374
4	0.195	0.275	0.441	1.0	0.579
5	0.240	0.369	0.374	0.579	1.0

Z=3

Y	X= 1	2	3	4	5
1	1.0	0.162	0.176	0.154	0.196
2	0.162	1.0	0.309	0.203	0.336
3	0.176	0.309	1.0	0.334	0.355
4	0.154	0.203	0.334	1.0	0.421
5	0.196	0.336	0.355	0.421	1.0

Z=4

Y	X= 1	2	3	4	5
1	1.0	0.228	0.246	0.162	0.218
2	0.228	1.0	0.352	0.343	0.412
3	0.246	0.352	1.0	0.425	0.377
4	0.162	0.343	0.425	1.0	0.552
5	0.218	0.412	0.377	0.552	1.0

XとYはともに教科，Zは年度であらわす。

表6.2 平方和の分解表

FACTOR	DIM	S.S.	DIM	S.S.
0	1	21.266	1	0.050
1	4	0.0	3	0.0
2	4	0.0	3	0.0
3	3	0.0	2	0.0
12	16	7.623	9	0.029
13	12	0.012	6	0.003
23	12	0.012	6	0.003
123	48	0.105	18	0.021

7. あとがき

多重分割表データの解析に応用することを主目的として、多重配列構造を持つ母数空間の構造を追求し、多重配列の特異値分解を提示してその性質を調べてきた。多重分割表への応用としては、平方根分解の結果を分割表におけるモデル選択に有効に使用できるところを示し、また、多重分割表における数量化手法と加法モデルとの関係を明らかにした。行列における特異値分解が、しばしば行列の固有値問題とひとしく、多変量解析の重要な理論的基礎となっているように、今後、多重配列構造を持つデータの解析とその理論に、多重配列の特異値分解が役立つことが期待される。例えば、1.3節で触れた Tucker による 3 相因子分析と Lazarsfeld による潜在構造分析は多重配列形式のデータの異なる種の分解を目的とするものである。補題 2.3 のおきの注意や (3.35) 式で示したように、多重配列 A はベクトル空間 $\Omega^{(i)}$ ($\forall i \in \bar{m}$) に適当な基底 $\{\tilde{u}_{k_i}^{(i)}, \forall k_i \in \bar{r}_i\}$ をとって

$$A_{\lambda_1 \lambda_2 \dots \lambda_m} = \sum_{k_1} \dots \sum_{k_m} \rho_{k_1 \dots k_m} \tilde{u}_{\lambda_1}^{k_1 (1)} \dots \tilde{u}_{\lambda_m}^{k_m (m)} \quad (7.1)$$

と表現することができる。この式の右辺は、 A のセル総数 n に等しいだけの項があり、その n 項のうち特殊なものだけを取り出して配列 A が分解されるというようなモデルは、一般に A のセル数 n が大きいとき

には、有利期待できないであろう。潜在構造分析のモデルは、 A を多重分割表のセル確率として

$$A_{\lambda_1 \dots \lambda_m} = \sum_k \rho_k u_{\lambda_1}^{k(1)} \dots u_{\lambda_m}^{k(m)}$$

と書けるという形式になっている [Goodman (1974) を参照]。このモデルでは k も一つの潜在的なグループに対応するものとして、グループ内では、セル確率が

$$u_{\lambda_1}^{k(1)} u_{\lambda_2}^{k(2)} \dots u_{\lambda_m}^{k(m)}$$

という形で表わされる、すなわち、グループごとには項目間が相互に独立であると仮定している。さらにグループの数も不明とし、これも推定の対象とする。現実にもそのようなモデルが意味があるかは別にしても、推定されるベクトル $\tilde{u}^{(i)}$ ($i \in \bar{m}$) に対する、本論での定義による因子効果などの意味も考察してみる必要がある。

また、主成分分析では、(7.1)での $k_1, \dots, k_m$ の動く範囲をできるだけ小さく、かつ、添字によらず一定にしておく、何らかの手段でベクトル $\tilde{u}^{(i)}$ を推定しようとする。Tuckerの方法では、判別指標ごとに何らかの形で単純な固有値問題としてベクトル $\tilde{u}^{(i)}$ が求まるように持つていく。これに対し、最小二乗的に

$$\sum_{\lambda_1, \dots, \lambda_m} \left(A_{\lambda_1 \dots \lambda_m} - \sum_{k_1, \dots, k_m} \rho_{k_1 \dots k_m} u_{\lambda_1}^{k_1(1)} \dots u_{\lambda_m}^{k_m(m)} \right)^2 \rightarrow \min.$$

を規準として，特異値分解のときに似て決定方程式を導いて解くという点とも不可能ではない．以上のように，潜在構造分析や主成分分析などと本論での特異値の実質的關係の追求は今後の課題の一つであろう．

さて，Riemann空間の構造を持つ母数空間の研究も含めて，本論では母数空間自体の構造に焦点を絞って置く．しかし，母数空間は統計的推測問題と一体の關係にあり，とくに，多重分割表では標本抽出過程の確率モデルと切り離して考えられけにはいかない．多重分割表の多くの研究は，多項分布モデル，ときにはポアソン分布モデルも当然のものととして用いる．それに対してあまりに複雑な確率モデルを考へても多くの成果は期待できないが，これほどの各種の手法の異なるモデルの下での信頼性の研究の観点からも，種々の確率モデルも考へる必要があらう．また，多重分割表の構造的なところから対数線形モデルや加法モデルは，分散分析モデルでいう母数モデル (ANOVA Model I) の考へにとづいていう．Fienberg (1975) も指摘しているように，多重分割表への変量モデルなどの導入も重要な課題であらう．筆者は化学工学における粘土の混合問題 [吉澤・新藤 (1977)] と

多重分割表との関連からこの種の課題の研究もはじめている。

なお、特異値分解は数値解析における多変数関数の近似問題とも関係がのろうと思われることなど、残された興味ある問題も少なくない。特異値分解の算法の数値解析的面からの報告も、プログラミング上の工夫について少し述べたばかりにとどめた。

解かれる問題より、あらたに生じる問題のほうが多いのは学問の世界の常である。これまでに得られた成果の上に、それぞれの課題を今後とも一層追求してゆきたいと考えている。

謝 辞

本研究にあたっては、更に多くの方々からの御指導や助言、
あるいは直接間接の協力を与えらるゝに、

東京大学名誉教授、現電気通信大学教授・恩師森口繁一先生には、在学中に始まり、先生の助手として、また東京大学データ処理センター(現教育用計算機センター)在職中の上長として、現在に至るまで絶えず御指導を頂いてきた。そもそも母数空間の研究は学部卒業研究として森口先生の示唆を頂いたのが始まりで今日まで続いているものである。統計学以外にも数値解析、あるいは計算機科学の面についても先生の御指導と忘れらるゝことはできない。

東京大学教授奥野忠一先生には、多変量解析の研究を通じて、また、本研究をまとめるにあたって、ひとかたならぬ御指導を受けらるゝ。

また、山梨大学工学部計算機科学科重永実、田中正次、赤尾洋二、野崎昭弘各教授をはじめ、藤本洋子助手、井内総枝官、大学院生志村健一君、その他全教職員の方々には、日頃、直接的には研究に対する貴重な意見、あるいは協同研究

者としての討論を通じて、間接的には研究に適した環境の提供によって、この5年間の気ままな研究を許して頂いた。

さらに、東京大学教授伊理正夫先生をはじめ、審査にあられた先生方からは、予備審査の段階より貴重な御意見を頂いた。東京大学工学部計数工学科教務室の方々には在学中より何かと御世話も頂き、今回もまた事務的な面倒をみて頂いた。

ここに記して、深甚の謝意を表する次第である。

参考文献

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. 2nd. International Symposium on Information Theory, B.N. Petrov and F. Csaki eds., Akademiai Kiado Budapest, 267-281.
- Bahadur, R.R. (1961). A representation of the joint distribution of responses to N dichotomous items. In H. Solomon ed., Studies in Item Analysis and Prediction. Stanford University Press.
- Bartlett, M.S. (1935). Contingency table interactions. J. Roy. Stat. Soc. Supplement, 2, 248-252.
- Bhattacharya, A. (1946). On a measure of divergences between two multinomial populations. Sankhya, 7, 401-406.
- Birch, M.W. (1963). Maximum likelihood in three-way contingency tables. J. Roy. Stat. Soc., B, 25, 220-233.
- Birch, M.W. (1964). The detection of partial association, I: The 2×2 case. J. Roy. Stat. Soc., B, 26, 313-324.
- Birch, M.W. (1965). The detection of partial association, II: The general case. J. Roy. Stat. Soc., B, 27, 111-124.
- Bishop, Y.M.M. (1967). Multidimensional contingency tables: cell estimates. Unpublished Doctoral Dissertation, Dept. of Stat. Harvard University, Cambridge, Massachusetts.
- Bishop, Y.M.M. (1969). Full contingency tables, logits, and split contingency tables. Biometrics, 25, 383-400.
- Bishop, Y.M.M. (1971). Effects of collapsing multidimensional contingency tables. Biometrics, 27, 545-562.
- Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975). Discrete Multivariate Analysis: Theory and Practice. M.I.T. Press, Cambridge, Massachusetts.
- Cox, D.R. (1972). The analysis of multivariate binary data. J. Roy. Stat. Soc., C, 21, 113-120.
- Cramer, H. (1946). Mathematical Methods of Statistics. Princeton University Press, 424.
- Darroch, J.N. (1962). Interactions in multi-factor contingency tables. J. Roy. Stat. Soc., B, 24, 251-263.
- Darroch, J.N. (1974). Multiplicative and additive interaction in contingency tables. Biomtrika, 61, 207-214.
- Fienberg, S.E. (1975). Comment on "The observational study—A review. J. Amer. Stat. Assoc., 70, 521-523.
- Fisher, R.A. (1922). On the interpretation of χ^2 from contingency tables and the calculation of p . J. Roy. Stat. Soc., 85, 87-94.
- Fisher, R.A. (1924). The conditions under which χ^2 measures the discrepancy between observation and hypothesis. J. Roy. Stat. Soc., 87, 442-450.
- Forsythe, G.E. and Moler, C. (1967). Computer Solution of Linear Algebraic Equations. Prentice Hall, Englewood Cliffs, N.J.
- Freeman, M.F. and Tukey, J.W. (1950). Transformations related to the angular and the square root. Ann. Math. Statist., 21, 607-611.
- Good, I.J. (1963). Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables. Ann. Math. Statist., 34, 911-934.
- Goodman, L.A. (1964). Interactions in multidimensional contingency tables. Ann. Math. Statist., 35, 632-646.
- Goodman, L.A. (1969). On partitioning of χ^2 and detecting partial association in three-way contingency tables. J. Roy. Stat. Soc., B, 31, 486-498.

- Goodman, L.A.(1970). The multivariate analysis of qualitative data: J. Roy. Stat. Soc., 65, 226-256.
- Goodman, L.A.(1971a). The analysis of multidimensional contingency tables: stepwise procedures and direct estimation methos for building models for multiple classifications. Technometrics, 13, 33-61.
- Goodman, L.A.(1971b). Partitioning of χ^2 , analysis of marginal contingency tables, and estimation of expected frequencies in multidimensional contingency tables. J. Amer. Statist. Assoc., 66, 339-344.
- Goodman, L.A.(1974). Exploratory latent structure analysis using identifiable and unidentifiable models. Biometrika, 61, 215-231.
- Grizzle, J.E. Starmer, C.F. and Koch, G.G.(1969). Analysis of categorical data by linear models. Biometrics, 25, 489-504.
- Haberman, S.J.(1974). The Analysis of Frequency Data. University of Chicago Press, Chicago, Il.
- Halmos, P.R.(1958). Finite Dimensional Vector Spaces. 2nd. ed., Van Nostrand, Princeton, N.J.
- Hellinger, E.(1909). Neue Begrundung der Theorie quadratischer Formen von unendichvielen Veranenderlichen. J. Reine Angew. Math., 136, 210-271.
- Hotelling, H.(1936). Relations between two sets of variables. Biometrika, 28, 321-377.
- 岩坪秀一(1971). 多次元質的データ構造の解析—重相関係数による数量化—
電子統計総合研究所報告 35. 7.
- 岩坪秀一(1974). 3-way 離散データも分類する二つの技法—相関係数による数量化—
行動計量学, 2, 54-65.
- Kendall, M.G. and Stuart, A.(1961). The Advanced Theory of Statistics, Vol.2. Charles Griffin and Co., London.
- Kettenring, J.R.(1971). Canonical analysis of several sets of variables. Biometrika, 58, 433-451.
- Killion, R.A. and Zahn, D.A.(1976). A bibliography of contingency table literature: 1900 to 1974. Int. Stat. Rev., 44, 71-112.
- Kullback, S.(1959). Information Theory and Statistics. John Wiley and Sons, New York.
- Lancaster, H.O.(1951). Complex contingency tables treated by the partition of χ^2 . J. Roy. Stat. Soc., B, 13, 242-249.
- Lancaster, H.O.(1960). On test of independence in several dimensions. J. Austral. Math. Soc., 1, 241-254.
- Lancaster, H.O.(1969). The Chi-squared Distribution. John Wiley and Sons, New York.
- Lazarsfeld, P.F.(1961). The algebra of dichotomous systems. In H. Solomon ed., Studies in Item Analysis and Prediction. Stanford University Press.
- Lasarsfeld, P.F. and Henry, N.W.(1968). Latent Structure Analysis. Houghton.
- Maung, K.(1942). Measurement of association in a contingency table with special reference to the pigmentation of hair and eye colours of Scottish school children. Ann. Eugen. London, 11, 189-223.
- McKinlay, S.M.(1975). The design and analysis of the observational study—A review. J. Amer. Statist. Assoc., 70, 503-520.
- Neyman, J.(1949). Contribution to the thory of the χ^2 test. Proc. of the First Berkeley Symposium, 239-274.
- 西里静彦(1977). 尺度法とその関連領域: 文献目録に見る最近の動向(1). 行動計量学, 4, 74-95.

- 奥野忠一, 芳賀敏郎, 久米均, 吉澤正 (1971). 多変量解析法. 日科技連出版, 430.
- Pearson, K.(1900a). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably supposed to have arisen from random sampling. Philos. Mag., Series 5, 50, 152-172.
- Pearson, K.(1900b). Mathematical contributions to the theory of evolution. VII: On the correlation of characters not quantitatively measureable. Philos. Trans. Roy. Soc. London, A, 195, 1-47.
- Pearson, K.(1904). Mathematical contributions to the theory of evolution. XIII: On the theory of contingency and its relation to association and normal correlation. Drap. Co. Mem., Biom. Ser. No.1.
- Plackett, R.L.(1969). Multidimensional contingency tables: A survey of models and methods. Bull. Inter. Statist. Inst., Vol.43, Book 1, 133-142.
- Plackett, R.L.(1974). Analysis of Categorical Data. Hafner, New York.
- Rao, C.R.(1961). Asymptotic efficiency and limiting information. Proc. Fourth Berkeley Symposium, 1, 531-546.
- Rao, C.R.(1962). Efficient estimates and optimum inference procedures in large samples(with discussion). J. Roy. Stat. Soc., B,24, 46-72.
- Rao, C.R.(1963). Criteria of estimation in large samples. Sankhya A,25, 189-206.
- Rao, C.R.(1973). Linear Statistical Inference and its Application. 2nd. ed., John Wiley and Sons, New York.
- Roy, S.N. and Kastenbaum, M.A.(1956). On the hypothesis of no interaction in multiway contingency table. Ann. Math. Statist., 27, 749-757.
- Simpson, E.H.(1951). The interpretation of interaction in contingency table. J. Amer. Statist. Assoc., 69, 971-976.
- 竹内啓 (1976). 情報統計量の分布とモデルの適切さの規準. 数理科学, 1976年3月号, 12-18
- Tucker, L.R.(1964). The extension of factor analysis to three dimensional matrices. Contributions to Mathematical Psychology. Holt, Rinehart and Winston, Inc.
- 吉澤正 (1974). 分割表における平方根分解. 第2回日本行動計量学会総会発表論文抄録集, 80-81.
- 吉澤正 (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, 1-11.
- 吉澤正 (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, 32-43.
- 吉澤正, 藤本洋子 (1976). 入学者選抜試験における選択科目の解析. 山梨大学工学部研究報告, 27, 108-115.
- 吉澤正, 新藤久和 (1977). 混合状態の記述とベータ二項分布. 品質, 7, 130-136.

- Yule, G.U.(1900). On the association of attributes in statistics:
With illustrations from the material from The Childhood Society.
Philos. Trans. Roy. Soc. London, A, 194, 257-319.
- Yule, G.U.(1912). On the methods of measuring association between
two attributes. J. Roy. Stat. Soc., 72, 579-642.

附録 1

Memorandum TYH-2

A Geometry of Parameter Space
and its Statistical Interpretation

Tadashi Yoshizawa

November 2, 1971

at Department of Statistics
Harvard University
Cambridge, Massachusetts, U.S.A.

0. Abstracts

The space which is composed of the parameters of a distribution, the parameter space, may be considered as a Riemannian space by introducing an appropriate metric under some conditions. Invariant quantities in a geometry have very important meanings in its application to various fields of science. It is shown that a necessary and sufficient condition for existence of covariance stabilizing transformation is that the Riemann-Christoffel curvature tensor calculated from the metric is zero. Some population spaces with constant Gaussian curvatures which are immersed in higher dimensional Euclidean spaces and the concept of distance in population spaces are discussed with examples. Finally the relations between this geometry and Fisher's information matrix or the other definitions of distance or divergence between two distributions are mentioned.

1. Introduction

In this paper we consider a parameter space as a Riemannian space by introducing a fundamental tensor of the metric and discuss the statistical meanings of various invariant quantities in the Riemannian space. The inverse of a covariance matrix of a real-valued random vector whose asymptotic distribution is a Normal distribution is used as a fundamental tensor of the metric in section 2.

In section 3, the Riemann-Christoffel curvature tensor which is a typical invariant quantity in a Riemannian space is interpreted. It is shown that the condition for existence of a covariance stabilizing transformation is that all components of the Riemann-Christoffel curvature tensor are zeros, and that a condition given by Holland (1971) in two dimensional case is equivalent to our condition.

In section 4, by calculating Gaussian curvatures of some parameter spaces, it is shown that parameter spaces of a multinomial distribution and one dimensional normal distribution have positive and negative constant Gaussian curvatures respectively. Since a Riemannian space with a constant curvature may be interpreted as a fundamental hyperquadric of a higher (by one) dimensional Euclidean space, the coordinates of the Euclidean space in which the Riemannian space is immersed are given.

In section 5, it is shown that a statistic with asymptotically constant variance is formed by a transformation associated with geodesics in the parameter space. Such

transformations are given with respect to the same examples as in section 4. Finally in section 6, the relation between this geometry and Fisher's information matrix is discussed.

Details of tiresome calculations will be given in the appendices.

The idea that a parameter space may be regarded as a Riemannian space by introducing an appropriate metric has been proposed by Rao (1945) and Yoshizawa (1962) using Fisher's information matrix. Rao gave a solution for geodesics in the case of a parameter space composed of the parent mean and standard deviation of a normal distribution in more complicated form than ours and tried to use it for testing in large samples without noticing its asymptotically constant variance. Yoshizawa, at the suggestion of Professor Moriguti, gave some examples of spaces with constant Gaussian curvatures and discussed their statistical meanings.

Recently Holland (1971) considered an asymptotic concept of a covariance stabilizing transformation and gave a necessary and sufficient condition for its existence. The author studied this problem again receiving impetus from Holland's paper.

2. Introduction of a fundamental tensor of the metric

Let x_n and θ be a real valued p -dimensional random vector with coordinates x_n^i and a p -dimensional parameter with coordinates θ^i , respectively. This vector parameter, θ , varies over D , an open, simply connected region of R^p . Finally, assume that $\sqrt{n}(x_n - \theta)$ has an asymptotic multivariate Normal distribution with zero mean vector and a non-singular covariance matrix, i.e.

$$(1) \quad \mathcal{L}[\sqrt{n}(x_n - \theta)] \rightarrow N(0, \Sigma(\theta))$$

where $\Sigma(\theta)$ is positive definite for all $\theta \in D$.

Let θ' be a new p -dimensional coordinate system transformed one to one by

$$(2) \quad \theta' = f(\theta).$$

It is easily seen that $\sqrt{n}(f(x_n) - f(\theta))$ has an asymptotic multivariate Normal distribution with zero mean vector and a non-singular covariance matrix, i.e.

$$(3) \quad \mathcal{L}[\sqrt{n}(f(x_n) - f(\theta))] \rightarrow N(0, \Sigma'(\theta))$$

where

$$(4) \quad \Sigma'(\theta) = \left(\frac{\partial \theta^i}{\partial \theta^k} \right) \Sigma(\theta) \left(\frac{\partial \theta^j}{\partial \theta^k} \right)$$

and $\left(\frac{\partial \theta^i}{\partial \theta^k} \right)$ means the Jacobian matrix of the transformation (2) under conditions that all partial derivatives $\partial \theta^i / \partial \theta^k$ exist and that the Jacobian is not zero (see Holland [1971]).

Let g^{ij} and g'_{ij} denote the elements of $\Sigma(\theta)$ and $\Sigma'(\theta)$ respectively. We may rewrite (4) as

$$(5) \quad g^{ij} = g^{kl} \frac{\partial \theta^i}{\partial \theta^k} \frac{\partial \theta^j}{\partial \theta^l}$$

where Einstein summation convention* is used. We shall use this convention hereafter. From (5) it is seen that g^{ij} is a contravariant tensor of the second order.

Let g_{ij} be components of the inverse matrix $\Sigma(\theta)$, i.e.

$$(6) \quad g^{ij} g_{kj} = \delta_k^i$$

where δ_k^i 's are Kronecker deltas. Since it is seen from (5) and (6) that the law of transformation of g_{ij} to g'_{ij} is

$$(7) \quad g'_{ij} = g_{kl} \frac{\partial \theta^k}{\partial \theta'^i} \frac{\partial \theta^l}{\partial \theta'^j}$$

g_{ij} is a covariant tensor. It is also positive definite.

Therefore we may take formally as the basis of the metric of a parameter space, a space of parameter θ , a real fundamental quadratic form

$$(8) \quad \phi = g_{ij} d\theta^i d\theta^j .$$

The tensor g_{ij} is called the fundamental tensor of the metric. If element of length ds is defined by

$$ds^2 = g_{ij} d\theta^i d\theta^j$$

noticing that g_{ij} is positive definite, from (7) it is seen that ds^2 is invariant under the transformation (2). This definition of ds may be acceptable as an extension of the concept of concentration matrix by Dempster [1969]. When θ

* When the same letter appears in any terms as a subscript and superscript, it is understood that this letter is summed up for all the values, say p , which this letter takes. k and l in (5) are the examples of such letters, called dummy indecies.

is a mean vector of a multivariate Normal distribution with constant Σ this distance is no more than Mahalanobis' generalized distance (Mahalanobis [1936]).

Invariant quantities in a geometry have very important meanings in its application. In the following chapters we will discuss statistical meanings of several geometrical concepts in parameter space as a Riemannian space.

3. Riemann-Christoffel curvature tensor and covariance stabilizing transformation

Holland [1971] defined a covariance stabilizing transformation f as a set of functions (2) which satisfies

$$(10) \quad g^{ij} \frac{\partial \theta^i}{\partial \theta^1} \frac{\partial \theta^j}{\partial \theta^2} = \delta_{\ell}^k.$$

We may rewrite this condition as

$$(11) \quad g_{ij} \frac{\partial \theta^i}{\partial \theta^1} \frac{\partial \theta^j}{\partial \theta^2} = \delta_{\ell}^k$$

from (6).

A space which has δ_{ℓ}^k as the fundamental tensor of the metric is called a Euclidean space. Therefore the condition (10) or (11) for existence of a covariance stabilizing transformation may be replaced by the condition that a Riemannian space be Euclidean. Since it is well known that the latter condition is that the Riemann-Christoffel curvature tensor vanishes (e.g. see Veblen [1933], pp 69-71), we obtain the following theorem.

Theorem 1: The necessary and sufficient condition that a covariance stabilizing transformation exists is that all components of the Riemann-Christoffel curvature tensor are equal to zero, i.e.

$$(12) \quad R_{hijk} = 0,$$

where

$$(13) \quad R_{hijk} = \frac{\partial}{\partial \theta^j} [ik, h] - \frac{\partial}{\partial \theta^k} [ij, h] + \{_{ij}^{\ell}\} [hk, \ell] - \{_{ik}^{\ell}\} [hj, \ell],$$

$$(14) \quad [ij, k] = \frac{1}{2} \left(\frac{\partial g_{ik}}{\partial \theta^j} + \frac{\partial g_{jk}}{\partial \theta^i} - \frac{\partial g_{ij}}{\partial \theta^k} \right),$$

$$(15) \quad \{_{ij}^{\ell}\} = g^{\ell k} [ij, k].$$

$[ij, k]$ and $\{_{ij}^{\ell}\}$ are called Christoffel 3-index symbols of the first and second kinds, respectively.

In the definition (6) we find that the components of Riemann-Christoffel curvature tensor satisfy the following identities:

$$(16) \quad \begin{aligned} R_{hijk} &= -R_{ihjk}, \\ R_{hijk} &= -R_{hikj}, \\ R_{hijk} &= R_{jkhi}, \end{aligned}$$

and

$$(17) \quad R_{hijk} + R_{hjki} + R_{hkij} = 0.$$

The number of independent components of Riemann-Christoffel curvature tensor is at most $p^2(p^2-1)/12$ due to the identities (16) and (17). In the two dimensional case

$$R_{1212} = R_{2121} = -R_{1221} = -R_{2112}$$

and the other components are zero. Therefore we may write the condition (12) in two dimensional case as

$$(18) \quad R_{1212} = 0.$$

The equivalence between (18) and the condition obtained by Holland [1971] will be shown in Appendix A.

4. Gaussian curvature and space with a constant curvature

Gaussian curvature defined as

$$(19) \quad K = \frac{R_{hijk}}{g_{hj}g_{ik} - g_{hk}g_{ij}}$$

has an important role in Riemannian geometry. In particular a p-dimensional Riemannian space with constant Gaussian curvature can be immersed in p+1 dimensional Euclidean space and can be interpreted as a fundamental hyperquadric of the Euclidean space. The hyperquadric is defined by

$$(20) \quad \sum_{\alpha=1}^{p+1} c_{\alpha} (z^{\alpha})^2 = \frac{1}{K},$$

where c_{α} 's are plus or minus one according to the character of the fundamental form and z^{α} 's are a set of solutions of the equations

$$(21) \quad \frac{\partial^2 z^{\alpha}}{\partial \theta^i \partial \theta^j} - \frac{\partial z^{\alpha}}{\partial \theta^h} \{^h_{ij}\} = -K g_{ij} z^{\alpha}, \quad (\alpha=1, \dots, p+1)$$

(See Eisenhart [1926]). The left-hand of the above equation (21) is called covariant differentiation with respect to a tensor g_{ij} and denoted by $z^{\alpha}_{,ij}$

Example 1: A parameter space of means of a p-dimensional Normal distribution with constant covariance matrix Σ .

The fundamental tensor of the metric is the inverse of Σ . It is obvious that this space is Euclidean since the fundamental tensor is constant and it may be diagonalized by simple transformation. Of course, K is zero in this case.

Example 2: The trinomial distribution

Let nT_n have a trinomial distribution. The parameter space, D , is given by:

$$D = \left\{ \begin{pmatrix} \theta^1 \\ \theta^2 \end{pmatrix} : \theta^1 > 0 \text{ and } \theta^1 + \theta^2 < 1 \right\}.$$

Standard theory implies that

$$\mathcal{L}[\sqrt{n}(T_n - \theta)] \rightarrow N(0, \Sigma(\theta))$$

where

$$\Sigma(\theta) = \begin{pmatrix} \theta^1(1-\theta^1) & -\theta^1\theta^2 \\ -\theta^1\theta^2 & \theta^2(1-\theta^2) \end{pmatrix}.$$

Therefore, the fundamental tensor of the space is given as follows:

$$(g_{ij}) = \Sigma(\theta)^{-1} = \begin{pmatrix} \frac{1-\theta^2}{\theta^1(1-\theta^1-\theta^2)} & \frac{1}{1-\theta^1-\theta^2} \\ \frac{1}{1-\theta^1-\theta^2} & \frac{1-\theta^1}{\theta^2(1-\theta^1-\theta^2)} \end{pmatrix}$$

Holland [1971] showed that there exists no covariance stabilizing transformation in this case and Yoshizawa [1962] obtained the Gaussian curvature as

$$K = \frac{1}{4} .$$

This space has a positive constant Gaussian curvature and may be regarded as a sphere in a three dimensional Euclidean space. The coordinates of the Euclidean space are obtained as a set of solutions as follows:

$$(22) \quad z^1 = 2\sqrt{\theta^1}, \quad z^2 = 2\sqrt{\theta^2}, \quad z^3 = 2\sqrt{\theta^3},$$

where $\theta^3 = 1 - \theta^1 - \theta^2$ (See Appendix B). From (20) the equation of the sphere is

$$(23) \quad (z^1)^2 + (z^2)^2 + (z^3)^2 = 4,$$

that is,

$$\theta^1 + \theta^2 + \theta^3 = 1.$$

This fact gives a convenient interpretation of the parameter space of a trinomial distribution.

The results obtained here can be extended to the parameter space of the multinomial distribution (See Appendix B).

Example 3: The Normal distribution $N(\mu, \sigma^2)$

Let θ^1 and θ^2 be the mean μ and variance σ^2 respectively. The fundamental tensor of the metric is given as follows:

$$g_{11} = \frac{1}{\theta^2}, \quad g_{12} = g_{21} = 0, \quad g_{22} = \frac{1}{2(\theta^2)^2}$$

The only one independent components of the Riemann-Christoffel curvature R_{1212} is

$$R_{1212} = - \frac{1}{4(\theta^2)^3},$$

and Gaussian curvature is

$$K = \frac{R_{1212}}{g_{11}g_{22} - g_{12}g_{21}} = - \frac{1}{2}.$$

Therefore no covariance stabilizing transformation exists.

Since this space has a negative constant Gaussian curvature it may be regarded as a hyperbolic surface in a three dimensional Euclidean space. The coordinates of the Euclidean space are given as

$$\begin{aligned} z^1 &= \frac{\theta^1}{\sqrt{\theta^2}} = \frac{\mu}{\sigma}, \\ (24) \quad z^2 &= \frac{(\theta^1)^2 - 2}{2\sqrt{2}\sqrt{\theta^2}} + \frac{\sqrt{\theta^2}}{\sqrt{2}} = \frac{\mu^2 - 2}{2\sqrt{2}\sigma} + \frac{\sigma}{\sqrt{2}}, \\ z^3 &= \frac{(\theta^1)^2 + 2}{2\sqrt{2}\sqrt{\theta^2}} + \frac{\sqrt{\theta^2}}{\sqrt{2}} = \frac{\mu^2 + 2}{2\sqrt{2}\sigma} + \frac{\sigma}{\sqrt{2}}, \end{aligned}$$

solving the equations (21) and comparing the fundamental forms of the Euclidean space and of the Riemannian space.

The equation of the hyperbolic surface is given by

$$(25) \quad (z^1)^2 + (z^2)^2 - (z^3)^2 = -2.$$

5. Distance and geodesics

The distance between two points in a parameter space has a meaning as a kind of measure for the difference or divergence between two distributions. The distance may be naturally defined as the arc of geodesic curve between two points due to the theory of Riemannian geometry. The elements of length ds is defined by (9) and the arc of geodesic curve is given by the solution of the following equations:

$$(26) \quad \frac{d^2 \theta^{\ell}}{ds^2} + \{^{\ell}_{jk}\} \frac{d\theta^j}{ds} \frac{d\theta^k}{ds} = 0.$$

These equations are Euler's equations of the integral

$$s = \int_{t_1}^t \sqrt{g_{ij} \frac{d\theta^i}{dt} \frac{d\theta^j}{dt}} dt$$

where t is a parameter which defines real curve. Along the geodesic we have

$$(27) \quad g_{ij} \frac{d\theta^i}{ds} \frac{d\theta^j}{ds} = 1.$$

If we put new coordinates associated with the geodesic passing through the particular point θ_0 such that

$$(28) \quad \theta^{,i} = \left(\frac{d\theta^i}{ds} \right)_0 s$$

the arc of the geodesic may be expressed as

$$(29) \quad s^2 = (g_{ij})_0 \theta^{,i} \theta^{,j}.$$

Notice that this form may be interpreted as an extension of Mahalanobis' generalized distance. Moreover if we transform θ^i to $\theta^{''i}$ such that $\theta^{''1}$ will express the arc of the geodesic, the fundamental form of $\theta^{''i}$ is reduced to

$$(30) \quad \phi = (d\theta^{''1})^2 + g_{\alpha\beta}'' d\theta^{''\alpha} d\theta^{''\beta} \quad (\alpha, \beta = 2, \dots, p)$$

(See Eisenhart [1926], p. 57). From the form (30) it is seen that if we use the transformation $\theta^{''1}$ we may get a transformation of random variables which obeys asymptotical Normal distribution with unit variance. Therefore

$$ns^2(x_n : \theta_o)$$

will obey Chi-square distribution of one degree of freedom asymptotically. Then we have

Theorem 2: Under the assumption (1) substituting θ_o in θ

$$(31) \quad ns^2(x_n : \theta_o)$$

asymptotically obeys Chi-square distribution of one degree of freedom, where $s(\theta : \theta_o)$ is the geodesic given as the solution (26) using the inverse of the asymptotic variance of $\sqrt{n} x_n$ as a fundamental tensor.

It is known that the equation (27) may reduce to

$$(32) \quad \frac{d^2 z^\alpha}{ds^2} = -kz^\alpha$$

with respect to the coordinates z^α of the Euclidean space in which a Riemannian space with a constant Gaussian curvature is immersed. There are two cases to be considered according to the sign of K .

1°. $K > 0$. In this case we have

$$(33) \quad \sum_{\alpha} c_{\alpha} (z^{\alpha} - z_0^{\alpha})^2 = \frac{4}{K} \sin^2 \frac{\sqrt{K}s}{2},$$

where z_0^{α} is the value of z^{α} at a point θ_0 . From (20) we may rewrite (33) as

$$(34) \quad K \sum_{\alpha} c_{\alpha} z^{\alpha} z_0^{\alpha} = \cos \sqrt{K} s.$$

2°. $K < 0$. Similarly we have

$$(35) \quad \sum_{\alpha} c_{\alpha} (z^{\alpha} - z_0^{\alpha})^2 = -\frac{4}{K} \sinh^2 \frac{\sqrt{-K}s}{2},$$

$$(36) \quad K \sum_{\alpha} c_{\alpha} z^{\alpha} z_0^{\alpha} = \cosh \sqrt{-K} s.$$

Example 4: The trinomial distribution of the example 2.

From the example 2, the arc of geodesic is given by:

$$(37) \quad s = 2 \cos^{-1} \left(\sqrt{\theta^1 \theta_0^1} + \sqrt{\theta^2 \theta_0^2} + \sqrt{\theta^3 \theta_0^3} \right),$$

substituting (22) and $K = 1/4$ in 34. In the case of multinomial distribution we have

$$(38) \quad s = 2 \cos^{-1} \sum_{\alpha} \sqrt{\theta^{\alpha} \theta_0^{\alpha}}$$

The half of s has been introduced as a measure of divergence by Bhattacharyya [1942].

From theorem 2 it is seen that the asymptotic variance of

$$(39) \quad 2 \cos^{-1} \sum_{\alpha} \sqrt{\hat{\theta}^{\alpha} \theta^{\alpha}}$$

is 1, where $\hat{\theta}^{\alpha}$ is the usual maximum likelihood estimator of θ^{α} .

Example 5: The Normal distribution of the example 3

Since the Gaussian curvature K is negative constant $-1/2$, substituting (24) in (33), we obtain the arc of geodesic as follows:

$$(40) \quad s = \sqrt{2} \cosh^{-1} \frac{(\mu - \mu_o)^2 + 2(\sigma^2 + \sigma_o^2)}{4\sigma\sigma_o}.$$

The statistic associated with the geodesic is

$$(41) \quad \sqrt{n} s(\bar{x}_n, S_n^2; \mu_o, \sigma_o^2) = \sqrt{2n} \cosh^{-1} \frac{(\bar{x} - \mu_o)^2 + 2(S_n^2 + \sigma_o^2)}{4\sqrt{S_n^2} \sigma_o}$$

and by the theorem 2 it has asymptotic unit variance, where $\bar{x}_n$ and S_n^2 are the sample mean and variance respectively.

6. Fisher's information

Fisher's information matrix in several parameters are defined by:

$$(42) \quad I^{ij} = E \left(\frac{\partial \log L}{\partial \theta^i} \frac{\partial \log L}{\partial \theta^j} \right)$$

where L is the likelihood function of parameters θ^i (Fisher [1921]). It is easily shown that I^{ij} is a covariant tensor under transformations of parameters. Rao [1945] and

Yoshizawa [1962] used I^{ij} as the fundamental tensor of a parameter space. The inverse of the matrix I^{ij} gives the lower bound of the variance of the estimators θ under some regularity condition by Rao-Cramer inequality. The fundamental tensors of the metric in the examples here are quite the same as Fisher's information matrix and they are rather easily calculated directly from (42) or from the following formula equivalent to (42) in the case that maximum likelihood estimators are used:

$$I^{ij} = -E \left(\frac{\partial^2 \log L}{\partial \theta^i \partial \theta^j} \right).$$

Anyway this geometry of parameter spaces is concerned with asymptotic variance and the limit of this method should be considered in this point.

Finally notice that the various definitions of distance or divergence between two distributions, e.g. Kullback's divergence [1959] and Matusita's [1955], are often locally equivalent to Fisher's information.

References

- Bhattacharyya, A. (1942) "On a measure of divergence between two multinomial populations" *Sankhyā*, Vol. 7, 401-6.
- Dempster, A.P. (1969) Continuous Multivariate Analysis Addison-Wesley, Reading, Massachusetts.
- Eisenhart, L.P. (1926) Riemann Geometry, Princeton University Press, Princeton.
- Eisenhart, L.P. (1940) An Introduction to Differential Geometry with Use of the Tensor Calculus, Princeton University Press, Princeton.
- Fisher, R.A. (1921) "On the mathematical foundations of theoretical statistics", *Phil. Trans. Roy. Soc. A*, 222, 309-368.
- Holland, P.W. (1971) "On the possibility of a multivariate extension of the variance stabilizing transformations", Memorandum NS-143 Harvard University, Cambridge, Massachusetts.
- Kullback, S. (1959) Information Theory and Statistics, John Wiley & Sons, New York.
- Mahalanobis, P.C. (1936) "On the generalized distance in Statistics", *Proc. Nat. Inst. Sc. Ind.*, 2, 49-55.
- Matusita, K. (1955) "Decision rules, based on the distance, for problems of fit, two samples, and estimation" *Ann. Math. Stat.*, Vol. 26, pp. 631-640.
- Rao, C.R. (1945) "Information and the accuracy attainable in the estimation of statistical parameters", *Bulletin of the Calcutta Math. Soc.*, Vol. 37, 81-91.
- Veblen, O. (1933) Invariants of Quadratic Differential Forms, Cambridge University Press, London.
- Yoshizawa, T. (1962) "Information-distances in parameter spaces" (in Japanese). Unpublished graduation thesis at University of Tokyo.

Appendix A: A proof of the equivalence between Holland's theorem and the theorem 1 in this paper.

Holland's condition for existence of a covariance stabilizing transformation in two dimensional case is

$$(A-1) \quad \frac{\partial}{\partial \theta^2} \left[\frac{1}{\det(A)} \left\{ a_{11} \left(\frac{\partial a_{11}}{\partial \theta^2} - \frac{\partial a_{12}}{\partial \theta^1} \right) + a_{21} \left(\frac{\partial a_{21}}{\partial \theta^2} - \frac{\partial a_{22}}{\partial \theta^1} \right) \right\} \right] \\ = \frac{\partial}{\partial \theta^1} \left[\frac{1}{\det(A)} \left\{ a_{12} \left(\frac{\partial a_{11}}{\partial \theta^2} - \frac{\partial a_{12}}{\partial \theta^1} \right) + a_{22} \left(\frac{\partial a_{21}}{\partial \theta^2} - \frac{\partial a_{22}}{\partial \theta^1} \right) \right\} \right],$$

where A is any matrix (2x2) that satisfies

$$(A-2) \quad A'A = \Sigma(\theta)^{-1}$$

Our condition (12) is reduced to

$$(A-3) \quad K = \frac{R_{1212}}{g} = 0$$

in two dimensional case and Gaussian curvature K may be written as follows:

$$(A-4) \quad K = \frac{1}{2\sqrt{g}} \left[\frac{\partial}{\partial \theta^1} \left(\frac{g_{12}}{g_{11}\sqrt{g}} \frac{\partial g_{11}}{\partial \theta^2} - \frac{1}{\sqrt{g}} \frac{\partial g_{22}}{\partial \theta^1} \right) \right. \\ \left. + \frac{\partial}{\partial \theta^2} \left(\frac{2}{\sqrt{g}} \frac{\partial g_{12}}{\partial \theta^1} - \frac{1}{\sqrt{g}} \frac{\partial g_{11}}{\partial \theta^2} - \frac{g_{12}}{g_{11}\sqrt{g}} \frac{\partial g_{11}}{\partial \theta^1} \right) \right],$$

(See Eisenhart (1940), p. 154 Ex. 10) We will prove that (A-4) is equivalent to (A-3) using the form (A-4).

We may assume that $A(\theta)$ is an upper triangular matrix without loosing generality. Then the components of A are expressed as follows:

$$\begin{aligned} a_{11} &= \sqrt{g_{11}} , \\ (A-5) \quad a_{12} &= \frac{g_{12}}{\sqrt{g_{11}}} , \quad a_{21} = 0, \\ a_{22} &= \sqrt{g_{22} - \frac{g_{12}^2}{g_{11}}} = \sqrt{\frac{g}{g_{11}}} . \end{aligned}$$

Partially differentiating the both sides of (A-5), we obtain

$$\begin{aligned} \frac{\partial a_{11}}{\partial \theta} &= \frac{1}{2\sqrt{g_{11}}} \frac{\partial g_{11}}{\partial \theta} , \\ (A-6) \quad \frac{\partial a_{12}}{\partial \theta} &= \frac{1}{\sqrt{g_{11}}} \frac{\partial g_{12}}{\partial \theta} - \frac{1}{2} \frac{g_{12}}{g_{11}\sqrt{g_{11}}} \frac{\partial g_{11}}{\partial \theta} , \\ \frac{\partial a_{22}}{\partial \theta} &= \frac{1}{2} \sqrt{\frac{g_{11}}{g}} \left(\frac{\partial g_{22}}{\partial \theta} - \frac{2g_{12}}{g_{11}} \frac{\partial g_{12}}{\partial \theta} + \left(\frac{g_{12}}{g_{11}} \right)^2 \frac{\partial g_{11}}{\partial \theta} \right) , \end{aligned}$$

where θ may take either θ^1 or θ^2 . Substituting (A-5) and (A-6) in (A-1), and using $\det(A) = \sqrt{g}$ we rewrite (A-1) as

$$\begin{aligned}
& \frac{\partial}{\partial \theta^2} \left[\frac{1}{\sqrt{g}} \sqrt{g_{11}} \left(\frac{1}{2\sqrt{g_{11}}} \frac{\partial g_{11}}{\partial \theta^2} - \frac{1}{\sqrt{g_{11}}} \frac{\partial g_{12}}{\partial \theta^1} + \frac{g_{12}}{2g_{11}\sqrt{g_{11}}} \frac{\partial g_{11}}{\partial \theta^1} \right) \right. \\
&= \frac{\partial}{\partial \theta^1} \left[\frac{1}{\sqrt{g}} \left\{ \frac{g_{12}}{\sqrt{g_{11}}} \left(\frac{1}{2\sqrt{g_{11}}} \frac{\partial g_{11}}{\partial \theta^2} - \frac{1}{\sqrt{g_{11}}} \frac{\partial g_{12}}{\partial \theta^1} + \frac{1}{2} \frac{g_{12}}{g_{11}\sqrt{g_{11}}} \frac{\partial g_{11}}{\partial \theta^1} \right) \right. \right. \\
&\quad \left. \left. - \sqrt{\frac{g}{g_{11}}} \frac{1}{2} \sqrt{\frac{g_{11}}{g}} \left(\frac{\partial g_{22}}{\partial \theta^1} - \frac{2g_{12}}{g_{11}} \frac{\partial g_{12}}{\partial \theta^1} + \left(\frac{g_{12}}{g_{11}} \right)^2 \frac{\partial g_{11}}{\partial \theta^1} \right) \right\} \right]
\end{aligned}$$

After simple algebra we get

$$\begin{aligned}
& - \frac{\partial}{\partial \theta^2} \left[\frac{1}{\sqrt{g}} \left(\frac{\partial g_{12}}{\partial \theta^1} - \frac{\partial g_{11}}{2\partial \theta^2} - \frac{g_{12}}{2g_{11}} \frac{\partial g_{11}}{\partial \theta^1} \right) \right] \\
&= \frac{\partial}{\partial \theta^1} \left[\frac{1}{2\sqrt{g}} \left(\frac{g_{12}}{g_{11}} \frac{\partial g_{11}}{\partial \theta^2} - \frac{\partial g_{22}}{\partial \theta^1} \right) \right] .
\end{aligned}$$

It is easily seen that the above equation is equal to $K=0$.

Appendix B: The multinomial distribution

Let the likelihood of a multinomial distribution be

$$(B-1) \quad L = \frac{n!}{\prod_{i=1}^{p+1} x_i!} \prod_{i=1}^{p+1} \theta_i^{x_i},$$

where θ 's are $p+1$ parameters and x 's are random variables. Notice that we use subscripts for parameters in order not to confuse superscripts and squares, etc. The first p parameters are used for the coordinates of the population space since

$$(B-2) \quad \sum \theta_i = 1$$

1°. Fundamental tensor of the metric.

By calculating Fisher's information matrix, the fundamental tensor of the metric is given as below:

$$(B-3) \quad (g_{ij}) = \begin{pmatrix} \frac{\theta_1 + \theta_{p+1}}{\theta_1 \theta_{p+1}} & \frac{1}{\theta_{p+1}} & \frac{1}{\theta_{p+1}} & \dots & \frac{1}{\theta_{p+1}} \\ \frac{1}{\theta_{p+1}} & \frac{\theta_1 + \theta_{p+1}}{\theta_2 \theta_{p+1}} & \frac{1}{\theta_{p+1}} & \dots & \frac{1}{\theta_{p+1}} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{1}{\theta_{p+1}} & \frac{1}{\theta_{p+1}} & \cdot & \cdot & \frac{\theta_p + \theta_{p+1}}{\theta_p \theta_{p+1}} \end{pmatrix}.$$

The covariance matrix of the usual estimators x_i/n , i.e., the inverse of (g_{ij}) is

$$(B-4) \quad (g^{ij}) = \begin{pmatrix} \theta_1(1-\theta_1) & -\theta_1\theta_2 & \dots & -\theta_1\theta_p \\ -\theta_1\theta_2 & \theta_2(1-\theta_2) & \dots & -\theta_2\theta_p \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ -\theta_p\theta_1 & -\theta_p\theta_2 & \dots & \theta_p(1-\theta_p) \end{pmatrix}.$$

2°. Christoffel's 3 index symbols.

From (B-3) Partial derivatives of g_{ij} are

$$(B-5) \quad \frac{\partial g_{ij}}{\partial \theta_k} = \begin{cases} \frac{1}{\theta_{p+1}^2} - \frac{1}{\theta_i^2}, & i=j=k, \\ \frac{1}{\theta_{p+1}^2}, & \text{otherwise,} \end{cases}$$

From the definition (14) and (15) Christoffel's 3 index symbols of the first kind are

$$(B-6) \quad [ij,k] = \begin{cases} \frac{1}{2} \left(-\frac{1}{\theta_i^2} + \frac{1}{\theta_{p+1}^2} \right), & i=j=k, \\ \frac{1}{2} \frac{1}{\theta_{p+1}^2}, & \text{otherwise,} \end{cases}$$

and Christoffel's 3 index symbols of the second kind are

$$(B-7) \quad \{\frac{\ell}{ij}\} = \begin{cases} \frac{\theta_{\ell}}{2\theta_{p+1}}, & i \neq j, \\ \frac{\theta_{\ell}}{2} \left(\frac{1}{\theta_i^2} + \frac{1}{\theta_{p+1}} \right), & i=j, \neq \ell, \\ \frac{\theta_i}{2\theta_{p+1}} - \frac{1-\theta_i}{2\theta_i}, & i=j=\ell. \end{cases}$$

3°. Riemann-Christoffel curvature tensor

In two dimensional case, i.e. in the case of the trinomial distribution, Riemann-Christoffel curvature tensor is from (13)

$$\begin{aligned} (B-8) \quad R_{1212} &= \frac{\partial}{\partial \theta_1} [22,1] - \frac{\partial}{\partial \theta_2} [21,1] + \{\frac{\ell}{21}\} [12,\ell] - \{\frac{\ell}{22}\} [11,\ell] \\ &= \frac{1}{\theta_3} - \frac{1}{\theta_3} + \frac{\theta_1}{2\theta_3} \frac{1}{2\theta_3^2} + \frac{\theta_2}{2\theta_3} \frac{1}{2\theta_3} \\ &\quad - \frac{\theta_1}{2} \left(\frac{1}{\theta_2} + \frac{1}{\theta_3} \right) \frac{1}{2} \left(-\frac{1}{\theta_1^2} + \frac{1}{\theta_3^2} \right) - \left(\frac{\theta_2}{2\theta_3} - \frac{1-\theta_2}{2\theta_2} \right) \frac{1}{2\theta_3^2} \\ &= \frac{1}{4\theta_1\theta_2\theta_3}. \end{aligned}$$

Therefore

$$(B-9) \quad K = \frac{R_{1212}}{g} = \frac{1}{4}.$$

In p dimensional case it is seen from the identities (16) that R_{hijk} is equal to zero if $h = i$ or $j = k$. Under the condition $h \neq i$ and $j \neq k$, noticing that

$$\frac{\partial}{\partial \theta_j} [ik,h] - \frac{\partial}{\partial \theta_k} [ij,h] = 0$$

R_{hijk} 's are calculated as follows:

$$(B-10) \left\{ \begin{array}{l} \circ h=j, i \neq k \\ R_{hihk} = \frac{1}{4\theta_{p+1}\theta_h} \\ \circ h \neq j, i=k \\ R_{hiji} = R_{ihij} = \frac{1}{4\theta_{p+1}\theta_i} \\ \circ h=j, i=k \\ R_{hihi} = \frac{1}{4} \left\{ \left(\frac{1}{\theta_i} + \frac{1}{\theta_{p+1}} \right) \left(\frac{1}{\theta_h} + \frac{1}{\theta_{p+1}} \right) - \frac{1}{\theta_{p+1}^2} \right\} \end{array} \right.$$

Therefore it is easily seen that Gaussian curvature

$$(B-11) \quad K = \frac{R_{hijk}}{g_{hj}g_{ik} - g_{hk}g_{ij}} = \frac{1}{4}.$$

4°. Coordinates of a Euclidean space in which the population space of multinomial distribution is immersed.

Let z^α denote the coordinates of the Euclidean space. z^α 's are given as a set of solutions of (21), i.e.

$$\frac{\partial^2 z^\alpha}{\partial \theta_i \partial \theta_j} - \sum_{h=1}^p \frac{\partial z^\alpha}{\partial \theta_h} \{ \frac{h}{ij} \} = - \frac{1}{4} g_{ij} z^\alpha.$$

It is easily seen that a set of solutions is given as

$$z^\alpha = c\sqrt{\theta_\alpha}, \quad \alpha=1, \dots, p+1.$$

Appendix C: The normal distribution $N(\mu, \sigma^2)$

Let θ^1 and θ^2 be μ and σ^2 respectively. We shall use θ_1 and θ_2 for θ^1 and θ^2 because of practical convenience in algebra.

1°. Fundamental tensor of the metric

Standard theory implies that

$$(C-1) \quad \Sigma(\theta) = \begin{pmatrix} \theta_2 & 0 \\ 0 & 2\theta_2^2 \end{pmatrix} = (g^{ij}) .$$

Therefore the fundamental tensor of the metric is

$$(C-2) \quad g_{11} = \frac{1}{\theta_2}, \quad g_{12} = g_{21} = 0, \quad g_{22} = \frac{1}{2\theta_2^2} .$$

2°. Christoffel's 3 index symbols.

Since partial derivatives by θ_1 and θ_2 are zeros except that

$$(C-3) \quad \frac{\partial g_{11}}{\partial \theta_2} = -\frac{1}{\theta_2^2}, \quad \text{and} \quad \frac{\partial g_{22}}{\partial \theta_2} = -\frac{1}{\theta_2^3} ,$$

from (14) and (15) Christoffel's 3 index symbols of the first and second kinds are as follows:

First kind:

$$(C-4) \quad \begin{aligned} [11,1] &= 0, \\ [11,2] &= \frac{1}{2\theta_2^2}, \\ [12,1] &= [21,1] = -\frac{1}{2\theta_2^2}, \\ [12,2] &= [21,2] = 0, \\ [21,1] &= 0, \\ [22,2] &= \frac{-1}{2\theta_2^3}, \end{aligned}$$

Second kind:

$$\begin{aligned}
 \{ \frac{1}{11} \} &= 0, \\
 \{ \frac{2}{11} \} &= 1, \\
 \{ \frac{1}{21} \} &= \{ \frac{1}{21} \} = - \frac{1}{2\theta_2}, \\
 \{ \frac{2}{12} \} &= \{ \frac{2}{21} \} = 0, \\
 \{ \frac{1}{22} \} &= 0, \\
 \{ \frac{2}{22} \} &= - \frac{1}{\theta_2}.
 \end{aligned}
 \tag{C-5}$$

3°. Riemann-Christoffel curvature tensor

Substituting (C-4) and (C-5) in (13), we obtain

$$\begin{aligned}
 R_{1212} &= \frac{\partial}{\partial \theta_1} [22,1] - \frac{\partial}{\partial \theta_2} [21,1] \\
 &\quad + \{ \frac{1}{21} \} [12,1] + \{ \frac{2}{21} \} [12,2] - \{ \frac{1}{22} \} [11,1] - \{ \frac{2}{22} \} [11,2] \\
 &= - \frac{1}{\theta_2^3} + \left(\frac{-1}{2\theta_2} \right) \left(\frac{-1}{2\theta_2^2} \right) - \left(\frac{1}{\theta_2} \right) \left(\frac{1}{2\theta_2^3} \right) \\
 &= - \frac{1}{4\theta_2^3}.
 \end{aligned}$$

Therefore no covariance stabilizing transformation exists.
Gaussian curvature becomes

$$K = \frac{R_{1212}}{g} = - \frac{1}{2},$$

where g is the determinant of (g_{ij}) .

4°. Coordinates of a Euclidean space in which the population space is immersed.

Let z^α denote the coordinates of the Euclidean space. From (21) z^α must satisfy the following equations:

$$\left\{ \begin{array}{l} \frac{\partial^2 z^\alpha}{\partial \theta_1^2} - \frac{\partial z^\alpha}{\partial \theta_2} = \frac{z^\alpha}{2\theta_2}, \\ \frac{\partial^2 z^\alpha}{\partial \theta_1 \partial \theta_2} + \frac{1}{2\theta_2} \frac{\partial z^\alpha}{\partial \theta_1} = 0, \\ \frac{\partial^2 z^\alpha}{\partial \theta_2^2} + \frac{1}{\theta_2} \frac{\partial z^\alpha}{\partial \theta_2} = \frac{z^\alpha}{4\theta_2^2}. \end{array} \right.$$

From the third equation it is seen that z^α must be of the form

$$\frac{f(\theta_1)}{\sqrt{\theta_2}} + g(\theta_1)\sqrt{\theta_2}.$$

where f and g are some functions of only θ_1 . Substituting it in the first and second equations, it is seen that $f(\theta_1)$ and $g(\theta_1)$ must satisfy

$$\frac{d^2 f}{d\theta_1^2} = g, \quad \frac{dg}{d\theta_1} = 0.$$

Therefore z^α must be of the form

$$\frac{a_\alpha \theta_1^2 + b_\alpha \theta_1 + d_\alpha}{\sqrt{\theta_2}} + 2a_\alpha \sqrt{\theta_2}.$$

Since the fundamental form of the Euclidean space is of the form

$$c_1(dz^1)^2 + c_2(dz^2)^2 + c_3(dz^3)^2$$

where c_1 , c_2 and c_3 are plus or minus one, comparing it with the fundamental form of the Riemannian space

$$\frac{1}{\theta_2}(d\theta_1)^2 + \frac{1}{2\theta_2^2}(d\theta_2)^2$$

we obtain that $c_1=c_2=-c_3=1$ and that

$$z^1 = \frac{\theta_1}{\sqrt{\theta_2}},$$

$$z^2 = \frac{\theta_1^2 - 2}{2\sqrt{2}\sqrt{\theta_2}} + \frac{\sqrt{\theta_2}}{\sqrt{2}},$$

$$z^3 = \frac{\theta_1^2 + 2}{2\sqrt{2}\sqrt{\theta_2}} + \frac{\sqrt{\theta_2}}{\sqrt{2}}.$$

It is easily seen that the coordinates satisfy (20), i.e.

$$(z^1)^2 + (z^2)^2 - (z^3)^2 = \frac{1}{K} = -2.$$

5°. Transformation by geodesic

It is seen from (24) that the arc of geodesic from (μ_0, σ_0^2) to (μ, σ^2) is

$$s = \sqrt{2} \cosh^{-1} \frac{(\mu - \mu_0) + 2(\sigma^2 + \sigma_0^2)}{4\sigma\sigma_0}.$$

By the law of transformation (5), if we use s as a transformation it follows that

$$\begin{aligned} g'^{11} &= \sigma^2 \left(\frac{\partial s}{\partial \mu} \right)^2 + 2\sigma^4 \left(\frac{\partial s}{\partial (\sigma^2)} \right)^2 \\ &= \sigma^2 \left(\frac{\partial s}{\partial \mu} \right)^2 + \frac{\sigma^2}{2} \left(\frac{\partial s}{\partial \sigma} \right)^2. \end{aligned}$$

Since partial derivatives s by μ and σ are

$$\begin{aligned} \frac{\partial s}{\partial \mu} &= \sqrt{2} \frac{1}{\sinh \frac{s}{\sqrt{2}}} \cdot \frac{2(\mu - \mu_0)}{4\sigma\sigma_0}, \\ \frac{\partial s}{\partial \sigma} &= \sqrt{2} \frac{1}{\sinh \frac{s}{\sqrt{2}}} \frac{4\sigma^2 - [(\mu - \mu_0)^2 + 2(\sigma^2 + \sigma_0^2)]}{4\sigma_0\sigma^2}, \end{aligned}$$

it is seen that

$$g'^{11} = 1.$$

This fact shows that theorem 2 is certainly valid in this case.

附録 2

Memorandum TYH-3

A Geometrical Interpretation of Location
and Scale Parameters

Tadashi Yoshizawa

November 9, 1971

at Department of Statistics
Harvard University
Cambridge, Massachusetts, U.S.A.

1. Introduction

Conditions for existence of a covariance stabilizing transformation and transformations of parameters which gives statistics with unit variance asymptotically were found in the memorandum TYH-2 [1971].

In this memorandum a geometrical interpretation of location and scale parameters is given after the Normal distribution's case exhibited in the memorandum TYH-2.

It is shown that the two dimensional space of location and scale parameters has negative constant Gaussian curvature if the space is considered as a Riemannian's using Fisher's information as the fundamental tensor of the metric, so that the space is expressed on a hyperbolic surface in three dimensional Euclidean space.

Moreover the geodesic distance between two points is given in the similar form as in the Normal case, and it gives statistics with unit variance asymptotically.

2. Fundamental tensor of the parameter space

Consider a frequency function

$$(1) \quad dF(x) = f\left(\frac{x-\alpha}{\beta}\right) d\left(\frac{x-\alpha}{\beta}\right), \quad \beta > 0,$$

where the parameter α locates the distribution and β is a scale parameter. We rewrite (1) as

$$(2) \quad dF = \exp\{g(y)\}dy = \exp\{g(y)\}dx/\beta$$

where $y = (x-\alpha)/\beta$, $g(y) = \log f(y)$.

In sample of size n the logarithm of the likelihood function is

$$(3) \quad \log L(x|\alpha, \beta) = \sum_{i=1}^n g(y_i) - n \log \beta.$$

(3) yields the likelihood equations

$$(4) \quad \frac{\partial \log L}{\partial \alpha} = -\frac{1}{\beta} \sum_{i=1}^n \left(\frac{\partial g}{\partial y} \right)_{y=y_i} = 0,$$

$$\frac{\partial \log L}{\partial \beta} = -\frac{1}{\beta} \left\{ \sum_{i=1}^n y_i \left(\frac{\partial g}{\partial y} \right)_{y=y_i} + n \right\} = 0.$$

Under regularity conditions, solution of (4) gives the ML estimators $\hat{\alpha}$ and $\hat{\beta}$. We now assume that for all permissible values of α and β

$$(5) \quad E \left(\frac{\partial \log L}{\partial \alpha} \right) = 0,$$

$$E \left(\frac{\partial \log L}{\partial \beta} \right) = 0.$$

Fisher's information matrix is given as follows:

$$(6) \quad -\frac{n}{\beta^2} E \begin{pmatrix} g'' & g''y \\ g''y & g''y^2 - 1 \end{pmatrix},$$

from which the variances and covariance may be determined by inversion (See Kendall & Stuart [1967], p. 63).

Consider the space of the location and scale parameters as a Riemannian's by introducing the fundamental tensor of the metric

$$(7) \quad (g_{ij}) = \frac{1}{\beta^2} \begin{pmatrix} a & c \\ c & b \end{pmatrix}$$

where $a = -E(g'')$, $b = E(g''y^2 - 1)$, and $c = E(g''y)$ from (6).

Notice that a , b , and c are constants, and that c is vanished if the distribution (1) is symmetric about α . Even for asymmetrical distributions, we can make the off-diagonal term in (7) zero by a simple change of origin such that

$$(8) \quad z = y - \frac{c}{a}.$$

This change of origin corresponds to the transformation of parameters such that

$$(9) \quad \begin{cases} \alpha' = \alpha + \frac{c}{a}\beta, \\ \beta' = \beta. \end{cases}$$

In this case the fundamental tensor of the metric (7) is transformed to

$$(10) \quad (g'_{ij}) = \frac{1}{\beta^2} \begin{pmatrix} a & 0 \\ 0 & \frac{G}{a} \end{pmatrix}$$

where $G = ab - c^2$, since the law of transformation of the fundamental tensor is given as

$$g'_{ij} = g_{k\ell} \frac{\partial \theta^k}{\partial \theta'^i} \frac{\partial \theta^\ell}{\partial \theta'^j}$$

where $\theta^1 = \alpha$, $\theta^2 = \beta$, $\theta'^1 = \alpha'$, $\theta'^2 = \beta'$ (See Yoshizawa [1971]).

3. Gaussian curvature

When the off-diagonal terms in the fundamental tensor are zeros, Gaussian curvature of the space is easily calculated by the following formula:

$$K = -\frac{1}{2} \frac{1}{\sqrt{g}} \left[\frac{\partial}{\partial \theta^1} \left(\frac{1}{\sqrt{g}} \frac{\partial g_{22}}{\partial \theta^1} \right) + \frac{\partial}{\partial \theta^2} \left(\frac{1}{\sqrt{g}} \frac{\partial g_{11}}{\partial \theta^2} \right) \right],$$

where $g = g_{11}g_{22}$. In the case of (10), g_{ij} does not depend on $\theta^1 (= \alpha)$. Then we obtain

$$\begin{aligned} K &= -\frac{1}{2} \frac{\beta^2}{\sqrt{G}} \cdot \frac{\partial}{\partial \beta} \left(\frac{\beta^2}{\sqrt{G}} \cdot \frac{\partial}{\partial \beta} \left(\frac{a}{\beta^2} \right) \right), \\ &= -\frac{1}{2} \frac{\beta^2}{\sqrt{G}} \frac{\partial}{\partial \beta} \left(-\frac{2a}{\sqrt{G}\beta} \right). \end{aligned}$$

So we find that

$$(11) \quad K = -\frac{a}{G}$$

Since a and G are positive, it is seen that the parameter space of location and scale parameters has a negative constant Gaussian curvature. Notice that Gaussian curvature is invariant under the transformation of the parameters.

Example 1: The Normal distribution $N(\mu, \sigma^2)$

Yoshizawa [1971] showed that the Gaussian curvature of the space is negative constant $-1/2$, using μ and σ^2 as coordinates. If we use μ and σ instead of μ and σ^2 , the fundamental tensor (7) is given as

$$(12) \quad (g_{ij}) = \frac{1}{\sigma^2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Of course, K is invariant under this transformation.

Example 2: The Gamma distribution

Consider the Gamma distribution

$$(13) \quad dF(x) = \frac{1}{\Gamma(p)} \left(\frac{x-\alpha}{\beta} \right)^{p-1} \exp \left\{ - \frac{(x-\alpha)}{\beta} \right\} d\left(\frac{x-\alpha}{\beta} \right) ,$$

$$\alpha \leq x < \infty; \beta > 0; p > 2.$$

In this case the fundamental tensor (7) is given as

$$(14) \quad (g_{ij}) = \frac{1}{\beta^2} \begin{pmatrix} \frac{1}{p-2} & 1 \\ 1 & p \end{pmatrix} .$$

From (11) it is seen that

$$(15) \quad K = - \frac{1}{2} ,$$

$$\text{since } G = \frac{p}{p-2} - 1 = 2/(p-2) .$$

4. An interpretation of the space as a hyperbolic surface

A two dimensional Riemannian space with negative constant Gaussian curvature may be expressed as a hyperbolic surface in a three dimensional Euclidean space. The coordinates of the Euclidean space and the expression of the hyperbolic surface can be obtained by the method used by Yoshizawa [1971].

The equations for the coordinates z are as follows:

$$(16) \quad \left\{ \begin{array}{l} \frac{\partial^2 z^\alpha}{\partial \alpha^2} - \frac{a^2}{Gy} \cdot \frac{\partial z^\alpha}{\partial \beta} = \frac{a^2}{Gy^2} z^\alpha, \\ \frac{\partial^2 z^\alpha}{\partial \alpha \partial \beta} + \frac{1}{\beta} \frac{\partial z}{\partial \alpha} = 0, \\ \frac{\partial^2 z}{\partial \beta^2} + \frac{1}{\beta} \frac{\partial z}{\partial \beta} = \frac{1}{\beta^2} z^\alpha. \end{array} \right.$$

Solving the above equation (16) and comparing the fundamental forms of the Euclidean space and of the Riemannian space, we obtain the following as a set of coordinates:

$$(17) \quad \begin{aligned} z^1 &= \frac{\sqrt{a\alpha}}{\beta}, \\ z^2 &= \frac{\frac{a}{2} \sqrt{\frac{a}{G}} \alpha^2 - \frac{1}{2} \sqrt{\frac{G}{a}}}{\beta} + \frac{1}{2} \sqrt{\frac{G}{a}} \beta, \\ z^3 &= \frac{\frac{a}{2} \sqrt{\frac{a}{G}} \alpha^2 + \frac{1}{2} \sqrt{\frac{G}{a}}}{\beta} + \frac{1}{2} \sqrt{\frac{G}{a}} \beta. \end{aligned}$$

Then the equation of the hyperbolic surface becomes

$$(18) \quad (z^1)^2 + (z^2)^2 - (z^3)^2 = -\frac{G}{a}.$$

Notice that the right hand of the above equation is the reciprocal of Gaussian curvature (11).

Since β is positive

$$z^3 > 0,$$

and at the bottom of the hyperbolic surface

$$z^1 = z^2 = 0,$$

$$z^3 = \sqrt{\frac{G}{a}},$$

this point corresponds $\alpha=0$ and $\beta=1$.

Since from (17)

$$(19) \quad z^3 - z^2 = \frac{1}{2\beta} \sqrt{\frac{\theta}{a}} ,$$

α changes on the curve cut by the plane (19) if β is fixed. It is easily seen that the curve is a parabolic.

5. Geodesic distance

The distance between two points is naturally measured by the arc of the geodesic which is easily given as

$$s = \sqrt{\beta} \cosh^{-1} \frac{a^2(\alpha - \alpha_0)^2 + G(\beta^2 + \beta_0^2)}{2G\beta\beta_0}$$

by the same method used in Yoshizawa [1971]. If we have maximum likelihood estimates $\hat{\alpha}$ and $\hat{\beta}$ of α and β respectively, we get a statistics with asymptotic unit variance such that

$$s = \sqrt{\frac{G}{a}} \cosh^{-1} \frac{a^2(\hat{\alpha} - \alpha)^2 + G(\hat{\beta}^2 + \beta^2)}{2G\hat{\beta}\beta}$$

where α takes the value α' of (9) if the change of origin (8) occurs. This fact is a direct result of the theorem 2 of Yoshizawa [1971].

Example 3: The Gamma distribution

From example 2, in the case of the Gamma distribution we have

$$a = \frac{1}{p-2} , \quad g = \frac{2}{p-2} .$$

Then we obtain

$$s = \sqrt{2} \cosh^{-1} \frac{\frac{1}{p-2}(\alpha - \alpha_0)^2 + 2(\beta^2 + \beta_0^2)}{4\beta\beta_0}$$

The fact that s is the same as in the case of the Normal distribution if p is three, is a reasonable result.

References

Kendall, M. G. & Stuart, A. [1967] The Advanced Theory of Statistics, 2nd. Edition, Hafner Publishing Company, New York.

Yoshizawa, T. [1972], "A Geometry of Parameter Space and Its Statistical Interpretation", Memorandum TYH-2, Harvard University, Cambridge, Massachusetts.

付録 3

多重分割表モデル選択についての補足

5.5 節数値例 b) モデルの段階的選択において述べた手続きについて補足しておく。122~125頁では表5.9のデータについて平方根分解により因子ごとの平方和への分解を表5.10に与え、その表を用いて逐次的に進めるモデル選択の手順を示した(図5.1, 表5.11参照)。ここでは、モデル{14/123}がAICRの観点からは最良であると結論した。まず、表5.11で*EPとしておいた対数線形モデルでの G^2 の値を我々のプログラムで求めた結果を下の表に示しておく。対数線形モデルについてもすべての項目を含む階層的モ

表付3.1 表5.11の補足

MODEL	D^2	G^2	DF	AICR	AIC
2/3/14	36.55	37.04	16	4.55	5.04
1/234	15.76	15.37	11	-6.24	-6.63
4/123	14.62	14.51	14	-13.38	-13.49
14/234	9.60	9.42	9	-8.40	-8.58
14/123	8.47	8.55	12	-15.53	-15.45
234/123	8.31	7.75	8	-7.69	-8.25
14/234/123	2.17	2.35	6	-9.83	-9.65

デルの中ではモデル{14/123}も選択するのが妥当であろう。

次に、選択されたモデル{14/123}でのセル期待値の推定値に

について検討しておこう。表5.9の4重分割表でのセル期待値を m_{ijkl} $i=1,2, j=1,2, k=1,2, l=1,2,3$ と書くと、対数線形モデルでのモデル $\{14/123\}$ は

$$m_{ijkl} = m_{i++k} \cdot m_{ijk+} / m_{i+++}$$

という関係が成り立つとも意味する。すなわち、才1項目を与えら
るとき、才4項目に対して才2と才3項目が独立であるというモデル
である。4.1節b) で述べた条件のもとで、そのモデルでのセル
期待値の最尤推定値は

$$\hat{m}_{ijkl} = x_{i++k} x_{ijk+} / x_{i+++}$$

で与えられる。ここで、 x_{ijkl} はもとの分割表のデータである。一方、
平方根分解でのセル期待値の推定値は、これを $\bar{m}_{ijkl}$ と書くと、

$$\bar{m}_{ijkl} = \frac{N}{\lambda^2} \sum_{\alpha \in K} \{B_{ijkl}(\alpha)\}^2$$

で与えられる。ここで、 $B_{ijkl}(\alpha)$ は平方根分解での α 因子効果の
 $ijkl$ 成分、 λ は (5.61) で定義した量、 N は分割表データの総
和 x_{++++} 、 K はモデル $\{14/123\}$ に対するクラス $\{ \emptyset, 1, 2, 3, 12, 13, 23, 14, 123 \}$ である。 $\hat{m}_{ijkl}$ と $\bar{m}_{ijkl}$ は 5.5 節の始めに
述べた汎用プログラムで計算される。表付3.2 にこれらの値とデ
ータ x_{ijkl} 、及び

$$C_{ijkl} = \frac{(\hat{m}_{ijkl} - x_{ijkl})^2}{\hat{m}_{ijkl}}$$

を並べて示す。 C_{ijkl} の総和はいわゆる χ^2 統計量であり、 D^2 や G^2 とほぼ等しい値を与える。 $\hat{m}_{ijkl}$ と $\bar{m}_{ijkl}$ は非常に近い値をとっている上に、 x_{ijkl} によく適合している。 本文で論じられているように (82頁, 114~116頁参照), 平方根分解によりモデル構成は確率論的解釈に難点があるが, 対数線形モデルの近似として考えることができ。 また, この例のように平方根分解によるモデル選択の結果を利用して対数線形モデルでのモデル選択を容易に行うことができる。

なお, データのモデルからの残差に相当する C_{ijkl} は異なるセルの間で非常に無相関というとはできないが, 異常なセル度数の検出に役立つことがある。 経験的には残差のヒストグラムから異常にはずれの値が検討を要するものとされるが, 理論的取り扱いは今後の課題である。

表 付3.2 セル期待値の推定

i (T)	j (U)	k (P)	l (S)	x_{ijkl}	$\bar{m}_{ijkl}$	$\hat{m}_{ijkl}$	c_{ijkl}
1	1	1	1	19	18.5	18.6	.009
2	1	1	1	57	49.0	48.9	1.338
1	2	1	1	29	29.7	29.3	.003
2	2	1	1	63	68.4	68.3	.416
1	1	2	1	29	34.4	33.5	.614
2	1	2	1	49	54.7	54.5	.547
1	2	2	1	27	22.5	22.5	.879
2	2	2	1	53	50.7	50.3	.145
1	1	1	2	23	22.4	22.5	.010
2	1	1	2	47	47.9	48.0	.022
1	2	1	2	33	35.1	35.5	.178
2	2	1	2	66	67.4	67.1	.018
1	1	2	2	47	40.1	40.6	.997
2	1	2	2	55	53.7	53.5	.043
1	2	2	2	23	27.1	27.3	.682
2	2	2	2	50	49.5	49.4	.008
1	1	1	3	24	25.4	24.9	.030
2	1	1	3	37	43.4	44.1	1.132
1	2	1	3	42	39.0	39.2	.204
2	2	1	3	68	61.7	61.6	.673
1	1	2	3	43	44.0	44.8	.074
2	1	2	3	53	48.9	49.1	.316
1	2	2	3	30	30.5	30.1	.001
2	2	2	3	42	44.8	45.3	.242

