

因子分析への行列集約アプローチ

足立 浩平*

A Matrix-Intensive Approach to Factor Analysis

Kohei Adachi*

現行の標準的な因子分析の定式化では、因子負荷量と独自分散は固定したパラメータであり、共通・独自因子は潜在的な確率変数として扱われる。これとは対照的に、共通因子および独自因子もパラメータと見なして、モデル部全てをパラメータ行列で表現する因子分析の定式化が近年になって提示されている。これを行列因子分析と名づけて、その諸性質を論じることが本稿の主題である。論及することには、行列因子分析の解法が、線形代数の定理だけに基づく点で明解であり、低階数近似としての主成分分析とは対照的に、因子分析をデータ行列の高階数近似と見なせる論拠を与えることが含まれる。さらに、行列因子分析の解と標準的な因子分析の解を比較する数値例を提示し、行列因子分析を発展させたスパース因子分析法にも言及する。

In the standard formulation of factor analysis (FA), factor loadings and unique variances are treated as fixed parameters, while common and unique factors are regarded as latent random variables. A very different formulation of FA has recently been presented in which common and unique factors are also treated as parameters and all model parts are expressed as parameter matrices. This is referred to as matrix factor analysis (MFA), whose properties are discussed in this paper. It is shown that the MFA algorithm is clearly described with concepts in linear algebra and allows FA to be viewed as higher rank approximation of a data matrix in contrast to principal component analysis as lower rank approximation. We further give numerical illustrations for comparing MFA solutions with the standard FA ones and discuss a sparse FA procedure derived from MFA.

キーワード: 探索的因子分析, 行列因子分析, 線形代数, 高階数近似, 因子の不確定性, スパース因子分析, 主成分分析

1. 序論

因子分析では、複数の観測変数の変動の多くは少数の共通因子によって説明され、共通因子によっては説明されずに残る各変数に独自の変動は、独自因子によって説明されると仮定される。こうしたモデルに基づき、共通因子と観測変数の関係を表す因子負荷量と、各変数に独自の変動の大きさを表す独自分散が、因子分析で推定されるパラメータとなる。因

* 大阪大学人間科学研究科：〒 565-0871 大阪府吹田市山田丘 1 番 2 号 (E-mail: adachi@hus.osaka-u.ac.jp).

因子分析の歴史は、共通因子を1つに限定した Spearman (1904) の着想に始まり、Thurstone (1935) による共通因子の複数化、そして、Rao (1955), Anderson and Rubin (1956) や Lawley and Maxwell (1971) などによる数理統計学的基礎の確立などを経て、現在に至る。近年も、市川 (2010), 豊田 (2012) や、改訂版である Mulaik (2010) などが発刊され、因子分析が多く研究者の興味を引き続けていることが窺える。なお、独自因子は想定されず、共通因子だけが考慮されるモデルを、広義に因子モデルと呼ぶことがあるが(例えば、奥井 (2014)), 本稿では、独自因子を考慮するモデルに基づく分析を指して、因子分析と呼ぶ。さらに、因子分析は、因子負荷量に分析者が制約を課す確認的(検証的)因子分析と、そうした制約のない探索的因子分析に区別されるが(足立 (2006), 狩野・三浦 (2002)), 本稿では、後者を指して単に因子分析と呼ぶ。

丘本 (1986) によれば、因子分析の定式化は、**変量モデル**と**母数モデル**に大別される。前者は、共通因子・独自因子を確率的に変動する潜在変数と見なすもので、変量モデルに基づく最尤法や最小二乗法が、因子分析の標準的な解法となっている(市川 (2010), Yanai and Ichikawa (2007))。従って、これらの解法による因子分析を、本論文では、現行の因子分析といった呼び方をする。以上の変量モデルで確率変数と見なされた共通因子を、固定された未知のパラメータ(母数)と見なすのが、母数モデルであるが、共通因子を他のパラメータと同時に推定できないことが知られ (Anderson and Rubin (1956), 丘本 (1986), 柳井他 (1990)), 母数モデルは、所与の因子負荷量と独自分散から共通因子の得点を推定するための理論的基礎の一つに留まっている(丘本 (1986))。

前段の変量・母数モデルとは異なる、因子分析の新たな定式化が近年になって登場し、これを考究することが、本論文の目的である。この定式化のモデル・解法を、それぞれ、**行列モデル・行列因子分析**と呼ぶことにする。行列モデルでは、母数モデルよりも、さらにパラメータと見なされるパーツが多くなる。すなわち、変量・母数モデルで確率変数として扱われた独自因子をも固定した未知パラメータと見なすのが、行列モデルの特徴である。従って、因子分析の**モデル部のすべてがパラメータ**とみなされて、それらを要素とする行列で表現され、パラメータの推定法の細部は全て線形(行列)代数の原理に基づくことが、行列モデル・行列因子分析の呼称の所以である。

著者の知る限り、行列因子分析が最初に現れるのは、Sočan (2003) がオランダ・Groningen 大学に提出した博士論文である。その中に、同大学教授の Henk A. L. Kiers が 2001 年に Sočan に提示したものとして、行列因子分析の解法が記される。これとは独立に、編著書の一部に de Leeuw (2004) が行列因子分析を記しているが、記述は簡単であり、解法は部分的にしか明示されていない。その後、数年を経た Unkel and Trendafilov (2010a, b) による論文が、行列因子分析を扱った最初の原著論文であると考えられる。彼らの 2010 年 a の論文は、因子分析の種々の定式化をレビューしたものであり、その中で行列モデルと解

法が明示されている。同著者の 2010 年 b の論文は、行列因子分析をロバストになるように変形したものである。その後、Adachi (2012) は、データ行列を必要とした行列因子分析のアルゴリズムを、変数間の共分散行列さえ与えられれば解が求められるように修正し、Adachi and Trendafilov (2014) は、負荷行列がスパースになるように (多くの 0 要素を持つように) 制約された行列因子分析を提案している。以上のように、多くの読者をもつ学術誌への公刊が 2010 年以降になることが、行列因子分析を「新たな定式化」と称する理由である。

行列因子分析は、母数モデルよりもパラメータが多いのにも関わらず、全てのパラメータの同時推定が可能である。しかし、現行の変量モデルの因子分析と同様に、因子は不確定、つまり、因子の得点は一意に定まらず、興味の対象となるのは因子負荷量と独自分散の解に限られる。この点で、行列因子分析が、現行の因子分析を上回る実用上のベネフィットは見られないが、前者の意義が、次に列挙するような理論的側面にあることを論じることが、本論文の目的となる。[1] 解法のステップが、理解しやすい線形代数の原理に基づき、因子分析の数学的成り立ちが把握しやすい。[2] 目的関数が、低階数近似としての主成分分析の目的関数を拡張した形をとる点で、因子分析と主成分分析の類似と相違を明確にする。[3] データ行列の低階数近似と見なせる主成分分析とは逆に、因子分析が高階数近似と見なし得ることを明瞭にする。[4] どのように因子が不確定であるかを、現行の因子分析とは違った方法で示す。[5] 今までの解法では見出せなかった因子分析の発展の方向を示唆する。

本論文の以下の章は、次のように構成される。まず、第 2 節で、現行の因子分析を概観した後、第 3 節では、主題となる行列モデル・行列因子分析を紹介し、その解法を第 4 節に記す。ただし、第 3 節と第 4 節では、前段に記した意義の [1] を明瞭にするために、以前の文献とは異なった解説を試みる。第 5 節では、行列モデルと変量モデルの解を数値的に比較する。第 6 節では、意義の [2], [3], [4] を論じる。第 7 節では意義 [5] の一例となるスパース因子分析を概説する。

2. 現行の因子分析の概要

因子分析の変量モデルを 2.1 節で概説した後、変量モデルに基づく解法を 2.2 節に記す。2.3 節で母数モデルを概説した後、2.4 節では因子分析のモデルが持つ 2 種の不定性を記す。

2.1 変量モデル

観測される p 個の変数を $p \times 1$ のベクトル $\mathbf{x}$ で表そう。ただし、その期待値 $E(\mathbf{x})$ は $p \times 1$ の零ベクトル $\mathbf{0}_p$ とする。変量モデルでは、 $\mathbf{x}$ に対して、潜在変数ベクトル $\mathbf{f}(m \times 1)$ と $\mathbf{u}(p \times 1)$ が、

$$\mathbf{x} = \mathbf{\Lambda}\mathbf{f} + \mathbf{\Psi}\mathbf{u} \quad (2.1)$$

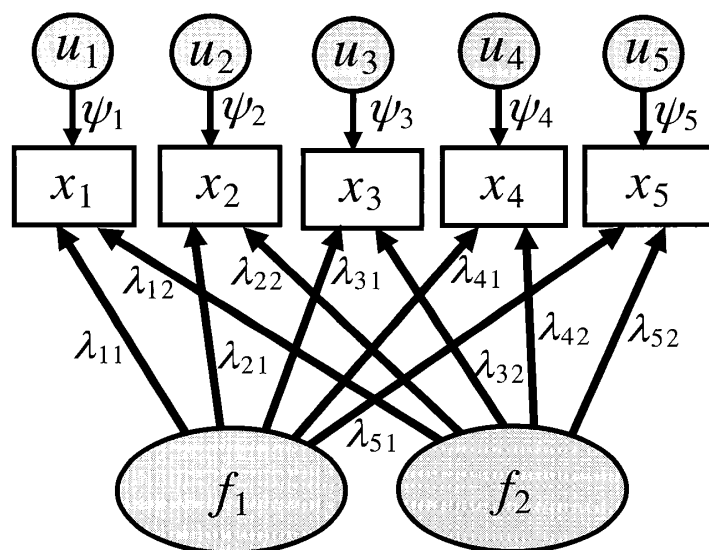


図1 因子分析のパス図による表現.

のように関係づけられる. ここで, $m < p$, Ψ は $p \times p$ の対角行列であり, Λ は $p \times m$ の行列で, その要素が**因子負荷量**である. 潜在変数は,

$$E(\mathbf{f}) = \mathbf{0}_m, \quad E(\mathbf{f}\mathbf{f}') = \mathbf{I}_m, \quad E(\mathbf{u}) = \mathbf{0}_p, \quad E(\mathbf{u}\mathbf{u}') = \mathbf{I}_p, \quad E(\mathbf{f}\mathbf{u}') = {}_m\mathbf{O}_p \quad (2.2)$$

を満たすと仮定される. ここで, $\mathbf{I}_m$ は $m \times m$ の単位行列, ${}_m\mathbf{O}_p$ は $m \times p$ の零行列を表す.

図1は, $p = 5$, $m = 2$ として, モデル (2.1), (2.2) を描いたものである. 図の下部のパス (矢印) は, $\mathbf{f} = [f_1, \dots, f_m]$ に前から $\Lambda = (\lambda_{jk})$ が乗じられて $\mathbf{x} = [x_1, \dots, x_p]'$ を説明する様子を表し, 上部のパスは, $\mathbf{u} = [u_1, \dots, u_p]'$ に Ψ が乗じられて $\mathbf{x}$ を説明する様子を表す. ここで, $\mathbf{f}$ の各要素は, $\mathbf{x}$ の p 変数すべてにパスを伸ばして, 全観測変数に共通して作用する要因と見なせるので, **共通因子**と呼ばれる. 一方, $\mathbf{u}$ に乗じられる Ψ が対角行列であるため, $\mathbf{u}$ と $\mathbf{x}$ の対応要素だけがパスで結ばれ, さらに, (2.2) より $\mathbf{u}$ の要素は互いに無相関であることから, $\mathbf{u}$ の要素は該当変数に独自の要因と見なせて, **独自因子**と呼ばれる.

(2.1) 式右辺の最後の項を $\mathbf{e} = \Psi\mathbf{u}$ において, $\mathbf{e}$ を誤差ベクトルと呼んで構わない. (2.2) より $E(\mathbf{e}) = \mathbf{0}_p$, $E(\mathbf{e}\mathbf{e}') = \Psi^2$ (対角行列) が導かれ, これらが示す誤差の変数間の無相関, つまり, 誤差が各変数に独自に作用するという仮定が, 因子分析の特徴である. こうした誤差の独自性から, その分散を表す Ψ^2 の対角要素は, **独自分散**と呼ばれる.

2.2 最小二乗法と最尤法

変量モデルのパラメータは, 負荷行列 Λ と独自分散を含む対角行列 Ψ^2 であるが, それらを推定するために, 変数間の共分散行列が重要な役割を果たす. この共分散行列は, モデル (2.1) と (2.2) に基づけば,

$$\Sigma = E(\mathbf{x}\mathbf{x}') = \Lambda\Lambda' + \Psi^2 \quad (2.3)$$

と表せ、データから算出される共分散行列の最尤推定値は、

$$\mathbf{S} = \frac{1}{n} \mathbf{X}' \mathbf{X} \quad (2.4)$$

と表せる。ここで、 $\mathbf{X}$ は $\mathbf{x}'$ の n 個の実現値を行として、列平均が 0 になるように中心化された n 個体 $\times p$ 変数のデータ行列であり、不偏性を重視すれば、(2.4) 右辺の分母の n は $n-1$ に代えられる。

現行の主要なパラメータ推定法に最小二乗法と最尤法がある。最小二乗法では、(2.3) と (2.4) の二乗誤差

$$h_{\text{LS}}(\mathbf{\Lambda}, \mathbf{\Psi}^2 | \mathbf{S}) = \|\mathbf{S} - \mathbf{\Lambda} \mathbf{\Lambda}' - \mathbf{\Psi}^2\|^2 = \text{tr}(\mathbf{S} - \mathbf{\Lambda} \mathbf{\Lambda}' - \mathbf{\Psi}^2)(\mathbf{S} - \mathbf{\Lambda} \mathbf{\Lambda}' - \mathbf{\Psi}^2)' \quad (2.5)$$

を最小にする $\mathbf{\Lambda}$ と $\mathbf{\Psi}^2$ が求められる。一方、最尤法は、(2.2) に $\mathbf{f}$, $\mathbf{u}$ の正規性の仮定を加えて導かれる対数尤度

$$h_{\text{ML}}(\mathbf{\Lambda}, \mathbf{\Psi}^2 | \mathbf{S}) \propto -\log |\mathbf{\Lambda} \mathbf{\Lambda}' + \mathbf{\Psi}^2| - \text{tr}\{\mathbf{S}(\mathbf{\Lambda} \mathbf{\Lambda}' + \mathbf{\Psi}^2)^{-1}\} \quad (2.6)$$

を最大化するものである。ソフトウェアの SAS や SPSS では、(2.5), (2.6) の最適化に、勾配法に基づくアルゴリズムが利用されるが (SPSS Inc. (1997), SAS Institute Inc. (2009)), (2.5) の最小化に特化した交互最小二乗法 (Harman and Jones (1966), Harman (1976)) や、(2.6) の最大化に特化した EM アルゴリズムも考案されている (Rubin and Thayer (1982), 豊田 (2012), Adachi (2013))。なお、(2.5) にウエイトを加えた目的関数を最小化する一般化最小二乗法もある (市川 (2010), 豊田 (2012))。

2.3 母数モデル

以上の変量モデルで確率変数と見なされた共通因子 $\mathbf{f}$ を、各個体に固有の固定されたパラメータ (母数) と見なすのが、母数モデルである。すなわち、個体 i の観測値および共通因子の得点を、それぞれ、 $\mathbf{x}_i$ および $\mathbf{f}_i$ と表すと、母数モデルは

$$\mathbf{x}_i = \mathbf{\Lambda} \mathbf{f}_i + \mathbf{\Psi} \mathbf{u}_i \quad (2.7)$$

と表せる。ここで、 $\mathbf{u}_i$ は個体 i の独自因子であるが、これは、確率変数と見なされて、 $\mathbf{u}_i$ および n 個体 $\times m$ 共通因子の行列 $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_n]'$ は、

$$\mathbf{1}_n' \mathbf{F} = \mathbf{0}_m', \quad \frac{1}{n} \mathbf{F}' \mathbf{F} = \mathbf{I}_m, \quad E(\mathbf{u}_i) = \mathbf{0}_p, \quad E(\mathbf{u}_i \mathbf{u}_i') = \mathbf{I}_p \quad (2.8)$$

を満たすと仮定される。ただし、 $\mathbf{1}_n$ は要素が全て 1 の $n \times 1$ のベクトルである。(2.7), (2.8) より (2.3) が導かれるので、最小二乗基準 (2.5) の最小化は、母数モデルの解法とも見なせる。一方、母数モデルの最尤解は存在しないことが知られる。すなわち、 $\mathbf{u}_i$ の正規性仮定

と (2.7), (2.8) より, 対数尤度

$$\begin{aligned} g_{\text{ML}}(\mathbf{F}, \mathbf{\Lambda}, \mathbf{\Psi}^2 \mid \mathbf{X}) &\propto -n \sum_{j=1}^p \log \psi_j^2 - \text{tr}(\mathbf{X} - \mathbf{F}\mathbf{\Lambda}')\mathbf{\Psi}^{-2}(\mathbf{X} - \mathbf{F}\mathbf{\Lambda}')' \\ &= -n \sum_{j=1}^p \log \psi_j^2 - \sum_{j=1}^p \frac{1}{\psi_j^2} \|\mathbf{x}_j - \mathbf{F}\boldsymbol{\lambda}_j\|^2 \end{aligned} \quad (2.9)$$

が導かれる. ここで, $\mathbf{x}_j$ は $\mathbf{X}$ の第 j 列, $\boldsymbol{\lambda}_j'$ は $\mathbf{\Lambda}$ の第 j 行, ψ_j は $\mathbf{\Psi}$ の第 j 対角要素を表すが, $\mathbf{x}_j = \mathbf{F}\boldsymbol{\lambda}_j'$, $\psi_j = 0$ となるように, $\mathbf{F}$, $\boldsymbol{\lambda}_j'$, ψ_j を選ぶことができ, (2.9) は無限大となる. すなわち, 共通因子行列 $\mathbf{F}$ と $\mathbf{\Lambda}$, $\mathbf{\Psi}^2$ を同時に推定することはできない (Anderson and Rubin (1956), 丘本 (1986), 柳井他 (1990)).

2.4 回転の不定性と因子の不確定性

$m \times m$ の正規直交行列を $\mathbf{T}$ と表すと, モデル (2.1) では $\mathbf{\Lambda}\mathbf{T}\mathbf{T}'\mathbf{f} = \mathbf{\Lambda}\mathbf{f}$, (2.7) では $\mathbf{\Lambda}\mathbf{T}\mathbf{T}'\mathbf{f}_i = \mathbf{\Lambda}\mathbf{f}_i$ のように, $\mathbf{\Lambda}\mathbf{T}$ を負荷行列とみなし, $\mathbf{f}\mathbf{T}$ または $\mathbf{f}_i\mathbf{T}$ を共通因子と見なしても, 条件 (2.2), (2.8) が満たされる. このように $\mathbf{T}$ を自由に選べるのが, 回転の不定性と呼ばれる. 適用目的に適した負荷行列 $\mathbf{\Lambda}\mathbf{T}$ を与える $\mathbf{T}$ を求める手続きは, 因子回転と総称されるが (Browne (2001)), 本論文では, 因子回転を数値例で使うだけで, それに関する議論はしない.

上記とは異なる不定性に, **因子の不確定性**がある. これは, 変量モデルにおいて, (2.1) と (2.2) を満たす $\mathbf{f}$ と $\mathbf{e}$ の組が一意でないことを表す (丘本 (1986)). この因子の不確定性は, 次節の行列因子分析の解にも見られ, 第6節で論じられる.

3. 行列モデル

前節の母数モデルで確率変数と見なされた独自因子 $\mathbf{u}_i$ を, 固定された未知パラメータと見なすのが, 行列モデルである. これを扱った文献では (Sočan (2003), de Leeuw (2004), Unkel and Trendafilov (2010a, b), Adachi (2012)), 行列因子分析の目的関数が最初に記されて, その基礎となるモデルの記述がないが, 本稿では, **行列モデル**を最初に記す. これは, n 個体 $\times p$ 変数の独自因子の行列を $\mathbf{U}$ と表すと,

$$\mathbf{X} \cong \mathbf{F}\mathbf{\Lambda}' + \mathbf{U}\mathbf{\Psi} = \mathbf{Z}\mathbf{A}' \quad (3.1)$$

と表せる. ここで,

$$\mathbf{A} = [\mathbf{\Lambda}, \mathbf{\Psi}] \quad (3.2)$$

は, 負荷行列 $\mathbf{\Lambda}$ と独自分散の平方根を含む対角行列 $\mathbf{\Psi}$ をブロックとする p 変数 $\times (m+p)$ 因子の行列であり, 図1のパスに付く係数を含むことから, $\mathbf{A}$ を**係数行列**と呼ぶことにする. また,

$$\mathbf{Z} = [\mathbf{F}, \mathbf{U}] \quad (3.3)$$

は、共通因子行列 $\mathbf{F}$ と独自因子行列 $\mathbf{U}$ をブロックとする n 個体 $\times (m+p)$ 因子の行列であり、これを**因子行列**と呼ぼう。この行列は、

$$\mathbf{1}'_n \mathbf{Z} = \mathbf{0}'_{m+p}, \quad \frac{1}{n} \mathbf{Z}' \mathbf{Z} = \mathbf{I}_{m+p} \quad (3.4)$$

と制約される。(3.4) の 2 つの制約は、(2.2) の行列版といえる 5 つの制約

$$\mathbf{1}'_n \mathbf{F} = \mathbf{0}'_m, \quad \frac{1}{n} \mathbf{F}' \mathbf{F} = \mathbf{I}_m, \quad \mathbf{1}'_n \mathbf{U} = \mathbf{0}'_p, \quad \frac{1}{n} \mathbf{U}' \mathbf{U} = \mathbf{I}_p, \quad \frac{1}{n} \mathbf{F}' \mathbf{U} = \mathbf{0}_p \quad (3.4')$$

と同値である。

モデル (3.1) と類似した因子分析の表現 $\mathbf{X} = \mathbf{F}\mathbf{\Lambda}' + \mathbf{U}\mathbf{\Psi}$ は、Mulaik (2010) や芝 (1979) に見られる。この表現では、左辺と右辺が等号で結ばれるのに対して、それらが (3.1) では近似記号 $\cong$ で結ばれる。この相違は、前者では、 $\mathbf{X}$ と $\mathbf{U}$ さらには $\mathbf{F}$ の行が確率変数と見なされているのに対して、後者の (3.1) では、両辺に現れる行列が全て固定した要素を持つと想定されているからである。一般に、任意の行列 $\mathbf{X}$ は、(3.4) を満たす $\mathbf{Z}$ と対角行列 $\mathbf{\Psi}$ をブロックとする $\mathbf{A}$ の転置の積 $\mathbf{Z}\mathbf{A}'$ には分解できないので、(3.1) の両辺は等号では結べないわけである。しかし、 $\mathbf{X}$ を近似する $\mathbf{Z}\mathbf{A}'$ を求めることは、(3.1) から直接導ける最小二乗基準

$$f(\mathbf{Z}, \mathbf{A}) = \|\mathbf{X} - (\mathbf{F}\mathbf{\Lambda}' + \mathbf{U}\mathbf{\Psi})\|^2 = \|\mathbf{X} - \mathbf{Z}\mathbf{A}'\|^2 \quad (3.5)$$

を制約条件 (3.4) のもとで最小化することで達成できる。この方法は最小二乗法に違いないが、2.2 節の (2.5) の最小化を最小二乗法と呼び続け、(3.5) の最小化を**行列因子分析**と呼ぶ。

4. 行列因子分析のアルゴリズム

行列因子分析のために最初に提示された Sočan (2003), de Leeuw (2004), Unkel and Trendafilov (2010a) の解法を、4.1 節に記す。これは、データ行列 $\mathbf{X}$ を要するものであったが、 $\mathbf{X}$ が与えられなくとも標本共分散行列 (2.4) だけで十分な Adachi (2012) の解法を、4.2 節に記す。4.3 節では、最小二乗法・最尤法とは異なる行列因子分析の特徴を論じる。

4.1 データ行列を要する解法

行列因子分析の解は、次のステップを収束するまで交互に反復する交互最小二乗法によって求められる。

[係数ステップ] $\mathbf{Z}$ を一定にして、(3.5) を最小にする $\mathbf{A}$ を求める。

[因子ステップ] $\mathbf{A}$ を一定にして、条件 (3.4) のもとで (3.5) を最小にする $\mathbf{Z}$ を求める。

まず、係数ステップの解を求めるために、 $\mathbf{\Psi}$ が対角行列であることと (3.4) を用いて、

(3.5) が

$$\begin{aligned}\|\mathbf{X} - \mathbf{Z}\mathbf{A}'\|^2 &= n \left\| \frac{1}{n} \mathbf{X}'\mathbf{Z} - \mathbf{A} \right\|^2 + \text{const} \\ &= n \left\| \frac{1}{n} \mathbf{X}'\mathbf{F} - \mathbf{\Lambda} \right\|^2 + n \left\| \frac{1}{n} \mathbf{X}'\mathbf{U} - \mathbf{\Psi} \right\|^2 + \text{const} \\ &= n \left\| \frac{1}{n} \mathbf{X}'\mathbf{F} - \mathbf{\Lambda} \right\|^2 + n \left\| \frac{1}{n} \text{diag}(\mathbf{X}'\mathbf{U}) - \mathbf{\Psi} \right\|^2 + \text{const}^* \quad (4.1)\end{aligned}$$

と書き換えられることに着目する. ここで, const と const^* は $\mathbf{A}$ に関係のない項を表し, $\text{diag}(\mathbf{X}'\mathbf{U})$ は, $\mathbf{X}'\mathbf{U}$ の対角要素を対角に配した対角行列である. (4.1) より, $\mathbf{A} = [\mathbf{\Lambda}, \mathbf{\Psi}]$ のブロックである負荷行列と独自分散行列の解は, それぞれ,

$$\mathbf{\Lambda} = \frac{1}{n} \mathbf{X}'\mathbf{F} \quad (4.2)$$

$$\mathbf{\Psi} = \frac{1}{n} \text{diag}(\mathbf{X}'\mathbf{U}) \quad (4.3)$$

によって与えられることがわかる.

次に, 因子ステップの解を求めよう. (3.4) を用いて (3.5) は $\|\mathbf{X} - \mathbf{Z}\mathbf{A}'\|^2 = \text{tr } \mathbf{X}'\mathbf{X} + n \text{tr } \mathbf{A}\mathbf{A}' - 2 \text{tr } \mathbf{X}'\mathbf{Z}\mathbf{A}'$ と展開されるので, 本ステップの課題は, 条件 (3.4) のもとでの $g(\mathbf{Z}) = \text{tr } \mathbf{X}'\mathbf{Z}\mathbf{A}' = \text{tr } \mathbf{Z}'\mathbf{X}\mathbf{A}$ の最大化に帰着する. この解は,

$$\mathbf{Z} = \sqrt{n}\mathbf{K}\mathbf{L}' = \sqrt{n}\mathbf{K}_1\mathbf{L}_1' + \sqrt{n}\mathbf{K}_2\mathbf{L}_2' \quad (4.4)$$

で与えられる. ここで, $\mathbf{K} = [\mathbf{K}_1, \mathbf{K}_2]$, $\mathbf{L} = [\mathbf{L}_1, \mathbf{L}_2]$ は, それぞれ, $n \times (p+m)$, $p \times (p+m)$ のブロック行列であり, そのブロック $\mathbf{K}_1$, $\mathbf{L}_1$ は, $n^{-1/2}\mathbf{X}\mathbf{A}$ の特異値分解

$$\frac{1}{\sqrt{n}}\mathbf{X}\mathbf{A} = \mathbf{K}_1\mathbf{\Theta}\mathbf{L}_1' \quad (4.5)$$

より求められ, $\mathbf{K}_2$, $\mathbf{L}_2$ は $\mathbf{K}_2'\mathbf{K}_1 = \mathbf{L}_2'\mathbf{L}_1 = \mathbf{0}_p$ を満たす正規直交行列である. すなわち, $\mathbf{\Theta}$ は $p \times p$ の対角行列であり, $\mathbf{K}'\mathbf{K} = \mathbf{L}'\mathbf{L} = \mathbf{I}_{p+m}$ である. (4.4) が解であることは, ten Berge (1983) の定理より, (3.4) の制約条件の 1 つ $n^{-1}\mathbf{Z}'\mathbf{Z} = \mathbf{I}_{m+p}$ のもとで導かれる不等式

$$g(\mathbf{Z}) = \text{tr } \mathbf{Z}'\mathbf{X}\mathbf{A} \leq \text{tr } \mathbf{\Theta} \quad (4.6)$$

の上限 $\text{tr } \mathbf{\Theta}$ を (4.4) が達成し, さらに, (3.4) のもう一つの条件 $\mathbf{1}_n'\mathbf{Z} = \mathbf{0}_{m+p}'$ も, (4.4) によって満たされることから確認される (Adachi (2012, Appendix 1)).

(4.4), (4.5) でわかるように, 解 (4.4) の右辺の $\sqrt{n}\mathbf{K}_1\mathbf{L}_1'$ は一意に定まるが, $\sqrt{n}\mathbf{K}_2\mathbf{L}_2'$ は一意でなく, $\mathbf{K}'\mathbf{K} = \mathbf{L}'\mathbf{L} = \mathbf{I}_{p+m}$ を満たす任意の $\mathbf{K}_2$, $\mathbf{L}_2$ を選ぶこととなる.

以上より, (4.2), (4.3), (4.4) を求めることを交互に反復すれば, 条件 (3.5) のもとでの (3.4) の最小化は達成される. すなわち, 行列因子分析では, 共通・独自因子, 負荷行列と

独自分散の同時推定ができるが、 $\mathbf{K}_2$, $\mathbf{L}_2$ が一意でないため、共通・独自因子は一意に定まらない。

4.2 共分散行列だけで十分な解法

前節のアルゴリズムは $n \times p$ のデータ行列 $\mathbf{X}$ を要するが、Adachi (2012) は、 $p \times p$ の共分散行列 (2.4) すなわち $\mathbf{S}$ さえあれば、負荷行列と独自分散の解を求められるアルゴリズムを考案している。因子分析のデータは $n > p$ となることが要請されるため、 $\mathbf{X}$ よりサイズが小さい $\mathbf{S}$ だけを要する点で、本節の方法はより経済的である。なお、Adachi (2012) は、(3.5) の目的関数に重み行列も加えた $\text{tr}(\mathbf{X} - \mathbf{Z}\mathbf{A}')\mathbf{W}^{-1}(\mathbf{X} - \mathbf{Z}\mathbf{A}')'$ を考慮しているが、本節では、 $\mathbf{W} = \mathbf{I}_p$ として Adachi (2012) の方法を記す。

$p \times (p + m)$ の変数と共通・独自因子の共分散行列を

$$\mathbf{C} = \frac{1}{n}\mathbf{X}'\mathbf{Z} = \frac{1}{n}\mathbf{X}'[\mathbf{F}, \mathbf{U}] \quad (4.7)$$

と表せば、(4.6) 左辺の $g(\mathbf{Z})$ は、 $\mathbf{C}$ の関数として

$$g(\mathbf{C}) = \text{tr} \mathbf{Z}'\mathbf{X}\mathbf{A} = n \text{tr} \mathbf{A}'\mathbf{C} \quad (4.8)$$

と書き換えられる。これが (4.6) 右辺の上限 $\text{tr} \mathbf{\Theta}$ を達成するときの (4.8) は、(4.5) から導かれる $n^{-1/2}\mathbf{X} = n^{-1/2}\mathbf{X}\mathbf{A}\mathbf{A}^+ = \mathbf{K}_1\mathbf{\Theta}_1\mathbf{L}_1'\mathbf{A}^+$ の転置に、後ろから (4.4) を乗じて、

$$\mathbf{C} = \mathbf{A}'^+\mathbf{L}_1\mathbf{\Theta}\mathbf{L}_1' \quad (4.9)$$

となる。ここで、 $\mathbf{A}'^+$ は $\mathbf{A}'$ のムーアペンローズ逆行列であり、 $\mathbf{A}$ の階数が p で $\mathbf{A}\mathbf{A}^+ = \mathbf{I}_p$ となることを仮定している。(4.9) に現れる $\mathbf{L}_1\mathbf{\Theta}\mathbf{L}_1'$ は、(4.5) から導かれる固有値分解

$$\mathbf{A}'\mathbf{S}\mathbf{A} = \mathbf{L}_1\mathbf{\Theta}^2\mathbf{L}_1' \quad (4.10)$$

によって求められる。以上より、(4.6) の上限達成という因子ステップの課題は、 $\mathbf{A}$ と標本共分散行列 $\mathbf{S}$ が与えられれば、(4.9) と (4.10) によって達成されることがわかる。

係数ステップでも $\mathbf{X}$ が不要であることは、(4.7) を使って、(4.2), (4.3) が、それぞれ、

$$\mathbf{\Lambda} = \mathbf{C}\mathbf{H}^{[m]}, \quad \mathbf{\Psi} = \text{diag}(\mathbf{C}\mathbf{H}_{[p]}) \quad (4.11)$$

と書き換えられることからわかる。ここで、 $\mathbf{H}^{[m]} = [\mathbf{I}_{m,m}\mathbf{O}_p]'$ ($(m+p) \times m$)、 $\mathbf{H}_{[p]} = [\mathbf{O}_m, \mathbf{I}_p]'$ ($(m+p) \times p$) であり、 $\mathbf{C}$ は (4.9) によって与えられる。さらに、達成される目的関数の値も、

$$n \text{tr} \mathbf{S} - n \text{tr}(\mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{\Psi}^2) = n \text{tr}(\mathbf{S} - \mathbf{A}\mathbf{A}') \quad (4.12)$$

と表せ、 $\mathbf{X}$ を要せずに算出される。これは、(3.4') を用いれば、(3.5) を $\text{tr} \mathbf{X}'\mathbf{X} + n \text{tr} \mathbf{\Lambda}\mathbf{\Lambda}' + n \text{tr} \mathbf{\Psi}^2 - 2 \text{tr} \mathbf{X}'\mathbf{F}\mathbf{\Lambda} - 2 \text{tr} \mathbf{X}'\mathbf{U}\mathbf{\Psi}$ と展開でき、この式に (4.2), (4.3), (2.4) を使って、

$\text{tr } \mathbf{X}'\mathbf{U}\Psi = \text{tr } \text{diag}(\mathbf{X}'\mathbf{U})\Psi$ に着目すれば導かれる. なお, (4.12) を $n \text{tr } \mathbf{S}$ で除した $1 - \text{tr } \mathbf{A}\mathbf{A}' / \text{tr } \mathbf{S}$ は, 0 以上 1 以下に正規化されて, 収束判定に便利な指標として使える.

以上より, 最適な $\mathbf{A}$ と Ψ を求めるための行列因子分析のアルゴリズムは, 次のように記される.

- [1] $\mathbf{A}$ を初期化する.
- [2] 固有値分解 (4.10) を行って, (4.9) を求める.
- [3] (4.11) によって, $\mathbf{A} = [\mathbf{\Lambda}, \Psi]$ を更新する.
- [4] 収束していれば終了し, 他のはときは [2] に戻る.

4.3 行列因子分析の解法の特徴

現行の因子分析と違って独自因子をパラメータと見なす行列因子分析では, 当然ながら, その解法の要所に独自因子に関わる行列が現れる点が, 2.2 節の解法には見られない特徴になる. そうした行列の中でも, **変数と共通・独自因子の共分散行列** (4.7) すなわち $\mathbf{C}$ が, 前節のアルゴリズムで重要な役割を果たす. このアルゴリズムの要点は, 因子ステップにおいて, 因子を求めるかわりに, $\text{tr } \mathbf{A}'\mathbf{C}$ を制約 (3.4) のもとで最大化する $\mathbf{C}$ を求めることであった. さらに, $\mathbf{C}$ が係数ステップでも重要な役割を果たすことは, (4.1) より, $\|n^{-1}\mathbf{X}'\mathbf{Z} - \mathbf{A}\|^2 = \|\mathbf{C} - \mathbf{A}\|^2$ の最小化が, 係数ステップの課題となることからわかる. すなわち, 行列因子分析のアルゴリズムは, $\|\mathbf{C} - \mathbf{A}\|^2$ の $\mathbf{A}$ に関する最小化と, $\text{tr } \mathbf{A}'\mathbf{C}$ の $\mathbf{C}$ に関する制約つき最大化の交互反復に帰着すると言える.

(4.3) に示されるように, **独自分散**の平方根が, 変数と独自因子の共分散によって与えられることも, 行列因子分析の特徴である. これとは違って, 現行の因子分析, すなわち, 2.2 節の最小二乗法・最尤法の両者では, $\Psi_{\text{ML}}^2 = \text{diag}(\mathbf{S} - \mathbf{\Lambda}_{\text{ML}}\mathbf{\Lambda}_{\text{ML}}')$ が独自分散の推定値となる. ここで, 行列因子分析の $\mathbf{\Lambda}$, Ψ と区別するため, それらに, 最尤法の略語である ML を添え字として付している. 上記の式のトレースを考えると,

$$\text{tr } \Psi_{\text{ML}}^2 = \text{tr}(\mathbf{S} - \mathbf{\Lambda}_{\text{ML}}\mathbf{\Lambda}_{\text{ML}}') \quad (4.13)$$

が得られるが, これと, 行列因子分析の目的関数が達成する値 (4.12) を比較すると, 独自分散の大小に関して解法間に一定の傾向があることが推察される. すなわち, (4.12) は最小二乗基準 (3.5) から導かれるので, 非負の値をとる, つまり, $\text{tr } \mathbf{S} - \text{tr}(\mathbf{\Lambda}\mathbf{\Lambda}' + \Psi^2) \geq 0$ であり, これは,

$$\text{tr } \Psi^2 \leq \text{tr } \mathbf{S} - \text{tr } \mathbf{\Lambda}\mathbf{\Lambda}' \quad (4.14)$$

と書き換えられる. この不等式と, (4.13) で両辺を結ぶ等号を見比べると, 行列因子分析の独自分散が, 最小二乗・最尤法のそれより, 小さい値をとりやすい傾向が推察される. ただし, (4.13), (4.14) の両辺はトレース, すなわち, 個々の独自分散ではなく, その合計に

関する式であり、さらに、右辺の Λ_{ML} と Λ が等しい保証はないので、上記の独自分散の大小関係は推察にすぎない。この推察の妥当性は、次節で経験的に検証される。

5. 数値例

現在普及する変量モデルとは趣を異にする行列因子分析が、本当に「因子分析らしい解」を出力するのかが、素朴な問いとなる。そこで、変量モデルから導かれる最小二乗法・最尤法と同等の解を、行列因子分析が与えるか否かを、本節で数値的に検証する。この同等性は、Adachi (2012) の数値例からも察せられるが、同等性を示す適用例が増えれば、勿論、その経験的妥当性は高まる。そこで、本節では、Adachi (2012) とは異なるデータへの適用例を提示し、さらに、シミュレーション研究も報告する。これらの数値例には、前節に記した独自分散の大小関係の検討も含まれる。なお、本節では、行列因子分析によって得られた因子負荷行列、独自分散の平方根の行列、その対角要素を単に Λ , Ψ , ψ_j と表すのに対して、変量モデルの最小二乗法によるそれらの解を Λ_{LS} , Ψ_{LS} , ψ_j^{LS} と、最尤法による解を Λ_{ML} , Ψ_{ML} , ψ_j^{ML} と表す。

5.1 実データへの適用

Web ページ <http://bm.osaka-u.ac.jp/data/big5/> から入手できる 25 変数間の相関行列を、適用例に用いる。この相関行列は、繁桝 (1998, p. 144) の性格特性に関する 25 項目 (変数) を、著者が $n = 190$ 名の男女大学生に提示して、各項目の特性を自身が有する程度を自己評定させた結果に基づく。表 1 の左の列が示すように、25 変数は、5 因子のいずれかに関係するものとして 5 群に分類されることが、性格特性研究の知見から知られている (柏木 (1997))。相関行列に、因子数を 5 とし、最小二乗法・最尤法・行列因子分析を適用し、解の負荷行列にはバリマックス回転を適用した。

表 1 は、行列因子分析の解を示す。太字にした絶対値の大きい因子負荷量は、前段に記した知見のとおり、25 変数が 5 つの因子に分類されることを示している。以上の解と最小二乗法・最尤法の解を比較するため、それらの解から行列因子分析の解を減じた値、すなわち、 $\Lambda_{LS} - \Lambda$, $\Psi_{LS}^2 - \Psi^2$ の対角要素と、 $\Lambda_{ML} - \Lambda$, $\Psi_{ML}^2 - \Psi^2$ の対角要素を表 2 に掲げる。表より、殆どの行列因子分析の因子負荷量は他の 2 方法の解と一致して、違ったとしても ± 0.01 の差であり、負荷量は方法間で同等といえる。独自分散の差も 0.04 以下であり、方法間でほぼ同等といえる。ただし、マイナスの差がないことに着目すべきである。すなわち、4.3 節で推察したとおり、行列因子分析によって得られる独自分散は、最小二乗法・最尤法の独自分散より小さくなっている。同じ傾向は、Adachi (2012, Table 2) の最尤法と行列因子分析の解にも窺える。

表1 行列因子分析の解: 絶対値の大きい負荷量は太字で表示.

群	変数	Λ					Ψ^2
1	心配性	0.71	-0.14	-0.13	0.22	-0.11	0.39
	傷つきやすい	0.74	-0.05	-0.04	0.01	0.09	0.43
	悲観的	0.67	-0.29	-0.13	-0.03	-0.27	0.37
	動揺する	0.56	0.05	-0.15	-0.08	0.32	0.53
	気苦労が多い	0.61	-0.09	-0.16	0.10	-0.31	0.47
2	社交的	-0.15	0.81	0.10	0.09	-0.12	0.27
	おしゃべり	0.05	0.82	-0.05	-0.02	0.08	0.30
	自発的	-0.09	0.72	0.18	0.10	-0.13	0.40
	陽気	-0.21	0.80	0.08	0.16	0.00	0.27
	目立ちたがり	0.07	0.64	0.22	-0.03	0.05	0.52
3	独創的	-0.09	0.03	0.85	-0.04	0.07	0.26
	冒険志向	-0.22	0.22	0.67	-0.01	0.20	0.41
	進歩的	-0.18	0.24	0.62	0.09	-0.09	0.50
	臨機応変	-0.27	0.31	0.35	0.20	-0.03	0.66
	想像的	0.18	0.09	0.45	0.11	0.34	0.62
4	おだやかな	-0.14	-0.10	0.15	0.59	-0.03	0.58
	やさしい	0.08	0.20	0.11	0.59	0.00	0.57
	利他的	0.12	0.01	-0.01	0.69	-0.15	0.46
	協力的	0.01	0.27	-0.12	0.63	-0.09	0.49
	共感的	0.11	0.13	-0.03	0.72	-0.20	0.39
5	慎重な	0.14	-0.03	0.07	0.22	-0.56	0.59
	信頼できる	-0.12	0.31	0.08	0.27	-0.56	0.49
	勤勉な	-0.04	0.05	-0.06	0.15	-0.77	0.37
	計画的	0.10	-0.02	-0.09	0.02	-0.69	0.49
	几帳面な	0.21	0.02	-0.18	0.04	-0.73	0.38

5.2 シミュレーション

Adachi (2012) は, 変量モデルによって発生させた人工データに, 行列因子分析を適用して, それが, 最小二乗法・最尤法と同程度の正確さで因子負荷量と独自分散の真値を再現することを確認している. ここでは, 真値再現の程度ではなく, 方法間の解の類似性, および, 行列因子分析がより小さい独自分散を与える傾向を, 人工データを使ったシミュレーションによって検証する.

データ生成の手続きは, 区間 $[a, b]$ の整数を定義域とする離散一様分布を $DU(a, b)$, $[\alpha, \beta]$ の (連続) 一様分布を $U(\alpha, \beta)$ と表すと, 次のように記述される. [1] 因子数 m を $DU(1, 5)$, 変量数 p を $DU(4m, 8m)$, そして, 個体数 n を $DU(8p, 12p)$ から抽出する. [2] Λ の要素を $U(-1.0, 1.0)$ から抽出し, Ψ^2 の対角要素を $U(0.05, 0.65)$ から抽出する. [3] $\mathbf{f}$ および $\mathbf{u}$ の各要素を独立に標準正規分布から抽出して, モデル (2.2) を満たす $\mathbf{x}'$ を n 個生成し, それ

表 2 最小二乗法・最尤法の解から行列因子分析の解を減じた値. 空欄は 0.00 を表す.

変数	最小二乗法と行列因子分析				最尤法と行列因子分析				
	$\Lambda_{LS} - \Lambda$			$\Psi_{LS}^2 - \Psi^2$	$\Lambda_{ML} - \Lambda$			$\Psi_{ML}^2 - \Psi^2$	
心配性	0.01	0.01	-0.01	0.01	-0.01			0.02	
傷つきやすい	-0.01		-0.01	0.01		-0.01		0.01	
悲観的				0.01	0.01	0.01	0.01		
動揺する			0.01	0.01		0.01		0.02	
気苦労が多い		-0.01		0.02	0.01	0.00	0.01	0.01	
社交的				0.02		0.01		0.02	
おしゃべり			0.01	0.02	-0.01			0.01	
自発的			0.01	0.02		-0.01		0.02	
陽気	0.01		-0.01	-0.01	0.01	-0.01	-0.01	0.03	
目立ちたがり		-0.01		0.03	-0.01		0.01	0.01	0.02
独創的		-0.01		0.01					0.01
冒険志向				0.01		0.01			
進歩的			-0.01	0.01	-0.01	-0.01		0.01	0.02
臨機応変				0.01					0.01
想像的			0.01	0.01	-0.01				0.02
おだやかな		-0.01	0.01	0.01		0.01	-0.01		0.03
やさしい		0.01	0.01	0.01	-0.01	-0.01	-0.01	-0.01	0.04
利他的	0.01	0.01	0.01	0.02					0.03
協力的			-0.01	0.03		0.01			0.02
共感的		-0.01	-0.01	-0.01	0.04	0.01			0.02
慎重な			-0.01	0.01		-0.01	0.01		0.02
信頼できる			-0.01	0.01	0.02	-0.01		0.01	0.01
勤勉な				0.01	0.01	-0.01		-0.01	
計画的				0.02	0.01		-0.01	0.01	0.02
几帳面な		-0.01	0.01	0.02		0.01			0.02

表 3 行列因子分析・最小二乗法・最尤法の解どうしの非類似度 (差の絶対値の平均).

要約統計量	因子負荷量			独自分散		
	行列-LS	行列-ML	LS-ML	行列-LS	行列-ML	LS-ML
	$d_{\Lambda}(\Lambda, \Lambda_{LS})$	$d_{\Lambda}(\Lambda, \Lambda_{ML})$	$d_{\Lambda}(\Lambda_{LS}, \Lambda_{ML})$	$d_{\Psi}(\Psi^2, \Psi_{LS}^2)$	$d_{\Psi}(\Psi^2, \Psi_{ML}^2)$	$d_{\Psi}(\Psi_{LS}^2, \Psi_{ML}^2)$
平均値	0.003	0.004	0.006	0.007	0.007	0.009
95 パーセンタイル	0.009	0.012	0.018	0.015	0.016	0.022

を $\mathbf{X}$ の行とする. [4] $\mathbf{X}$ の列を標準化して, 変数間の相関行列 $\mathbf{S}$ を求める. 以上の [1]~[4] を繰り返して, 計 1000 個の $\mathbf{S}$ を求めた. 各 $\mathbf{S}$ に, 最小二乗法・最尤法・行列因子分析を適用した.

各データについて, 二つの方法間の解の非類似度を, 対応する行列要素間の差の絶対値の平均によって定義した. すなわち, 最尤法と行列因子分析の間の因子負荷量, 独自分散の非類似度を, それぞれ, $d_{\Lambda}(\Lambda, \Lambda_{ML}) = (pm)^{-1} \|\Lambda - \Lambda_{ML}\|_{l1}$ および $d_{\Psi}(\Psi^2, \Psi_{ML}^2) = \|\Psi^2 - \Psi_{ML}^2\|_{l1}$ とした. ここで, $\|\cdot\|_{l1}$ は $l1$ ノルムを表す. 他の方法間の解の非類似度 $d_{\Lambda}(\Lambda, \Lambda_{LS})$, $d_{\Lambda}(\Lambda_{LS}, \Lambda_{ML})$, $d_{\Psi}(\Psi^2, \Psi_{LS}^2)$, $d_{\Psi}(\Psi_{LS}^2, \Psi_{ML}^2)$ も同様に定義される. ここで,

表 4 独自分散の大小関係の頻度 (%).

行列-LS		行列-ML		LS-ML	
$\psi_j < \psi_j^{\text{LS}}$	$\psi_j > \psi_j^{\text{LS}}$	$\psi_j < \psi_j^{\text{ML}}$	$\psi_j > \psi_j^{\text{ML}}$	$\psi_j^{\text{LS}} < \psi_j^{\text{ML}}$	$\psi_j^{\text{LS}} > \psi_j^{\text{ML}}$
85.4	14.6	84.7	15.3	49.7	50.3

回転の不定性がある負荷行列の非類似度については、 $d_{\Lambda}(\cdot, \cdot)$ の後ろの行列に近づくように前の行列にプロクラステス回転を適用した。計 1000 個の非類似度の平均と 95 パーセンタイルを表 3 の左に示す。例えば、 $d_{\Lambda}(\mathbf{\Lambda}, \mathbf{\Lambda}_{\text{ML}})$, $d_{\Psi}(\mathbf{\Psi}^2, \mathbf{\Psi}_{\text{ML}}^2)$ の 95 パーセンタイルは、1000 個のデータのうち 950 個以上について、最尤法と行列因子分析の負荷行列間の要素の平均相違が 0.012 以下で、独自分散の平均相違が 0.016 以下と微小であったことを表す。さらに、行列因子分析と最尤法・最小二乗法との相違は、最尤法と最小二乗法の相違よりも、むしろ小さいことが窺える。以上より、行列因子分析と普及する 2 つの解法の類似性が確認できる。

次に、解法間の独自分散の大小関係をみるため、計 1000 個のデータから得られた延べ 17177 個 $\times$ 3 方法の独自分散の大小を、解法間で比較した結果を表 4 に示す。例えば、表の $\psi_j < \psi_j^{\text{ML}}$ および $\psi_j > \psi_j^{\text{ML}}$ の欄は、17177 回の比較機会のうちの 84.7% で行列因子分析の独自分散が最尤法のそれより小さく、逆の関係であったのは 15.3% の機会にすぎなかったことを示す。最小二乗法と行列因子分析との大小比較も同様の結果を示し、行列因子分析がより小さい独自分散を示す傾向が確認されると同時に、これが常に成り立つ性質でないことも明らかになった。なお、表 4 の右から、最小二乗法と最尤法の独自分散には大小関係がないこともわかる。

6. 高階数近似としての因子分析

因子分析と同様に、 n 個体 $\times p$ 変数のデータ行列 $\mathbf{X}$ を対象にして、 p 変数の変動を少数の成分に縮約することを目的とする方法に、Pearson (1901), Hotelling (1933) が導入した**主成分分析**がある (Okamoto (1969), Jolliffe (2002)). 因子分析と主成分分析の関係は、種々の側面から議論されてきたが (例えば, Sato (1987), Bentler and Kano (1990), Ogasawara (2000)), 両者の類似と相違を明瞭にするのは、種々の定式化の中でも、主成分分析を、

$$f_{\text{PC}}(\mathbf{F}, \mathbf{\Lambda}) = \|\mathbf{X} - \mathbf{F}\mathbf{\Lambda}'\|^2 \quad (6.1)$$

を最小にする $n \times m$ の主成分得点 $\mathbf{F}$ 、および、 $p \times m$ の負荷行列 $\mathbf{\Lambda}$ を求めることと定式化するアプローチである (Eckart and Young (1936), 足立・村上 (2011), Takane (2014)). ここで、 $m \leq \min(n, p)$ である。(6.1) より、主成分分析は、データ行列 $\mathbf{X}$ を階数が m の行列 $\mathbf{F}\mathbf{\Lambda}'$ で近似する**低階数近似問題**と見なせる。行列因子分析の目的関数 (3.5) と主成分分

析の (6.1) を見比べると明らかであるが、主成分分析の目的関数に $\mathbf{U}\Psi$ を加えて拡張したものが、因子分析であると捉えられる。言い方を変えれば、 $\mathbf{X}$ に $\mathbf{F}\mathbf{A}'$ をフィットさせる、つまり、変数の変動を共通因子・成分で説明しようとする点までは、因子分析と主成分分析は同様であるが、前者を後者から区別するのは、独自因子の項 $\mathbf{U}\Psi$ の存在である。この視点からすると、因子分析を主成分分析から峻別する特徴は、共通因子よりも、むしろ独自因子の存在であると言えよう。

さらに、低階数近似としての主成分分析とは対照的に、因子分析を等階数近似、さらには、高階数近似とも見なせることを、以下に論じる。

まず、因子分析がデータ行列 $\mathbf{X}$ の**等階数近似**と見なせることは、(3.5) が $\mathbf{X}$ に $\mathbf{Z}\mathbf{A}'$ をフィットさせていることからわかる。すなわち、 $\mathbf{X}$ と $\mathbf{A}'$ が列フルランクであれば、 $\mathbf{X}$ の階数は p であり、(3.4) より $\mathbf{Z}$ の階数は $p+m > p$ であることから、 $\mathbf{Z}\mathbf{A}'$ の階数は $\mathbf{X}$ と等しく p である。等階数の行列どうしのフィッティングであるのに、(3.5) が 0 にならないのは、 $\mathbf{A} = [\mathbf{A}, \mathbf{\Psi}]$ のブロックの $\mathbf{\Psi}$ が対角行列、すなわち、 $\mathbf{\Psi}$ の非対角要素が 0 に制約されていることによる。

次に、因子分析を**高階数近似**と称することができる理由は、(3.5) が

$$\|\mathbf{X} - \mathbf{Z}\mathbf{A}'\|^2 = \|\mathbf{Z} - \mathbf{X}\mathbf{A}\|^2 + c^\# \quad (6.2)$$

と書き換えられることにある。ここで、 $c^\# = \|\mathbf{X}\|^2 + n\|\mathbf{A}\|^2 - \|\mathbf{X}\mathbf{A}\|^2 - n(p+m)$ は、 $\mathbf{Z}$ に無関係な項を表す。すなわち、所与の $\mathbf{A}$ に対して、階数が p の行列 $\mathbf{X}\mathbf{A}$ を、より高階数 $m+p$ の行列 $\mathbf{Z}$ で近似する問題として因子分析をとらえることができる。

著者の知る限り、因子分析を高階数近似と称した過去の論文は見出せないが、この呼称は新奇なものではない。変量モデルでも、共通因子と独自因子の高次元性が因子の不確定性を導くことは、古くから知られている(丘本(1986), 柳井他(1990))。例えば、McDonald(1979) は、 $\mathbf{z} = [\mathbf{f}, \mathbf{\Psi}\mathbf{u}]$, $\mathbf{H} = [\mathbf{A}, \mathbf{I}_p]$ とすると (2.1) は $\mathbf{x} = \mathbf{H}\mathbf{z}$ と書き替えられ、 $\mathbf{H}$ が $p \times (m+p)$ であることから方程式 $\mathbf{x} = \mathbf{H}\mathbf{z}$ を満たす $\mathbf{z}$ の解は一意でないことを示している。ただし、(6.2) のように、因子を求めるための目的関数の形で、因子分析を高階数近似として見なせることを示すのは、独自因子もパラメータと見なす行列因子分析に特有のものである。

共通・独自因子は不確定、つまり、因子の解は無数にあるが、それら無数の解どうしが一定の関係を持つことをみるために、行列因子分析の解 (4.4) を思い出そう。その右辺の一意に定まる第 1 項は、(4.5) を用いて、

$$\mathbf{Z}_1 = \sqrt{n}\mathbf{K}_1\mathbf{L}'_1 = \mathbf{X}\mathbf{A}\mathbf{L}_1\mathbf{\Theta}^{-1}\mathbf{L}'_1 \quad (6.3)$$

のように $\mathbf{X}$ の関数となる. Adachi (2012) は, (6.3) と共通・独自因子の解 (4.5) の間に,

$$\|\mathbf{Z} - \mathbf{Z}_1\|^2 = \|\sqrt{n}\mathbf{K}\mathbf{L}' - \sqrt{n}\mathbf{K}_1\mathbf{L}_1'\|^2 = n(\text{tr } \mathbf{L}'\mathbf{L} - \text{tr } \mathbf{L}_1'\mathbf{L}_1) = nm \quad (6.4)$$

が成り立つことを示している. 行列 $\mathbf{Z}$ および $\mathbf{Z}_1$ の列を縦に積んだベクトルを, それぞれ, $\text{Vec}(\mathbf{Z})$, $\text{Vec}(\mathbf{Z}_1)$ と表すと, (6.4) は, 解に対応する $\text{Vec}(\mathbf{Z})$ が, $\mathbf{X}$ の関数として一意に定まる $\text{Vec}(\mathbf{Z}_1)$ を中心に, 半径 $(nm)^{1/2}$ の円を描くことを表す. 以上の行列因子分析の性質に類似した事実が, 変量モデルでは, Mulaik (1976) の円錐の名で知られる (丘本 (1986), 柳井他 (1990)). すなわち,

$$E(\mathbf{s}) = \mathbf{0}_m, \quad E(\mathbf{s}\mathbf{s}') = \mathbf{I}_m - \mathbf{\Lambda}'\mathbf{\Sigma}^{-1}\mathbf{\Lambda}, \quad E(\mathbf{x}\mathbf{s}') = {}_p\mathbf{O}_m \quad (6.5)$$

を満たす任意の確率変数ベクトル $\mathbf{s}$ を用いて定義される $\mathbf{f} = \mathbf{\Lambda}'\mathbf{\Sigma}^{-1}\mathbf{x} + \mathbf{s}$, $\mathbf{u} = \mathbf{\Psi}^{-1}\mathbf{\Sigma}^{-1}\mathbf{x} + \mathbf{\Psi}^{-1}\mathbf{\Lambda}\mathbf{s}$ は, 変量モデル (2.1), (2.2) を満たすことが知られ, 共通因子ベクトル $\mathbf{f}$ が, $\mathbf{f}_1 = \mathbf{\Lambda}'\mathbf{\Sigma}^{-1}\mathbf{x}$ に $\mathbf{s}$ を加えた円周を構成する. この変量モデルの円と, 上述の行列モデルの円の違いは, 前者が共通因子だけに関わり, 円周と中心との隔たり $\mathbf{s}$ が確率的に変動するのに対して, 後者は, 共通因子・独自因子の両者からなる行列の性質を述べ, 円の半径は定数 $nm^{1/2}$ である点である.

7. スパース因子分析

行列因子分析が, 現行の定式化では見出しにくい因子分析の発展の方向を示す例として, 負荷行列がスパースであると制約する因子分析の研究開発がある. ここで, スパースな行列とは, その要素の多くが 0 であるが, どの要素が 0 であるかは特定されない行列を指す. こうした研究の背景には, Jolliffe *et al.* (2003) による最初の提案から現在まで, スパースなウェイト行列を求めるために, 主成分分析を修正した多くの手法が提案されてきた動向がある (Trendafilov (2014)). ここで, ウェイト行列とは, 変数を重みづける行列, すなわち, (6.1) における主成分得点 $\mathbf{F}$ は $\mathbf{F} = \mathbf{X}\mathbf{W}$ と表せるが, $\mathbf{W}$ を指す. 上記の修正手法は **スパース主成分分析** と総称されるが, それらの多くは, 非零の重みを罰則するペナルティ関数を従来の主成分分析の定式化に組み込むものであり, 代表的なペナルティ関数の一つに, ウェイト行列 $\mathbf{W}$ の l_1 ノルムがある (Zou *et al.* (2006)). ウェイト行列の要素の多くが 0 になれば, 非零要素に着目するだけの容易な解釈が可能になり, こうした解は因子分析の負荷行列にも望まれる. 従って, Yoshida and West (2010) や Hirose and Yamamoto (2014a, b) の方法を先駆として, スパースな負荷行列を求めるための **スパース因子分析** の開発が, 近年になってトレンドになりつつある.

上記の Hirose and Yamamoto (2014a, b) の方法は, 変量モデルに基づく尤度にペナルティ関数を組み込むアプローチであり, Yoshida and West (2010) の方法は, ベイズ法に基

づき、事前分布がペナルティ関数の役割を果たす。これらのスパース因子分析に対して、行列モデルに基づき、ペナルティ関数を使わないアプローチを Adachi and Trendafilov (2014) が提案している。この方法は、負荷行列の l_0 ノルム $\|\mathbf{\Lambda}\|_{l_0}$ つまり非零の負荷量の数を

$$\|\mathbf{\Lambda}\|_{l_0} = K \quad (\text{所与の整数}) \quad (7.1)$$

とする制約条件を行列因子分析に加えるだけで定式化され、目的関数 (3.5) の**平方和の分割**

$$f = \|\mathbf{X} - \mathbf{F}\mathbf{B}' - \mathbf{U}\mathbf{\Psi} - (\mathbf{F}\mathbf{\Lambda}' - \mathbf{F}\mathbf{B}')\|^2 = \|\mathbf{X} - \mathbf{F}\mathbf{B}' - \mathbf{U}\mathbf{\Psi}\|^2 + n\|\mathbf{\Lambda} - \mathbf{B}\|^2 \quad (7.2)$$

を利用することが要点になる。ここで、 $\mathbf{B} = n^{-1}\mathbf{X}'\mathbf{F}$ は p (変数) $\times$ m (共通因子) の共分散行列であり、条件 (3.4') を分割の導出に利用している。(7.2) の右辺で $\mathbf{\Lambda}$ と関わるのは単純な関数 $g(\mathbf{\Lambda}) = \|\mathbf{\Lambda} - \mathbf{B}\|^2$ だけであるため、 $\mathbf{F}$, $\mathbf{U}$, $\mathbf{\Psi}$ を一定として (7.2) を最小にする $\mathbf{\Lambda}$ を求める問題は、 $g(\mathbf{\Lambda})$ の最小化に帰着する。ただし、 $\mathbf{\Lambda}$ は (7.1) のように制約される。この制約つき最小化が容易であることは、 $\mathbf{\Lambda} = (\lambda_{jk})$, $\mathbf{B} = (b_{jk})$ とおけば、 $g(\mathbf{\Lambda})$ が

$$g(\mathbf{\Lambda}) = \sum_{j,k} (\lambda_{jk} - b_{jk})^2 = \sum_{(j,k) \in \mathbf{N}} b_{jk}^2 + \sum_{(j,k) \in \mathbf{N}^\perp} (\lambda_{jk} - b_{jk})^2 \geq \sum_{(j,k) \in \mathbf{N}} b_{jk}^2, \quad (7.3)$$

の不等式を満たすことからわかる。ここで、 $\mathbf{N}$ は $\lambda_{jk} = 0$ に対応する要素番号 (j, k) の集合、 $\mathbf{N}^\perp$ は $\lambda_{jk} \neq 0$ に対応する (j, k) の集合を表す。(7.3) より、 b_{jk}^2 ($j = 1, \dots, p; k = 1, \dots, m$) を降順に並べて、上位 K 個の b_{jk}^2 と要素番号が同じ λ_{jk} を $\lambda_{jk} = b_{jk}$ 、他の λ_{jk} を 0 とすれば、 $g(\mathbf{\Lambda})$ が下限 $\sum_{\mathbf{N}} b_{jk}^2$ を達成して、かつ、下限が最小化されることがわかる。この方法による $\mathbf{\Lambda}$ の更新、および、4.2 節に記した $\mathbf{\Psi}$ の更新の交互反復によって、スパース行列因子分析の $\mathbf{\Lambda}$ と $\mathbf{\Psi}$ の最適解が得られる。また、4.2 節のアルゴリズムと同様に、 $\mathbf{X}$ がなくとも $\mathbf{S}$ さえあれば十分で、反復中の $\mathbf{F}$, $\mathbf{U}$ の更新は不要である (Adachi and Trendafilov (2014))。

前段のスパース因子分析は、行列因子分析の目的関数 (3.5) が、(7.2) のように、 $\mathbf{\Lambda}$ の単純な関数 $g(\mathbf{\Lambda}) = \|\mathbf{\Lambda} - \mathbf{B}\|^2$ と $\mathbf{\Lambda}$ に無関係な項に分割できることが基礎となっている。こうした分割を現行の最小二乗法や最尤法の目的関数 (2.5), (2.6) から導くことは難しく、Adachi and Trendafilov (2014) のスパース因子分析は、行列因子分析の波及的効用を示す例といえる。

8. 結論

本論文では、列中心化された n 個体 $\times$ p 変数のデータ行列 $\mathbf{X}$ に対して、 $\min_{\mathbf{Z}, \mathbf{\Lambda}} \|\mathbf{X} - \mathbf{Z}\mathbf{A}'\|^2 = \|\mathbf{X} - (\mathbf{F}\mathbf{\Lambda}' + \mathbf{U}\mathbf{\Psi})\|^2$ と定式化される行列因子分析を論じた。ここで、 $\mathbf{Z} = [\mathbf{F}, \mathbf{U}]$ は共通因子行列 $\mathbf{F}$ と独自因子行列 $\mathbf{U}$ をブロックとする n 個体 $\times$ $(p + m)$ 因子の行列、 $\mathbf{A} = [\mathbf{\Lambda}, \mathbf{\Psi}]$ は負荷行列 $\mathbf{\Lambda}$ と独自分散の平方根を含む対角行列 $\mathbf{\Psi}$ からなる p 変数 $\times$ $(m + p)$

因子の行列であり、 $\mathbf{Z}$ は列中心化されて、 $n^{-1}\mathbf{Z}'\mathbf{Z} = \mathbf{I}_{m+p}$ を満たすと制約される。上記の目的関数 $\|\mathbf{X} - \mathbf{Z}\mathbf{A}'\|^2$ は、モデル部をデータ行列 $\mathbf{X}$ にフィットさせる形をとるが、第4節に記したように、変数 $\times$ 変数の標本共分散行列さえ与えられれば、最適化できる。その過程で、変数 $\times$ (共通・独自) 因子の共分散行列 $\mathbf{C} = n^{-1}\mathbf{X}'\mathbf{Z}$ が重要な役割を果たす。すなわち、行列因子分析のアルゴリズムは、 $\|\mathbf{C} - \mathbf{A}\|^2$ の $\mathbf{A}$ に関する最小化と、 $\text{tr } \mathbf{A}'\mathbf{C}$ の $\mathbf{C}$ に関する制約つき最大化の交互反復に帰着する。

第5節の数値例では、行列因子分析は、現在普及する最小二乗法・最尤法に比べて、小さい独自分散を示す傾向があるものの、ほぼ同等の解を与えることが示された。この結果は、行列因子分析が、因子分析の定式化のファミリーに含められることを、経験的に確認させる。さらに、行列因子分析は、第6節で論じたように、新たな理論的視座を提供する。その一つは、低階数近似としての主成分分析の目的関数 $\|\mathbf{X} - \mathbf{F}\mathbf{A}'\|^2$ に独自因子に関する項 $\mathbf{U}\mathbf{\Psi}$ を加えた $\|\mathbf{X} - (\mathbf{F}\mathbf{A}' + \mathbf{U}\mathbf{\Psi})\|^2$ が目的関数になる点で、因子分析と主成分分析の相違は、独自因子の有無に集約されるという視座である。また、因子分析が、低階数近似とは逆に高階数近似問題であるという視点も与える。すなわち、 $\mathbf{A}$ が所与のとき、行列因子分析が $\min_{\mathbf{Z}} \|\mathbf{Z} - \mathbf{X}\mathbf{A}'\|^2 + (\mathbf{Z}$ に無関係な項) と表せ、データ行列の線形関数である階数 p の $\mathbf{X}\mathbf{A}$ を、階数 $p+m$ の行列 $\mathbf{Z}$ で近似する問題と捉えられる。さらに、(6.4) のように $\mathbf{X}$ の関数として定義される $\mathbf{Z}_1$ と $\mathbf{Z}$ の解は、 $\|\mathbf{Z} - \mathbf{Z}_1\|^2 = nm$ のように一定の距離を保ち、因子の不確定性に関する新たな表現を与える。

行列因子分析を発展させた方法として、第7節に、ペナルティ関数を使わないスパース因子分析法を紹介したが、その他にも、行列因子分析を出発点とした因子分析の新展開が期待できる。

謝辞

本稿は、2014年3月に第8回日本統計学会春季集会(同志社大学)において、著者が企画したセッション「行列分解による多変量データ解析」の中の演題の一つとして、著者が行った講演に加筆したものである。セッション企画をお勧めいただいた集会のプログラム委員長の宿久洋先生、編集委員長の谷崎久志先生、講演後に質問・コメントをいただいた先生、匿名の査読者の先生に感謝申し上げます。そして、多変量データ解析法への行列集約アプローチに著者が興味を抱く契機を、著書を通して与えていただいた(故)柳井晴夫先生に謝意を表します。

参考文献

- 足立浩平(2006).『多変量データ解析法—心理・教育・社会系のための入門—』ナカニシヤ出版.
- Adachi, K. (2012). Some contributions to data-fitting factor analysis with empirical comparisons to covariance-fitting factor analysis, *J. Jpn. Soc. Comp. Statist.*, **25**, 25–38.

- Adachi, K. (2013). Factor analysis with EM algorithm never gives improper solutions when sample covariance and initial parameter matrices are proper, *Psychometrika*, **78**, 380–394.
- 足立浩平, 村上 隆 (2011). 『非計量多変量解析法—主成分分析から多重対応分析へ—』朝倉書店.
- Adachi, K. and Trendafilov, N. T. (2014). Sparse orthogonal factor analysis, *Advances in Latent Variables: Methods, Models and Applications*, pp. 227–239, Springer.
- Anderson, T. W. and Rubin, H. (1956). Statistical inference in factor analysis, *Proc. Third Berkeley Symp. Math. Statist. Prob.*, vol. 5, (ed. J. Neyman), pp. 111–150, University of California Press, Berkeley, CA.
- Bentler, P. M. and Kano, Y. (1990). On the equivalence of factors and components, *Multivariate Behav. Res.*, **25**, 67–74.
- Browne, M. (2001). An overview of analytic rotation in exploratory factor analysis, *Multivariate Behav. Res.*, **36**, 111–150.
- de Leeuw, J. (2004). Least squares optimal scaling of partially observed linear systems, *Recent Developments of Structural Equation Models: Theory and Applications*, (eds. K. van Montfort, J. Oud and A. Satorra), pp. 121–134, Kluwer Academic Publishers, Dordrecht.
- Eckart, C. and Young, G. (1936). The approximation of one matrix by another of lower rank, *Psychometrika*, **1**, 211–218.
- Harman, H. H. (1976). *Modern Factor Analysis*, 3rd ed., Chicago, The University of Chicago Press.
- Harman, H. H. and Jones, W. H. (1966). Factor analysis by minimizing residuals (Minres), *Psychometrika*, **31**, 351–369.
- Hirose, K. and Yamamoto, M. (2014a). Sparse estimation via nonconcave penalized likelihood in factor analysis model, *Statist. Comput.*, <http://dx.doi.org/10.1007/s11222-014-9475-z> (to appear).
- Hirose, K. and Yamamoto, M. (2014b). Estimation of an oblique structure via penalized likelihood factor analysis, *Comput. Statist. Data Anal.*, **79**, 120–132.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components, *J. Educ. Psychol.*, **26**, 417–441, 498–520.
- 市川雅教 (2010). 『因子分析』朝倉書店.
- Jolliffe, I. T. (2002). *Principal Component Analysis*, 2nd ed., Springer, New York.
- Jolliffe, I. T., Trendafilov, N. T. and Uddin, M. (2003). A modified principal component technique based on the LASSO, *J. Comput. Graph. Statist.*, **12**, 531–547.
- 狩野 裕, 三浦麻子 (2002). 『AMOS, EQS, LISREL によるグラフィカル多変量解析 (増補版)—目で見る共分散構造分析—』現代数学社.
- 柏木繁男 (1997). 『性格の評価と表現—特性 5 因子論からのアプローチ—』有斐閣.
- Lawly, D. N. and Maxwell, A. E. (1971). *Factor Analysis as a Statistical Method*, 2nd ed., London, Butterworths.
- McDonald, R. P. (1979). The simultaneous estimation of factor loadings and scores, *Brit. J. Math. Statist. Psychol.*, **32**, 212–228.
- Mulaik, S. A. (1976). Comments on “The measurement of factorial indeterminacy”, *Psychometrika*, **41**, 249–262.
- Mulaik, S. A. (2010). *Foundations of Factor Analysis*, 2nd ed., CRC Press, Boca Raton.
- Ogasawara, H. (2000). Some relationships between factors and components, *Psychometrika*, **65**, 167–185.
- Okamoto, K. (1969). Optimality of principal components, *Multivariate Analysis*, vol. II, (ed. P. R. Krishnaiah), pp. 673–687, Academic Press, New York.
- 丘本 正 (1986). 『因子分析の基礎』日科技連.
- 奥井 亮 (2014). 「因子モデルに関する近年の計量経済学研究の進展」『日本統計学会誌』**43**, 247–273.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space, *Phil. Mag.*, **2**, 559–572.
- Rao, C. R. (1955). Estimation and tests of significance in factor analysis, *Psychometrika*, **20**, 93–111.
- Rubin, D. B. and Thayer, D. T. (1982). EM algorithms for ML factor analysis, *Psychometrika*, **47**, 69–76.
- SAS Institute Inc. (2009). *SAS/STAT® 9.2 Users Guide*, version, 2nd ed., Cary NC, SAS Institute Inc.

- Sato, M. (1987). A study of an identification problem and substitute use of principal component analysis in factor analysis, *Horoshima Mathematical Journal*, **39**, 479–527.
- 芝 祐順 (1979). 『因子分析法』 (第2版), 東京大学出版会.
- 繁榘算男 (1998). 『心理測定法』 日本放送出版協会.
- Sočan, G. (2003). *The Incremental Value of Minimum Rank Factor Analysis*, PhD Thesis, University of Groningen, Groningen.
- Spearman, C. (1904). ‘General intelligence’ objectively determined and measured, *Amer. J. Psychol.*, **15**, 201–293.
- SPSS Inc. (1997). *SPSS 7.5 Statistical Algorithm*, SPSS Inc., Chicago.
- Takane, Y. (2014). *Constrained Principal Component Analysis and Related Techniques*, CRC Press, Boca Raton.
- ten Berge, J. M. F. (1983). A generalization of Kristof’s theorem on the trace of certain matrix products, *Psychometrika*, **48**, 519–523.
- Thurstone, L. L. (1935). *The Vectors of Mind*, University of Chicago Press, Chicago.
- 豊田秀樹 (2012). 『因子分析入門—Rで学ぶ最新データ解析—』 東京図書.
- Trendafilov, N. T. (2014). From simple structure to sparse components: A review, *Comput. Statist.*, **29**, 431–454.
- Unkel, S. and Trendafilov, N. T. (2010a). Simultaneous parameter estimation in exploratory factor analysis: An expository review, *Int. Statist. Rev.*, **78**, 363–382.
- Unkel, S. and Trendafilov, N. T. (2010b). A majorization algorithm for simultaneous parameter estimation in robust exploratory factor analysis, *Comput. Statist. Data Anal.*, **54**, 3348–3358.
- Yanai, H. and Ichikawa, M. (2007). Factor analysis, *Handbook of statistics vol. 26: Psychometrics*, (eds. C. R. Rao and S. Sinharay), pp. 257–296, Elsevier, Amsterdam.
- 柳井晴夫, 繁榘算男, 前川眞一, 市川雅教 (1990). 『因子分析—その理論と方法—』 朝倉書店.
- Yoshida, R. and West, M. (2010). Bayesian learning in sparse graphical factor models via variational mean-field annealing, *J. Mach. Learn. Res.*, **11**, 1771–1798.
- Zou, D. M., Hastie, T. and Tibshirani, R. (2006). Sparse principal component analysis, *J. Comput. Graph. Statist.*, **15**, 265–286.