因子分析へのグレブナー基底に基づくアプローチ

A Gröbner basis-based approach to factor analysis

九州大学　深作亮也 *1
Ryoya Fukasaku
Kyushu University

九州大学　廣瀬慧 *2
Kei Hirose
Kyushu University

長崎大学　加葉田雄太朗 *3
Yutaro Kabata
Nagasaki University

広島大学　寺本圭佑 *4
Keisuke Teramoto
Hiroshima University

1　はじめに

因子分析は, 実際のデータから直接観測できないような「共通因子」と呼ばれる潜在変数を推定する統計解析手法である. そして, この共通因子の推定結果が, 実証分析等の応用の場において重要な役割を果たす. 因子分析は, 100 年以上前の 1904 年に C. Spearman 氏によって考えられた手法であり, もともとは心理学や社会科学に応用されてきたが, 最近は生命科学, 経済学, マーケティング, 脳科学等でも応用されている. 本稿では, 因子分析モデルの最尤解の候補全てをグレブナー基底によって厳密に計算していく. そして, 因子分析における最尤法の計算課題「不適解問題」を厳密な計算結果によって紐解いていく.

不適解問題は因子分析を応用する実際の現場で発生してしまい, モデルの解釈に疑問を残す課題である. これは, 最尤法による分散パラメータが負となる計算課題 (ただし, 本稿では, 真に零となる分散パラメータを含む推定値も不適解と呼ぶことにする: 詳細 節 4) であり, 解空間の次元との関わりが指摘されている [6]. しかしながら, これはあくまで近似的・経験的にわかってきたものであり, 近似計算では未知な部分が未だ多い. 本研究では, グレブナー基底や代数円柱分解などの計算代数手法によって, 代数解を厳密に計算し, 不適解問題の解明を目指す. とくに, 本稿では, 因子分析モデルの最尤解の候補をグレブナー基底や代数円柱分解によって厳密に算出する.

*1 〒 819-0395 福岡県福岡市西区元岡 744　E-mail: fukasaku@math.kyushu-u.ac.jp
*2 〒 819-0395 福岡県福岡市西区元岡 744　E-mail: hirose@imi.kyushu-u.ac.jp
*3 〒 852-8521 長崎県長崎市文教町 1-14　E-mail: kabata@nagasaki-u.ac.jp
*4 〒 739-8526 広島県東広島市鏡山 1-3-1　E-mail: kteramoto@hiroshima-u.ac.jp

本稿の構成は次の通りである: 2 節では因子分析モデルにおける最尤法の理論背景を紹介する. 3 節では因子負荷行列の不定性について触れる. 4 節では不適解問題を説明し, グレブナー基底の活用について考える. 5 節ではグレブナー基底を用いた代数計算による実験結果を報告する.

2 因子分析モデルと最尤法

因子分析モデルでは, $i = 1, \ldots, p$, $j = 1, \ldots, k$ (ただし $k < p$) に対し, 観測変数 X_i とその平均 μ_i, 及び, 二種類の潜在変数 ε_i, F_j が次の関係式を満たすと仮定する:

$$X_i - \mu_i = \ell_{i1}F_1 + \cdots + \ell_{ik}F_k + \varepsilon_i,$$

潜在変数 ε_i は各 X_i に独自に影響し, 潜在変数 F_j は全 X_i に共通で影響し, ℓ_{ij} は X_i への F_j の負荷を表す. そして, 因子分析モデルでは, 各 ε_i が**独自因子**, 各 F_j が**共通因子**, 各 ℓ_{ij} が**因子負荷量**と呼ばれる.

観測変数ベクトル $\boldsymbol{X} = (X_1, \cdots, X_p)^{\mathrm{T}}$, 観測変数の平均ベクトル $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_p)^{\mathrm{T}}$, 独自因子ベクトル $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_p)^{\mathrm{T}}$, 共通因子ベクトル $\boldsymbol{F} = (F_1, \cdots, F_k)^{\mathrm{T}}$, 及び因子負荷行列 $L = (\ell_{ij})$ を用いることにより, 上の関係式は次のように表される:

$$\boldsymbol{X} - \boldsymbol{\mu} = L\boldsymbol{F} + \boldsymbol{\varepsilon}.$$

また, 本稿では, 共通因子ベクトル $\boldsymbol{F}$ と独自因子ベクトル $\boldsymbol{\varepsilon}$ が, 各々, 次の多変量正規分布に従うと仮定する:

$$\boldsymbol{F} \sim N_k(\boldsymbol{0}, I), \quad \boldsymbol{\varepsilon} \sim N_k(\boldsymbol{0}, \Psi), \tag{1}$$

ただし, Ψ を対角行列 $\mathrm{diag}(\psi_1, \ldots, \psi_p)$ とし, **独自分散行列**と呼ぶ. さらに, $\boldsymbol{F}, \boldsymbol{\varepsilon}$ は無相関とする, つまり,

$$\mathrm{Cov}(\boldsymbol{F}, \boldsymbol{\varepsilon}) = 0. \tag{2}$$

仮定 (1), (2) の下で, 観測変数 $\boldsymbol{X}$ は次の p 次元正規分布に従う (参照: [1, 2.1 節]):

$$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, LL^{\mathrm{T}} + \Psi)$$

従って, 多変量観測データ $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in \mathbb{R}^p$ が与えられたときに, 尤度関数 f は次のように表される:

$$\begin{aligned} f &= \prod_{n=1}^{N} (2\pi)^{-\frac{p}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}_n - \boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu})\right) \\ &= (2\pi)^{-\frac{Np}{2}} \det(\Sigma)^{-\frac{N}{2}} \exp\left(-\frac{1}{2}\sum_{n=1}^{N}(\boldsymbol{x}_n - \boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu})\right), \end{aligned}$$

ただし, $\Sigma = LL^{\mathrm{T}} + \Psi$ であり, 尤度関数の導出に Σ の正則性が仮定されることに注意する. こうして, 対数尤度関数 $l(\boldsymbol{\mu}, L, \Psi | \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$ は

$$\begin{aligned} l(\boldsymbol{\mu}, L, \Psi | \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) &= -\frac{Np}{2}\log(2\pi) - \frac{N}{2}\log(\det(\Sigma)) - \frac{1}{2}\sum_{n=1}^{N}(\boldsymbol{x}_n - \boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}_n - \boldsymbol{\mu}) \\ &= -\frac{Np}{2}\log(2\pi) - \frac{N}{2}\log(\det(\Sigma)) - \frac{1}{2}\left(N \cdot \mathrm{tr}(\Sigma^{-1}S) + N(\bar{\boldsymbol{x}} - \boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu})\right) \end{aligned}$$

となる, ここで, $\bar{\boldsymbol{x}}, S$ は次のように, 各々, 標本平均ベクトル, 標本分散共分散行列である:

$$\bar{\boldsymbol{x}} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{x}_n,$$
$$S = \frac{1}{N}\sum_{n=1}^{N}(\boldsymbol{x}_n - \bar{\boldsymbol{x}})(\boldsymbol{x}_n - \bar{\boldsymbol{x}})^{\mathrm{T}}.$$

最尤法では, 与えられた $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in \mathbb{R}^p$ に対して上の対数尤度関数 $l(\boldsymbol{\mu}, L, \Psi | \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$ が最大となるパラメータ $\hat{\boldsymbol{\mu}}, \hat{L}, \hat{\Psi}$ を求める. まず, $\boldsymbol{\mu}$ の最尤推定値は $\hat{\boldsymbol{\mu}} = \bar{\boldsymbol{x}}$ で与えられる. 従って, 以降では, 一般性を失うことなく $\boldsymbol{\mu} = \boldsymbol{0}$ とした上で, データの中心化を施した $\boldsymbol{x}_n - \bar{\boldsymbol{x}}$ を用いる. 次に, $\hat{\boldsymbol{\mu}} = \bar{\boldsymbol{x}}$ を代入した上の対数尤度関数

$$l(\hat{\boldsymbol{\mu}}, L, \Psi | \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) = -\frac{Np}{2}\log(2\pi) - \frac{N}{2}\log(\det(\Sigma)) - \frac{N}{2}\mathrm{tr}(\Sigma^{-1}S)$$

を最大にする Ψ, L を求める. この関数の Ψ, L による微分を考えると, 次を得る (参照: [1, 3.1.1 節]):

$$\frac{\partial l}{\partial \Psi} = 0 \quad \Leftrightarrow \quad \Psi = \mathrm{diag}(S - LL^{\mathrm{T}}), \tag{3}$$
$$\frac{\partial l}{\partial L} = 0 \quad \Leftrightarrow \quad \Sigma^{-1}(S - \Sigma)\Sigma^{-1}L = 0. \tag{4}$$

こうして, 最尤推定量 $\hat{\mu}, \hat{\Psi}, \hat{L}$ の計算が, (3), (4) の右式を満たす因子負荷行列 L の計算に帰着される.

注意 1
関係式 *(3)* の右式は, 独自分散対角行列 Ψ の正則性を仮定した上で得られているという事実に, 注意する. また, 関係式 *(3)* によって, 各 ψ_i は次の形をとる変数 ℓ_{ij} の実係数多変数多項式として表すことができる:

$$\psi_i = s_{ii} - \sum_{j=1}^{k}\ell_{ij}^2 \in \mathbb{R}[\ell_{ij} : 1 \leq i \leq p, 1 \leq j \leq k].$$

注意 2
観測変数分散共分散行列 Σ と独自分散対角行列 Ψ だけでなく, 標本分散共分散行列 S も正則性を仮定する. このとき, 次の等価関係たちが得られる:

$$(4) \quad \Leftrightarrow \quad (S - \Sigma)\Sigma^{-1}L = 0 \tag{5}$$
$$\Leftrightarrow \quad (S - \Sigma)\Psi^{-1}L = 0 \tag{6}$$
$$\Leftrightarrow \quad (S - \Sigma)S^{-1}L = 0. \tag{7}$$

まず, 注意 *1* より, Ψ は次のように変数 ℓ_{ij} の実係数多変数多項式による対角行列として表される:

$$\Psi = \mathrm{diag}\left(s_{11} - \sum_{j=1}^{k}\ell_{1j}^2, \ldots, s_{pp} - \sum_{j=1}^{k}\ell_{pj}^2\right) \in \mathbb{R}[\ell_{ij} : 1 \leq i \leq p, 1 \leq j \leq k]^{p \times p}.$$

次に, $\Sigma = LL^{\mathrm{T}} + \Psi$ なので, 上の表現より, Σ の (i_1, i_2) 成分は変数 ℓ_{ij} の次の実係数多変数多項式である:

$$\begin{aligned}\Sigma_{i_1 i_2} &= \sum_{j=1}^{k}\ell_{i_1 j}\ell_{i_2 j} + \delta_{i_1 i_2}(s_{i_1 i_2} - \sum_{j=1}^{k}\ell_{i_1 j}\ell_{i_2 j}) \\ &= \begin{cases} s_{i_1 i_2} & (i_1 = i_2) \\ \sum_{j=1}^{k}\ell_{i_1 j}\ell_{i_2 j} & (i_1 \neq i_2) \end{cases}.\end{aligned}$$

関係式 *(5)* が Σ の逆行列を含むので, *(5)* は非常に複雑な表現の実係数有理関数たちにより表されてしまう. 次に, 関係式 *(6)* は Ψ の逆行列を含むことから, 関係式 *(5)* と同様に, *(6)* は実係数有理関数たちで表される. その分子多項式たちは通分で複雑化し, 結果として, 関係式 *(6)* は複雑な連立方程式たちで表されてしまう. 一方, 関係式 *(7)* は標本分散共分散行列を用いているだけなので, 有理関数は一切現れないことに注意する. つまり, 関係式 *(7)* は変数 ℓ_{ij} の実係数多変数多項式のみで表されるような連立方程式となる.

最尤法では, 関係式 *(3), (6)* 等を用いることによって漸化式を与えて近似計算アルゴリズムを設計する. 近似計算よりも計算資源を要求し易いグレブナー基底計算は簡略な入力でも膨大な計算資源が要求される. もちろん, 複雑な多項式より簡略な多項式の方がグレブナー基底計算では少ない計算資源で済むことが多い. そのため, 本稿では, 有理関数が現れない, 標本分散共分散行列 S の逆行列を含む関係式 *(7)* を用いていく. つまり, 与えられた標本分散共分散行列 $S \in \mathbb{R}^{p\times p}$ に対して, 次を満たす全ての実係数行列 L を計算する:

$$(S - (LL^{\mathrm{T}} + \mathrm{diag}(S - LL^{\mathrm{T}})))S^{-1}L = 0. \tag{8}$$

最尤法の漸化式の導出において独自分散対角行列 Ψ の正則性を仮定していたので, 本稿では実験のために, 次のように, その正則性で出力結果を区別する.

表 1: 出力結果の分類

タイプ $\Psi_{\mathrm{nonregular}}$	非正則な (零である独自分散 ψ_i を持つ) 独自分散行列 Ψ
タイプ Ψ_{regular}	正則な (零である独自分散 ψ_i を持たない) 独自分散行列 Ψ

3 回転の不定性

因子負荷行列は一般に一意ではない. 実際, 任意の直交行列 T に対し $L^* = LT, \boldsymbol{F}^* = T^{\mathrm{T}}\boldsymbol{F}$ とおくことで

$$\boldsymbol{X} - \boldsymbol{\mu} = L\boldsymbol{F} + \boldsymbol{\varepsilon} = L^*\boldsymbol{F}^* + \boldsymbol{\varepsilon}, \quad \boldsymbol{X} \sim N_p(\boldsymbol{\mu}, LL^{\mathrm{T}} + \Psi) = N_p(\boldsymbol{\mu}, (L^*)(L^*)^{\mathrm{T}} + \Psi)$$

のように, 異なる因子分析モデルから同一の観測変数の分布が得られてしまう. このような不定性は**回転の不定性**として知られており, 一意性を与える条件として, 例えば, 次の命題の十分条件等が知られている [2].

命題 1
次のいずれかを仮定すると, 列の正負の反転を除いて, 因子負荷行列 L は一意に定まる:

1. L の上三角部分の成分が全て零であり, かつ, 対角成分は全て非零である,
2. $L^{\mathrm{T}}\Psi L$ は成分が全て異なる対角行列である.

本稿で報告する実験においては, 上の命題の条件 1 の前半を仮定する. つまり, 連立方程式 (8) を満たして,

$$L = \left(\begin{array}{ccccc}
\ell_{11} & & & & \\
\ell_{21} & \ell_{22} & & \mathrm{O} & \\
\vdots & \vdots & \ddots & & \\
\ell_{k-1,1} & \cdots & \cdots & \ell_{k-1,k-1} & \\
\ell_{k1} & \cdots & \cdots & \ell_{k,k-1} & \ell_{kk} \\
\hline
\ell_{k+1,1} & \cdots & \cdots & \ell_{k+1,k-1} & \ell_{k+1,k} \\
\ell_{k+2,1} & \cdots & \cdots & \ell_{k+2,k-1} & \ell_{k+2,k} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\ell_{p1} & \cdots & \cdots & \ell_{p,k-1} & \ell_{pk}
\end{array}\right)$$

のような形を持つ全ての実係数行列 L を計算する．ここでは, 上の命題の条件 1 の後半の条件は仮定しない.

4　不適解問題とグレブナー基底

不適解問題は, 因子分析モデルに最尤法を適用した際に零未満の独自分散が推定される計算課題である. その最尤法の漸化式は独自分散対角行列の正則性を仮定し得られていた．また, 因子分析が元々は心理学で利用されていたという背景から零となってしまう独自分散も, 因子分析モデルの解釈に疑問を残してしまう. そこで, 本稿では零以下の独自分散が推定されたとき, つまり, 次を満たすとき, 推定値 L, Ψ を**不適解**と呼ぶ:

$$\exists i \in \{1, 2, \dots, p\}\ (\psi_i \leq 0).$$

また, 独自分散行列 Ψ の最尤推定量は関係式 (3) の右式で与えられていたので, 標本分散共分散行列 S に対して次を満たす因子負荷行列 L は**不適**ということにして, そうでない場合は**非不適**ということにする:

$$\exists i \in \{1, 2, \dots, p\}\ \left(s_{ii} - \sum_{j=1}^{k} \ell_{ij}^2 \leq 0 \right).$$

ここで, [6] で分類された不適解を紹介する:

- 識別不能性: 回転の不定性を除いても, 因子負荷行列 L が一意でない場合など,
- 標本変動: 真の独自因子の分散は全て正であるが, 負の独自因子の推定値が計算された場合など,
- モデルの不適合: モデルに合わない変数を含む場合など.

まず, 識別不能性に関連して, 次の結果が知られている [2].

命題 2
回転の不定性を除いて因子負荷行列 L が一意ならば, 任意の正則行列 A に対し LA の各列は三個以上の非零成分を持つ.

注意 3
上の命題の対偶によって, ある正則行列 A に対して三個以上の非零成分を持たない列を LA が含む場合に, 回転の不定性を除いても因子負荷行列 L は一意でないことがわかる．例えば, A として単位行列を取れば, 三個以上の非零成分を持たない列を L が含む場合, 回転の不定性を除いても因子負荷行列 L は一意でない.

次に, 標本変動については共通性 $(\ell_{i1}^2 + \ell_{i2}^2 + \cdots \ell_{ik}^2)/(\ell_{i1}^2 + \ell_{i2}^2 + \cdots \ell_{ik}^2 + \psi_i)$ が小さい場合に起こりやすいことが知られている．そこで, 本稿の実験では, 識別不能性が起こり易いタイプ S_1, 標本変動が起こり易いタイプ S_2, どちらも起こりにくいタイプ S_3 の標本分散共分散行列 S を, 各々, 次の因子負荷行列 L_1, L_2, L_3 を真に持つ標本として, 標本数 $N = 100$, 正規分布 $N(0, LL^{\mathrm{T}} + \Psi)$ の下で, 生成する.

表 2: 入力のための分類

タイプ S_1	因子負荷行列 L_1 を真に持つ, 識別不能性が起こりやすい標本分散共分散行列 S
タイプ S_2	因子負荷行列 L_2 を真に持つ, 標本変動が起こりやすい標本分散共分散行列 S
タイプ S_3	因子負荷行列 L_3 を真に持つ, 識別不能性・標本変動が起こりにくい標本分散共分散行列 S

$$L_1 = \begin{pmatrix} 0.9 & 0 \\ 0.8 & 0 \\ 0.7 & 0 \\ 0 & 0.8 \\ 0 & 0.7 \end{pmatrix}, \quad L_2 = \begin{pmatrix} 0.5 & 0 \\ 0.5 & 0 \\ 0.5 & 0.5 \\ 0 & 0.5 \\ 0 & 0.5 \end{pmatrix}, \quad L_3 = \begin{pmatrix} 0.9 & 0 \\ 0.8 & 0 \\ 0.6 & 0.7 \\ 0 & 0.8 \\ 0 & 0.9 \end{pmatrix}.$$

グレブナー基底によって, 各標本分散共分散行列を与えた際の因子負荷行列 L の候補を具体的に計算する. 特に, その計算では, 標本 $S \in \mathbb{R}^{p \times p}$ を与え, 連立方程式 (8) を次の多変数多項式イデアルとして扱う:

$$\begin{aligned} I &= \langle f : f \in (S - (LL^{\mathrm{T}} + \mathrm{diag}(S - LL^{\mathrm{T}})))S^{-1} \rangle \\ &= \left\{ \sum_{f \in (S - (LL^{\mathrm{T}} + \mathrm{diag}(S - LL^{\mathrm{T}})))S^{-1}} f h_f : h_f \in \mathbb{R}[\ell_{ij} : 1 \le i \le p, 1 \le j \le k] \right\} \end{aligned}$$

準素イデアル分解により I の解空間を自動で既約な解空間に分解できるが, 残念ながら, その計算は重い. そこで, 和イデアルや飽和イデアルを用いることで, 連立方程式 (8) の解空間を手動で分解することを考える.

まず, タイプ $\Psi_{\mathrm{nonregular}}, \Psi_{\mathrm{regular}}$ の解空間を分解するため次の和イデアル J と飽和イデアル K を考える:

$$\begin{aligned} J &= I + \langle \det(\mathrm{diag}(S - LL^{\mathrm{T}})) \rangle \\ &= \{ f_1 + f_2 : f_1 \in I, f_2 \in \langle \det(\mathrm{diag}(S - LL^{\mathrm{T}})) \rangle \}, \\ K &= I : \langle \det(\mathrm{diag}(S - LL^{\mathrm{T}})) \rangle^{\infty} \\ &= \left\{ f : \forall g \in \langle \det(\mathrm{diag}(S - LL^{\mathrm{T}})) \rangle \exists m \in \mathbb{N} \, (f g^m \in I) \right\}. \end{aligned}$$

ここでは, 連立方程式 (8) を与えた際に用いた関係式 (3) の右式を思い出し, 次の関係に注意して欲しい.

注意 4
和イデアル J の複素解空間 $\mathbb{V}_{\mathbb{C}}(J) = \{ L = (\ell_{ij}) \in \mathbb{C}^{p \times k} \mid \forall f \in J \, (f(\ell_{ij} : i, j) = 0) \}$ は連立方程式

$$(8) \wedge \det(\mathrm{diag}(S - LL^{\mathrm{T}})) = 0 \tag{9}$$

の解空間に一致する *(参照: [4, 4 章 3 節, 定理 4])*.

注意 5
飽和イデアル K の複素解空間 $\mathbb{V}_{\mathbb{C}}(K) = \{ L = (\ell_{ij}) \in \mathbb{C}^{p \times k} \mid \forall f \in K \, (f(\ell_{ij} : i, j) = 0) \}$ は連立方程式

$$(8) \wedge \det(\mathrm{diag}(S - LL^{\mathrm{T}})) \neq 0 \tag{10}$$

の解空間のザリスキ閉包 *(*つまり, 等式制約のみの連立方程式の解集合として表すことができる, 与えられた集合を含む包含関係最小の集合*)* に一致する *(参照: [4, 4 章 4 節, 定理 10])*.

上のモンテカルロシミュレーションのように, 本稿の実験では $p = 5, k = 2$ なる標本分散共分散行列を扱う. つまり, 与えられた標本分散共分散行列 S に対して (8) を満たす, 上三角部分が零である次の L を求めたい:

$$L = \begin{pmatrix} \ell_{11} & 0 \\ \ell_{21} & \ell_{22} \\ \ell_{31} & \ell_{32} \\ \ell_{41} & \ell_{42} \\ \ell_{51} & \ell_{52} \end{pmatrix}.$$

因子負荷行列 L の上三角部分を零とした仮定によって, (8) の導出に矛盾しない解空間 $\mathbb{V}_{\mathbb{C}}(K)$ を持つイデアル K は零次元イデアル (つまり, $\mathbb{V}_{\mathbb{C}}(K)$ は有限集合; 参照: [4, 5 章 3 節, 定理 6]) になり易いと思われる. しかし, その生成元は複雑になり得るので, 連立方程式 (10) を L の小行列式の振る舞いに基づいて分解する:

$$(10) \wedge \ell_{11} = 0 \tag{11}$$
$$(10) \wedge \ell_{22} = 0 \quad \wedge \ell_{11} \neq 0 \tag{12}$$
$$(10) \wedge \ell_{32} = 0 \quad \wedge \ell_{11}\ell_{22} \neq 0 \tag{13}$$
$$(10) \wedge \ell_{42} = 0 \quad \wedge \ell_{11}\ell_{22}\ell_{32} \neq 0 \tag{14}$$
$$(10) \wedge \ell_{52} = 0 \quad \wedge \ell_{11}\ell_{22}\ell_{32}\ell_{42} \neq 0 \tag{15}$$
$$(10) \wedge m_1 = 0 \quad \wedge \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52} \neq 0 \tag{16}$$
$$(10) \wedge m_2 = 0 \quad \wedge \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1 \neq 0 \tag{17}$$
$$(10) \wedge m_3 = 0 \quad \wedge \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2 \neq 0 \tag{18}$$
$$(10) \wedge m_4 = 0 \quad \wedge \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2m_3 \neq 0 \tag{19}$$
$$(10) \wedge m_5 = 0 \quad \wedge \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2m_3m_4 \neq 0 \tag{20}$$
$$(10) \wedge m_6 = 0 \quad \wedge \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2m_3m_4m_5 \neq 0 \tag{21}$$
$$(10) \quad \wedge \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2m_3m_4m_5m_6 \neq 0 \tag{22}$$

ただし $m_1 = \ell_{21}\ell_{32} - \ell_{31}\ell_{22}$, $m_2 = \ell_{21}\ell_{42} - \ell_{41}\ell_{22}$, $m_3 = \ell_{21}\ell_{52} - \ell_{51}\ell_{22}$, $m_4 = \ell_{31}\ell_{42} - \ell_{41}\ell_{32}$, $m_5 = \ell_{31}\ell_{52} - \ell_{51}\ell_{32}$, $m_6 = \ell_{41}\ell_{52} - \ell_{51}\ell_{42}$. 上のイデアルは零次元イデアルの場合が多いと予想されるので, 根基イデアル計算も難しくないと思われる. そこで, 代数解を計算し易くするため, 和イデアルと飽和イデアルで上に関する根基イデアルを考える:

$$K_1 = \sqrt{K + \langle \ell_{11} \rangle} \tag{11'}$$
$$K_2 = \sqrt{(K + \langle \ell_{22} \rangle) : \langle \ell_{11} \rangle^\infty} \tag{12'}$$
$$K_3 = \sqrt{(K + \langle \ell_{32} \rangle) : \langle \ell_{11}\ell_{22} \rangle^\infty} \tag{13'}$$
$$K_4 = \sqrt{(K + \langle \ell_{42} \rangle) : \langle \ell_{11}\ell_{22}\ell_{32} \rangle^\infty} \tag{14'}$$
$$K_5 = \sqrt{(K + \langle \ell_{52} \rangle) : \langle \ell_{11}\ell_{22}\ell_{32}\ell_{42} \rangle^\infty} \tag{15'}$$
$$K_6 = \sqrt{(K + \langle m_1 \rangle) : \langle \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52} \rangle^\infty} \tag{16'}$$
$$K_7 = \sqrt{(K + \langle m_2 \rangle) : \langle \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1 \rangle^\infty} \tag{17'}$$
$$K_8 = \sqrt{(K + \langle m_3 \rangle) : \langle \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2 \rangle^\infty} \tag{18'}$$
$$K_9 = \sqrt{(K + \langle m_4 \rangle) : \langle \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2m_3 \rangle^\infty} \tag{19'}$$
$$K_{10} = \sqrt{(K + \langle m_5 \rangle) : \langle \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2m_3m_4 \rangle^\infty} \tag{20'}$$
$$K_{11} = \sqrt{(K + \langle m_6 \rangle) : \langle \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2m_3m_4m_5 \rangle^\infty} \tag{21'}$$
$$K_{12} = \sqrt{K : \langle \ell_{11}\ell_{22}\ell_{32}\ell_{42}\ell_{52}m_1m_2m_3m_4m_5m_6 \rangle^\infty} \tag{22'}$$

ただし, イデアル $A \subset \mathbb{R}[\ell_{ij} : 1 \leq i \leq p, 1 \leq j \leq k]$ の根基イデアルは次で定義されることに注意する:

$$\sqrt{A} = \{f \in \mathbb{R}[\ell_{ij} : 1 \leq i \leq p, 1 \leq j \leq k] : \exists m \in \mathbb{N}\ (f^m \in A)\}$$

以降では, 集合 $T \subset \mathbb{C}^{p\times k}$ のザリスキ閉包 (つまり, 等式制約のみの連立方程式の解集合として表すことができる, T を含む包含関係最小の集合) を $\overline{T}$ で表す. このとき, 上と同様に, 次の関係に注意して欲しい.

注意 6

イデアル $K_1, K_2, \ldots, K_{12}$ の解空間は, 各々, 次のザリスキ閉包と一致する.

$$\mathbb{V}_{\mathbb{C}}(K_1) = \overline{\{L \in \mathbb{C}^{p\times k} \mid (11)\}}, \quad \mathbb{V}_{\mathbb{C}}(K_2) = \overline{\{L \in \mathbb{C}^{p\times k} \mid (12)\}}, \quad \ldots, \quad \mathbb{V}_{\mathbb{C}}(K_{12}) = \overline{\{L \in \mathbb{C}^{p\times k} \mid (22)\}}.$$

我々の実装では, 多くの標本分散共分散行列 S において, イデアル $K_1, K_2, \ldots, K_{12}$ の辞書式順序のグレブナー基底を計算でき, そのグレブナー基底から代数解を具体的に計算することができたが, イデアル K_2 の辞書式順序のグレブナー基底から代数解を計算する際にメモリや時間の消費が大きい標本もいくつかあった.

ここで, 方程式 (12) は (13) から (21) に比べて不等式制約が少なく, 解空間が広いことに注意して欲しい. また, モンテカルロシミュレーションで計算した標本分散共分散行列の真の因子負荷行列は $\ell_{22} = 0$ である. つまり, 方程式 (12) の解空間が複雑であることが, 上のようなメモリや計算時間の浪費の原因であることが推測できるので, 方程式 (12) を更に分解する:

$$\begin{cases} (12) \wedge \ell_{32} = 0 \wedge \ell_{42} = 0 \wedge \ell_{52} = 0 \\ (12) \wedge \ell_{32} = 0 \wedge \ell_{42} = 0 \wedge \ell_{52} \neq 0 \\ (12) \wedge \ell_{32} = 0 \wedge \ell_{42} \neq 0 \wedge \ell_{52} = 0 \\ (12) \wedge \ell_{32} = 0 \wedge \ell_{42} \neq 0 \wedge \ell_{52} \neq 0 \\ (12) \wedge \ell_{32} \neq 0 \wedge \ell_{42} = 0 \wedge \ell_{52} = 0 \\ (12) \wedge \ell_{32} \neq 0 \wedge \ell_{42} = 0 \wedge \ell_{52} \neq 0 \\ (12) \wedge \ell_{32} \neq 0 \wedge \ell_{42} \neq 0 \wedge \ell_{52} = 0 \\ (12) \wedge \ell_{32} \neq 0 \wedge \ell_{42} \neq 0 \wedge \ell_{52} \neq 0 \end{cases} \tag{23}$$

更なる手動での分解 (23) によって, メモリや計算時間の浪費は格段に軽減され, 次節で報告する最尤推定量の候補を計算することができた.

イデアルの次元, 根基イデアル, 飽和イデアルの計算アルゴリズムはグレブナー基底によって設計可能であり (詳細: [3], [4]), 多くの数式処理システム (例えば, Maple, Magma, Singular 等) にも実装されている. また, 本稿の実験で零次元イデアルが得られた場合には, 最尤推定量となる因子負荷行列の候補を代数的数として具体的に計算したいので, 計算が重くなりやすい辞書式順序のグレブナー基底を計算する必要がある. 零次元イデアルの場合, FGLM アルゴリズムは, 与えられた単項式順序のグレブナー基底から, 他の単項式順序のグレブナー基底へ, 効率的に変換できる (詳細: [5]). そこで, 比較的計算が簡単と知られている全次数逆辞書式順序のグレブナー基底を計算し, 零次元の場合は FGLM アルゴリズムで辞書式順序のグレブナー基底に変換した上で, 代数的数として表現された実代数解を計算することが効率的と思われる. なお, 我々の実装では, グレブナー基底, イデアルの次元, 根基イデアル, 飽和イデアル等の代数構造の計算に Magma を用い, 実代数解の計算などに Mathematica の Solve や CylindricalDecomposition を用いている.

5　グレブナー基底を用いた実験

前節の通り, タイプ S_1, S_2, S_3 の標本分散共分散行列を 10 個生成し, 10 回のモンテカルロシミュレーションを行う. そして, 以下ではタイプ S_i の j 番目の標本を $S_i^{(j)}$ のように表すことにする. まず, タイプ $\Psi_{\mathrm{nonregular}}$ では, 全て (タイプ S_1, S_2, S_3, 各 10 個) の標本分散共分散行列 (計 30 個) に対してイデアル J の解空間 (つまり, (9) を満たす解空間) が複素空間で次元 1 であった. そのため, タイプ $\Psi_{\mathrm{nonregular}}$ に関する因子負荷行列の候補は実空間でも複雑であることが推測される. 次に, タイプ Ψ_{regular} については, イデアル K の解空間 (つまり, (10) を満たす解空間) は, 多くの場合, 実空間で零次元であったが, 一次元の場合もあった. 実解の詳細について, 次の 3 タイプに分けて報告する:

1. $\ell_{11}=0$ であるような実解 (つまり, 全ての共通因子から観測変数 X_1 への負荷が零である候補),

2. $\ell_{11}\neq 0$ であり, $\ell_{22}=\ell_{32}=\ell_{42}=\ell_{52}=0$ であるような実解 (つまり, 共通因子 F_2 から全ての観測変数への負荷が零である候補),

3. 上を除いた実解.

統計解析ソフトウェア R による計算結果と, 上の実解タイプ 1, 2, 3 の個数を比較する. R では, psych パッケージの fa 関数を用いて最尤推定値を算出した. 以下の表では

- R の計算結果が不適解であるか,
- タイプ 1, 2, 3 の実解の個数 (括弧内: 非不適な実解の個数),
- R の結果が不適である場合に, 非不適かつ R で得られた最尤推定値よりも尤度の高い実解が存在するか, 及び, 不適かつ R で得られた最尤推定値よりも尤度の高い実解が存在するか,

を記している. ただし, 実解の個数は列の正負の反転による行列の違いを無視したうえでカウントしている. 個数が無限個の場合は ∞ で表しているが, このような場合は円柱代数分解を用いて実解空間を計算した.

入力	R	タイプ 1	タイプ 2	タイプ 3	高尤度の非不適実解	高尤度の不適実解
$S_1^{(1)}$	非不適解	1 (1)	4 (3)	5 (3)	-	-
$S_1^{(2)}$	不適解	1 (1)	7 (6)	8 (1)	存在しない	存在しない
$S_1^{(3)}$	不適解	1 (1)	6 (6)	6 (2)	存在しない	存在しない
$S_1^{(4)}$	不適解	1 (1)	6 (5)	5 (3)	存在しない	存在しない
$S_1^{(5)}$	不適解	1 (1)	6 (4)	6 (3)	存在する	存在しない
$S_1^{(6)}$	不適解	1 (1)	2 (2)	3 (3)	存在しない	存在しない
$S_1^{(7)}$	不適解	1 (1)	6 (3)	10 (0)	存在しない	存在する
$S_1^{(8)}$	非不適解	1 (1)	4 (4)	3 (3)	-	-
$S_1^{(9)}$	非不適解	1 (1)	4 (4)	4 (1)	-	-
$S_1^{(10)}$	非不適解	1 (1)	4 (4)	1 (1)	-	-

入力	R	タイプ 1	タイプ 2	タイプ 3	高尤度の非不適実解	高尤度の不適実解
$S_2^{(1)}$	非不適解	1 (1)	6 (3)	12 (8)	-	-
$S_2^{(2)}$	非不適解	1 (1)	2 (2)	5 (2)	-	-
$S_2^{(3)}$	不適解	1 (1)	3 (3)	5 (3)	存在しない	存在する
$S_2^{(4)}$	不適解	1 (1)	6 (4)	10 (5)	存在しない	存在しない
$S_2^{(5)}$	非不適解	1 (1)	4 (4)	6 (3)	-	-
$S_2^{(6)}$	不適解	1 (1)	3 (3)	4 (0)	存在しない	存在する
$S_2^{(7)}$	不適解	1 (1)	2 (2)	4 (2)	存在する	存在する
$S_2^{(8)}$	非不適解	1 (1)	5 (3)	4 (2)	-	-
$S_2^{(9)}$	非不適解	1 (1)	4 (2)	6 (3)	-	-
$S_2^{(10)}$	非不適解	∞	4 (1)	6 (4)	-	-

入力	R	タイプ 1	タイプ 2	タイプ 3	高尤度の非不適実解	高尤度の不適実解
$S_3^{(1)}$	非不適解	1 (1)	6 (6)	1 (1)	-	-
$S_3^{(2)}$	非不適解	1 (1)	6 (3)	7 (1)	-	-
$S_3^{(3)}$	非不適解	1 (1)	6 (4)	4 (3)	-	-
$S_3^{(4)}$	非不適解	1 (1)	6 (4)	3 (2)	-	-
$S_3^{(5)}$	非不適解	∞	6 (5)	3 (1)	-	-
$S_3^{(6)}$	非不適解	1 (1)	6 (3)	5 (1)	-	-
$S_3^{(7)}$	非不適解	1 (1)	9 (5)	5 (2)	-	-
$S_3^{(8)}$	非不適解	1 (1)	6 (3)	8 (1)	-	-
$S_3^{(9)}$	非不適解	1 (1)	4 (2)	10 (2)	-	-
$S_3^{(10)}$	非不適解	1 (1)	7 (5)	6 (1)	-	-

タイプ 1 の実解は $\ell_{11} = 0$ によって観測変数 X_1 に全ての共通因子 F_1, F_2 が全く影響していないことから因子負荷行列の候補として特殊である. また, タイプ 2 の実解も $\ell_{22} = \cdots = \ell_{52} = 0$ によって全ての観測変数 X_1, X_2, X_3, X_4, X_5 に共通因子 F_2 が全く影響していないので因子負荷行列の候補としては特殊である.

こうした特徴からタイプ S_1, S_2, S_3 の標本分散共分散行列でのタイプ 3 の非不適実解の個数を比較したい. 上の表を見ると, タイプ S_3 の標本分散共分散行列において計算された非不適実解の個数は他よりも少ない. 従って, 最尤法がタイプ S_3 のような標本分散共分散行列では最尤推定量に収束しやすいと考えられる.

また, R で不適解が計算されたタイプ S_1, S_2 の標本分散共分散行列が持つ高尤度の不適実解を比較する. 上の表を見ると, タイプ S_2 は S_1 よりも高尤度の不適実解を持ちやすく, Ψ の一部がマイナスな候補に最適解が存在する.

本稿の実験は, Magma での計算時間が 2, 3 週間程度, Mathematica での計算は数分程度の入力が多い. つまり, 解空間を具体的に得るために行なっている Magma での代数計算に多くの時間が費やされているが, 準素イデアル分解のような自動でのイデアルの分解ではなく, 前節で触れたような手動でのイデアルの分解によって, 具体的に解を計算することができている.

謝　辞

本研究は科研費 (20K19745), 九州大学マス・フォア・インダストリ研究所共同利用・共同研究拠点 (2020 年度カテゴリー: 特異点論の視点からの異分野探訪, 20200025) の助成・支援を受けたものである.

参 考 文 献

[1] 市川雅教 (2010). 因子分析 (シリーズ行動計量の科学), 朝倉書店.

[2] Anderson, T. W. and Rubin, H. (1956). Statistical inference in factor analysis, In J. Neyman (Ed.), Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 5, 111-150. Berkeley: University of California Press.

[3] Becker, T. and Weispfenning V. (1993). Gröbner Bases, A Computational Approach to Commutative Algebra, Springer New York, NY.

[4] Cox, D. A., and Little, J., and O'She, aD. (2015). Ideals, Varieties, and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra, 4th edition, Springer Cham.

[5] Faugere, J, and Gianni, P., and Lazard, P., and Mora, T. (1993). Efficient computation of zero-dimensional Gröbner bases by change of ordering. Journal of Symbolic Computation, 16, 329-344.

[6] Van Driel, O. P. (1978). On various causes of improper solutions in maximum likelihood factor analysis. Psychometrika, 43(2), 225-243.