

日本テスト学会論文賞記念講演

測定モデルユーザーのための 次元性評価法総論

堀 一輝（熊本県立大学）

牧野 直道（大学入試センター）

※ 採択時の所属はベネッセ教育総合研究所

本講演について

2

- 堀・牧野（2024）の解説
- 出版後に公開された関連研究をフォローアップ
- Rでの分析例を追加
 - 川端他（2018）第8章（確認的因子分析）のサンプルデータを使用
 - 1,000人 × 12項目、4因子を想定

- 次元性 dimensionality (≡ 因子数、成分数、次元数) の決め方に関するレビュー論文
 - 2010年以降に提案された相対的に新しい研究を中心にレビュー
- 探索的因子分析（または主成分分析）と項目反応理論に限定
 - 潜在クラス分析および認知診断モデルは扱わない
- I. 固有値パターン、II. モデル選択、III. ネットワークサイコメトリクス、IV. 機械学習の4アプローチに大別される

表Ⅰ 測定モデルのための次元性評価法の分類とRでの実装手法⁴

表 1 測定モデルのための次元性評価法の分類と R での実装

手法	探索的手法		検証的手法		R での実装 (パッケージ名::関数名)
	測定モデル	測定モデル	測定モデル	測定モデル	
	要適用	適用不要	要適用	適用不要	
<u>固有値のパターンにもとづく手法</u>					
	仮説なし		仮説あり 例:一次元性の確認		
カイザー・ガットマン基準		✓			psych::kaiser, EFAtools::GKC
平行分析/修正平行分析	✓	✓ ^a			psych::fa.parallel, EFAtools::PARALLEL
比較データ法	✓				EFAtools::CD, RGenData::EFACompData
後続固有値十分性検定	✓	✓ ^a			Rnest::nest ^c , latentFactorR::NEST
経験カイザー基準		✓			EFAtools::EKC, semTools::efa.ekc
<u>モデル選択にもとづく手法</u>					
モデル適合度指標	✓ ^b		✓ ^b		psych::vss, EFAtools::SMT
凸包法	✓				EFAtools::HULL, EFA.dimensions::DIMTESTS
<u>ネットワーク・サイコメトリクスアプローチ</u>					
探索的グラフ分析		✓ ^a			EGAnet::EGA, latentFactorR::estimate_dimensions
<u>機械学習アプローチ</u>					
ファクター・フォレスト	✓	✓ ^a			factor.forest ^d , latentFactorR::factor_forest
標本外予測誤差	✓				fspe::fspe, latetFactorR::estimate_dimensions

^a 共通性の推定を行わない場合, ^b 決定方式に依存する, ^c <https://github.com/quantmeth/Rnest> から入手可能,

^d <https://osf.io/mvrau> から入手可能

I. 固有値パターンにもとづく手法

5

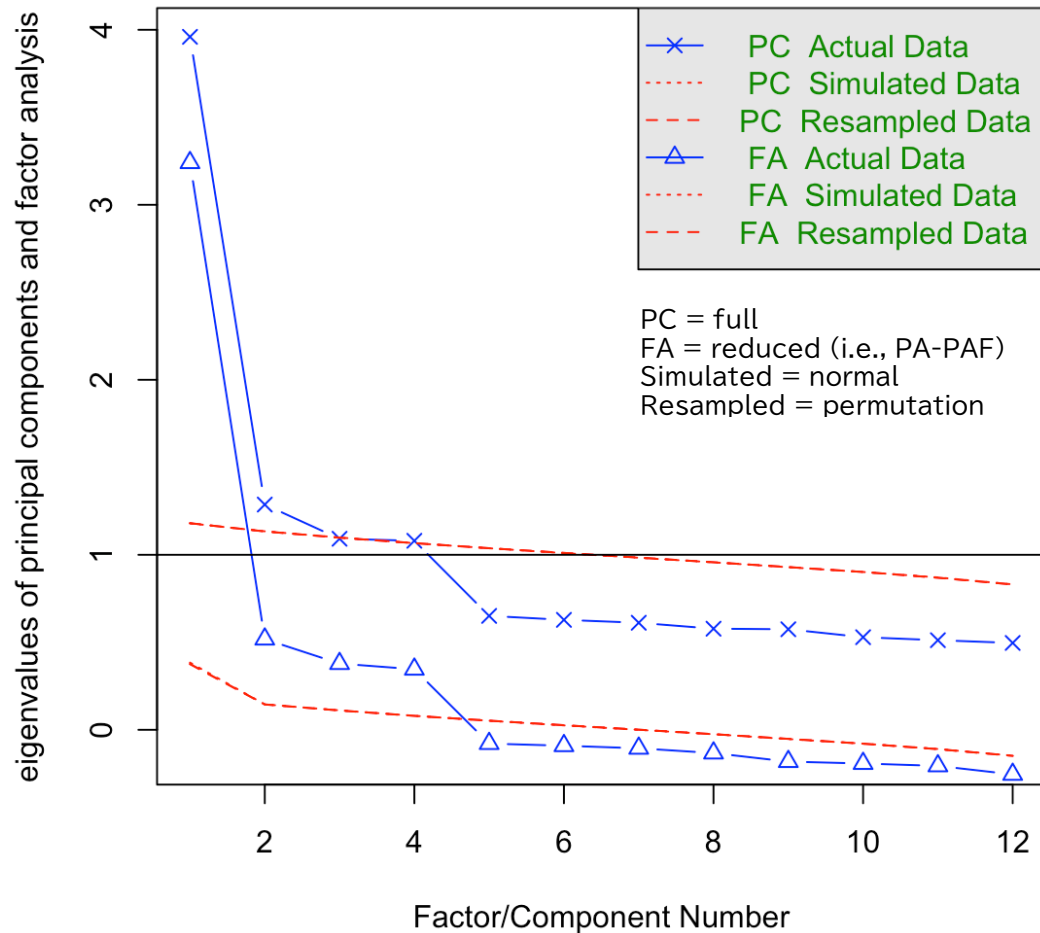
- 平行分析 (Parallel Analysis (PA) ; Horn, 1965)
 - カイザー・ガットマン基準 (Guttman, 1954; Kaiser, 1960) の拡張
 - ランダムな相関行列 (0因子モデル) を多数生成し、固有値の参照分布を得る
 - 参照分布の要約値と観測固有値を比較
→ 因子数 = 要約値より大きい観測固有値数
- 派生法の改良点
 - 要約統計量：平均 → パーセンタイル点 (95%, 99%)
 - 固有値分解の対象：相関行列 (PA-PCA) → **縮小相関行列** (PA-PAF)
 - 共通性の推定をするかどうか (PAF=Principal Axis Factoring)
 - 人工データの生成：0因子モデル → **k-1 因子**モデル (**修正平行分析**; RPA)

Rでの実装例（平行分析; PA-PCA, PA-FA）

6

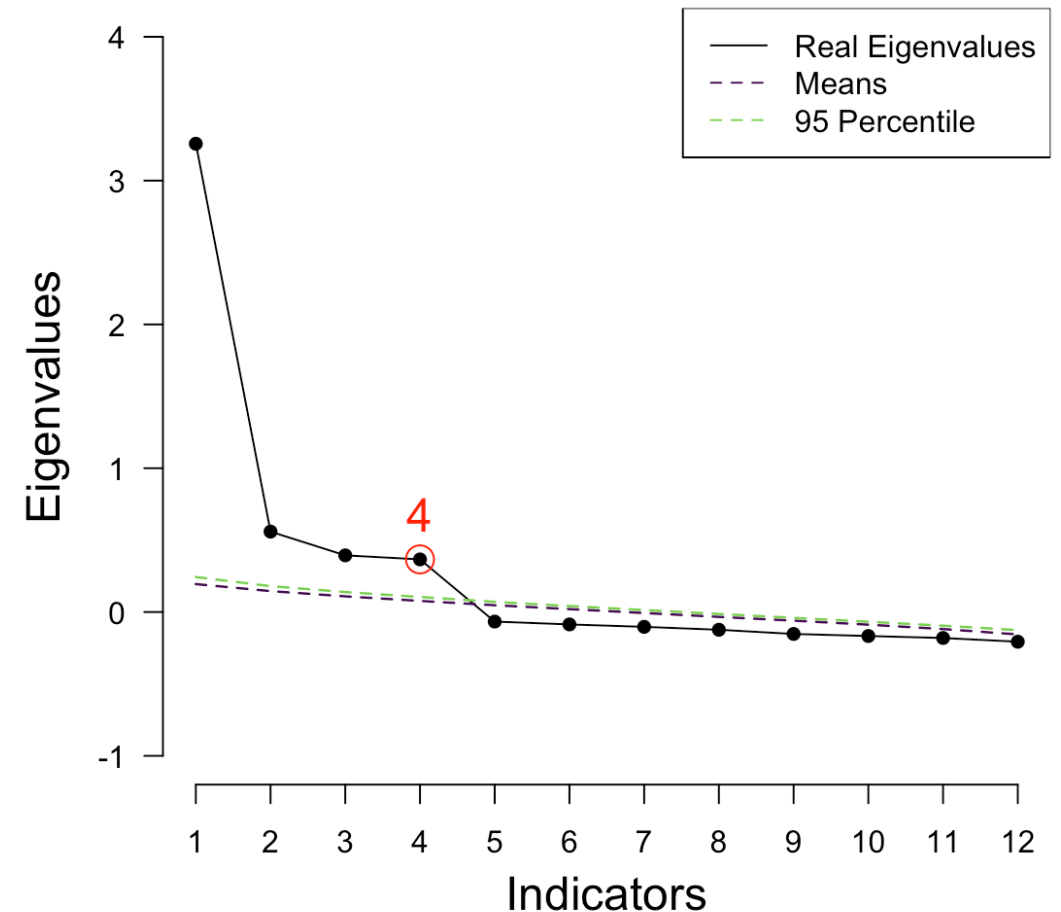
- psych::fa.parallel

Parallel Analysis Scree Plots



- EFAtools::PARALLEL

N Factors with Decision Rule 'percentile' and Eigen Type 'SMC'



後続固有値十分性検定 (NEST; Achim, 2017)⁷

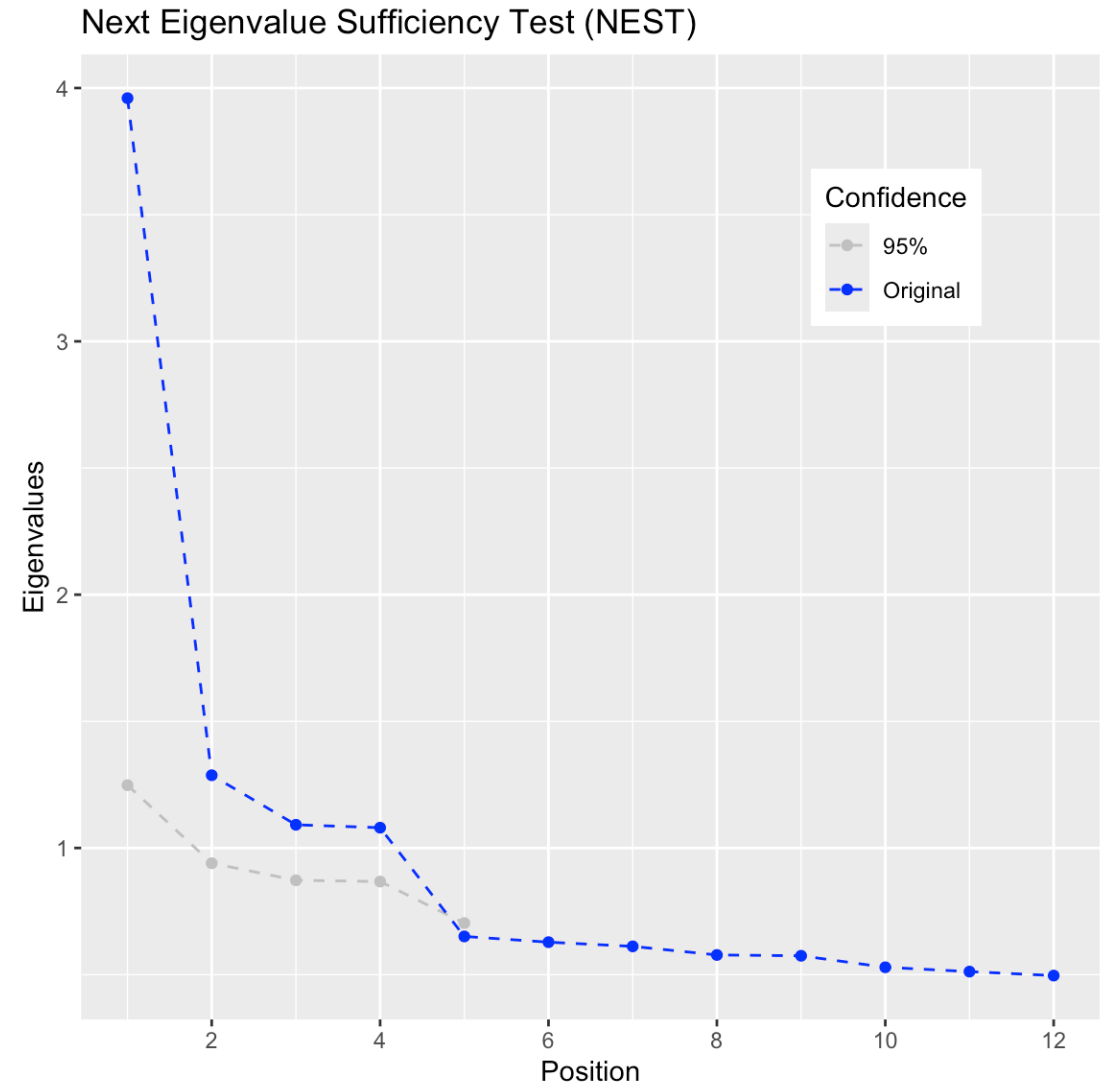
- Next Eigenvalue Sufficiency Test

- RPA-PAFから**共通性の推定法**を改良
- SMC → 反復主因子法（観測） & SMC（人工）
- Rnestパッケージ（Caron, 2025）では、人工データでの共通性の推定方法が拡充（最尤法、最小ランク因子分析、正則化共通因子分析）

- 本研究が出版された2024年以降、徐々に使われ始めている

- Caron (2025)
- Bae and Hong (2024)
- Brandenburg (2024)
- Goretzko (2025)

- Rnest::nest



比較データ法 (Ruscio & Roche, 2012)

8

- Comparison Data (CD)
- EFAtools::CD
- $k-1$ 因子モデルと k 因子モデルからランダムデータを発生させる
 - GenData (Ruscio & Kaczetow, 2008) によって項目間相関や項目周辺分布が観測データに近いものを生成できるパラメタを推定する
- **固有値列全体** を使って比較する
 - 観測固有値列と人工固有値列のRMSRを使用
 - **k 因子 vs. $k-1$ 因子**
 - 有意水準 $\alpha = .30$ でマン・ホイットニーのU検定
 - 有意なら k を1つ大きくする
 - 非有意なら $k-1$ 因子を採用する

N factors suggested by comparison data analysis: 4

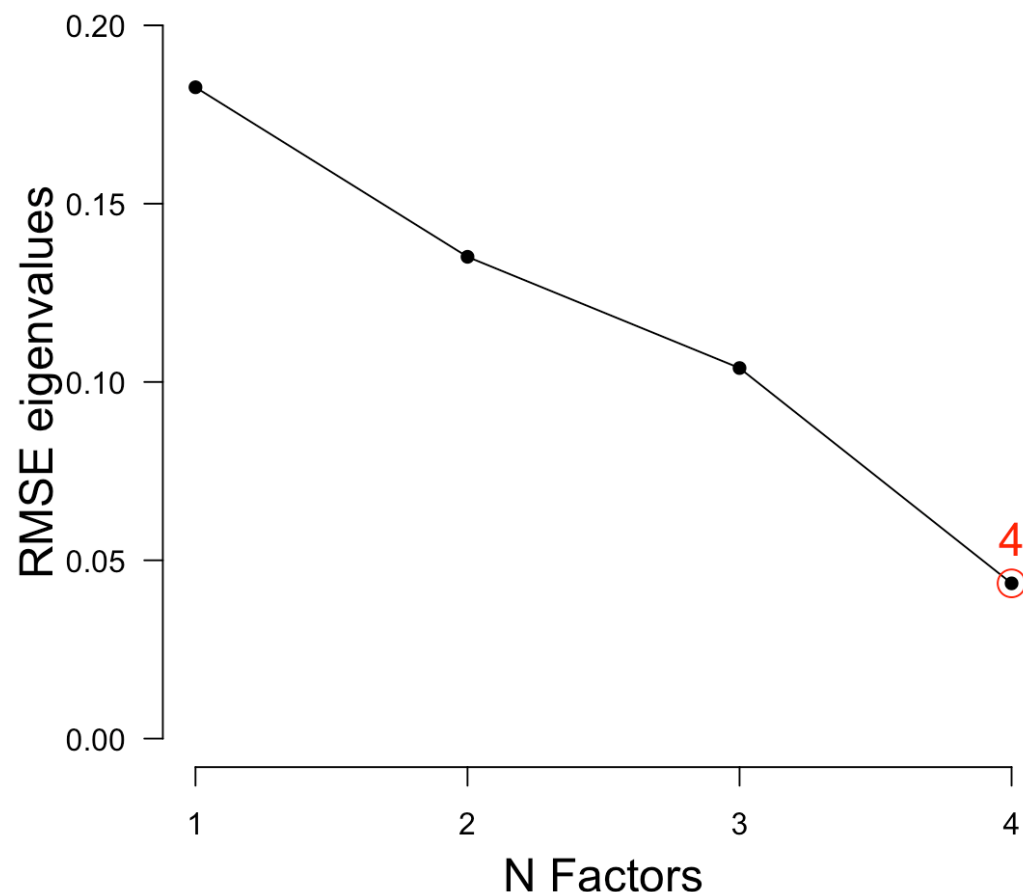


表 3 平行分析関連手法のまとめ

- 平行分析とその派生方法は
 - ① 人工データの生成モデル,
 - ② 共通性の推定方法（観測データ・人工データ）,
 - ③ 決定方式の3観点から整理できる

表 3 平行分析と関連手法の設定の比較

手法	人工データ 生成モデル	共通性の推定 (観測データ)	共通性の推定 (人工データ)	決定方式
平行分析	ゼロ因子モデル	なし	なし	観測固有値と人工固有値の平均を比較
修正平行分析	$k - 1$ 因子モデル	SMC	SMC	観測固有値と人工固有値の 95%点を比較
比較データ法	$k - 1$ 因子モデル k 因子モデル	なし	なし	$k - 1$ 因子モデルと k 因子モデルの RMSR を比較
後続固有値十分性検定	$k - 1$ 因子モデル	反復主因子法	SMC	観測固有値と人工固有値の 95%点を比較

SMC=平方重相関係数, 95%点=95 パーセンタイル点

II. モデル選択にもとづく手法

10

- SEMのモデル適合度指標や情報量規準によって次元数を評価
 - 適合度指標： χ^2 統計量、CFI、TLI、RMSEA、SRMR
 - 情報量規準：AIC、BIC、WAIC、WBIC
- 3つの決定方式：1～ m 因子モデルで分析し、
 - a. 適合度が最も良いモデルの因子数を採用する
 - b. 適合度が「基準」を満たすモデルの中で、最も小さい因子数を採用する
 - c. 因子数が小さい方から順に適合度を比較し、改善が見られなくなったモデルの因子数を採用する（例：0因子 vs. 1因子 → 1因子 vs. 2因子 → …）

モデル適合度にもとづく方法の近年の動向

||

- 固定の基準値を使うことへの批判
 - 例：CFI > 0.96, RMSEA < 0.05 (Hu & Bentler, 1999)
- 同等性検定の考えにもとづく基準値の修正 → b方式
 - Yuan et al. (2016), Montoya & Edwards (2021)
- 適合度の差に基準値を設定する方法 → c方式
 - Finch (2020), Bae & Hong (2024)
- Dynamic fit index cutoffsを利用する方法 → c方式
 - McNeish & Wolf (2023) → [dynamic](#)パッケージ、[Shiny app](#)
 - 一因子モデルの適合度についての論文で、次元性評価としての論文ではない。

Dynamic Fit Index Cutoffs



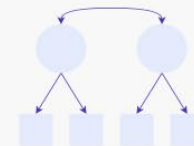
The applications on this website compute dynamic fit index (DFI) cutoffs for latent variable models that are tailored to your specific model. These cutoffs can be used as *one* piece of evidence of validity to help gauge if the misspecifications in your model are trivial or substantial.

**New! Measurement Models**

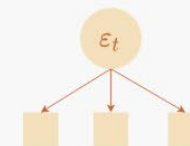
One-factor, bi-factor, categorical data, oh my! Want to compute DFI cutoffs for a measurement model? You can do it here! All you need is your dataset.

**One-factor CFA**

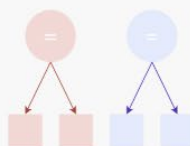
To compute DFI cutoffs for one-factor confirmatory factor analysis models, you need your standardized factor loadings and sample size.

**Multi-Factor CFA**

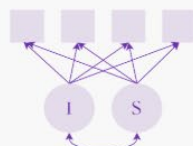
To compute DFI cutoffs for multi-factor confirmatory factor analysis models, you need your standardized factor loadings and sample size.

**Equivalence Testing**

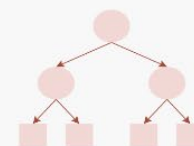
To compute fit index values using the equivalence testing method, you will need five pieces of information from your output. Click the app to learn more.

**Measurement Invariance**

Not available yet

**Growth Models**

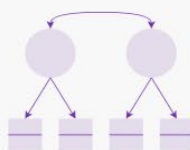
Not available yet

**Higher-Order Models**

Check out the Measurement Models app or the R package on Github!

**Bi-Factor Models**

Check out the Measurement Models app or the R package on Github!

**Categorical CFA**

Check out the Measurement Models app or the R package on Github!

**Reliability Representativeness**

This application is a tool to assess how well a reliability coefficient generalizes across an entire score distribution.

<https://dynamicfit.app/connect/#/welcome>

凸包法 (Lorenzo-Seva et al., 2011)

13

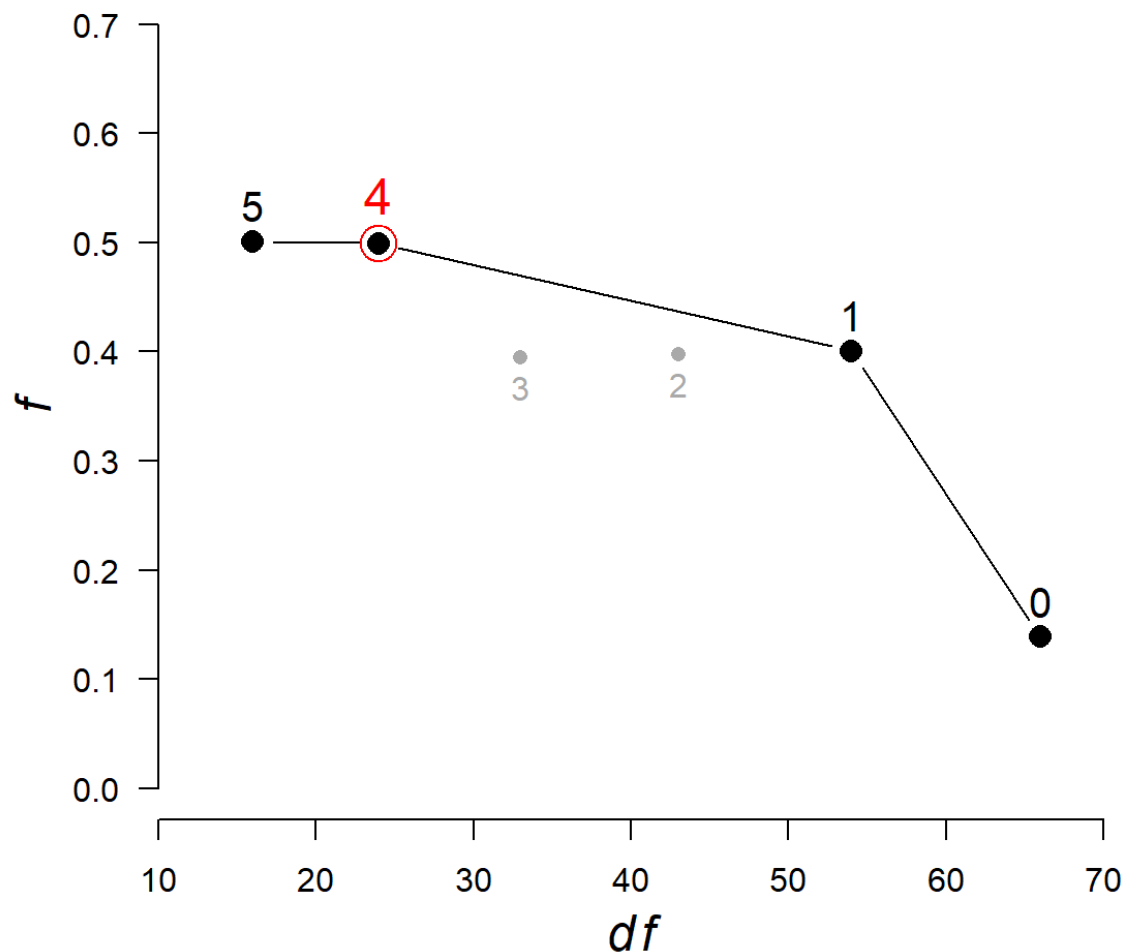
- 自由度 (df) vs. 適合度 (f) の散布図で凸包 (convex hull) を見つける
 - 内点 (e.g., 2因子・3因子モデル) は除外
 - CFIやCAF (≡KMO指標) などを使う
 - 除外アルゴリズムについては本文参照
- 残った点の中で、以下の st_i を最大にする点の因子数を採用する

$$st_i = \frac{(f_i - f_{i-1}) / (df_i - df_{i-1})}{(f_{i+1} - f_i) / (df_{i+1} - df_i)}$$

- st_i = 隣接する線分の傾きの比
- 傾きが最も急に落ちる点を採用する

- EFAtools::HULL

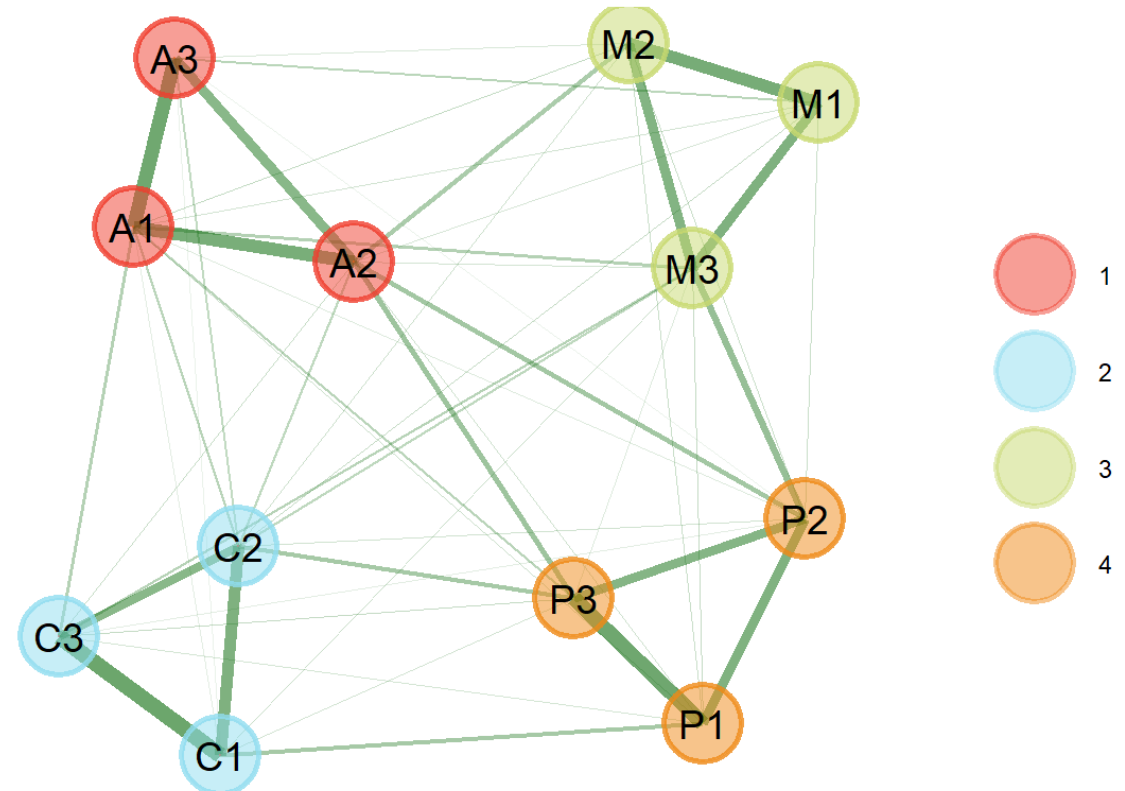
Hull Method with PAF estimation and CAF



III. ネットワークサイコメトリクス

14

- 観測変数の（因果的な）相互の結びつきとして構成概念を捉えるアプローチ
 - ノード＝変数、エッジ＝（非ゼロな）偏相関係数
 - ガウシアングラフィカルモデル（GGM）
- EGAnet::EGA
- 測定モデルの潜在変数（因子）＝ネットワークモデルのコミュニティ
 - 因子数＝コミュニティ数
- 探索的グラフ分析（EGA）
 1. ネットワークの推定
 2. コミュニティの検出



III. ネットワークサイコメトリクス

15

• 1. ネットワークの推定法

• グラフィカルLASSO + EBIC

- Golino & Epskamp (2017)

• 精度行列のベイズ推定

- Shi et al. (2024)
- 共役事前分布（ウィシャート分布）→ 解析的
- ジェフリーズ事前分布 → サンプリング的
- 90%CI (HDI) とROPEによってスパース化

• 2. コミュニティの検出法

- モジュラリティーが基準として使われる
- 一因子かどうかを別途確認する必要有

• Walktrapアルゴリズム

- Golino & Epskamp (2017)

• TMFGアルゴリズム

- triangulated maximally filtered graph
- Golino et al. (2020)

• Louvainアルゴリズム

- Christensen et al. (2024)

IV. 機械学習アプローチ

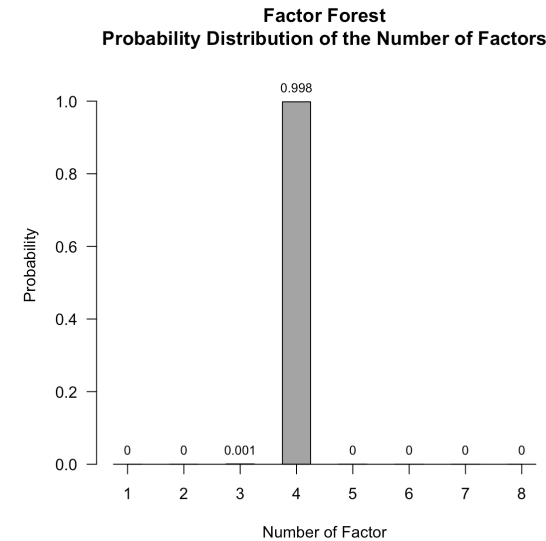
16

- ファクターフォレスト
 - Goretzko & Bühner (2020, 2022)
- シミュレーション研究と同様に、様々な因子数とパラメタで合成データを大量に生成し、学習データとする

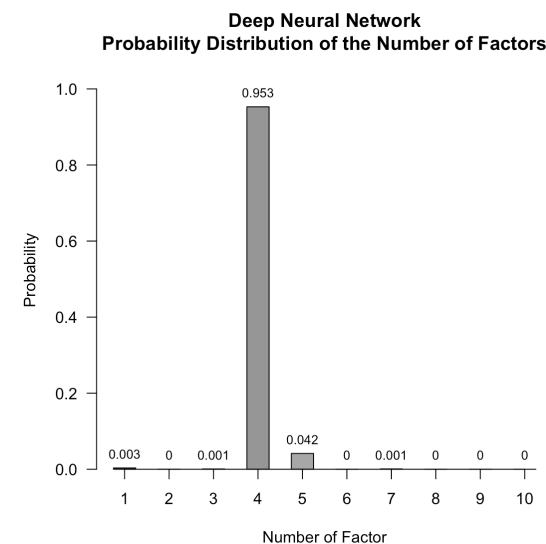
→ 本来は教師なし学習問題である次元性評価が教師あり学習問題（分類）になる

- 498,971セットの合成データ
 - 181個の特徴量 + 他の手法の予測値
 - 予測モデルにはXGBoostを利用
- DNNを使った実装も (Qin & Guo, 2025)

• EFAfactors::FF



• EFAfactors::DNN_predictor



IV. 機械学習アプローチ

17

- **予測誤差**を最小化（≡汎化能力を最大化）する因子数を採用する方法 (Haslbeck & van Bork, 2024)
- 交差検証（cross validation; CV）によって予測誤差を評価する
 - Haslbeck & van Bork (2024)は項目反応についての予測誤差を最小化する因子数を採用（K-fold CV）
 - Browne & Cudeck (1989)は共分散行列についての予測誤差を最小化する因子数を採用（ホールドアウト）
 - 豊田（1986）は対数尤度を最大にする因子数を採用する（ホールドアウト）
- 因子分析モデルで独立変数にあたる因子得点 $\mathbf{f}$ には不定性があるため、因子分析の結果は直接的には使用できない
 - ネットワークサイコメトリクスのな「因子」を考える
- 訓練データから求めた分散共分散行列の推定値 $\hat{\Sigma}$ を使って、各観測変数を他の観測変数に回帰した際の偏回帰係数 $\hat{\mathbf{B}}$ を求め、テストデータ $\mathbf{Z}$ と合わせて予測誤差 $\hat{\mathbf{E}}$ を求める
 - $\hat{\mathbf{E}} = \mathbf{Z} - \mathbf{Z}\hat{\mathbf{B}} = \mathbf{Z}\hat{\mathbf{K}}\hat{\mathbf{D}} = \mathbf{Z}\hat{\Sigma}^{-1}[\text{diag}(\hat{\Sigma}^{-1})]^{-1}$
 - fspeパッケージ (Haslbeck & van Bork, 2023)

- 平行分析に類似した手法や設定の多さ
→ 詳細の把握や理解が困難
 - 本研究の表2で整理
 - 各実装では引数として設定可能な場合と、特定の設定で固定の場合がある

表 2 平行分析の設定点とその選択肢

設定点	選択肢
固有値分解の対象 (観測データ)	標本相関行列(積率相関, 四分相関, 多分相関) 縮小相関行列(積率相関, 四分相関, 多分相関)
共通性の推定法 (観測データ)	PCA(推定しない), SMC, PAF, MRFA
人工データ生成法	無相関(正規乱数, 一様乱数), $m-1$ 因子モデル, ノンパラメトリックブートストラップ, 行並び替え, GenData
固有値分解の対象 (人工データ)	相関行列(積率相関, 四分相関, 多分相関) 縮小相関行列(積率相関, 四分相関, 多分相関)
共通性の推定法 (人工データ)	PCA(推定しない), SMC, PAF, MRFA
人工データ参照値	平均値, 95 パーセンタイル点, 99 パーセンタイル点, 固有値列
因子数の判定方法	観測値が参照値より大きいのか, 観測固有値列と人工固有値列の RMSR
四分相関の推定法	一段階最尤法, 二段階最尤法, 直接近似, オッズ比近似
平滑化法	再スケール法, 非対角要素調整法, 交互射影法, リッジ法, 置換法, 最小二乗近似, 逐次再スケール法, 平滑化なし
頻度ゼロの修正方法	Yates の修正, 修正なし

PCA=主成分分析, SMC=平方重相関係数, PAF=反復主因子法, MRFA=最小ランク因子分析

- 適合度指標の値が良いことの意味
 - 真の次元数 vs. 最適な次元数
 - 近真度 vs. 一般化可能性
 - 研究の目的に合致した「指標」を使うべき
- 「コミュニティ」と「因子」
 - どちらも構成概念を測定しようとしている
 - そもそもNWサイコメトリクスは潜在変数モデルとは質的に異なるものとして提案されたのでは？
 - Guttmanのイメージ理論に近い？
- 合成データの利用による教師あり学習への問題の転換
- 深層学習技術と測定（次元性）

- Root Deterioration per Restriction (RDR) の利用
 - RMSEA関連の適合度指標 (Brown & Du Toit, 1992; Bae & Hon, 2024)
- XGBoost以外の機械学習の利用 (Chen et al., 2025)
- 包括的レビュー論文 (Goretzko, 2025)
 - 本論文とかなり近い内容のレビュー論文
 - 本研究が外したやや古い手法も含めている
 - スパース因子分析 (正則化EFA) を手法の一つに含めている

- Bae, S., & Hong, S. (2024). Comparison of factor retention methods in exploratory factor analysis: RMSEA, root deterioration per restriction and parallel analysis. *Sage Open*, 14(4). <https://doi.org/10.1177/21582440241301840>
- Brandenburg, N. (2024). Factor retention in ordered categorical variables: Benefits and costs of polychoric correlations in eigenvalue-based testing. *Behavioral Research Methods*, 56, 7241–7260. <https://doi.org/10.3758/s13428-024-02417-0>
- Caron, P.-O. (2025). A comparison of the next eigenvalue sufficiency test to other stopping rules for the number of factors in factor analysis. *Educational and Psychological Measurement*, Advance online publication. <https://doi.org/10.1177/00131644241308528>
- Chen, C., D'hondt, R., Vens, C., & Van den Noortgate, W. (2025). Factor retention in exploratory multidimensional item response theory. *Educational and Psychological Measurement*, Advance online publication. <https://doi.org/10.1177/00131644241306680>
- Finch, H. (2025). A comparison of methods for determining the number of factors to retain in exploratory factor analysis for categorical indicator variables. *Psychology International*, 7(1), 3. <https://doi.org/10.3390/psycholint7010003>
- Goretzko, D. (2025). How many factors to retain in exploratory factor analysis? A critical overview of factor retention methods. *Psychological Methods*, Advance online publication. <https://doi.org/10.1037/met0000733>
- 川端 一光・岩間 徳兼・鈴木 雅之 (2018). Rによる多変量解析入門—データ分析の実践と理論— オーム社
- Qin, H., & Guo, L. (2025). *EFAfactors: Determining the number of factors in exploratory factor analysis* [R package Version 1.2.0]. <https://doi.org/10.32614/CRAN.package.EFAfactors>
- Shi, D., Christensen, A. P., Day, E. A., Golino, H. F., & Garrido, L. E. (2024). Exploring estimation procedures for reducing dimensionality in psychological network modeling. *Multivariate Behavioral Research*, Advance online publication. <https://doi.org/10.1080/00273171.2024.2395941>
- 豊田 秀樹 (1986). 交差妥当化による共通因子数の推定法 教育心理学研究, 34(3), 252–256. https://doi.org/10.5926/jjep1953.34.3_252

- “dimensionality of the complete latent space”
 - Lord & Novick (1968) p.359
 - complete (完備) :
 $\theta = (\theta_1, \dots, \theta_k)'$ で条件づけたときの項目反応の分布が対象母集団で全て等しい
 - 局所独立の仮定と、 θ が張る潜在空間が完備であることは同値 (p.361)
- 本質的次元性 (Stout, 1990)
 - 本質的独立 (と弱単調性) を満たすのに必要な最小の θ の次元数
 - 本質的独立：条件つき項目共分散の平均が項目数について漸近的にゼロ
 - 独立を無相関に緩めた「弱局所独立」を次元性の定義に採用

- IRTモデルでは明示的に局所独立が仮定される

$$\Pr[\mathbf{U}|\boldsymbol{\theta}] = \prod_{j=1}^p (\Pr[U_j = 1|\boldsymbol{\theta}])^{u_j} (1 - \Pr[U_j = 1|\boldsymbol{\theta}])^{1-u_j}$$

- 因子分析では本質的独立（条件つき無相関）が仮定されている

$$\text{Var}[\mathbf{x}|\mathbf{f}] = \boldsymbol{\Psi}^2 = \text{diag}(\psi_1^2, \dots, \psi_p^2)$$

【補足】経験カイザー基準 (Braeken & van Assen, 2017)²³

- Empirical Kaiser Criterion (EKC)
- 数理統計学的アプローチによるカイザー・ガットマン基準の拡張
- ゼロ因子モデル下での第一固有値の標本分布が漸近的にMarchenko-Pastur分布に従うことを利用して第一参照値を決定

$$\lambda_{EKC,1} = \left(1 + \sqrt{\frac{p}{N}}\right)^2$$

- 固有値の和に制約があることを考慮しつつ、第 j 参照値を第一参照値に比例する形で決定

$$\lambda_{EKC,j} = \max \left(\frac{p - \sum_{j'=1}^{j-1} \lambda_{j'}}{p - j - 1} \cdot \lambda_{EKC,1}, 1 \right)$$

- 四分相関/多分相関 + 連続変数向け手法
 - 推定法の選択、非負定値ではないときの平滑化手法、ゼロ頻度への対応
 - 周辺分布（項目反応の分布）を揃えた方が精度が高い
→ 正規/一様乱数より、リサンプリング・並び替えの方が望ましい
- 人工データ生成モデル/分析モデルをIRTに変更
- 特に対応しない
 - 連続変数で学習したファクターフォレストをそのままカテゴリーカルデータに適用しても、他の手法より高精度となった (Goretzko & Bühner, 2022)