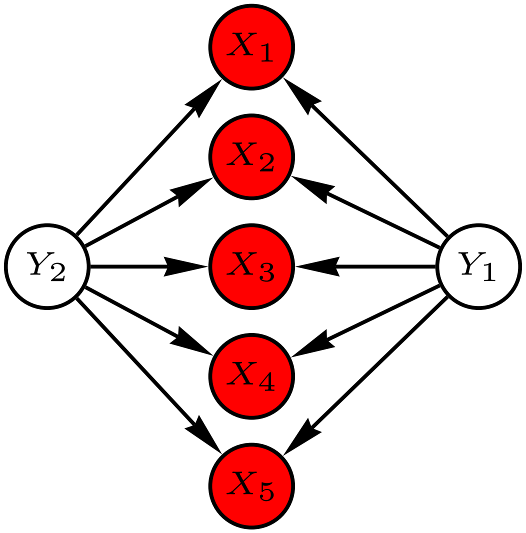
Finiteness in Factor Analysis

Mathias Drton

Department of Statistics

The University of Chicago



Outline

1. Testing for complete independence in high-dimensions
2. Factor analysis
3. Finiteness in factor analysis
4. Equations beyond off-diagonal minors
5. Conclusion

1. Complete independence in high dimensions

- Multivariate normal random vector

$$X = (X_1, \ldots, X_p)^t \sim \mathcal{N}_p(\mu, \Sigma).$$

- Test complete independence

$$H_0 : X_1 \perp\!\!\!\perp X_2 \perp\!\!\!\perp \ldots \perp\!\!\!\perp X_p \iff [\sigma_{ij} = 0 \quad \forall i < j]$$

  based on a sample $X^{(1)}, \ldots, X^{(n+1)} \in \mathbb{R}^p$.

- Likelihood ratio test rejects if $\det(R)$ is too large, where $R = (r_{ij})$ is the sample correlation matrix. (Alternative: saturated model)

- If $p > n$: LRT without power since $\det(R) = 0$ always.

Schott's (2005) test

- Simple and intuitively plausible statistic

$$T_{p,n} = \left(\sum_{1 \leq i < j \leq p} r_{ij}^2 \right) - \frac{p(p-1)}{2n}.$$

- Assume n and p grow large simultaneously forming a sequence $(n_h, p_h)_{h=1}^{\infty}$ such that

$$\lim_{h\to\infty} n_h = \lim_{h\to\infty} p_h = \infty, \qquad \lim_{h\to\infty} p_h/n_h = \gamma \in (0,\infty).$$

- **Schott's Theorem**:

  Under complete independence H_0,

$$T_{p_h,n_h} \longrightarrow_d \mathcal{N}(0,\gamma^2).$$

Schott's test as quadratic form

- Complete independence is defined by the constraints

$$\sigma_{ij} = 0 \quad \forall\, 1 \leq i < j \leq p.$$

- Estimate constraints based on sample covariance matrix $S = (s_{ij})$ as

$$s_{ij} \stackrel{?}{\approx} 0 \quad \forall\, 1 \leq i < j \leq p.$$

- Let $s = (s_{12}, s_{13}, \ldots, s_{p-1,p})^t \in \mathbb{R}^{p(p-1)/2}$. Under complete independence H_0, the covariance matrix $\mathrm{Var}_\Sigma[s]$ is the diagonal matrix with diagonal entries $\sigma_{ii}\sigma_{jj}/n$.

- Quadratic form

$$Q = s^t\, \mathrm{Var}_S[s]^{-1} s = n \cdot \sum_{1 \leq i < j \leq p} r_{ij}^2.$$

Finiteness in complete independence model

- Let $F_{p,0}$ be the set of diagonal covariance matrices.
- Complete independence hypothesis is $H_0 : \Sigma \in F_{p,0}$
- Obvious finiteness property:

  For all $p \geq p_0 = 2$,

$$\Sigma \in F_{p,0} \iff \quad \Sigma_{A \times A} \in F_{p_0,0} \quad \forall\, A \in \left\{ \begin{matrix} p \\ p_0 \end{matrix} \right\},$$

  where

$$\left\{ \begin{matrix} p \\ m \end{matrix} \right\} = \big\{ A \subseteq \{1, \ldots, p\} \ : \ |A| = m \big\}.$$

- Independence determined by bivariate marginal distributions

2. Factor analysis

- Observe scores of $n = 100$ students on $p = 4$ exams:

X_1 : *algebra* $\quad$ X_3 : *statistics*

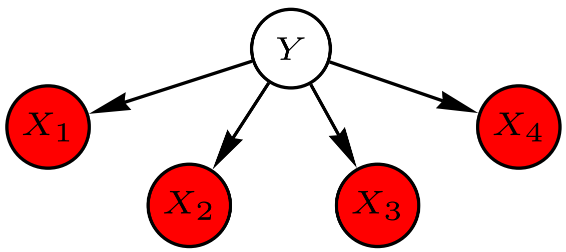
X_2 : *analysis* $\quad$ X_4 : *physics*

- Correlation matrix $R =$

$$\begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ X_4 \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 1.00 & 0.72 & 0.63 & 0.54 \\ 0.72 & 1.00 & 0.56 & 0.48 \\ 0.63 & 0.56 & 1.00 & 0.42 \\ 0.54 & 0.48 & 0.42 & 1.00 \end{pmatrix} \end{array}$$

- A hidden variable Y ("general math ability") may explain correlations:

$$X_1 \perp\!\!\!\perp X_2 \perp\!\!\!\perp X_3 \perp\!\!\!\perp X_4 \mid Y$$



Fictive example of factor analysis

Explanation by one hidden variable:

$$R = \begin{pmatrix} 0.19 & 0 & 0 & 0 \\ 0 & 0.36 & 0 & 0 \\ 0 & 0 & 0.51 & 0 \\ 0 & 0 & 0 & 0.64 \end{pmatrix} + \begin{pmatrix} 0.81 & 0.72 & 0.63 & 0.54 \\ 0.72 & 0.64 & 0.56 & 0.48 \\ 0.63 & 0.56 & 0.49 & 0.42 \\ 0.54 & 0.48 & 0.42 & 0.36 \end{pmatrix}$$

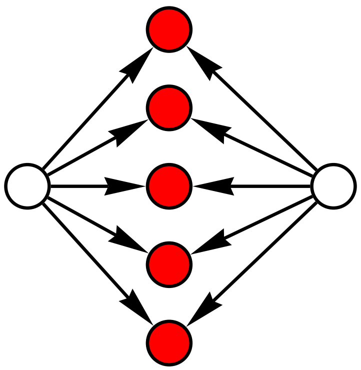
$$= \begin{pmatrix} 0.19 & 0 & 0 & 0 \\ 0 & 0.36 & 0 & 0 \\ 0 & 0 & 0.51 & 0 \\ 0 & 0 & 0 & 0.64 \end{pmatrix} + \begin{pmatrix} 0.9 \\ 0.8 \\ 0.7 \\ 0.6 \end{pmatrix} \begin{pmatrix} 0.9 \\ 0.8 \\ 0.7 \\ 0.6 \end{pmatrix}^t$$

Factor analysis model

- # observed variables: $p \in \mathbb{N}$;

  # hidden factors: $m \in \mathbb{N}$; $(m \ll p)$

- **Factor analysis model**



$$\{\mathcal{N}_p(\mu, \Sigma) \ : \ \mu \in \mathbb{R}^p, \ \Sigma \in F_{p,m}\}$$

with covariance matrix parameter space

$$F_{p,m} = \{\Delta + \Lambda\Lambda^t \ : \ \Delta > 0 \text{ diag.}, \ \Lambda \in \mathbb{R}^{p \times m}\}$$

- Model dimension:

$$\dim(F_{p,m}) = \min\left\{p(m+1) - \binom{m}{2}, \binom{p+1}{2}\right\}$$

3. Finiteness in one-factor model

- **Theorem** (Finiteness for $m = 1$):

  If $p \geq p_1 = 4$, then $\Sigma \in F_{p,1}$ if and only if

$$\Sigma_{A \times A} \in F_{p_1,1} \quad \forall A \in \binom{p}{p_1}.$$

- Simple induction proof

- For $m = 0$ (complete indendence) and $m = 1$, it holds that

$$p_m = 2(m+1).$$

- One-factor structure determined by 4-dim. marginal distributions

Finiteness in two-factor model

- **Theorem** (Finiteness for $m = 2$):

  There exists a $p_2 \in \mathbb{N}$ such that if $p \geq p_2$, then $\Sigma \in F_{p,2}$ if and only if

$$\Sigma_{A \times A} \in F_{p_2,2} \quad \forall A \in \begin{Bmatrix} p \\ p_2 \end{Bmatrix}.$$

- Work on progress:

  current state $p_2 = 8$ but I believe in $p_2 = 2(m+1) = 6$.

- Induction proof not as simple anymore

  (trouble if tetrads = off-diagonal 2×2-minors are zero)

Off-diagonal minors

- Off-diagonal minor: a subdeterminant $\det(\Sigma_{A\times B})$ with $A \cap B = \emptyset$.
- Example:

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \boldsymbol{\sigma_{13}} & \boldsymbol{\sigma_{14}} \\ \sigma_{12} & \sigma_{22} & \boldsymbol{\sigma_{23}} & \boldsymbol{\sigma_{24}} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_{44} \end{pmatrix}$$

Off-diagonal 2×2-minor:

$$\det(\Sigma_{12\times 34}) = \sigma_{13}\sigma_{24} - \sigma_{14}\sigma_{23} \qquad \textit{tetrad}$$

- Tetrads appear in Spearman's original work; Software TETRAD
- Off-diagonal $(m+1) \times (m+1)$-minors exist if $p \geq 2(m+1)$.

Vanishing off-diagonal minors

- **Theorem** (improves on Fiedler & Markham, 1987):

  Let Σ be a $p \times p$-matrix with $p \geq 2(m+1)$. If

  (i) all $(m+1) \times (m+1)$ off-diagonal minors of Σ are zero, and

  (ii) all $m \times m$ off-diagonal minors of Σ are non-zero,

  then $\Sigma = \Delta + \Gamma$ for a diagonal Δ and Γ of rank m.

- **Open statistical problem:**

  Test goodness-of-fit of factor analysis models in high dimensions using the sample off-diagonal minors?

- Covariance matrix of sample minors known (D-Massam-Olkin, 2006)

4. Equations beyond off-diagonal minors

- The factor analysis parameter space $F_{p,m}$ is a semi-algebraic set

  i.e., $\exists$ polynomials $\{f_{ke}, g_{ku}\}$ in $\mathbb{R}[\sigma_{ij} \mid i \leq j]$:

$$F_{p,m} = \bigcup_{k=1}^{K} \{\Sigma \in \mathbb{R}^{p\times p}_{\text{sym}} \mid f_{k1}(\Sigma) = 0, \dots, f_{ke_k}(\Sigma) = 0, \\ g_{k1}(\Sigma) > 0, \dots, g_{ku_k}(\Sigma) > 0\}$$

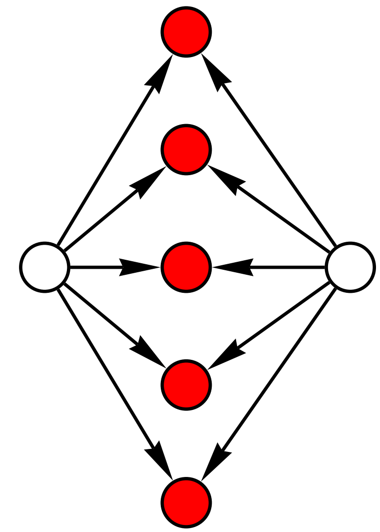
- **Model invariants** are equality relations among parameters:

$$f \in \mathbb{R}[\sigma_{ij} \mid i \leq j] : \; f(\Sigma) = 0 \quad \forall\, \Sigma \in F_{p,m}$$

- Invariants encode geometric information about model (e.g. singularities)

- If $\Sigma \in F_{p,m}$ then $\Sigma =$ diag $+$ (rank $\leq m$) and thus off-diagonal $(m + 1) \times (m + 1)$-minors vanish over $F_{p,m}$.

Pentads (Kelley, 1935)



- Example ($p = 5$ variables, $m = 2$ factors):

$$\text{codim}(F_{5,2}) = 1$$

- No off-diag $(m+1) \times (m+1)$-minors since

$$p = 2m + 1 < 2(m+1)$$

- **Pentad:**

$$\begin{aligned} f = \quad & \sigma_{12}\sigma_{13}\sigma_{24}\sigma_{35}\sigma_{45} - \sigma_{12}\sigma_{13}\sigma_{25}\sigma_{34}\sigma_{45} - \sigma_{12}\sigma_{14}\sigma_{23}\sigma_{35}\sigma_{45} \\ & + \sigma_{12}\sigma_{14}\sigma_{25}\sigma_{34}\sigma_{35} + \sigma_{12}\sigma_{15}\sigma_{23}\sigma_{34}\sigma_{45} - \sigma_{12}\sigma_{15}\sigma_{24}\sigma_{34}\sigma_{35} \\ & + \sigma_{13}\sigma_{14}\sigma_{23}\sigma_{25}\sigma_{45} - \sigma_{13}\sigma_{14}\sigma_{24}\sigma_{25}\sigma_{35} - \sigma_{13}\sigma_{15}\sigma_{23}\sigma_{24}\sigma_{45} \\ & + \sigma_{13}\sigma_{15}\sigma_{24}\sigma_{25}\sigma_{34} - \sigma_{14}\sigma_{15}\sigma_{23}\sigma_{25}\sigma_{34} + \sigma_{14}\sigma_{15}\sigma_{23}\sigma_{24}\sigma_{35} \end{aligned}$$

Linear eliminants

- Requires $p \geq 2m+1$
- Two $(m+1) \times (m+1)$-minors involving exactly one diag σ_{ii}:

$$f_1 = c_{11}(\Sigma) \cdot \sigma_{ii} + c_{10}(\Sigma), \quad c_{1i}(\Sigma) \text{ off-diag}$$

$$f_2 = c_{21}(\Sigma) \cdot \sigma_{ii} + c_{20}(\Sigma), \quad c_{2i}(\Sigma) \text{ off-diag}$$

- **Linear elimination** of σ_{ii} yields

$$f = \det \begin{pmatrix} c_{10}(\Sigma) & c_{11}(\Sigma) \\ c_{20}(\Sigma) & c_{21}(\Sigma) \end{pmatrix} = c_{10}(\Sigma) \cdot c_{21}(\Sigma) - c_{11}(\Sigma) \cdot c_{20}(\Sigma)$$

  that vanishes on $F_{p,m}$.

- Generalization of this approach $\longrightarrow$ **multilinear resultants**

Pentad as linear eliminant

Two 3×3-minors involving exactly one diag σ_{11}:

$$\det(\Sigma_{123\times 145}) = \det\begin{pmatrix} \sigma_{24} & \sigma_{25} \\ \sigma_{34} & \sigma_{35} \end{pmatrix} \bar{\sigma}_{11} + \det\begin{pmatrix} 0 & \sigma_{14} & \sigma_{15} \\ \sigma_{12} & \sigma_{24} & \sigma_{25} \\ \sigma_{13} & \sigma_{34} & \sigma_{35} \end{pmatrix} = 0$$

$$\det(\Sigma_{124\times 135}) = \det\begin{pmatrix} \sigma_{23} & \sigma_{25} \\ \sigma_{34} & \sigma_{45} \end{pmatrix} \bar{\sigma}_{11} + \det\begin{pmatrix} 0 & \sigma_{14} & \sigma_{15} \\ \sigma_{12} & \sigma_{23} & \sigma_{25} \\ \sigma_{14} & \sigma_{34} & \sigma_{45} \end{pmatrix} = 0$$

Pentad: $f =$

$$\det\begin{pmatrix} 0 & \sigma_{14} & \sigma_{15} \\ \sigma_{12} & \sigma_{24} & \sigma_{25} \\ \sigma_{13} & \sigma_{34} & \sigma_{35} \end{pmatrix} \det(\Sigma_{24\times 35}) - \det\begin{pmatrix} 0 & \sigma_{14} & \sigma_{15} \\ \sigma_{12} & \sigma_{23} & \sigma_{25} \\ \sigma_{14} & \sigma_{34} & \sigma_{45} \end{pmatrix} \det(\Sigma_{23\times 45})$$

Ideal of invariants

- Polynomial ring

$$R = \mathbb{R}[\sigma_{ij} \mid 1 \leq i \leq j \leq p] = \mathbb{R}[\sigma_{11}, \ldots, \sigma_{1p}, \sigma_{22}, \ldots, \sigma_{pp}]$$

- Ideal of invariants

$$I_{p,m} = \{f \in R \mid f(\Sigma) = 0 \quad \forall \Sigma \in F_{p,m}\}$$

- **Theorem:** *Let $M_{p,m} \subseteq R$ be the ideal generated by the $(m+1) \times (m+1)$-minors of $\Sigma = (\sigma_{ij})$. Then*

$$I_{p,m} = M_{p,m} \cap \mathbb{R}[\sigma_{ij} \mid i < j].$$

- Hilbert's basis theorem

$$\exists f_1, \ldots, f_k \in R : \quad I_{p,m} = \langle f_1, \ldots, f_k \rangle$$

Variety versus model

- The factor analysis model $F_{p,m}$ is image of real polynomial map.
- Strict inclusions:

$$F_{p,m} \subset \text{clos}(F_{p,m}) \subset V_{\text{pd}}(I_{p,m}) \subset V_{\mathbb{R}}(I_{p,m}) \subset V(I_{p,m})$$

- E.g. for $F_{3,1}$, the ideal $I_{3,1} = \{0\}$ but

$$\begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix} \notin F_{3,1} = \left\{ \begin{pmatrix} \delta_{11} + \lambda_1^2 & \boldsymbol{\lambda_1 \lambda_2} & \boldsymbol{\lambda_1 \lambda_3} \\ \lambda_1 \lambda_2 & \delta_{22} + \lambda_2^2 & \boldsymbol{\lambda_2 \lambda_3} \\ \lambda_1 \lambda_3 & \lambda_2 \lambda_3 & \delta_{33} + \lambda_3^2 \end{pmatrix} \right\}$$

- Matrix above is not in $\text{clos}(F_{3,1})$:

$$\sigma_{12}\sigma_{13} = \lambda_1^2 \sigma_{23} \quad \Longrightarrow \quad \lambda_1^2 = -1$$

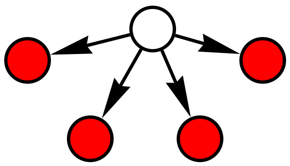
Tetrads form Gröbner basis

- Consider the set of $2\binom{p}{4}$ tetrads

$$\mathcal{T}_p = \{\underline{\sigma_{ij}\sigma_{k\ell}} - \sigma_{ik}\sigma_{j\ell},\ \underline{\sigma_{i\ell}\sigma_{jk}} - \sigma_{ik}\sigma_{j\ell} \mid 1 \le i < j < k < \ell \le p\}$$

- **Theorem (De Loera/Sturmfels/Thomas, 1995):** *If $p \le 3$, then $I_{p,1} = \{0\}$. If $p \ge 4$, then the set $\mathcal{T}_p$ is the reduced Gröbner basis of the ideal $I_{p,1}$ wrto a certain monomial order.*

- Example ($p = 4$ variables, $m = 1$ factor):



$$I_{4,1} = \langle \underline{\sigma_{12}\sigma_{34}} - \sigma_{13}\sigma_{24}, \underline{\sigma_{14}\sigma_{23}} - \sigma_{13}\sigma_{24} \rangle$$

Betti numbers of minimal generators

	$m = 1$	$m = 2$		$m = 3$			
p	deg 2	deg 5	deg 3	deg 8	deg 7	deg 7	deg 4
4	2	—	—	—	—	—	—
5	10	1	0	—	—	—	—
6	30	6	5	—	—	—	—
7	70	21	35	21	0	15	0
8	140	56	140	168	140	120	14
9	252	126	420	756	1386	540	126
	tetrad	*pentad*	*minor*	*ideal-theor.*	*ideal-theor.*	*septad*	*minor*

Finiteness?

Codimension and degree

	$m=1$		$m=2$		$m=3$		$m=4$		$m=5$	
p	codim	deg	cod	deg	cod	deg	cod	deg	cod	deg
3	0	1	0	1	0	1	0	1	0	1
4	2	4	0	1	0	1	0	1	0	1
5	5	11	1	5	0	1	0	1	0	1
6	9	26	4	45	0	1	0	1	0	1
7	14	57	8	259	3	91	0	1	0	1
8	20	120	13	1232	7	1368	2	98	0	1
9	27	247	19	5319	12	14232	6	??	**1**	**54**

Singularities in one-factor analysis ($m = 1$)

- **Theorem ($m = 1$):** *The singular locus of $V(I_{p,1})$ is equal to*

$$V_{\text{sing}}(I_{p,1}) = \{\Sigma \mid \textit{at most one off-diag entry } \sigma_{ij}, i < j, \textit{ is non-zero}\}.$$

- In $F_{p,1}$, local identifiability fails exactly at parameters (Δ, Λ) for which

$$\Sigma = \Delta + \Lambda\Lambda^t \in V_{\text{sing}}(I_{p,1}) \cap F_{p,1}.$$

- Example:

$$\begin{pmatrix} 2 & 1 & & \\ 1 & 2 & & \\ & & 2 & \\ & & & \ddots \end{pmatrix} = \begin{pmatrix} 2-\lambda^2 & & & \\ & 2-1/\lambda^2 & & \\ & & 2 & \\ & & & \ddots \end{pmatrix} + \begin{pmatrix} \lambda \\ 1/\lambda \\ 0 \\ \vdots \end{pmatrix} \begin{pmatrix} \lambda \\ 1/\lambda \\ 0 \\ \vdots \end{pmatrix}^t$$

5. Conclusion

- Factor analysis:

  Finiteness properties $\longleftrightarrow$ *Goodness-of-fit in high dimensions*

- Open problem: Tests involving off-diagonal minors?

- Some literature (see `ArXiv`):

  - D/Massam/Olkin (2006). *Moments of Minors of Wishart Matrices*.
  - D/Sturmfels/Sullivant (2005). *Algebraic Factor Analysis: Tetrads, Pentads and Beyond*. Probab Theory Related Fields, accepted.
  - Schott (2005). *Testing for complete independence in high dimensions*. Biometrika 92: 951–956