Algebraic Sparse Factor Analysis

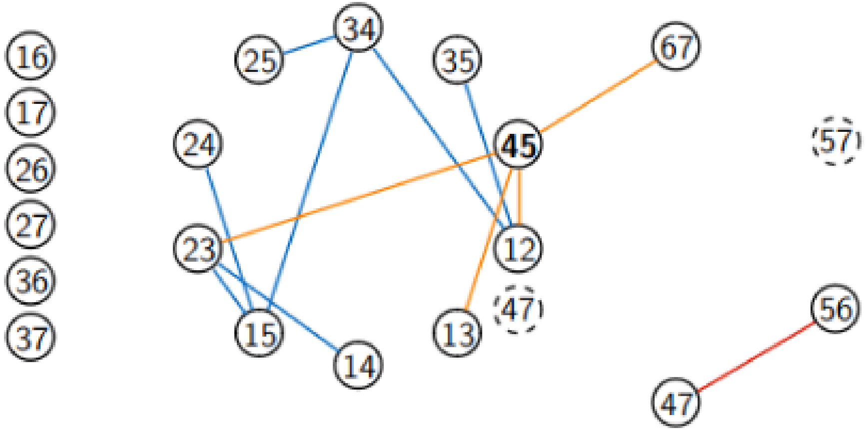
International Congress on Mathematical Software 2024

Classical Algebraic Geometry and Modern Computer Algebra: Innovative Software Design and its Applications

İrem Portakal (🍊)

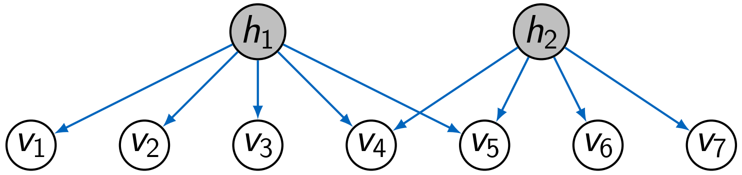
Max Planck Institute for Mathematics in the Sciences Leipzig

(joint work with Mathias Drton, Alexandros Grosdos and Nils Sturma)



Factor Analysis Models

Graph $G = (V \cup \mathcal{H}, E)$, only edges from latent to observed variables. Usually, $V = [p]$. The factor analysis model postulates that the observed variables are linear functions of the factors and noise, i.e., $X_V = \Lambda X_{\mathcal{H}} + \varepsilon$.



$$\Lambda = \begin{pmatrix} 0 & \Lambda_{V\mathcal{H}} \\ 0 & 0 \end{pmatrix} \qquad \Omega = \begin{pmatrix} \Omega_{VV} & 0 \\ 0 & \Omega_{\mathcal{H}\mathcal{H}} \end{pmatrix}$$

- Covariance matrix:

$$\mathrm{Cov}\left[\begin{pmatrix} X_V \\ X_{\mathcal{H}} \end{pmatrix}\right] = (I - \Lambda)^{-1}\Omega(I - \Lambda)^{-\top} = \begin{pmatrix} \Omega_{VV} + \Lambda_{V\mathcal{H}}\Omega_{\mathcal{H}\mathcal{H}}\Lambda_{V\mathcal{H}}^{\top} & \Lambda_{V\mathcal{H}}\Omega_{\mathcal{H}\mathcal{H}} \\ \Omega_{\mathcal{H}\mathcal{H}}\Lambda_{V\mathcal{H}}^{\top} & \Omega_{\mathcal{H}\mathcal{H}} \end{pmatrix}$$

- *Observed* covariance matrix (projection):

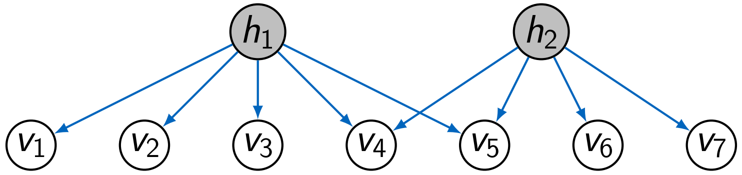
$$\mathrm{Cov}[X_V] = \Omega_{VV} + \Lambda_{V\mathcal{H}}\Omega_{\mathcal{H}\mathcal{H}}\Lambda_{V\mathcal{H}}^{\top}$$

- *Observed* covariance model:

$$F_G = \{\Sigma := \widetilde{\Omega} + \tilde{\Lambda}\tilde{\Lambda}^{\top} \in \mathbb{R}^{|V|\times|V|} : \widetilde{\Omega} > 0 \text{ diagonal}, \ \tilde{\Lambda} \in \mathbb{R}^{E_{V\mathcal{H}}}\}$$

Factor Analysis Models

Graph $G = (V \cup \mathcal{H}, E)$, only edges from latent to observed variables. Usually, $V = [p]$. The factor analysis model postulates that the observed variables are linear functions of the factors and noise, i.e., $X_V = \Lambda X_{\mathcal{H}} + \varepsilon$.



$$\Lambda = \begin{pmatrix} 0 & \Lambda_{V\mathcal{H}} \\ 0 & 0 \end{pmatrix} \qquad \Omega = \begin{pmatrix} \Omega_{VV} & 0 \\ 0 & \Omega_{\mathcal{H}\mathcal{H}} \end{pmatrix}$$

- Covariance matrix:

$$\mathrm{Cov}\left[\begin{pmatrix} X_V \\ X_{\mathcal{H}} \end{pmatrix}\right] = (I - \Lambda)^{-1}\Omega(I - \Lambda)^{-\top} = \begin{pmatrix} \Omega_{VV} + \Lambda_{V\mathcal{H}}\Omega_{\mathcal{H}\mathcal{H}}\Lambda_{V\mathcal{H}}^\top & \Lambda_{V\mathcal{H}}\Omega_{\mathcal{H}\mathcal{H}} \\ \Omega_{\mathcal{H}\mathcal{H}}\Lambda_{V\mathcal{H}}^\top & \Omega_{\mathcal{H}\mathcal{H}} \end{pmatrix}$$

- *Observed* covariance matrix (projection):

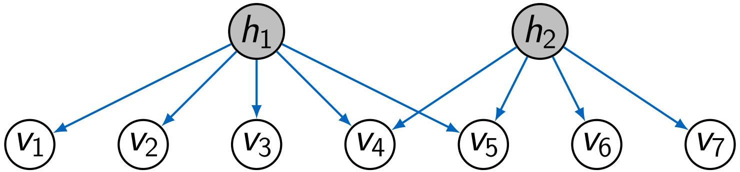
$$\mathrm{Cov}[X_V] = \Omega_{VV} + \Lambda_{V\mathcal{H}}\Omega_{\mathcal{H}\mathcal{H}}\Lambda_{V\mathcal{H}}^\top$$

- *Observed* covariance model:

$$F_G = \{\Sigma := \widetilde{\Omega} + \widetilde{\Lambda}\widetilde{\Lambda}^\top \in \mathbb{R}^{|V|\times|V|} : \widetilde{\Omega} > 0 \text{ diagonal}, \ \widetilde{\Lambda} \in \mathbb{R}^{E_{V\mathcal{H}}}\}$$

Goals: $\dim(F_G)$? $I(F_G) \subseteq \mathbb{R}[\sigma_{ij}, i < j]$?

Example: Factor Analysis Models



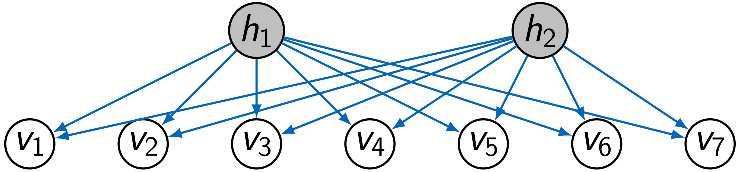
Parameter matrix:

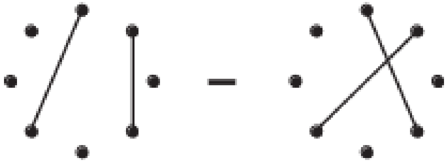
$$\tilde{\Lambda} = \begin{pmatrix} \lambda_{11} & \lambda_{21} & \lambda_{31} & \lambda_{41} & \lambda_{51} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{42} & \lambda_{52} & \lambda_{62} & \lambda_{72} \end{pmatrix}^{\top}, \qquad \tilde{\Omega} = \mathrm{diag}(\omega_{11}, \omega_{22}, \omega_{33}, \omega_{44}, \omega_{55}, \omega_{66}, \omega_{77}).$$

Observed covariance matrix:

$$\Sigma = \begin{pmatrix} \omega_{11}+\lambda_{11}^2 & \lambda_{11}\lambda_{21} & \lambda_{11}\lambda_{31} & \lambda_{11}\lambda_{41} & \lambda_{11}\lambda_{51} & 0 & 0 \\ \lambda_{11}\lambda_{21} & \omega_{22}+\lambda_{21}^2 & \lambda_{21}\lambda_{31} & \lambda_{21}\lambda_{41} & \lambda_{21}\lambda_{51} & 0 & 0 \\ \lambda_{11}\lambda_{31} & \lambda_{21}\lambda_{31} & \omega_{33}+\lambda_{31}^2 & \lambda_{31}\lambda_{41} & \lambda_{31}\lambda_{51} & 0 & 0 \\ \lambda_{11}\lambda_{41} & \lambda_{21}\lambda_{41} & \lambda_{31}\lambda_{41} & \omega_{44}+\lambda_{41}^2+\lambda_{42}^2 & \lambda_{41}\lambda_{51}+\lambda_{42}\lambda_{52} & \lambda_{42}\lambda_{62} & \lambda_{42}\lambda_{72} \\ \lambda_{11}\lambda_{51} & \lambda_{21}\lambda_{51} & \lambda_{31}\lambda_{51} & \lambda_{41}\lambda_{51}+\lambda_{42}\lambda_{52} & \omega_{55}+\lambda_{51}^2+\lambda_{52}^2 & \lambda_{52}\lambda_{62} & \lambda_{52}\lambda_{72} \\ 0 & 0 & 0 & \lambda_{42}\lambda_{62} & \lambda_{52}\lambda_{62} & \omega_{66}+\lambda_{62}^2 & \lambda_{62}\lambda_{72} \\ 0 & 0 & 0 & \lambda_{42}\lambda_{72} & \lambda_{52}\lambda_{72} & \lambda_{62}\lambda_{72} & \omega_{77}+\lambda_{72}^2 \end{pmatrix}.$$

Previous work: *Full* Factor Analysis Models





Dimension: [Drton, Sullivant, Sturmfels, 2007]

$$\dim(F_G) = \min\{p(|\mathcal{H}|+1) - \binom{|\mathcal{H}|}{2}, \binom{p+1}{2}\},$$

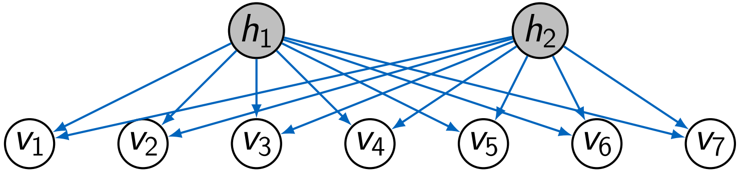
where $|V| = p$.

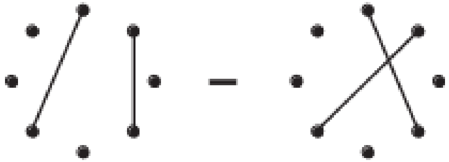
Ideal: [Drton, Sullivant, Sturmfels, 2007]

$$I(F_G) = M_{p,|\mathcal{H}|} \cap \mathbb{R}[\sigma_{ij}, i < j],$$

where $M_{p,|\mathcal{H}|} \subseteq \mathbb{R}[\sigma_{ij}, i \leq j]$ is the ideal generated by all $(|\mathcal{H}|+1) \times (|\mathcal{H}|+1)$ minors of a symmetric $p \times p$ matrix.

Previous work: *Full* Factor Analysis Models





Dimension: [Drton, Sullivant, Sturmfels, 2007]

$$\dim(F_G) = \min\{p(|\mathcal{H}|+1) - \binom{|\mathcal{H}|}{2}, \binom{p+1}{2}\},$$

where $|V| = p$.

Ideal: [Drton, Sullivant, Sturmfels, 2007]

$$I(F_G) = M_{p,|\mathcal{H}|} \cap \mathbb{R}[\sigma_{ij}, i < j],$$

where $M_{p,|\mathcal{H}|} \subseteq \mathbb{R}[\sigma_{ij}, i \leq j]$ is the ideal generated by all $(|\mathcal{H}|+1) \times (|\mathcal{H}|+1)$ minors of a symmetric $p \times p$ matrix.

Gröbner basis:

$|\mathcal{H}| = 1$:

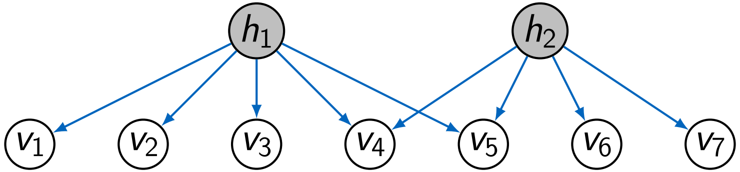
$$G_{p,1} = \{\sigma_{ij}\sigma_{kl} - \sigma_{ik}\sigma_{jl}, \sigma_{il}\sigma_{jk} - \sigma_{ik}\sigma_{jl} \\ 1 \leq i < j < k < l \leq p\}.$$

$|\mathcal{H}| = 2$:

- $\overline{F_G}$ secant variety of the 1-factor model.
- Delightful strategy. [Sullivant, 2009]

New Project: *Sparse* Factor Analysis Models

At least one edge is missing and under Zero Upper Triangular Assumption (ZUTA): ZUTA ensures that the rows and columns of the factor loading matrix Λ can be permuted such that the upper triangle of the matrix is zero. Generalizes k-pure children assumption.



- $C(V,2) := \{\{v,w\} : v,w \in V, v \neq w\}$.
- $\text{jpa}(\{u,v\}) = \{h \in \mathcal{H} : h \in \text{pa}(u) \cap \text{pa}(v)\}$: set of *joint parents* of a pair $\{u,v\} \in C(V,2)$.
- For any latent node $h \in \mathcal{H}$, $C(V,2)_h = \{\{v,w\} \in C(V,2) : h \in \text{jpa}(\{v,w\})\}$.

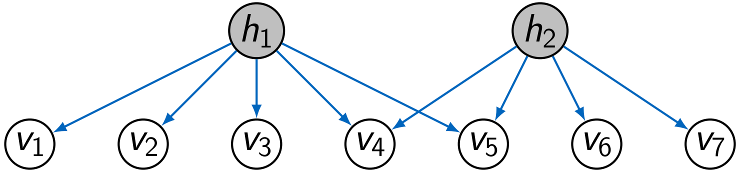
Dementation:

<u>Theorem</u>[DG🍊S]: Let $G = (V \cup \mathcal{H}, D)$ be a factor analysis graph such that ZUTA is satisfied. If there is a ZUTA-conform pairwise disjoint collection of subsets $\mathcal{A} = (A_h)_{h\in\mathcal{H}}$ of $C(V,2)_h$ where $|A_h| \leq |\,\text{ch}(h)|$ such that the sum of cardinalities $\sum_{h\in\mathcal{H}} |A_h|$ is maximal, then

$$\dim(F(G)) = |V| + \sum_{h\in\mathcal{H}} |A_h|.$$

New Project: *Sparse* Factor Analysis Models

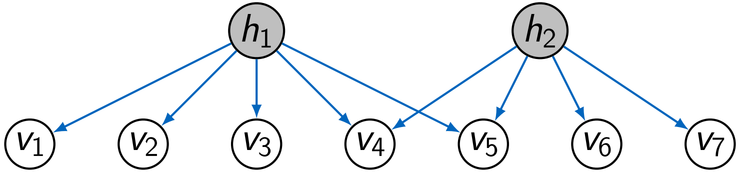
Example:



$A_{h_1} = \{\{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_1, v_4\}, \{v_2, v_4\}\}$
$A_{h_2} = \{\{v_4, v_5\}, \{v_5, v_6\}, \{v_5, v_7\}, \{v_6, v_7\}\}$
$\dim(F(G)) = |V| + \sum_{h \in \mathcal{H}} |A_h| = 7 + 5 + 4 = 16.$

New Project: *Sparse* Factor Analysis Models

Example:

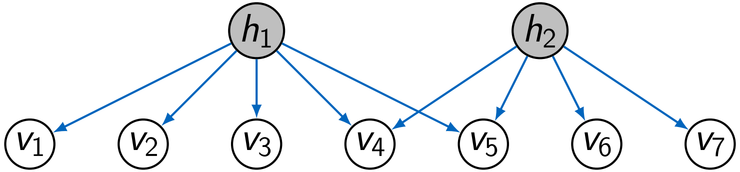


$A_{h_1} = \{\{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_1, v_4\}, \{v_2, v_4\}\}$
$A_{h_2} = \{\{v_4, v_5\}, \{v_5, v_6\}, \{v_5, v_7\}, \{v_6, v_7\}\}$
$\dim(F(G)) = |V| + \sum_{h \in \mathcal{H}} |A_h| = 7 + 5 + 4 = 16.$

Corollary[DG🍊S]: If $p \geq 5$ and $|\mathcal{H}| = 2$ where any latent node has at least 3 children, then $\dim(F_G) = p + |E|$.

New Project: *Sparse* Factor Analysis Models

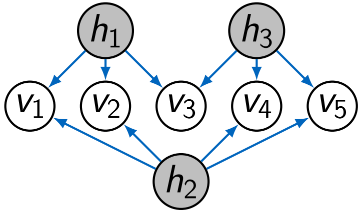
Example:



$A_{h_1} = \{\{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_1, v_4\}, \{v_2, v_4\}\}$
$A_{h_2} = \{\{v_4, v_5\}, \{v_5, v_6\}, \{v_5, v_7\}, \{v_6, v_7\}\}$
$\dim(F(G)) = |V| + \sum_{h\in\mathcal{H}} |A_h| = 7 + 5 + 4 = 16.$

Corollary[DG🍊S]: If $p \geq 5$ and $|\mathcal{H}| = 2$ where any latent node has at least 3 children, then $\dim(F_G) = p + |E|$.

No ZUTA but MathRepo: $\dim(F_G) = 14$.



Joins

What about $I(F_G)$?

- If $\text{pa}(u) \cap \text{pa}(v) = \emptyset$, then $\sigma_{uv} = 0$ (degree one monomials).
- Let $U, W \subseteq V$ be disjoint sets s.t. $|U| = |W| = 2$. $|\text{pa}(U) \cap \text{pa}(W)| \leq 1$, then $\det(\Sigma_{U,W}) = 0$(tetrads).

Theorem[DG🍊S]: Let $|\mathcal{H}| = 2$.

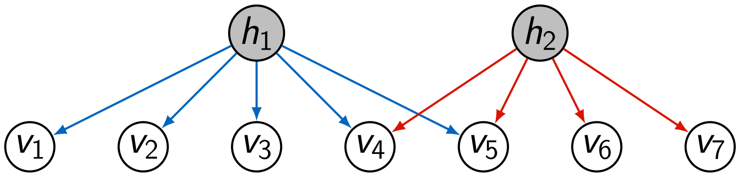
$$\mathbb{V}(I(F_G)) = \mathbb{V}(\{M_{p,2} + M_{\leq 1}(G)\} \cap \mathbb{R}[\sigma_{ij}, i < j]),$$

where $M_{\leq 1}(G)$ is the ideal generated by all degree one monomials and tetrads corresponding to G.

Join of Varieties:

$$W_1 * W_2 = \overline{\{\lambda w_1 + (1 - \lambda) w_2 : w_1 \in W_1, w_2 \in W_2, \lambda \in \mathbb{R}\}}$$

(Sparse) Factor Models = Joins of (Sparse) One-Factor Models: For two ideals I_1, I_2, one defines the join ideal $I_1 * I_2$ such that $I_1 * I_2 = I(\mathbb{V}(I_1) * \mathbb{V}(I_2))$. $I(F_G) = I_1 * I_2$.



Delightful Strategy for Gröbner Basis

Observation:

$\mathrm{in}_\prec(I_1 * I_2) \subseteq \mathrm{in}_\prec(I_1) * \mathrm{in}_\prec(I_2)$ for any term order $\prec$. If equality holds, then $\prec$ is *delightful* for I_1, I_2.

"Delightful" strategy: [Sturmfels, Sullivant, 2006]

Find $G \subseteq I_1 * I_2$ such that $\langle \mathrm{in}_\prec(g) | g \in G \rangle = \mathrm{in}_\prec(I_1) * \mathrm{in}_\prec(I_2)$. Then G is a Gröbner basis.

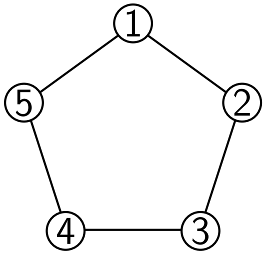
1. Find delightful term order $\prec$.
2. Understand $\mathrm{in}_\prec(I_1) * \mathrm{in}_\prec(I_2)$.
3. Define polynomials $g \in G$ with correct initial terms.

Circular Term Order

Circular order:

Any block-term order $\prec$ such that $\sigma_{uv} \succ \sigma_{wz}$ whenever the circular distance between u and v is smaller than the circular distance of w and z.

Example:



- $\sigma_{15} \succ \sigma_{24}$,
- $\sigma_{34} \succ \sigma_{25}$.

- $\sigma_{15}\sigma_{24} - \sigma_{25}\sigma_{14}$, $\sigma_{45}\sigma_{12} - \sigma_{25}\sigma_{14}$
- $\sigma_{34}\sigma_{25} - \sigma_{35}\sigma_{24}$, $\sigma_{45}\sigma_{23} - \sigma_{35}\sigma_{24}$.

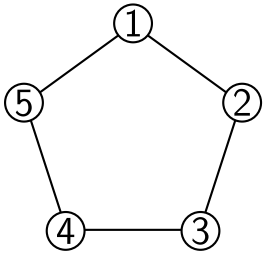
The ideal of the one-factor analysis model is the toric edge ideal of the complete graph on $A \subseteq [p]$ vertices with $B := [p] \backslash A$ isolated vertices. Its initial ideal is a monomial edge ideal.

Circular Term Order

Circular order:

Any block-term order $\prec$ such that $\sigma_{uv} \succ \sigma_{wz}$ whenever the circular distance between u and v is smaller than the circular distance of w and z.

Example:



- $\sigma_{15} \succ \sigma_{24}$,
- $\sigma_{34} \succ \sigma_{25}$.
- $\sigma_{15}\sigma_{24} - \sigma_{25}\sigma_{14}$, $\sigma_{45}\sigma_{12} - \sigma_{25}\sigma_{14}$
- $\sigma_{34}\sigma_{25} - \sigma_{35}\sigma_{24}$, $\sigma_{45}\sigma_{23} - \sigma_{35}\sigma_{24}$.

The ideal of the one-factor analysis model is the toric edge ideal of the complete graph on $A \subseteq [p]$ vertices with $B := [p] \backslash A$ isolated vertices. Its initial ideal is a monomial edge ideal.

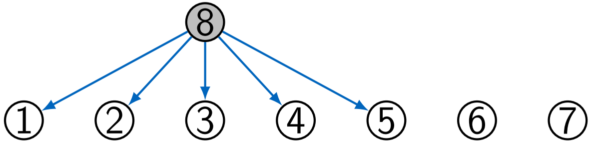
Gröbner basis: Let $B \subset [p]$ be the set of isolated vertices of the one-factor analysis graph and $A \sqcup B = V$.

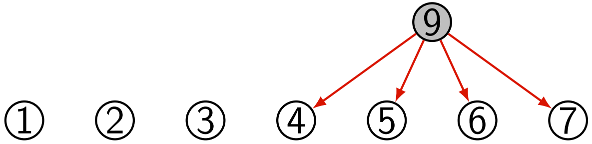
The set of degree-one monomials and tetrads

$$\{\sigma_{ij} \mid i \in B \text{ or } j \in B\} \cup \{\underline{\sigma_{ij}\sigma_{kl}} - \sigma_{ik}\sigma_{jl}, \underline{\sigma_{il}\sigma_{jk}} - \sigma_{ik}\sigma_{jl} \mid 1 \leq i < j < k < l \leq |A|\}.$$

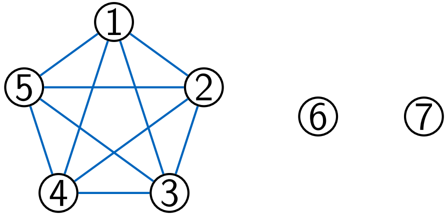
is a reduced Gröbner basis for the sparse one-factor analysis model with respect to any circular term order.

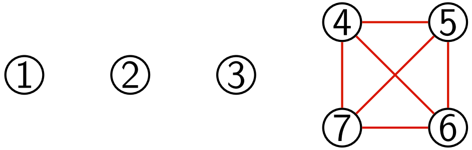
Non-Crossing Edge Graphs



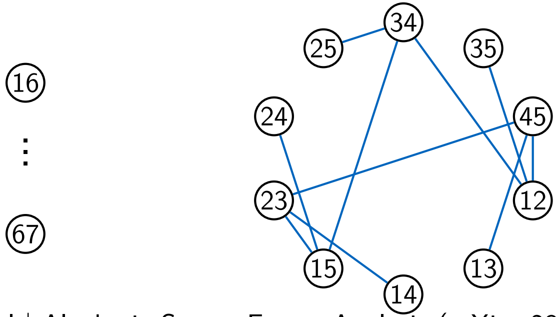


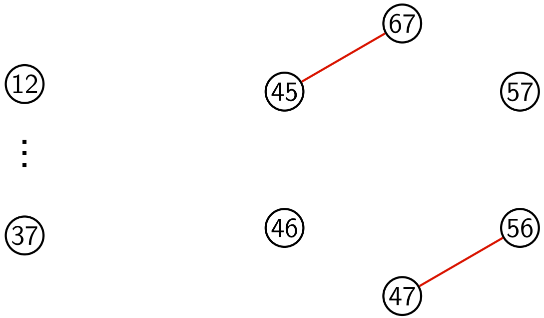
Complete graph on $v \in V$ s.t. $\mathrm{pa}(v) \neq \emptyset$:





Non-crossing edge graph:

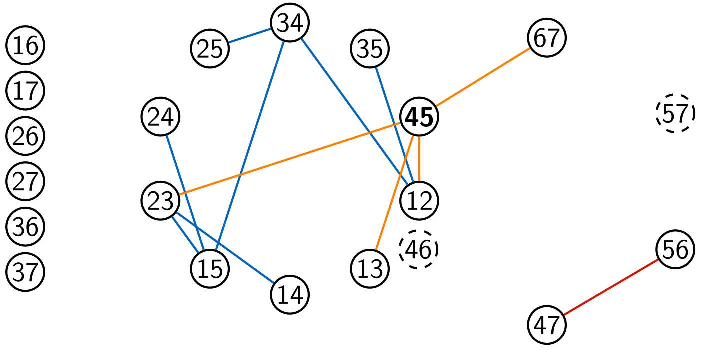




Gröbner Basis for "Overlap = 2"

$in_{\prec}(I_1) * in_{\prec}(I_2)$ is given by the edge ideal of the hypergraph obtained by gluing the non-crossing edge graphs.

Glued Hypergraph



Gröbner basis of $I_1 * I_2 = I(F_G)$

- $\sigma_{16}, \sigma_{17}, \sigma_{26}, \sigma_{27}, \sigma_{36}, \sigma_{37}.$
- $\sigma_{47}\sigma_{56} - \sigma_{57}\sigma_{46},$
  $\sigma_{12}\sigma_{34} - \sigma_{13}\sigma_{24},$
  $\sigma_{14}\sigma_{23} - \sigma_{13}\sigma_{24},$
  $\sigma_{12}\sigma_{35} - \sigma_{13}\sigma_{25},$
  $\sigma_{15}\sigma_{23} - \sigma_{13}\sigma_{25},$
  $\sigma_{15}\sigma_{24} - \sigma_{14}\sigma_{25},$
  $\sigma_{15}\sigma_{34} - \sigma_{14}\sigma_{35},$
  $\sigma_{25}\sigma_{34} - \sigma_{24}\sigma_{35}.$
- $\sigma_{67}\sigma_{12}\sigma_{45} - \sigma_{67}\sigma_{24}\sigma_{15} - \sigma_{12}\sigma_{47}\sigma_{56},$
  $\sigma_{67}\sigma_{13}\sigma_{45} - \sigma_{67}\sigma_{34}\sigma_{15} - \sigma_{13}\sigma_{47}\sigma_{56},$
  $\sigma_{67}\sigma_{23}\sigma_{45} - \sigma_{67}\sigma_{34}\sigma_{25} - \sigma_{23}\sigma_{47}\sigma_{56}.$

Conclusion

Theorem [DG🍊S]: The generators of a Gröbner basis for $I_{A_1,B_1,1} * I_{A_2,B_2,1}$ with respect to any circular term order for sparse two-factor analysis models where $A_1 \cap A_2 = \{j_1, j_2\}$ comes in three types:

1. Degree-one monomial: σ_{ik} is a generator, where $\mathrm{pa}(i) \cap \mathrm{pa}(j) = \emptyset$.
2. Tetrads: The binomial generators of the Gröbner basis of $I_{A_1,B_1,1}$ and $I_{A_2,B_2,1}$ with respect to any circular order that do not contain $\sigma_{j_1 j_2}$.
3. Hexads: Consider $i_1, i_2 \in A_1 \backslash \{j_1, j_2\}$ and $k_1, k_2 \in A_2 \backslash \{j_1, j_2\}$. Then

$$\underline{\sigma_{k_1 k_2} \sigma_{i_1 i_2} \sigma_{j_1 j_2}} - \sigma_{k_1 k_2} \sigma_{j_1 i_2} \sigma_{j_2 i_1} - \sigma_{i_1 i_2} \sigma_{j_1 k_2} \sigma_{j_2 k_1},$$

   is a degree three generator, where $\{i_1, i_2\}, \{j_1, j_2\}$ and $\{j_1, j_2\}, \{k_1, k_2\}$ are non-crossing edges of the complete graphs on the vertices $A_1 \backslash \{j_1, j_2\}$ and $A_2 \backslash \{j_1, j_2\}$ respectively.

Question: Is there a 2-delightful term order for sparse two-factor analysis models?

Conjecture: The ideal of the sparse two-factor analysis model corresponding to graph G is generated by off-diagonal 3×3-minors, pentads, and the polynomials in $M_{\leq 1}(G)$.