Model selection in the class of Gaussian models invariant under a subgroup of the symmetric group

Bartosz Kołodziejek

Warsaw University of Technology

Graphical Models: Conditional Independence and Algebraic Structures

Collaboration: Piotr Graczyk (Angers), Hideyuki Ishi (Osaka), Hélène Massam (Toronto)

October 23-25, 2019

Agenda

1. Colored graphical models
2. Main technical results
    1. Block decomposition of colored spaces
    2. Short intro to representation theory
    3. Gamma integrals
    4. Structure constants
    5. Specification to cyclic groups
3. RCOP-Wishart laws
4. Bayesian model selection
    1. Small p example - Frets' heads
    2. Arbitrary p - within cyclic groups (simulations)
5. Future work

Agenda

1. **Colored graphical models**
2. Main technical results
    1. Block decomposition of colored spaces
    2. Short intro to representation theory
    3. Gamma integrals
    4. Structure constants
    5. Specification to cyclic groups
3. RCOP-Wishart laws
4. Bayesian model selection
    1. Small p example - Frets' heads
    2. Arbitrary p - within cyclic groups (simulations)
5. Future work

- In order to make Graphical Gaussian Models a viable modeling tool when the number of variables outgrows the number of observations, $p >> n$, Højsgaard and Lauritzen (2008) propose models which impose equality restrictions on certain entries of precision matrix or partial correlation matrix.
- Such models can be represented by **colored graphs**: colored vertices and edges code the equality of entries of the matrix.
- Three types of restrictions on graphical Gaussian models are:
  1. RCON models
  2. RCOR models
  3. **RCOP models** - restrictions on covariance matrix are generated by a permutation subgroup

$$(\mathrm{RCOP}) \subsetneq (\mathrm{RCON}) \cap (\mathrm{RCOR})$$

RCOP colored spaces

- We will consider only **full graphs**.
- For a subgroup $\Gamma \subset \mathfrak{S}_p$, we define the space of symmetric matrices invariant under Γ, or the **colored space**,

$$\boxed{\mathcal{Z}_\Gamma := \left\{ x \in \mathrm{Sym}(p;\mathbb{R})\,;\; x_{ij} = x_{\sigma(i)\sigma(j)} \text{ for all } \sigma \in \Gamma \right\},}$$

and the cone of positive definite matrices in $\mathcal{Z}_\Gamma$,

$$\boxed{\mathcal{P}_\Gamma := \mathcal{Z}_\Gamma \cap \mathrm{Sym}^+(p;\mathbb{R}).}$$

- Equivalently,

$$\mathcal{Z}_\Gamma = \left\{ x \in \mathrm{Sym}(p;\mathbb{R})\,;\; R(\sigma)\cdot x = x \cdot R(\sigma) \text{ for all } \sigma \in \Gamma \right\},$$

where $R(\sigma)$ denotes the (permutation) matrix of σ.

Example

Let $p = 3$ and $\Gamma = \langle (123) \rangle$. We have $R((123)) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ and $R((123)) \cdot x = x \cdot R((123))$ implies

$$x = \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix}.$$

Thus,

$$\mathcal{Z}_{\langle (123) \rangle} = \left\{ \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix} ; a, b \in \mathbb{R} \right\}.$$

It is easily seen that we also have

$$\mathcal{Z}_{\mathfrak{S}_3} = \mathcal{Z}_{\langle (123) \rangle}.$$

- Same colored space can be generated by different subgroups.
- Let us define

$$\Gamma^* = \left\{ \sigma^* \in \mathfrak{S}_p \,;\, x_{ij} = x_{\sigma^*(i)\sigma^*(j)} \text{ for all } x \in \mathcal{Z}_\Gamma \right\}.$$

  Clearly, Γ is a subgroup of Γ^* and Γ^* is the unique largest subgroup of $\mathfrak{S}_p$ such that $\mathcal{Z}_{\Gamma^*} = \mathcal{Z}_\Gamma$.
- Wielandt (1969) and Siemons (1982, 1983)

Lemma

If $\mathcal{Z}_{\langle\sigma_0\rangle} = \mathcal{Z}_{\langle\sigma\rangle}$ for some $\sigma_0, \sigma \in \mathfrak{S}_p$, then $\langle\sigma_0\rangle = \langle\sigma\rangle$.

Agenda

1. Colored graphical models
2. Main technical results
   1. **Block decomposition of colored spaces**
   2. Short intro to representation theory
   3. Gamma integrals
   4. Structure constants
   5. Specification to cyclic groups
3. RCOP-Wishart laws
4. Bayesian model selection
   1. Small p example - Frets' heads
   2. Arbitrary p - within cyclic groups (simulations)
5. Future work

Notation

- We write $B^{\oplus r}$ for $I_r \otimes B$, that is,

$$B^{\oplus r} = \begin{pmatrix} B & & \\ & \ddots & \\ & & B \end{pmatrix}$$

- Let $M_{\mathbb{K}}$ be a real matrix representations of space $\mathrm{Herm}(r; \mathbb{K})$ for $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$.
  - $M_{\mathbb{R}} = \mathrm{Id}_{\mathrm{Sym}(r;\mathbb{R})}$.
  - For $z = a + b\,i \in \mathbb{C}$ define $M_{\mathbb{C}}(z) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

    $r \times r$ complex matrix can be realized as a $(2r) \times (2r)$ real matrix by setting the correspondence

$$\mathrm{Herm}(r; \mathbb{C}) \ni \left(z_{i,j}\right)_{1 \leqslant i,j \leqslant r} \simeq \left(M_{\mathbb{C}}(z_{i,j})\right)_{1 \leqslant i,j \leqslant r} \in \mathrm{Sym}(2r; \mathbb{R}).$$

  - Similarly we can define $M_{\mathbb{H}} \colon \mathrm{Herm}(r; \mathbb{H}) \to \mathrm{Sym}(4r; \mathbb{R})$.

Theorem

The space $\mathcal{Z}_\Gamma$ coincides with

$$\left\{ U_\Gamma \begin{pmatrix} M_{\mathbb{K}_1}(x_1)^{\oplus k_1/d_1} & & \\ & \ddots & \\ & & M_{\mathbb{K}_L}(x_L)^{\oplus k_L/d_L} \end{pmatrix} U_\Gamma^\top \; ; \; \begin{array}{r} x_i \in \mathrm{Herm}(r_i;\mathbb{K}_i) \\ i = 1,2,\ldots,L \end{array} \right\},$$

where

- *U_Γ is an orthogonal matrix,*
- *$(k_i, d_i, r_i)_{i=1}^L$ are the* ***structure constants****, which depend on Γ,*
- *$M_{\mathbb{K}}(x)$ is the real symmetric matrix representation of a Hermitian matrix x with values in $\mathbb{K}$,*
- *$\mathbb{K}_i \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$.*

Andersson (1975), Andersson and Madsen (1998)

- If $X \in \mathcal{Z}_\Gamma$ is as above, let $\phi_i(X) := x_i \in \mathrm{Herm}(r_i;\mathbb{K}_i)$.

Example

For $p = 3$ and $\Gamma = \langle (123) \rangle$, we have

$$\mathcal{Z}_\Gamma = \left\{ \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix} \; ; \, a, b \in \mathbb{R} \right\}.$$

Let $U_\Gamma = \begin{pmatrix} 1/\sqrt{3} & \sqrt{2/3} & 1 \\ 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \end{pmatrix}$.

Then,

$$\begin{aligned} U_\Gamma^\top \mathcal{Z}_\Gamma U_\Gamma &= \left\{ \begin{pmatrix} a+2b & 0 & 0 \\ 0 & a-b & 0 \\ 0 & 0 & a-b \end{pmatrix} \; ; \, a, b \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} x_1 & \\ & x_2^{\oplus 2} \end{pmatrix} \; ; \, x_1, x_2 \in \mathbb{R} \right\}. \end{aligned}$$

We have

$$(r_1, r_2) = (d_1, d_2) = (1, 1), \qquad (k_1, k_2) = (1, 2).$$

Example

For $p = 4$ and $\Gamma = \langle (12) \rangle$, we have

$$\mathcal{Z}_\Gamma = \left\{ \begin{pmatrix} a & b & c & d \\ b & a & c & d \\ c & c & e & f \\ d & d & f & g \end{pmatrix} ; a, b, c, d, e, f, g \in \mathbb{R} \right\}.$$

Let $U_\Gamma = \begin{pmatrix} 1/2 & 1/2 & 0 & 1/\sqrt{2} \\ 1/2 & 1/2 & 0 & -1/\sqrt{2} \\ 1/2 & -1/2 & 1/\sqrt{2} & 0 \\ 1/2 & -1/2 & -1/\sqrt{2} & 0 \end{pmatrix}$.

Then,

$$U_\Gamma^\top \mathcal{Z}_\Gamma U_\Gamma = \left\{ \begin{pmatrix} A & B & C & 0 \\ B & D & E & 0 \\ C & E & F & 0 \\ 0 & 0 & 0 & G \end{pmatrix} ; A, B, C, D, E, F, G \in \mathbb{R} \right\}$$

and $(r_1, r_2) = (3, 1)$, $(d_1, d_2) = (1, 1)$, $(k_1, k_2) = (1, 1)$,

$$\mathcal{Z}_\Gamma \simeq \mathrm{Sym}(3; \mathbb{R}) \oplus \mathbb{R}.$$

Agenda

1. Colored graphical models
2. Main technical results
    1. Block decomposition of colored spaces
    2. **Short intro to representation theory**
    3. Gamma integrals
    4. Structure constants
    5. Specification to cyclic groups
3. RCOP-Wishart laws
4. Bayesian model selection
    1. Small p example - Frets' heads
    2. Arbitrary p - within cyclic groups (simulations)
5. Future work

Sketch of the main argument

- $R\colon \Gamma \mapsto \mathrm{GL}(p;\mathbb{R})$ satisfies

$$\boxed{R(\sigma\circ\sigma') = R(\sigma)\cdot R(\sigma'), \qquad \sigma,\sigma'\in\mathfrak{S}_p.}$$

- In other words, R is a representation of group Γ.
- Observe that for any $\sigma\in\mathfrak{S}_p$

$$R(\sigma)\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}.$$

- The space $W_0 = \mathbb{R}(1,1,\ldots,1)^\top$ is a Γ invariant subspace for any subgroup Γ, that is, $\forall\sigma\in\Gamma$,

$$\forall w\in W_0 \qquad R(\sigma)w\in W_0.$$

- Similarly for $W_0^\perp$.

Sketch of the main argument

- Let orthogonal matrix U_Γ be constructed from a basis of W_0 (first column) and a basis of $W_0^\perp$. Then,

$$U_\Gamma^\top R(\sigma) U_\Gamma = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$$

- Recall that

$$\mathscr{Z}_\Gamma = \{\, x \in \mathrm{Sym}(p;\mathbb{R})\,;\; R(\sigma)\cdot x = x \cdot R(\sigma) \text{ for all } \sigma \in \Gamma \,\}.$$

- Then $U_\Gamma^\top \mathscr{Z}_\Gamma U_\Gamma$ coincides with

$$\{\, y \in \mathrm{Sym}(p;\mathbb{R})\,;\; [U_\Gamma^\top R(\sigma) U_\Gamma] \cdot y = y \cdot [U_\Gamma^\top R(\sigma) U_\Gamma] \,\}.$$

- Block decomposition of $U_\Gamma^\top R(\sigma) U_\Gamma$ implies block decomposition of $y \in U_\Gamma^\top \mathscr{Z}_\Gamma U_\Gamma$.
- In general, there exist proper Γ-invariant subspaces of $W_0^\perp$. Finding them is a very hard task.
- Structure constants arise from block decomposition of R.

Agenda

1. Colored graphical models
2. Main technical results
   1. Block decomposition of colored spaces
   2. Short intro to representation theory
   3. **Gamma integrals**
   4. Structure constants
   5. Specification to cyclic groups
3. RCOP-Wishart laws
4. Bayesian model selection
   1. Small p example - Frets' heads
   2. Arbitrary p - within cyclic groups (simulations)
5. Future work

Gamma integrals, part 1

- $\mathcal{P}_\Gamma = \mathcal{Z}_\Gamma \cap \mathrm{Sym}^+(p;\mathbb{R})$, $\Omega_i := \mathrm{Herm}^+(r_i;\mathbb{K}_i)$, $i = 1,\ldots,L$.
- For $Y \in \mathcal{P}_\Gamma$ define $\varphi_\Gamma(Y) = \prod_{i=1}^L (\det \phi_i(Y))^{-\dim \Omega_i / r_i}$.
- Let

$$I_1 := \int_{\mathcal{P}_\Gamma} \mathrm{Det}\,(X)^\lambda \, e^{-\mathrm{Tr}[Y\cdot X]} \varphi_\Gamma(X)\,\mathrm{d}X.$$

Theorem

The integral I_1 converges if and only if

- $\lambda > \max_{i=1,\ldots,L} \left\{ \frac{(r_i-1)d_i}{2k_i} \right\}$ *and*
- $Y \in \mathrm{Sym}^+(p,\mathbb{R})$.

If $Y \in \mathcal{P}_\Gamma$, then

$$I_1 = \frac{e^{-A_\Gamma\lambda + B_\Gamma} \prod_{i=1}^L \Gamma_{\Omega_i}(k_i\lambda)}{\mathrm{Det}\,(Y)^\lambda}$$

with A_Γ and B_Γ depending explicitly on structure constants only.

Gamma integrals, part 2

- Define

$$I_2 := \int_{\mathcal{P}_\Gamma} \mathrm{Det}\,(X)^\lambda \, e^{-\mathrm{Tr}[Y \cdot X]} \, \mathrm{d}X.$$

Theorem

The integral I_2 converges if and only if

- $\lambda > \max_{i=1,\ldots,L} \left\{ -\frac{1}{k_i} \right\}$ *and*
- $Y \in \mathrm{Sym}^+(p, \mathbb{R})$.

If $Y \in \mathcal{P}_\Gamma$, then

$$I_2 = e^{-A_\Gamma \lambda - B_\Gamma} \prod_{i=1}^{L} \Gamma_{\Omega_i}\left(k_i\,\lambda + \frac{\dim \Omega_i}{r_i}\right) \frac{\varphi_\Gamma(Y)}{\mathrm{Det}\,(Y)^\lambda}.$$

Agenda

1. Colored graphical models
2. Main technical results
    1. Block decomposition of colored spaces
    2. Short intro to representation theory
    3. Gamma integrals
    4. **Structure constants**
    5. Specification to cyclic groups
3. RCOP-Wishart laws
4. Bayesian model selection
    1. Small p example - Frets' heads
    2. Arbitrary p - within cyclic groups (simulations)
5. Future work

How to find structure constants?

- If $X \in \mathcal{Z}_\Gamma$, then

$$X = U_\Gamma \begin{pmatrix} M_{\mathbb{K}_1}(x_1)^{\oplus k_1/d_1} & & \\ & \ddots & \\ & & M_{\mathbb{K}_L}(x_L)^{\oplus k_L/d_L} \end{pmatrix} U_\Gamma^\top$$

  for some $x_i \in \mathrm{Herm}(r_i; \mathbb{K}_i)$, $i = 1, 2, \ldots, L$.
- Recall that $\phi_i(X) := x_i$.
- **In order to compute Gamma integrals on $\mathcal{P}_\Gamma$** we need to find $(r_i, d_i, k_i)_{i=1}^L$ and polynomials $\det \phi_i(X)$.

- In view of decomposition of $\mathcal{Z}_\Gamma$, we have

$$\boxed{\mathrm{Det}\,(X) = \prod_{i=1}^{L} (\det \phi_i(X))^{k_i}, \qquad X \in \mathcal{Z}_\Gamma.}$$

- Assume that we have an irreducible factorization

$$\boxed{\mathrm{Det}\,(X) = \prod_{j=1}^{L'} f_j(X)^{a_j}, \qquad X \in \mathcal{Z}_\Gamma,}$$

  where
  - each a_j is a positive integer,
  - each $f_j(X)$ is an irreducible polynomial of $X \in \mathcal{Z}_\Gamma$,
  - $f_i \neq f_j$ if $i \neq j$.
- **Basing on results in Jordan algebras, we can deduce that**
  - $L = L'$,
  - for each j, there exists i such that $f_j(X)^{a_j} = (\det \phi_i(X))^{k_i}$.
  - $k_i = a_j$ and r_i is the degree of $f_j(X) = \det \phi_i(X)$,
  - $\dim \Omega_i = r_i + d_i r_i \frac{(r_i - 1)}{2} = \mathrm{rank}[\mathrm{Hess}(\log f_j)(I)]$

Example

Let $\Gamma = \langle \sigma_1, \sigma_2 \rangle$ be a subgroup of $\mathfrak{S}_{16}$ generated by two permutations

$$\begin{aligned}
\sigma_1 &= (1,2,5,6)(3,4,7,8)(9,10,13,14)(11,12,15,16),\\
\sigma_2 &= (1,3,5,7)(2,8,6,4)(9,11,13,15)(10,16,14,12).
\end{aligned}$$

The space $\mathcal{Z}_\Gamma$ consists of matrices of the form

$$\left(\begin{array}{cccccccc|cccccccc}
\alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_2 & \alpha_3 & \alpha_4 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 & \gamma_5 & \gamma_6 & \gamma_7 & \gamma_8 \\
\alpha_2 & \alpha_1 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_5 & \alpha_4 & \alpha_3 & \gamma_6 & \gamma_1 & \gamma_8 & \gamma_3 & \gamma_2 & \gamma_5 & \gamma_4 & \gamma_7 \\
\alpha_3 & \alpha_4 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_2 & \gamma_7 & \gamma_4 & \gamma_1 & \gamma_6 & \gamma_3 & \gamma_8 & \gamma_5 & \gamma_2 \\
\alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_5 & \gamma_8 & \gamma_7 & \gamma_2 & \gamma_1 & \gamma_4 & \gamma_3 & \gamma_6 & \gamma_5 \\
\alpha_5 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \gamma_5 & \gamma_6 & \gamma_7 & \gamma_8 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\
\alpha_2 & \alpha_5 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & \alpha_4 & \alpha_3 & \gamma_2 & \gamma_5 & \gamma_4 & \gamma_7 & \gamma_6 & \gamma_1 & \gamma_8 & \gamma_3 \\
\alpha_3 & \alpha_4 & \alpha_5 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_1 & \alpha_2 & \gamma_3 & \gamma_8 & \gamma_5 & \gamma_2 & \gamma_7 & \gamma_4 & \gamma_1 & \gamma_6 \\
\alpha_4 & \alpha_3 & \alpha_2 & \alpha_5 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & \gamma_4 & \gamma_3 & \gamma_6 & \gamma_5 & \gamma_8 & \gamma_7 & \gamma_2 & \gamma_1 \\
\hline
\gamma_1 & \gamma_6 & \gamma_7 & \gamma_8 & \gamma_5 & \gamma_2 & \gamma_3 & \gamma_4 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_2 & \beta_3 & \beta_4 \\
\gamma_2 & \gamma_1 & \gamma_4 & \gamma_7 & \gamma_6 & \gamma_5 & \gamma_8 & \gamma_3 & \beta_2 & \beta_1 & \beta_4 & \beta_3 & \beta_2 & \beta_5 & \beta_4 & \beta_3 \\
\gamma_3 & \gamma_8 & \gamma_1 & \gamma_2 & \gamma_7 & \gamma_4 & \gamma_5 & \gamma_6 & \beta_3 & \beta_4 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_2 \\
\gamma_4 & \gamma_3 & \gamma_6 & \gamma_1 & \gamma_8 & \gamma_7 & \gamma_2 & \gamma_5 & \beta_4 & \beta_3 & \beta_2 & \beta_1 & \beta_4 & \beta_3 & \beta_2 & \beta_5 \\
\gamma_5 & \gamma_2 & \gamma_3 & \gamma_4 & \gamma_1 & \gamma_6 & \gamma_7 & \gamma_8 & \beta_5 & \beta_2 & \beta_3 & \beta_4 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\
\gamma_6 & \gamma_5 & \gamma_8 & \gamma_3 & \gamma_2 & \gamma_1 & \gamma_4 & \gamma_7 & \beta_2 & \beta_5 & \beta_4 & \beta_3 & \beta_2 & \beta_1 & \beta_4 & \beta_3 \\
\gamma_7 & \gamma_4 & \gamma_5 & \gamma_6 & \gamma_3 & \gamma_8 & \gamma_1 & \gamma_2 & \beta_3 & \beta_4 & \beta_5 & \beta_2 & \beta_3 & \beta_4 & \beta_1 & \beta_2 \\
\gamma_8 & \gamma_7 & \gamma_2 & \gamma_5 & \gamma_4 & \gamma_3 & \gamma_6 & \gamma_1 & \beta_4 & \beta_3 & \beta_2 & \beta_5 & \beta_4 & \beta_3 & \beta_2 & \beta_1
\end{array}\right).$$

Example

$$\begin{aligned}
&\operatorname{Det}(X)\\
&= \Big((\gamma_1-\gamma_5)^2+(\gamma_2-\gamma_6)^2+(\gamma_3-\gamma_7)^2+(\gamma_4-\gamma_8)^2-(\alpha_1-\alpha_5)(\beta_1-\beta_5)\Big)^4\\
&\cdot\Big((\gamma_1-\gamma_2-\gamma_3+\gamma_4+\gamma_5-\gamma_6-\gamma_7+\gamma_8)^2-(\alpha_1-2(\alpha_2+\alpha_3-\alpha_4)+\alpha_5)(\beta_1-2(\beta_2+\beta_3-\beta_4)+\beta_5)\Big)\\
&\cdot\Big((\gamma_1-\gamma_2+\gamma_3-\gamma_4+\gamma_5-\gamma_6+\gamma_7-\gamma_8)^2-(\alpha_1-2(\alpha_2-\alpha_3+\alpha_4)+\alpha_5)(\beta_1-2(\beta_2-\beta_3+\beta_4)+\beta_5)\Big)\\
&\cdot\Big((\gamma_1+\gamma_2-\gamma_3-\gamma_4+\gamma_5+\gamma_6-\gamma_7-\gamma_8)^2-(\alpha_1+2(\alpha_2-\alpha_3-\alpha_4)+\alpha_5)(\beta_1+2(\beta_2-\beta_3-\beta_4)+\beta_5)\Big)\\
&\cdot\Big((\gamma_1+\gamma_2+\gamma_3+\gamma_4+\gamma_5+\gamma_6+\gamma_7+\gamma_8)^2-(\alpha_1+2(\alpha_2+\alpha_3+\alpha_4)+\alpha_5)(\beta_1+2(\beta_2+\beta_3+\beta_4)+\beta_5)\Big),
\end{aligned}$$

. Thus,

$$r=(2,2,2,2,2),\quad k=(4,1,1,1,1),\quad d=(4,1,1,1,1).$$

This in turn implies

$$\mathcal{Z}_\Gamma \simeq \mathrm{Herm}(2;\mathbb{H})\oplus\mathrm{Sym}(2;\mathbb{R})\oplus\mathrm{Sym}(2;\mathbb{R})\oplus\mathrm{Sym}(2;\mathbb{R})\oplus\mathrm{Sym}(2;\mathbb{R}).$$

Agenda

1. Colored graphical models
2. Main technical results
    1. Block decomposition of colored spaces
    2. Short intro to representation theory
    3. Gamma integrals
    4. Structure constants
    5. **Specification to cyclic groups**
3. RCOP-Wishart laws
4. Bayesian model selection
    1. Small p example - Frets' heads
    2. Arbitrary p - within cyclic groups (simulations)
5. Future work

Cyclic groups

- In the case of **cyclic** Γ the orthogonal matrix U_Γ can be constructed **explicitly**, and we obtain the structure constants r_i, k_i and d_i **easily**.
- Let us consider $\Gamma = \langle \sigma \rangle$ with

$$\sigma = \underbrace{(i_1 \ldots)}_{p_1}\underbrace{(i_2 \ldots)}_{p_2} \ldots \underbrace{(i_C \ldots)}_{p_C}$$

Theorem

Let $\Gamma = \langle \sigma \rangle$ be a cyclic group of order N. For $\alpha = 0, 1, \ldots, \lfloor \frac{N}{2} \rfloor$ set

$$r^*_\alpha = \#\{\, c \in \{1, \ldots, C\}\,;\, \alpha\, p_c \text{ is a multiple of } N \,\}$$

$$d^*_\alpha = \begin{cases} 1 & (\alpha = 0 \text{ or } N/2) \\ 2 & (\text{otherwise}). \end{cases}$$

Then, $L = \#\{\, \alpha\,;\, r^*_\alpha > 0 \,\}$,

$$r = (r^*_\alpha;\, r^*_\alpha > 0) \quad \text{and} \quad k = d = (d^*_\alpha;\, r^*_\alpha > 0).$$

Let $(e_i)_{i=1}^p$ denote the standard basis of $\mathbb{R}^p$.

Theorem

The orthogonal matrix U_Γ ***can constructed*** *by arranging column vectors* $v_k^{(c)}$ *in an* ***appropriate*** *order, where* $v_1^{(c)}, \ldots, v_{p_c}^{(c)} \in \mathbb{R}^p$ *by*

$$v_1^{(c)} := \sqrt{\frac{1}{p_c}} \sum_{k=0}^{p_c-1} e_{\sigma^k(i_c)},$$

$$v_{2\beta}^{(c)} := \sqrt{\frac{2}{p_c}} \sum_{k=0}^{p_c-1} \cos\left(\frac{2\pi\beta k}{p_c}\right) e_{\sigma^k(i_c)} \qquad (1 \leqslant \beta < p_c/2),$$

$$v_{2\beta+1}^{(c)} := \sqrt{\frac{2}{p_c}} \sum_{k=0}^{p_c-1} \sin\left(\frac{2\pi\beta k}{p_c}\right) e_{\sigma^k(i_c)} \qquad (1 \leqslant \beta < p_c/2),$$

$$v_{p_c}^{(c)} := \sqrt{\frac{1}{p_c}} \sum_{k=0}^{p_c-1} \cos(\pi k) e_{\sigma^k(i_c)} \qquad (\textit{if } p_c \textit{ is even}).$$

Example

- Let us consider $\sigma = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 & 5 \end{pmatrix} \begin{pmatrix} 6 \end{pmatrix} \in \mathfrak{S}_6$.
- We have $p_1 = 3$, $p_2 = 2$, $p_3 = 1$ and $N = LCM(p_1, p_2, p_3) = 6$.
- We count $r_0^* = 3$, $r_1^* = 0$, $r_2^* = 1$, $r_3^* = 1$, so that,

$$r = (3, 1, 1) \qquad \text{and} \qquad d = k = (1, 2, 1).$$

- Thus, $\mathcal{Z}_\Gamma \simeq \mathrm{Sym}(3; \mathbb{R}) \oplus \mathrm{Herm}(1; \mathbb{C}) \oplus \mathrm{Sym}(1; \mathbb{R})$.
- Moreover,

$$U_\Gamma = \begin{pmatrix} 1/\sqrt{3} & 0 & 0 & \sqrt{2/3} & 0 & 0 \\ 1/\sqrt{3} & 0 & 0 & -\sqrt{1/6} & 1/\sqrt{2} & 0 \\ 1/\sqrt{3} & 0 & 0 & -\sqrt{1/6} & -1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 0 & 0 & 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 & 0 & 0 & -1/\sqrt{2} \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Agenda

1. Colored graphical models
2. Main technical results
    1. Block decomposition of colored spaces
    2. Short intro to representation theory
    3. Gamma integrals
    4. Structure constants
    5. Specification to cyclic groups
3. **RCOP-Wishart laws**
4. Bayesian model selection
    1. Small p example - Frets' heads
    2. Arbitrary p - within cyclic groups (simulations)
5. Future work

RCOP-Wishart laws

- Let $\pi_\Gamma \colon \mathrm{Sym}(p;\mathbb{R}) \to \mathcal{Z}_\Gamma$ be the projection

$$\pi_\Gamma(x) = \frac{1}{|\Gamma|}\sum_{\sigma\in\Gamma} R(\sigma)\cdot x\cdot R(\sigma)^\top$$

- Let $\Sigma \in \mathcal{P}_\Gamma \subset \mathrm{Sym}^+(p;\mathbb{R})$ and let $Z_1,\ldots,Z_n$ be iid from $\mathrm{N}_p(0,\Sigma)$. Define

$$W_n = \pi_\Gamma\left(\sum_{i=1}^n Z_i\cdot Z_i^\top\right).$$

Theorem

The law of W_n is absolutely continuous if and only if

$$n \geqslant n_0 := \max_{i=1,\ldots,L}\left\{\frac{r_i d_i}{k_i}\right\}.$$

If $n \geqslant n_0$, then its density function is given by

$$\frac{\mathrm{Det}\,(X)^{n/2}\,e^{-\frac{1}{2}\mathrm{Tr}[X\cdot\Sigma^{-1}]}}{\mathrm{Det}\,(2\Sigma)^{n/2}\,\Gamma_{\mathcal{P}_\Gamma}\left(\frac{n}{2}\right)}\varphi_\Gamma(X)\mathbf{1}_{\mathcal{P}_\Gamma}(X).$$

Agenda

1. Colored graphical models
2. Main technical results
    1. Block decomposition of colored spaces
    2. Short intro to representation theory
    3. Gamma integrals
    4. Structure constants
    5. Specification to cyclic groups
3. RCOP-Wishart laws
4. **Bayesian model selection**
    1. Small p example - Frets' heads
    2. Arbitrary p - within cyclic groups (simulations)
5. Future work

- Bayesian model search on all colored spaces seems at the moment intractable. There are two big **obstacles**:
  1. Lattice structure of $\{\, \mathcal{Z}_\Gamma \,;\, \Gamma \subset \mathfrak{S}_p \,\}$ (or equivalently, of $\{\Gamma^*\}$) is very complicated and it seems very hard to propose a consistent approach for travelling through the space of colors.
  2. It is in general **impossible** to find structure constants for arbitrary group Γ. How about Γ^*?
- We are making a small step forward and we propose a model selection procedure **restricted to cyclic colorings**, that is, when $\Gamma = \langle\sigma\rangle$ for $\sigma \in \mathfrak{S}_p$. This smaller space has much better combinatorial description.

Bayesian model selection

- We take prior on Γ to be uniform on all (cyclic) subgroups of $\mathfrak{S}_p$.
- Let $K = \Sigma^{-1}$. We will assume that $K|\Gamma$ is the Diaconis-Ylvisaker conjugate prior for K, that is,

$$f_{K|\Gamma}(k) = \frac{1}{I_\Gamma(\delta, D)} \mathrm{Det}\,(k)^{(\delta-2)/2}\, e^{-\frac{1}{2}\mathrm{Tr}[D\cdot k]} \mathbf{1}_{\mathcal{P}_\Gamma}(k).$$

- We assume that $Z_1, \ldots, Z_n$ given $\{K, \Gamma\}$ are i.i.d. $\mathrm{N}_p(0, K^{-1})$ random vectors with $K \in \mathcal{P}_\Gamma$.
- Then, it is easily seen that

$$\mathbb{P}(\Gamma | Z_1, \ldots, Z_n) \propto \frac{I_\Gamma(\delta + n, D + \sum_{i=1}^n Z_i \cdot Z_i^\top)}{I_\Gamma(\delta, D)}.$$

- For small p we calculate all possibilities.
- For big p we run Metropolis-Hastings algorithm.

- In a Bayesian framework, the classical approach for choosing between two models is to compute their posterior probability density and choose the model with the highest posterior probability.
- We look for

$$\hat{\Gamma} = \arg\max_{\Gamma} \frac{I_\Gamma(\delta + n, D + \sum_{i=1}^n Z_i \cdot Z_i^\top)}{I_\Gamma(\delta, D)}.$$

Agenda

1. Colored graphical models
2. Main technical results
    1. Block decomposition of colored spaces
    2. Short intro to representation theory
    3. Gamma integrals
    4. Structure constants
    5. Specification to cyclic groups
3. RCOP-Wishart laws
4. Bayesian model selection
    1. Small p example - Frets' heads
    2. Arbitrary p - within cyclic groups (simulations)
5. Future work

Structure constants for $p = 4$

- There are 22 different RCOP colorings.
- Up to conjugacy (renumbering of vertices), there are 8 different conjugacy classes.

Group	(k_i)	(r_i)	(d_i)
$\Gamma_1^* = \{\mathrm{id}\}$	(1)	(4)	(1)
$\Gamma_2^* = \langle(12)\rangle$, $\Gamma_3^* = \langle(13)\rangle$ $\Gamma_4^* = \langle(14)\rangle$, $\Gamma_5^* = \langle(23)\rangle$ $\Gamma_6^* = \langle(24)\rangle$, $\Gamma_7^* = \langle(34)\rangle$	(1,1)	(3,1)	(1,1)
$\Gamma_8^* = \langle(123),(12)\rangle$, $\Gamma_9^* = \langle(124),(12)\rangle$ $\Gamma_{10}^* = \langle(134),(13)\rangle$, $\Gamma_{11}^* = \langle(234),(23)\rangle$	(1,2)	(2,1)	(1,1)
$\Gamma_{12}^* = \langle(12)(34)\rangle$, $\Gamma_{13}^* = \langle(13)(24)\rangle$ $\Gamma_{14}^* = \langle(14)(23)\rangle$	(1,1)	(2,2)	(1,1)
$\Gamma_{15}^* = \langle(1234),(13)\rangle$, $\Gamma_{16}^* = \langle(1243),(14)\rangle$ $\Gamma_{17}^* = \langle(1324),(12)\rangle$	(1,1,2)	(1,1,1)	(1,1,1)
$\Gamma_{18}^* = \langle(12),(34)\rangle$, $\Gamma_{19}^* = \langle(13),(24)\rangle$ $\Gamma_{20}^* = \langle(14),(23)\rangle$	(1,1,1)	(2,1,1)	(1,1,1)
$\Gamma_{21}^* = \langle(12)(34),(14)(23)\rangle$	(1,1,1,1)	(1,1,1,1)	(1,1,1,1)
$\Gamma_{22}^* = \mathfrak{S}_4$	(1,3)	(1,1)	(1,1)

Hasse diagram

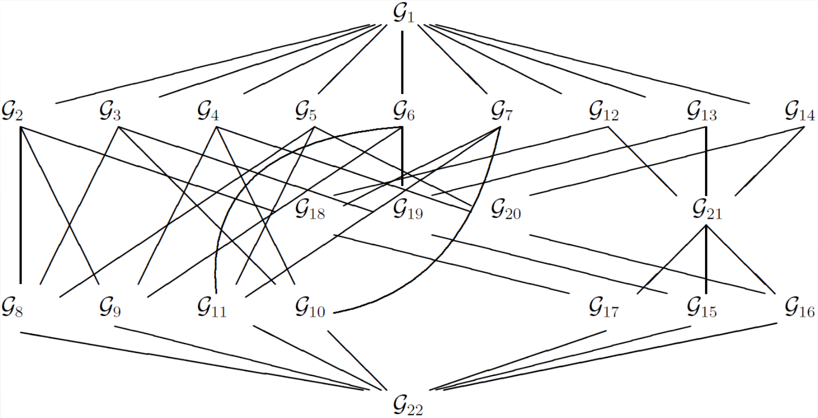


Figure borrowed from Gehrmann (2011).

Frets' heads

- The head dimensions (length L_i and breadth B_i, $i = 1, 2$) of 25 pairs of first and second sons were measured.
- $n = 25$, $p = 4$, $V = (L_1, B_1, L_2, B_2)$. We take $\delta = 3$.

$$\sum_{i=1}^{n} Z_i \cdot Z_i^\top = \begin{pmatrix} 2287.04 & 1268.84 & 1671.88 & 1106.68 \\ 1268.84 & 1304.64 & 1231.48 & 841.28 \\ 1671.88 & 1231.48 & 2419.36 & 1356.96 \\ 1106.68 & 841.28 & 1356.96 & 1080.56 \end{pmatrix}.$$

- Posterior probabilities:

D	Best model		2nd best		3rd best	
I_4	Γ^*_{22}	(95.2%)	Γ^*_{16}	(2.5%)	Γ^*_{17}	(1.3%)
$50 I_4$	Γ^*_{19}	(33.8%)	Γ^*_{13}	(29.6%)	Γ^*_{8}	(13.3%)
$100 I_4$	Γ^*_{13}	(39.6%)	Γ^*_{19}	(29.8%)	Γ^*_{8}	(7.2%)
$1000 I_4$	Γ^*_{1}	(38.9%)	Γ^*_{13}	(10.5%)	Γ^*_{3}	(10.3%)

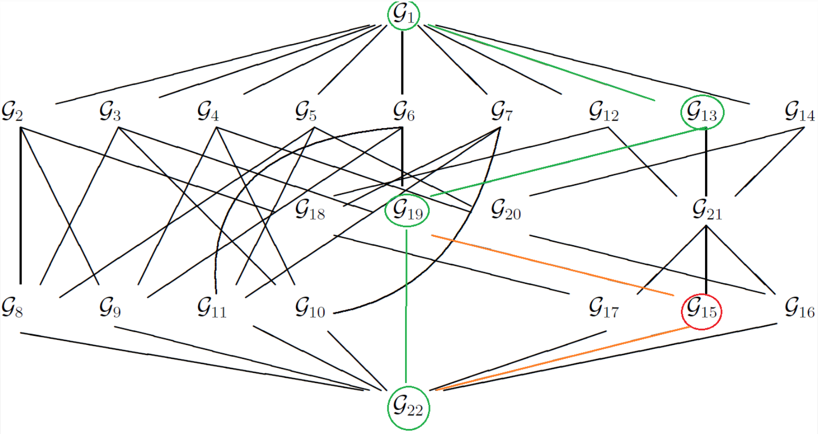


Figure borrowed from Gehrmann (2011).

D	Best model		2nd best		3rd best	
I_4	Γ_{22}^*	(95.2%)	Γ_{16}^*	(2.5%)	Γ_{17}^*	(1.3%)
$50I_4$	Γ_{19}^*	(33.8%)	Γ_{13}^*	(29.6%)	Γ_{8}^*	(13.3%)
$100I_4$	Γ_{13}^*	(39.6%)	Γ_{19}^*	(29.8%)	Γ_{8}^*	(7.2%)
$1000I_4$	Γ_{1}^*	(38.9%)	Γ_{13}^*	(10.5%)	Γ_{3}^*	(10.3%)

- For different values of $D = dI_4$, the only models with highest posterior probability are:
  - $\Gamma_{22}^* = \mathfrak{S}_4$,
  - $\Gamma_{19}^* = \langle(13),(24)\rangle$,
  - $\Gamma_{13}^* = \langle(13)(24)\rangle$,
  - $\Gamma_{1}^* = \{\mathrm{id}\}$.
- Recall the enumeration of vertices $(1,2,3,4) = (L_1, B_1, L_2, B_2)$. The invariance with respect to the transposition (13) means that L_1 is exchangeable with L_2 and, similarly, the invariance with respect to the transposition (24) implies exchangability of B_1 and B_2. Both together correspond to the fact that sons should be exchangable in some way.

Agenda

1. Colored graphical models
2. Main technical results
   1. Block decomposition of colored spaces
   2. Short intro to representation theory
   3. Gamma integrals
   4. Structure constants
   5. Specification to cyclic groups
3. RCOP-Wishart laws
4. Bayesian model selection
   1. Small p example - Frets' heads
   2. Arbitrary p - within cyclic groups (simulations)
5. Future work

- First, we introduce a Markov chain on the space of all permutations:

$$\sigma_t = \sigma_{t-1} \circ x_t, \qquad (x_t)_t \text{ are i.i.d. transpositions.}$$

- $(\sigma_t)_t$ induces a Markov chain on the space of cyclic groups, $(\langle \sigma_t \rangle)_t$, but we loose uniformity: it may happen that

$$\langle \sigma_{t-1} \circ x_t \rangle = \langle \sigma_{t-1} \circ x_t' \rangle \text{ for } x_t \neq x_t'.$$

- We choose the proposal distribution g to be proportional to the number of possible transitions from $\langle \sigma \rangle$ to $\langle \sigma' \rangle$, that is,

$$g\left(\langle \sigma' \rangle \,|\, \langle \sigma \rangle\right) := \frac{\#\left\{\, (i,j) \in \mathfrak{S}_p \,;\, \sigma' = \sigma \circ (i,j) \,\right\}}{\binom{p}{2}}.$$

Metropolis-Hastings algorithm

Starting from a cyclic group $\Gamma_0 = \langle \sigma_0 \rangle$, repeat the following two steps for $t = 1, 2, \ldots$:

1. Sample x_t uniformly from the set of all transpositions and set

$$\sigma_t = \sigma_{t-1} \circ x_t;$$

2. Accept the move $\Gamma_t = \langle \sigma_t \rangle$ with probability

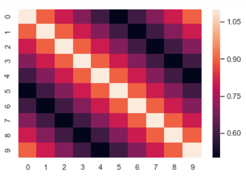
$$\min \left\{ 1, \frac{I_{\langle \sigma_t \rangle}(\delta + n, D + U) \cdot I_{\langle \sigma_{t-1} \rangle}(\delta, D)}{I_{\langle \sigma_t \rangle}(\delta, D) \cdot I_{\langle \sigma_{t-1} \rangle}(\delta + n, D + U)} \cdot \frac{g\left(\langle \sigma_{t-1} \rangle \mid \langle \sigma_t \rangle\right)}{g\left(\langle \sigma_t \rangle \mid \langle \sigma_{t-1} \rangle\right)} \right\}$$

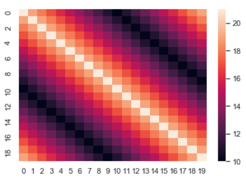
   If the move is rejected, set $\Gamma_t = \Gamma_{t-1}$ and $\sigma_t = \sigma_{t-1}$.

p	#subgroups of $\mathfrak{S}_p$	$\#\mathcal{Z}_\Gamma$	#cyclic groups
1	1	1	1
2	2	2	2
3	6	5	5
4	30	22	17
5	156	93	67
6	1 455	739	362
7	11 300	4 508	2039
8	151 221	?	14 170
9	1 694 723	?	109 694
10	29 594 446	?	976 412
18	$7 \cdot 10^{18}$	?	$7 \cdot 10^{14}$

Simulations

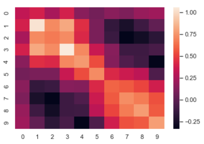
- We choose $p = 10, 20$
- $n = p$, $\delta = 3$, $D = I_p$.
- Let Σ_0 be a symmetric circular matrix of the form





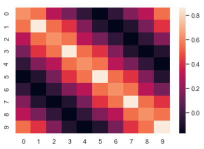
- We sample $Z_1, \ldots, Z_n$ from $\mathrm{N}_p(0, \Sigma_0)$, where Σ_0 is a symmetric circular matrix
- Σ_0 is invariant under $\Gamma_0 = \langle(1, 2, \ldots, p)\rangle$.
- We start Metropolis-Hastings algorithm with $\Gamma_0 = \{\mathrm{id}\}$ and iterate 500 000 times.

$p = 10$

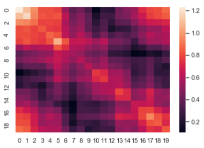
- There are $\approx 9 \cdot 10^7$ cyclic subgroups of $\mathfrak{S}_p$.
- $\underline{Z} \cdot \underline{Z}^\top / n$ equals 

- Acceptance rate $= 1.0\%$

$\langle \sigma \rangle$	$\hat{P}(\Gamma = \langle \sigma \rangle \vert \underline{Z})$
$\langle (0,2,4,6,8)(1,3,5,7,9) \rangle$	48.1%
$\langle (0,1,2,3,4,5,6,7,8,9) \rangle$	21.2%
$\langle (0,2,7,9,1,5,8,3,4,6) \rangle$	6.1%
$\langle (0,2,4,6,7)(1,3,5,8,9) \rangle$	4.3%
$\langle (0,6,4,5,1,9)(2,7)(3,8) \rangle$	1.3%

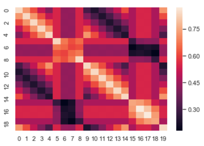
- First model is 

$p = 20$

- There are $\approx 2 \cdot 10^{17}$ cyclic subgroups of $\mathfrak{S}_p$.
- $\underline{Z} \cdot \underline{Z}^\top / n$ equals 

- Acceptance rate $= 0.29\%$

$\langle\sigma\rangle$	$\hat{P}(\Gamma = \langle\sigma\rangle \mid \underline{Z})$
$\langle(0-4, 9-14, 19)(5,8)(6,7)(15,18)(16,17)\rangle$	34.6%
$\langle(0-5, 8-15, 18, 19)(6,7)(16,17)\rangle$	26.1%
$\langle(0-15, 17, 18, 16, 19)\rangle$	16.6%
$\langle(0-5, 8-15, 18, 19)(16,17)\rangle$	4.4%
$\langle(0-5, 8-15, 18, 19)(6,7)\rangle$	1.9%

- First model is 

Agenda

1. Colored graphical models
2. Main technical results
    1. Block decomposition of colored spaces
    2. Short intro to representation theory
    3. Gamma integrals
    4. Structure constants
    5. Specification to cyclic groups
3. RCOP-Wishart laws
4. Bayesian model selection
    1. Small p example - Frets' heads
    2. Arbitrary p - within cyclic groups (simulations)
5. **Future work**

Future work

- We are able to compute gamma integrals for all RCOP models within **decomposable graphs**.
- We produce examples **outside RCOP**, for which we are still able to compute gamma integrals.
- Traveling through the space of models within **colored decomposable graphs**.

Thank you for your attention

Andersson, S. A. (1975) Invariant normal models. *Ann. Statist.*, **3**, 132–154.

Andersson, S. A. and Madsen, J. (1998) Symmetry and lattice conditional independence in a multivariate normal distribution. *Ann. Statist.*, **26**, 525–572.

Gehrmann, H. (2011) Lattices of graphical Gaussian models with symmetries. *Symmetry*, **3**, 653–679.,

Højsgaard, S. and Lauritzen, S. L. (2008) Graphical Gaussian models with edge and vertex symmetries. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **70**, 1005–1027.

Siemons, J. (1982) On partitions and permutation groups on unordered sets. *Arch. Math. (Basel)*, **38**, 391–403.

Siemons, J. (1983) Automorphism groups of graphs. *Arch. Math. (Basel)*, **41**, 379–384.

Wielandt, H. (1969) *Permutation groups through invariant relations and invariant functions*. Lect. Notes Dept. Math. Ohio St. Univ. Columbus.

Graczyk, P., Ishi. H, Kołodziejek, B., Massam, H. (2019) *Model selection in the space of Gaussian models invariant by symmetry*.
In preparation.