Topology of Graphs Representing Graphical Gaussian Models with Symmetry Constraints

Helene Gehrmann, gehrmann@stats.ox.ac.uk

Department of Statistics, University of Oxford, 1 South Parks Road, Oxford, OX1 3TG, United Kingdom

Abstract

Gaussian models with symmetries have been of interest for a long time (e.g. Votaw, 1948; Andersson, 1975) whereas the combination with conditional independence restrictions is more recent (e.g. Hylleberg et al., 1993). Højsgaard and Lauritzen (2008) introduced three types of graphical Gaussian models with symmetry constraints on their parameters, which can all be represented by vertex and edge coloured graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. They identified two sets of coloured graphs which lead to desirable statistical properties of the represented model. We specify two further such sets together with their implications for the represented models, and show that all four sets are lattices with respect to model inclusion. Computing the join of two models requires the computation of a supremum graph which we have implemented in our computer program *GraphCheck*. We believe the found structure can be effectively exploited in the study of the corresponding models. One instance of this is the Edwards-Havránek model selection procedure for lattices (Edwards and Havránek, 1987), for which we have developed an algorithm for one of the sets, to be described elsewhere.

Preliminaries

1. Graphs: $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ shall be denoting a coloured graph with vertex set V, edge set E and vertex and edge colourings $\mathcal{V}$ and $\mathcal{E}$ respectively. $\mathcal{G} = (V, E)$ denotes the corresponding uncoloured graph.
2. Partitions: For two partitions $P_1(S)$ and $P_2(S)$ of a set S we say that $P_1(S)$ is *finer* than $P_2(S)$, or equivalently that $P_2(S)$ is *coarser* than $P_1(S)$, if every set in $P_2(S)$ is a union of sets in $P_1(S)$.
3. Lattices: A set L is a *lattice* if for any $a, b \in L$ there is a unique smallest element in L which is larger than both a and b, denoted by $a \vee b$, and a unique largest element in L which is smaller than both, denoted by $a \wedge b$.

Graphical Gaussian Models with Symmetries

Graphical Gaussian models are concerned with the distribution of a multivariate random vector $Y = (Y_\alpha)_{\alpha \in V}$ following a $\mathcal{N}_{|V|}(\mu, \Sigma)$ distribution and are represented by undirected graphs. A graph $\mathcal{G} = (V, E)$ represents the model with concentration matrix $K = \Sigma^{-1}$ lying inside $\mathcal{S}^+(\mathcal{G})$, the set of (symmetric) positive definite matrices which satisfy

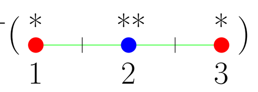
$$\alpha\beta \notin E \implies k_{\alpha\beta} = 0 \quad \text{for} \quad \alpha, \beta \in V.$$

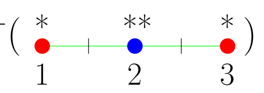
Højsgaard and Lauritzen (2008) introduced three types of graphical Gaussian models with symmetry constraints on their parameters, all represented by vertex and edge coloured graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$:

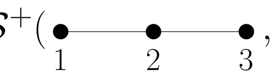
1. **RCON**: Equality between specified elements of the concentration matrix, $K \in \mathcal{S}^+(\mathcal{V}, \mathcal{E})$
2. **RCOR**: Equality between specified partial correlations, $K \in \mathcal{R}^+(\mathcal{V}, \mathcal{E})$
3. **RCOP**: Restrictions that are generated by permutation symmetry, $K \in \mathcal{S}^+(\mathcal{G}, \Gamma)$ with

$$\mathcal{S}^+(\mathcal{G}, \Gamma) = \mathcal{S}^+(\mathcal{G}) \cap \{K : G(\sigma) K G(\sigma)^{-1} = K \ \forall \sigma \in \Gamma\}.$$

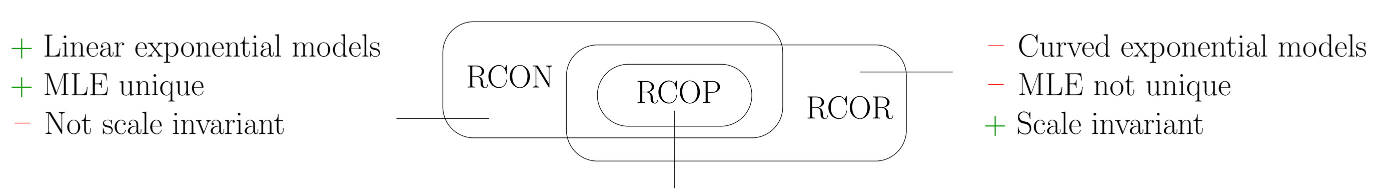
For example,

$\mathcal{R}^+($  $) = \{K \in \mathcal{S}^+(\mathcal{G}) : k_{11} = k_{33}, \rho_{12|3} = \rho_{23|1}\},$

$\mathcal{S}^+($  $) = \{K \in \mathcal{S}^+(\mathcal{G}) : k_{11} = k_{33}, k_{12} = k_{23}\},$

$= \mathcal{S}^+($  $, \{Id, (13)\}).$

They showed:



All '+' from above, and + considerable computational savings in computation of MLE

Graph Colourings

Definition 1 *Let the set of vertex and edge coloured graphs with vertex set V be denoted by $\mathcal{C}_V$. We shall say that the colouring of $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is*

- ***edge regular*** *if in $\mathcal{E}$ every pair of equally coloured edges connects the same vertex colour classes in $\mathcal{V}$, the set of graphs with such colourings on vertex set V being denoted by $\mathcal{B}_V$,*
- ***vertex regular*** *if in all subgraphs $\mathcal{G}^{u,u',u''}$ of $\mathcal{G}$ which are induced by the vertex colour classes $u, u' \in \mathcal{V}$ (allowing $u = u'$) and the edges inside $u'' \in \mathcal{E}$ connecting u to u' the vertex degree is invariant within each vertex colour class, their set being denoted by $\mathcal{P}_V$,*
- ***colour regular*** *if it is both edge regular and vertex regular, $\mathcal{R}_V = \mathcal{B}_V \cap \mathcal{P}_V$, and*
- ***permutation generated*** *by group Γ if $\Gamma \subseteq Aut(\mathcal{G})$, $\mathcal{V}$ is given by the orbits of Γ in V and $\mathcal{E}$ is a union of orbits of Γ in $V \times V$. Equivalently, $(\mathcal{V}, \mathcal{E})$ is permutation generated, by $Aut(\mathcal{V}, \mathcal{E})$, if $Aut(\mathcal{V}, \mathcal{E})$ acts transitively on each colour class in $(\mathcal{V}, \mathcal{E})$, $\Pi_{V,\Gamma} \subset \mathcal{R}_V$, $\bigcup_\Gamma \Pi_{V,\Gamma} = \Pi_V$.* □

Proposition 1 (Højsgaard and Lauritzen) $\mathcal{S}^+(\mathcal{V}, \mathcal{E}) = \mathcal{R}^+(\mathcal{V}, \mathcal{E})$ *if and only if $(\mathcal{V}, \mathcal{E})$ is edge regular, i.e. if and only if $\mathcal{G} \in \mathcal{B}_V$.* □

Proposition 2 (Gehrmann and Lauritzen) *Let $\mathcal{M}$ be a partition of V. Then restricting μ to lie inside*

$$\Omega(\mathcal{M}) = \{(\mu_\alpha)_{\alpha \in V} \in \mathbb{R}^V : \mu_\beta = \mu_\alpha \text{ whenever } \beta \text{ is in the same set as } \alpha \text{ in } \mathcal{M}\}$$

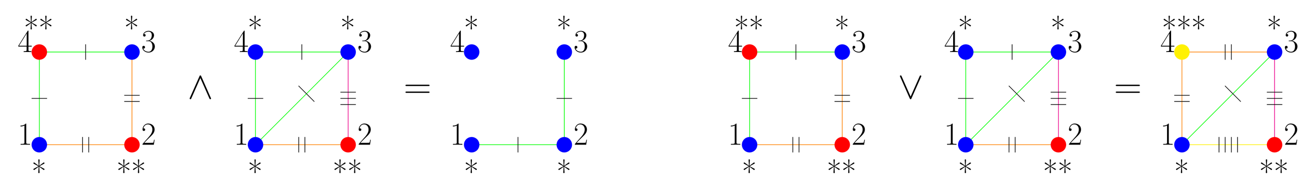
in the RCON and RCOR models determined by $(\mathcal{V}, \mathcal{E})$ gives equality between the maximum likelihood and least squares estimators for μ if and only if $\mathcal{M}$ is finer than or equal to $\mathcal{V}$ and $(\mathcal{M}, \mathcal{E})$ is vertex regular, i.e. if and only if $\mathcal{G}_\mu = (\mathcal{M}, \mathcal{E}) \in \mathcal{P}_V$. Then $\hat{\mu}$ is obtained through appropriate averaging. □

Proposition 3 *$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ represents the RCOP model $\mathcal{S}^+(\mathcal{G}, \Gamma)$ if and only if $(\mathcal{V}, \mathcal{E})$ is permutation generated by Γ, i.e. if only if $\mathcal{G} \in \Pi_{V,\Gamma}$.* □

Topology of $\mathcal{C}_V$, $\mathcal{B}_V$, $\mathcal{P}_V$, $\mathcal{R}_V$ and Π_V

Proposition 4 *The set $\mathcal{C}_V$ is a complete (non-distributive) lattice with respect to model inclusion, with meet and join operations given below.*

$$\mathcal{G}_1 \wedge \mathcal{G}_2 = (\mathcal{V}_1 \vee \mathcal{V}_2, \mathcal{E}_1' \vee \mathcal{E}_2') \qquad \mathcal{G}_1 \vee \mathcal{G}_2 = (\mathcal{V}_1 \wedge \mathcal{V}_2, \mathcal{E}_1'' \wedge \mathcal{E}_2'')$$

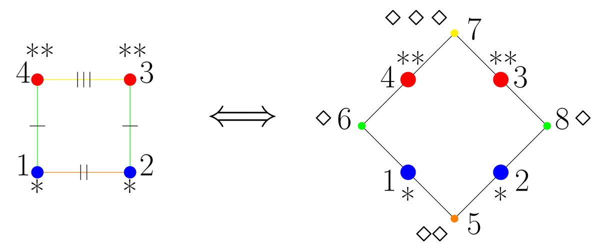


Fewer edges, coarser colouring ⇒ smaller model — More edges, finer colouring ⇒ larger model □

Proposition 5 *The sets $\mathcal{B}_V$, $\mathcal{P}_V$, $\mathcal{R}_V$ and Π_V are (generally non-distributive) lattices, with meet operation as in $\mathcal{C}_V$ and join operations given by*

$$\mathcal{G}_1 \vee_S \mathcal{G}_2 = \sup_S(\mathcal{G}_1 \vee \mathcal{G}_2) \quad \text{for} \quad S \in \{\mathcal{B}_V, \mathcal{P}_V, \mathcal{R}_V, \Pi_V\}.$$ □

1. $(\mathcal{V}, \mathcal{E})$ colour regular
2. $(\mathcal{V}, \mathcal{E})$ invariant under σ



1. $(\mathcal{V} \cup \mathcal{N}, E_F)$ *equitable* wrt $\phi(\mathcal{G})$
2. $(\mathcal{V} \cup \mathcal{N}, E_F)$ invariant under σ_ϕ

Figure 1: Relationships between $\mathcal{G}$ and corresponding *coloured factor graph* $\phi(\mathcal{G})$

Fact I: For every vertex coloured graph $\mathcal{G} = (\mathcal{V}, E)$ there exists a coarsest refinement $r(\mathcal{V})$ of $\mathcal{V}$ which is equitable with respect to $\mathcal{G}$ (McKay, 1981).

Fact II: For every vertex coloured graph $\mathcal{G} = (\mathcal{V}, E)$ there exists a unique coarsest refinement of $\mathcal{V}$ which is invariant under $Aut(\mathcal{V}, E)$, given by the orbits of $Aut(\mathcal{V}, E)$ in V.

Sketch proof of Proposition 5: For $\mathcal{B}_V$, $\mathcal{P}_V$ and $\mathcal{R}_V$,

1. All three sets are stable under $\wedge$ in $\mathcal{C}_V$.
2. The same does not hold for $\vee$, however 1. in Figure 1 together with Fact I for $\mathcal{G} \in \mathcal{C}_V$ give

$$\sup_{\mathcal{R}}(\mathcal{G}) = \phi^{-1}((r(\mathcal{V} \cup \mathcal{N}), E_F)) = (\mathcal{V}_{\sup}, \mathcal{E}_{\sup}), \qquad \sup_{\mathcal{B}} = (\mathcal{V}, \mathcal{E}_{\sup}), \qquad \sup_{\mathcal{P}}(\mathcal{G}) = (\mathcal{V}_{\sup}, \mathcal{E}).$$

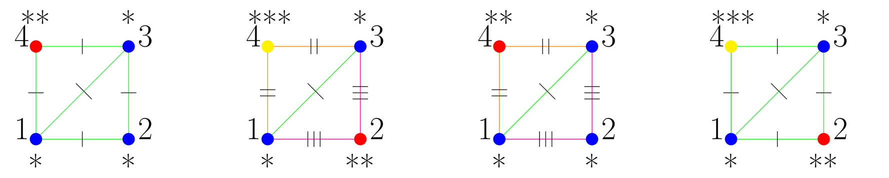
For Π_V,

1. It is a Group Theoretic fact (e.g. Chapter 1 in Schmidt (1994)) that the set of subgroups of $S_{|V|}$ forms a complete lattice with $\Gamma_1 \wedge \Gamma_2 = \Gamma_1 \cap \Gamma_2$ and $\Gamma_1 \vee \Gamma_2 = \{\sigma_1\sigma_2 : \sigma_1 \in \Gamma_1, \sigma_2 \in \Gamma_2\}$.
2. Π_V is stable under $\wedge$, with $\mathcal{G}_1 \wedge \mathcal{G}_2$ being generated by $\Gamma_1 \vee \Gamma_2$.
3. The same is not true for $\vee$, however 2. in Figure 1 together with Fact II for $\mathcal{G} \in \mathcal{C}_V$ give

$$\sup_{\Pi}(\mathcal{G}) = \phi^{-1}((\mathcal{V}_\Pi \cup \mathcal{N}_\Pi, E_F))$$

where $\mathcal{V}_\Pi \cup \mathcal{N}_\Pi$ denotes the coarsest refinement of $\mathcal{V} \cup \mathcal{N}$ satisfying the condition in Fact II. ■

The following examples were obtained with our program *GraphCheck* which is based on Brendan McKay's computer program `nauty` (`http://cs.anu.edu.au/~bdm/nauty/`). On a 32-bit processor `nauty` can handle graphs with up to 2^{30} vertices, giving a bound of $|V| + |E| < 2^{30}$ for *GraphCheck*.



$\mathcal{G}$ — $\sup_{\mathcal{R}}(\mathcal{G}) = \sup_{\Pi}(\mathcal{G})$ — $\sup_{\mathcal{B}}(\mathcal{G})$ — $\sup_{\mathcal{P}}(\mathcal{G})$

The lattice structure of the four sets qualifies them for the Edwards-Havránek model selection procedure (Edwards and Havránek, 1987), which is based on the principle that once a model is rejected (accepted) all of its submodels (supermodels) are rejected (accepted). We have developed the corresponding algorithm for $\mathcal{B}_V$, to be described elsewhere.

Summary

1. We have shown that the four sets $\mathcal{B}_V$, $\mathcal{P}_V$, $\mathcal{R}_V$ and Π_V form lattices. In all four sets,
   (i) the meet operation agrees with the standard meet by model inclusion,
   (ii) the join requires finding the supremum of the model inclusion join inside the given set. Its computation is implemented in our program *GraphCheck*.
2. The lattice structure of the four sets qualifies them for the Edwards-Havránek model selection procedure (Edwards and Havránek, 1987), for which we have developed an algorithm for $\mathcal{B}_V$, to be described elsewhere.

We believe the lattice structure to have further statistical consequences and shall be considering conjugate priors and likelihood ratio tests for the four considered sets, with particular interest in Π_V.

Acknowledgements

I would like to tank my supervisor Professor Steffen Lauritzen for his advice and guidance on the subject.

References

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

Edwards, D. and Havránek, T. (1987). A Fast Model Selection Procedure for Large Families of Models. *Journal of the American Statistical Association*, 82, No. 397:205–213.

Højsgaard, S. and Lauritzen, S. (2008). Graphical Gaussian models with edge and vertex symmetries. *Journal of Royal Statistical Society, Series B*, 70, Part 5:1005–1027.

Hylleberg, B., Jensen, M., and Ørnbøl, E. (1993). Graphical symmetry models. Master's thesis, Aalborg University.

McKay, B. (1981). Practical Graph Isomorphism. *Congressus Numerantium*, 30:45–87.

Schmidt, R. (1994). *Subgroup Lattices of Groups*. Walter de Gruyter & Co.

Votaw, D. F. (1948). Testing compound symmetry in a normal multivariate distribution. *Ann. Math. Statist.*, 19:447473.