Estimation of Means in Graphical Gaussian Models with Symmetries

Helene Gehrmann

Department of Statistics, University of Oxford

gehrmann@stats.ox.ac.uk

The Second CREST-SBM International Conference
"Harmony of Gröbner bases and the modern industrial society"
Osaka, Japan

29 June 2010

Graphical Gaussian Models

If $\mathcal{G} = (V, E)$ is an undirected graph and $Y = (Y_\alpha)_{\alpha \in V}$ is a random variable taking values in $\mathbb{R}^{|V|}$, the *graphical Gaussian model* for Y with graph $\mathcal{G}$ is given by assuming that Y follows a Gaussian distribution which obeys the (global) Markov property with respect to $\mathcal{G}$.

(Global) Markov Property: For $A, B, S \subset V$,

$$A \perp_{\mathcal{G}} B \mid S \Rightarrow Y_A \perp\!\!\!\perp Y_B \mid Y_S$$

where $\perp_{\mathcal{G}}$ denotes graph separation.

Graphical Gaussian Models

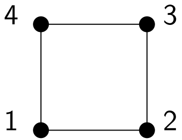
If $\mathcal{G} = (V, E)$ is an undirected graph and $Y = (Y_\alpha)_{\alpha \in V}$ is a random variable taking values in $\mathbb{R}^{|V|}$, the *graphical Gaussian model* for Y with graph $\mathcal{G}$ is given by assuming that Y follows a Gaussian distribution which obeys the (global) Markov property with respect to $\mathcal{G}$.

(Global) Markov Property: For $A, B, S \subset V$,

$$A \perp_{\mathcal{G}} B \mid S \Rightarrow Y_A \perp\!\!\!\perp Y_B \mid Y_S$$

where $\perp_{\mathcal{G}}$ denotes graph separation.

E.g.



$$(Y_1, Y_2, Y_3, Y_4) \sim N(\mu, \Sigma)$$

$$Y_1 \perp\!\!\!\perp Y_3 | (Y_2, Y_4)$$
$$Y_2 \perp\!\!\!\perp Y_4 | (Y_1, Y_3)$$

Graphical Gaussian Models

If $(Y_\alpha)_{\alpha \in V} \sim \mathcal{N}(\mu, \Sigma)$ and concentration matrix $K = \Sigma^{-1} = (k_{\alpha\beta})_{\alpha,\beta \in V}$,

$$Y_\alpha \perp\!\!\!\perp Y_\beta | (Y_{V \setminus \{\alpha,\beta\}}) \quad \iff \quad k_{\alpha\beta} = 0$$

Graphical Gaussian model satisfies Markov Property $\iff$ K satisfies

$$\alpha \not\sim \beta \text{ in } \mathcal{G} \implies k_{\alpha\beta} = 0$$

where $\sim$ stands for 'connected by an edge'.

Graphical Gaussian Models

If $(Y_\alpha)_{\alpha \in V} \sim \mathcal{N}(\mu, \Sigma)$ and concentration matrix $K = \Sigma^{-1} = (k_{\alpha\beta})_{\alpha,\beta \in V}$,

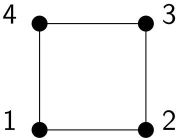
$$Y_\alpha \perp\!\!\!\perp Y_\beta | (Y_{V \setminus \{\alpha,\beta\}}) \quad \Longleftrightarrow \quad k_{\alpha\beta} = 0$$

Graphical Gaussian model satisfies Markov Property $\quad \Longleftrightarrow \quad$ K satisfies

$$\alpha \not\sim \beta \text{ in } \mathcal{G} \Longrightarrow k_{\alpha\beta} = 0$$

where $\sim$ stands for 'connected by an edge'.

E.g.

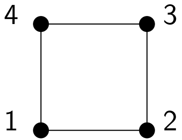


$$K = \begin{pmatrix} k_{11} & k_{12} & 0 & k_{14} \\ k_{21} & k_{22} & k_{23} & 0 \\ 0 & k_{23} & k_{33} & k_{34} \\ k_{14} & 0 & k_{34} & k_{44} \end{pmatrix}$$

Graphical Gaussian Models with Symmetries

Højsgaard and Lauritzen (2008) introduced models with symmetry restrictions, represented by vertex and edge coloured graphs $(\mathcal{V}, \mathcal{E})$:

RCON models: Symmetry restrictions on concentrations

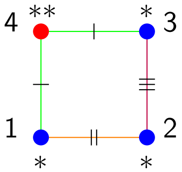


$$K = \begin{pmatrix} k_{11} & k_{12} & 0 & k_{14} \\ k_{21} & k_{22} & k_{23} & 0 \\ 0 & k_{23} & k_{33} & k_{34} \\ k_{14} & 0 & k_{34} & k_{44} \end{pmatrix}$$

Graphical Gaussian Models with Symmetries

Højsgaard and Lauritzen (2008) introduced models with symmetry restrictions, represented by vertex and edge coloured graphs $(\mathcal{V}, \mathcal{E})$:

RCON models: Symmetry restrictions on concentrations



$$K = \begin{pmatrix} a & d & 0 & c \\ d & a & e & 0 \\ 0 & e & a & c \\ c & 0 & c & b \end{pmatrix}$$

$$\mathcal{V} = \{\{1, 2, 3\}, \{4\}\}$$

$$\mathcal{E} = \{\{14, 34\}, \{12\}, \{23\}\}$$

Constraints on the Mean Vector

Højsgaard and Lauritzen (2008) assume: $(Y_\alpha)_{\alpha \in V} \sim N(\mu, \Sigma)$ with $\mu = 0$!

For a given RCON model, we are going to characterize all nice linear constraints on μ which ensure equality between maximum likelihood estimator of μ,

$$\hat{\mu} = \max_{\mu} L(\mu, K; y)$$

and least squares estimators of μ,

$$\mu^* = \min_{\mu} \sum_{\alpha \in V} (Y_\alpha - \mu_\alpha)^2$$

which guarantees that $\hat{\mu}$ exists (note the likelihood depends on unknown K) and is given by appropriate averages.

nice = all restrictions satisfied by zero vector

Chan and Godsil (1989) applied to graphical Gaussian models characterises all valid equality constraints, we are going to give a generalisation.

Constraints on the Mean Vector

Theorem (Kruskal): For $(Y_\alpha)_{\alpha \in V} \sim N(\mu, \Sigma)$ with mean μ lying inside a linear manifold Ω, $\hat{\mu} = \mu^*$ if and only if Ω is invariant under $K = \Sigma^{-1}$, i.e. if and only if

$$K\Omega \subset \Omega.$$

(Kruskal, 1968; Haberman, 1975; Eaton, 1983)

Constraints on the Mean Vector

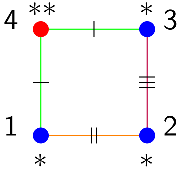
Theorem (Kruskal): For $(Y_\alpha)_{\alpha \in V} \sim N(\mu, \Sigma)$ with mean μ lying inside a linear manifold Ω, $\hat{\mu} = \mu^*$ if and only if Ω is invariant under $K = \Sigma^{-1}$, i.e. if and only if

$$K\Omega \subset \Omega.$$

(Kruskal, 1968; Haberman, 1975; Eaton, 1983)

For RCON models,

$$K = \sum_{u \in \mathcal{V} \cup \mathcal{E}} \theta_u T^u.$$



$$T^{\{1,2,3\}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Constraints on the Mean Vector

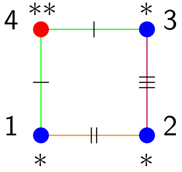
Theorem (Kruskal): For $(Y_\alpha)_{\alpha \in V} \sim N(\mu, \Sigma)$ with mean μ lying inside a linear manifold Ω, $\hat{\mu} = \mu^*$ if and only if Ω is invariant under $K = \Sigma^{-1}$, i.e. if and only if

$$K\Omega \subset \Omega.$$

(Kruskal, 1968; Haberman, 1975; Eaton, 1983)

For RCON models,

$$K = \sum_{u \in \mathcal{V} \cup \mathcal{E}} \theta_u T^u.$$



$$T^{\{14,34\}} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Stability under Generator Matrices T^u

Proposition 1 (G.): Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the dependence graph of an RCON model with linear mean space Ω. Then

$$K\Omega \subseteq \Omega \ \ \forall K \text{ inside the model} \quad \Longleftrightarrow \quad T^u\Omega \subseteq \Omega \ \ \forall u \in \mathcal{V} \cup \mathcal{E}.$$

Stability under Generator Matrices T^u

Proposition 1 (G.): Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the dependence graph of an RCON model with linear mean space Ω. Then

$$K\Omega \subseteq \Omega \quad \forall K \text{ inside the model} \quad \Longleftrightarrow \quad T^u\Omega \subseteq \Omega \quad \forall u \in \mathcal{V} \cup \mathcal{E}.$$

$u \in \mathcal{V}$: $T^u\Omega \subseteq \Omega \quad \Longleftrightarrow \quad \Omega = \oplus_{v\in\mathcal{V}} \Omega_v, \quad \Omega_v \leq \mathbb{R}^v$

$$T^{\{1,2,3\}}\mu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ 0 \end{pmatrix}$$

Stability under Generator Matrices T^u

Proposition 1 (G.): Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the dependence graph of an RCON model with linear mean space Ω. Then

$$K\Omega \subseteq \Omega \;\; \forall K \text{ inside the model} \quad \Longleftrightarrow \quad T^u\Omega \subseteq \Omega \;\; \forall u \in \mathcal{V} \cup \mathcal{E}.$$

$u \in \mathcal{V}$: $T^u\Omega \subseteq \Omega \quad \Longleftrightarrow \quad \Omega = \bigoplus_{v\in\mathcal{V}} \Omega_v, \quad \Omega_v \leq \mathbb{R}^v$

$u \in \mathcal{E}$: For 'nice' Ω, i.e. $\Omega_v = 0$ allowed, we only need to consider the (u, v, w)-components of $\mathcal{G}$, represented by $T^{[u,v,w]} \in \mathbb{R}^{v \cup w}$.

$$T^{[u,v,w]}_{\alpha\beta} = \begin{cases} T^u_{\alpha\beta} & \alpha \in v, \beta \in w \text{ or } \alpha \in v, \beta \in w \\ 0 & \text{otherwise} \end{cases}$$

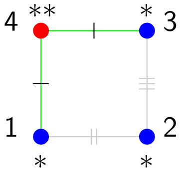
Stability under Generator Matrices T^u

Proposition 1 (G.): Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the dependence graph of an RCON model with linear mean space Ω. Then

$$K\Omega \subseteq \Omega \quad \forall K \text{ inside the model} \quad \Longleftrightarrow \quad T^u\Omega \subseteq \Omega \quad \forall u \in \mathcal{V} \cup \mathcal{E}.$$

$u \in \mathcal{V}$: $T^u\Omega \subseteq \Omega \quad \Longleftrightarrow \quad \Omega = \oplus_{v\in\mathcal{V}} \Omega_v, \quad \Omega_v \leq \mathbb{R}^v$

$u \in \mathcal{E}$: For 'nice' Ω, i.e. $\Omega_v = 0$ allowed, we only need to consider the (u, v, w)-components of $\mathcal{G}$, represented by $T^{[u,v,w]} \in \mathbb{R}^{v \cup w}$.



$$T^{[\{13,14\},\ \{1,2,3\},\ \{4\}]} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

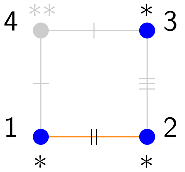
Stability under Generator Matrices T^u

Proposition 1 (G.): Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the dependence graph of an RCON model with linear mean space Ω. Then

$$K\Omega \subseteq \Omega \quad \forall K \text{ inside the model} \quad \Longleftrightarrow \quad T^u\Omega \subseteq \Omega \quad \forall u \in \mathcal{V} \cup \mathcal{E}.$$

$u \in \mathcal{V}$: $T^u\Omega \subseteq \Omega \quad \Longleftrightarrow \quad \Omega = \oplus_{v\in\mathcal{V}} \Omega_v, \quad \Omega_v \leq \mathbb{R}^v$

$u \in \mathcal{E}$: For 'nice' Ω, i.e. $\Omega_v = 0$ allowed, we only need to consider the (u, v, w)-components of $\mathcal{G}$, represented by $T^{[u,v,w]} \in \mathbb{R}^{v\cup w}$.



$$T^{[\{12\},\ \{1,2,3\},\ \{1,2,3\}]} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Stability under Generator Matrices T^u

Proposition 2 (G.): For 'nice' Ω,

$$T^u\Omega \subseteq \Omega \quad \Longleftrightarrow \quad T^{[u,v,w]}(\Omega_v \oplus \Omega_w) \subseteq (\Omega_v \oplus \Omega_w)$$

for all $u \in \mathcal{E}$, $v, w \in \mathcal{V}$.

Stability under Generator Matrices T^u

Proposition 2 (G.): For 'nice' Ω,

$$T^u\Omega \subseteq \Omega \quad \Longleftrightarrow \quad T^{[u,v,w]}(\Omega_v \oplus \Omega_w) \subseteq (\Omega_v \oplus \Omega_w)$$

for all $u \in \mathcal{E}$, $v, w \in \mathcal{V}$.

In general, if A is a symmetric matrix then a space S is stable under A if and only if S is a direct sum of subspaces of the eigenspaces E^A_λ of A, which are in fact orthogonal, i.e.

$$AS \subseteq S \quad \Longleftrightarrow \quad S = \oplus_\lambda R_\lambda, \quad R_\lambda \leq E^A_\lambda.$$

Thus we require

$$\begin{aligned} \Omega_v, \Omega_w &\leq \mathbb{R}^v, \mathbb{R}^w \\ \Omega_v \oplus \Omega_w &= \oplus_\lambda A_\lambda, \quad A_\lambda \leq E_\lambda^{[u,v,w]} \quad \text{for all } u \in \mathcal{E}. \end{aligned}$$

Stability under Component Generator Matrices $T^{[u,v,w]}$

Fact (e.g. West, 1999): A graph G is bipartite if and only if the eigenvalues of its adjacency matrix A come in pairs: whenever λ is an eigenvalue, so is $-\lambda$.

Stability under Component Generator Matrices $T^{[u,v,w]}$

Fact (e.g. West, 1999): A graph G is bipartite if and only if the eigenvalues of its adjacency matrix A come in pairs: whenever λ is an eigenvalue, so is $-\lambda$.

Proposition 3 (G.): $\Omega_v, \Omega_w \leq \mathbb{R}^v, \mathbb{R}^w$ are stable under $T^{[u,v,w]}$ if and only if

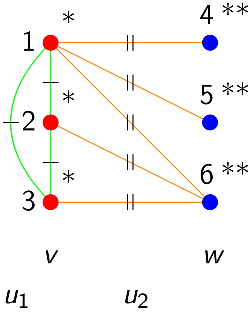
$$\Omega_v = \oplus_{\lambda \geq 0} (A_\lambda)_v \quad \text{and} \quad \Omega_w = \oplus_{\lambda \geq 0} (A_\lambda)_w$$

with $A_\lambda \leq E_\lambda^{[u,v,w]}$.

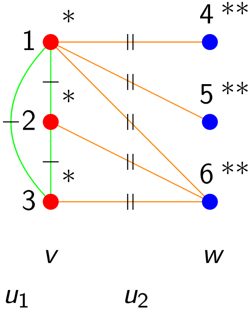
For $\lambda \neq 0$, $T^{[u,v,w]}(\Omega_v) \subseteq \Omega_w$ and vice versa.

For $\lambda = 0$, $T^{[u,v,w]}(\Omega_v) = 0 \in \Omega_w$ and vice versa.

Example



Example



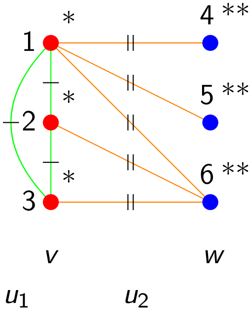
$$(A_\lambda)_v \leq \left\langle \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\rangle, \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

$$(A_\lambda)_v \leq \left\langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\rangle, \delta_1 \left\langle \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\rangle, \delta_2 \left\langle \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

$$(A_\lambda)_w \leq \left\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\rangle, \delta_1 \left\langle \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\rangle, \delta_2 \left\langle \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\rangle$$

Example

(i) $\Omega_v = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle$



(i) $\Omega_w = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle$

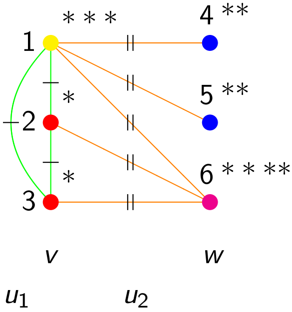
(i) $\mu_1 = 0, \mu_2 = -\mu_3$

(i) $\mu_4 = -\mu_5, \mu_6 = 0$

Example

(i) $\Omega_v = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle$

(ii) $\Omega_v = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle^{\perp}$



(i) $\Omega_w = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle$

(ii) $\Omega_w = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle^{\perp}$

(i) $\mu_1 = 0, \mu_2 = -\mu_3$

(ii) $\mu_1 \in \mathbb{R}, \mu_2 = \mu_3$

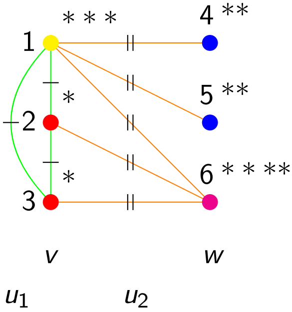
(i) $\mu_4 = -\mu_5, \mu_6 = 0$

(ii) $\mu_4 = \mu_5, \mu_6 \in \mathbb{R}$

Example

(i) $\Omega_v = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle$

(ii) $\Omega_v = \langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \rangle^\perp$



(i) $\Omega_w = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle$

(ii) $\Omega_w = \langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \rangle^\perp$

(i) $\mu_1 = 0, \mu_2 = -\mu_3$

(ii) $\mu_1 \in \mathbb{R}, \mu_2 = \mu_3$

(i) $\mu_4 = -\mu_5, \mu_6 = 0$

(ii) $\mu_4 = \mu_5, \mu_6 \in \mathbb{R}$

Particular application for equality constraints: design of experiments with non-trivial concentration structure.

Thank You!

References

Chan, A. and Godsil, C. (1989). Symmetry and eigenvectors. In Hahn, G. and Sabidussi, G., editors, *Graph Symmetry. Algebraic mehods and applications*, volume 497 of *C: Mathematical and Physical Sciences*, pages 75 – 106. NATO Scientific Affairs Division, Kluwer Academic Publishers.

Eaton, M. (1983). *Multivariate Statistics: A Vector Space Approach*. John Wiley & Sons Inc.

Haberman, S. (1975). How much do Gauss-Markov and least-square estimates differ? A coordinate-free approach. *The Annals of Statistics*, 3 No.4:982 – 990.

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.

Kruskal, W. (1968). When are Gauss-Markov and least squares estimators identical? A coordinate-free approach. *The Ann. Math. Statist.*, 39:70 – 75.

West, B. D. (1999). *Introduction to Graph Theory*. Prentice Hall.