Variance function of Wishart exponential families in graphical models.

Salha Mamane
University of Angers

with
Piotr Graczyk, University of Angers
Hideyuki Ishi, Nagoya University

Summer School on Mathematical Methods of Statistics
Angers, 21 June 2016

Motivation

- Graphical model: a statistical model + graph representation. Variable = Vertex. No edge = conditional independence.
- Example: $A_n : 1 - 2 - 3 - \ldots - n$.
  For all $|i - j| > 1$, $X_i \perp X_j | (X_k)_{k \neq i,j}$.
- In the case of a Gaussian graphical mode with covariance matrix Σ and concentration matrix $K = \Sigma^{-1}$, the conditional independence constraints are equivalent to $K_{ij} = 0$ for all non adjacent i and j.

Motivation

- Graphical model: a statistical model + graph representation. Variable = Vertex. No edge = conditional independence.
- Example: $A_n : 1 - 2 - 3 - \ldots - n$.
  For all $|i - j| > 1$, $X_i \perp X_j | (X_k)_{k \neq i,j}$.
- In the case of a Gaussian graphical mode with covariance matrix Σ and concentration matrix $K = \Sigma^{-1}$, the conditional independence constraints are equivalent to $K_{ij} = 0$ for all non adjacent i and j.

Motivation

- Graphical model: a statistical model + graph representation. Variable = Vertex. No edge = conditional independence.
- Example: $A_n : 1 - 2 - 3 - \ldots - n$.
  For all $|i-j| > 1$, $X_i \perp X_j | (X_k)_{k \neq i,j}$.
- In the case of a Gaussian graphical mode with covariance matrix Σ and concentration matrix $K = \Sigma^{-1}$, the conditional independence constraints are equivalent to $K_{ij} = 0$ for all non adjacent i and j.

- This implies that the elements Σ_{ij} are not free parameters. The model can be parametrised by an incomplete symmetric matrix $\pi(\Sigma)$ with entries corresponding to non-adjacent vertices left out and such that the submatrices corresponding to complete subsets of vertices are positive definite. This set of incomplete matrices is denoted Q_G.

Example

For the graph $1 - 2 - 3 - 4$:
The elements of Q_{A_4} are incomplete matrices of the form

$$x = \begin{pmatrix} x_{11} & x_{12} & * & * \\ x_{12} & x_{22} & x_{23} & * \\ * & x_{23} & x_{33} & x_{34} \\ * & * & x_{34} & x_{44} \end{pmatrix}$$

such that $\begin{vmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{vmatrix}$, $\begin{vmatrix} x_{22} & x_{23} \\ x_{23} & x_{33} \end{vmatrix}$ and $\begin{vmatrix} x_{33} & x_{34} \\ x_{34} & x_{44} \end{vmatrix}$ are positive definite.

Objective

Derive the variance function the exponential family of Wishart distributions constructed on the cone Q_{A_n},

- The mean function is given by

$$m_{\underline{s}}^{(M)}(y) = \pi\left(\sum_{i=1}^{M-1}(s_i - s_{i+1})[(y_{\{1:i\}})^{-1}]^0 + s_M y^{-1} + \sum_{i=M+1}^{n}(s_{i+1} - s_i)\,[(y_{\{i:n\}})^{-1}]^0\right)$$

- The inverse mean map is given by

$$\begin{aligned}\psi_{\underline{s}}^{(M)}(m) \quad &= \quad \sum_{k=1}^{M-1} s_k(m^{-1}_{\{k:k+1\}})^0 + \sum_{k=M+1}^{n} s_k(m^{-1}_{\{k-1:k\}})^0 \qquad (1)\\ &- \sum_{k=2}^{M-1} s_{k-1}(m^{-1}_{\{k,k\}})^0 - (s_{M-1} - s_M + s_{M+1})(m^{-1}_{\{M,M\}})^0\\ &- \sum_{k=M+1}^{n-1} s_{k+1}(m^{-1}_{\{k,k\}})^0.\end{aligned}$$

- The mean function is given by

$$m_{\underline{s}}^{(M)}(y) = \pi\left(\sum_{i=1}^{M-1}(s_i - s_{i+1})[(y_{\{1:i\}})^{-1}]^0 + s_M y^{-1} + \sum_{i=M+1}^{n}(s_{i+1} - s_i)\,[(y_{\{i:n\}})^{-1}]^0\right)$$

- The inverse mean map is given by

$$\begin{aligned}\psi_{\underline{s}}^{(M)}(m) \quad = \quad & \sum_{k=1}^{M-1} s_k (m^{-1}_{\{k:k+1\}})^0 + \sum_{k=M+1}^{n} s_k (m^{-1}_{\{k-1:k\}})^0 \\ & - \sum_{k=2}^{M-1} s_{k-1}(m^{-1}_{\{k,k\}})^0 - (s_{M-1} - s_M + s_{M+1})(m^{-1}_{\{M,M\}})^0 \\ & - \sum_{k=M+1}^{n-1} s_{k+1}(m^{-1}_{\{k,k\}})^0. \end{aligned} \quad (1)$$

Some notations

Let $m \in Q_G$; $\hat{m} \in S_n^+$ is the unique symmetric positive definite matrix verifying $\pi(\hat{m}) = m, \quad \hat{m}^{-1} \in P_G$.
For all $m \in Q_G$, let $M_A = [(\hat{m}_A^{-1})^{-1}]^0$.
For all $A \in S_d$, let $\rho(A)u = AuA$ and write $\mathbb{P} = \pi \circ \rho$.

Theorem 1

The variance function of the exponential family of Wishart distributions on Q_{A_n} is given by

$$
\begin{aligned}
V(m) = & \sum_{i=1}^{M-1} (s_i - s_{i+1}) \, \mathbb{P} \left(\sum_{j=1}^{i-1} \left(\frac{1}{s_j} - \frac{1}{s_{j+1}} \right) M_{1:j} + \frac{1}{s_i} M_{1:i} \right) \\
& + s_M \, \mathbb{P} \left(\sum_{j=1}^{n-1} \left(\frac{1}{s_j} - \frac{1}{s_{j+1}} \right) M_{1:j} + \frac{1}{s_n} \hat{m} \right) \\
& + \sum_{i=M+1}^{n} (s_i - s_{i-1}) \, \mathbb{P} \left(\frac{1}{s_i} M_{i:n} + \sum_{j=i+1}^{n} \left(\frac{1}{s_j} - \frac{1}{s_{j-1}} \right) M_{j:n} \right) .
\end{aligned} \tag{2}
$$

Theorem 2

The variance function of the Wishart exponential family $\gamma_{\underline{s},y}^{(M)}$ *is*

$$V(m) = \left[\frac{1}{s_1} + \frac{1}{s_n} - \frac{1}{s_M}\right]\mathbb{P}(\hat{m}) + \sum_{k=1}^{M-1}\left[\frac{1}{s_{k+1}} - \frac{1}{s_k}\right]\mathbb{P}(\hat{m} - M_{1:k}) + \sum_{k=M+1}^{n}\left[\frac{1}{s_{k-1}} - \frac{1}{s_k}\right]\mathbb{P}(\hat{m} - M_{k:n}).$$

IDEA OF PROOFS

The proof of the first result is based on writing $V(m) = v(\psi_{\underline{s}}^{(M)}(m))$, where

$$v(y) = \sum_{i=1}^{M-1} (s_i - s_{i+1})\, \mathbb{P}\left[\left((y_{\{1:i\}})^{-1}\right)^0\right] + s_M\, \mathbb{P}(y^{-1}) + \sum_{i=M+1}^{n} (s_i - s_{i-1})\, \mathbb{P}\left[\left((y_{\{i:n\}})^{-1}\right)^0\right].$$

Then we use a Choleski decomposition of $y = \psi_{\underline{s}}^{(M)}(m) = TT^T$, where the “triangular” matrix T has the same pattern of zeros as y.

The second result is proved by induction.