Graphic Lasso: Scaled membership with intercept

Span(U) orthogonal Span (intercept)

4' for each l: sum_{k} mu_{kl} = 0

Jiaxin Hu

April 5, 2021

K=5, r=1
k=1,…, 4: Theta_0+ mu1 Theta
k=5: Theta_0 + mu2 Theta
mu_1 , mu_2 >0

==> Theta0' = Theta0+mu1Theta
k=1,… 4: Theta0'+0*0
k=5,: Theta0' + 1*differnce

1 Identifiability

Consider the model

$$\Omega^k = \Theta^0 + \sum_{l=1}^{r} u_{kl}\Theta^l, \quad k \in [K]. \tag{1}$$

[intercept, U] [Theta_0, Theta]
insufficient:
rank([intercept, U]) = r+1

Let $U = [\![u_{kl}]\!] \in \mathbb{R}^{K\times r}$ be the membership matrix and u_l denote the l-th column of U. Let $I_l = \{k : u_{kl} \neq 0\}$.

Lemma 1 (Identifiability of scaled membership model with intercept)**.** *Suppose the parameter (U, Θ^l) satisfies the following condition.*

1. *$\Theta^0, \Theta^1, ..., \Theta^l$ are positive definite with bounded singular values, i.e., $0 < \tau_1 \leq \min_{l=0,1,...r} \varphi_{\min}(\Theta^l) \leq \max_{l=0,1,...r} \varphi_{\max}(\Theta^l) \leq \tau_2 < \infty$.*
2. *$\Theta^l, l = 0, 1, ..., r$ are irreducible in the sense that $\Theta^l \neq C\Theta^{l'}$ for any pair l, l' and for any constant C.*
3. *The columns of U are non-overlap, with $\|u_l\|_F = 1$.*
4. *For all $l \in [r]$, there exists at least one $u_{kl}, k \in I_l$ has different values with other non-zero entries, i.e., there exists a pair $k, k' \in I_l$ such that $u_{k'l} \neq u_{kl}$.* span(U) != Span (intercept)

Then, the parameters in model (1) are identifiable.

[intercept, U] ~ [Intercept, U+span(intercept)]
[Theta0, Theta] ~ S^{-1}[Theta0, Theta]

Proof. We prove the identifiability by cases. Suppose there exists another set of parameter $(\tilde U, \tilde\Theta^l)$ such that

$$\Omega^k = \tilde\Theta^0 + \sum_{l=1}^{r} \tilde u_{kl}\tilde\Theta^l,$$

with corresponding $\tilde I_l$.

1. Assume there exists a $l \in [r]$ such that $I_l \neq \tilde I_l$. Without the loss of generality, we assume $|I_l| \geq |\tilde I_l|$. Then, there exist a pair $k, k' \in I_l$ while $k \in \tilde I_l, k' \in \tilde I_{l'}$. Then, we have

$$\Theta^0 + u_{kl}\Theta^l = \tilde\Theta^0 + \tilde u_{kl}\tilde\Theta^l$$
$$\Theta^0 + u_{k'l}\Theta^l = \tilde\Theta^0 + \tilde u_{k'l'}\tilde\Theta^{l'},$$

which implies that

$$\Theta^0 - \tilde\Theta^0 = -u_{kl}\Theta^l + \tilde u_{kl}\tilde\Theta^l$$
$$= -\frac{u_{k'l}}{u_{kl}}\left[u_{kl}\Theta^l + \frac{\tilde u_{k'l'}u_{kl}}{u_{k'l}}\tilde\Theta^{l'}\right] \tag{2}$$

if \tilde Theta^{l'} = 0 and \tilde \Theta^l = c \Theta^l, the equation (2) holds and does not contradict to the assumptions.

The equation (2) is valid since $u_{kl} \neq 0$, for all $k \in I_l$. Note that by condition 2 $\tilde{\Theta}^l \neq C\tilde{\Theta}^{l'}$ for any constant C. Thus, there exists an index $i \in [p]$ such that the i-th column of $\tilde{\Theta}^l, \tilde{\Theta}^{l'}$, denoted $\tilde{v}_i, \tilde{v}_i'$, are linearly independent. Let v_i denote the i-th column of Θ^l. Then we the following two vectors are also linearly independent.

$$-u_{kl}v_i + \tilde{u}_{kl}\tilde{v}_i, \quad \text{and} \quad u_{kl}v_i + \frac{\tilde{u}_{k'l'}u_{kl}}{u_{k'l}}\tilde{v}_i',$$

which contradicts to the equation (2). Hence, for all $l \in [r]$, we have $I_l = \tilde{I}_l$.

2. Assume for all $l \in [r]$, we have $I_l = \tilde{I}_l$.

   (a) Assume there exists a $l \in [r]$ such that $\Theta^l \neq C\tilde{\Theta}^l$ for any constant C. Then, there exists an index $i \in [p]$ such that the columns $v_i, \tilde{v}_i$ of $\Theta^l, \tilde{\Theta}^l$ respectively are linearly independent.

   Also, by condition 4, there exist a pair $k, k' \in I_l$ such that $u_{kl} - u_{k'l} \neq 0$. Then, we have

$$\begin{aligned} \Theta^0 + u_{kl}\Theta^l &= \tilde{\Theta}^0 + \tilde{u}_{kl}\tilde{\Theta}^l \\ \Theta^0 + u_{k'l}\Theta^l &= \tilde{\Theta}^0 + \tilde{u}_{k'l}\tilde{\Theta}^l, \end{aligned}$$

   which implies that

$$\Theta^0 - \tilde{\Theta}^0 = -u_{kl}\Theta^l + \tilde{u}_{kl}\tilde{\Theta}^l = -u_{k'l}\Theta^l + \tilde{u}_{k'l}\tilde{\Theta}^l,$$

   and

$$[u_{k'l} - u_{kl}]v_i = [\tilde{u}_{k'l} - \tilde{u}_{kl}]\tilde{v}_i. \tag{3}$$

   Since $u_{k'l} - u_{kl} \neq 0$, the equation (3) contradicts to the assumption that v_i and $\tilde{v}_i$ are linear independent and $\Theta^l \neq C\tilde{\Theta}^l$. Therefore, for all $l \in [r]$, we have $\Theta^l = C\tilde{\Theta}^l$ for some constant C.

   (b) Assume for all $l \in [r]$, we have $\Theta^l = C\tilde{\Theta}^l$ for some constant C. Then, for any pair $l, l' \in [r]$, we have

$$\begin{aligned} \Theta^0 + u_{kl}\Theta^l &= \tilde{\Theta}^0 + \tilde{u}_{kl}\tilde{\Theta}^l \\ \Theta^0 + u_{k'l'}\Theta^{l'} &= \tilde{\Theta}^0 + \tilde{u}_{k'l'}\tilde{\Theta}^{l'}, \end{aligned}$$

   which implies that

$$\Theta^0 - \tilde{\Theta}^0 = (\tilde{u}_{kl} - Cu_{kl})\tilde{\Theta}^l = (\tilde{u}_{k'l'} - Cu_{k'l'})\tilde{\Theta}^{l'}. \tag{4}$$

   The equation (4) contradicts to the condition 2 if the left hand side $\Theta^0 - \tilde{\Theta}^0$ is not equal to 0. Hence, we have $\Theta^0 = \tilde{\Theta}^0$, and thus $\tilde{u}_{kl} = Cu_{kl}$ for all $k \in I_l$. By condition 3, we have $\|u_l\|_F = \|\tilde{u}_l\|_F = C\|u_l\|_F = 1$. Then, we have $C = 1$ and thus $\Theta^l = C\tilde{\Theta}^l, u_{kl} = \tilde{u}_{kl}$.

Therefore, we have shown that the $(\tilde{U}, \tilde{\Theta}^l) = (U, \Theta^l)$ and the parameters satisfying the condition 1-4 are identifiable. $\square$

2 Accuracy rate

Consider a simple case of model (1) when $r = 1$. The optimization problem is stated below

$$
\begin{aligned}
\min_{\{u,\Theta\}} \quad & \mathcal{L}(u,\Theta) = \sum_{k=1}^{K} \langle S^k, \Omega^k \rangle - \log\det(\Omega^k), \\
s.t. \quad & \Omega^k = \Theta^0 + u_k \Theta, \quad k \in [K], \\
& u_k > 0, \|u\|_F^2 = 1, \\
& u_{k'} \neq u_k, \quad \text{for some } k, k' \in [K] \\
& \Theta^0, \Theta \text{ are positive definite with, and} \\
& \tau_1 < \min\{\varphi_{\min}(\Theta^0), \varphi_{\min}(\Theta)\} \leq \max\{\varphi_{\max}(\Theta^0), \varphi_{\max}(\Theta)\} < \tau_2, \tau_1, \tau_2 > 0
\end{aligned}
$$