Graphic Lasso: Scaled membership with intercept

Jiaxin Hu

April 8, 2021

1 Identifiability

Consider the model

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

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 $I_l = \{k : u_{kl} \neq 0\}$ for $l \in [r]$ and $I_0 = \{k : u_{kl} = 0, l \in [r]\}$.

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 $l \in [r]$, we have $\sum_{k=1}^{K} u_{kl} = 0$.*

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

Intuition

Consider the linear regression model

$$\mathbb{E}[Y] = X\beta, \tag{2}$$

where $Y \in \mathbb{R}^K$, $X \in \mathbb{R}^{K\times(r+1)}$, and $\beta = (\beta_0, \beta_1, ..., \beta_r)^T$. The linear regression model (2) is a scalar analogy of model (1). Let $Y_k = \Omega^k, k = 1, ..., K$, $X = [1_K, U]$, and $\beta_i = \Theta^i, i = 0, 1, ..., r$. Then, we obtain the model (1).

Hence, we may get some intuitions of the identifiability problem from the simple model (2).

1. By the textbook, when X is fixed, the necessary and sufficient condition to identify β is that X has full rank. Thus, we rule out the case that $|I_0| = 0$ and $u_{kl} = u_{k'l}$, for all $k, k' \in I_l$ and $l \in [r]$.
2. Unlike the linear regression model in textbook, X and β are both unknown, and we also want to know the identifiability of X. Suppose there exist another pair of parameters $\tilde{X}, \tilde{\beta}$ such that $X\beta = \tilde{X}\tilde{\beta}$. Then, we must have $C(X) = C(\tilde{X})$.

Though the first column of $X, \tilde{X}$ are both 1_K, it is possible $X \neq \tilde{X}$ and $C(X) = C(\tilde{X})$. A simple example is that $X = [1_K, e_1]$ and $\tilde{X} = [1_K, 1_K + e_1]$. In previous note, the old condition 4 only guarantees the full rankness. Therefore, we need stronger condition to identify the unique X from the unique column space $C(X)$.

The new condition 4 requires the matrix X to be an orthogonal matrix. The following proof will show that the orthogonality and non-overlapping is sufficient to identify X from $C(X)$.

Proof. Suppose $\{\tilde{U}, \tilde{\Theta}^l\}$ also satisfy the model (1). By condition 4, we have

$$\begin{aligned}\sum_{k=1}^{K} \Omega^k &= K\Theta^0 + \sum_{k=1}^{K}\sum_{l=1}^{r} u_{kl}\Theta^l = K\Theta^0 \\ &= K\tilde{\Theta}^0 + \sum_{k=1}^{K}\sum_{l=1}^{r} \tilde{u}_{kl}\tilde{\Theta}^l = K\tilde{\Theta}^0,\end{aligned}$$

which implies that $\Theta^0 = \tilde{\Theta}^0$.

1. Suppose $I_l \neq \tilde{I}_l$ for some $l = 0, 1, ..., r$. Then, there exist a pair $k, k' \in I_l$ but $k \in \tilde{I}_l$ and $k' \in \tilde{I}'_l$. That is

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

   which implies that $\tilde{\Theta}^l = C\tilde{\Theta}^{l'}$ for some constant C. This contradicts to the condition 2.

2. Suppose $I_l = \tilde{I}_l$ for all $l = 0, 1, ..., r$. Then, for all k, l we have

$$u_{kl}\Theta^l = \tilde{u}_{kl}\tilde{\Theta}^{l'},$$

   which implies $\Theta^l = C\tilde{\Theta}^l$ for some constant C. Thus, we have $u_{kl}C = \tilde{u}_{kl}$. By condition 3, note that $\|u_l\|_F = C\|\tilde{u}_l\| = 1$. We have $C = 1$. Therefore, we have $U = \tilde{U}$ and $\Theta^l = \tilde{\Theta}^l$.

□

2 Accuracy rate

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

$$\begin{aligned}\min_{\{u,\Theta\}} \quad & \mathcal{L}(u, \Theta) = \sum_{k=1}^{K} \langle S^k, \Omega^k \rangle - \log\det(\Omega^k), \\ s.t. \quad & \Omega^k = \Theta^0 + u_k\Theta, \quad k \in [K], \\ & u_k > 0, \|u\|_F^2 = 1, \\ & \Theta^0, \Theta \text{ are positive definite with, and} \\ & \tau_1 < \min\{\varphi_{\min}(\Theta^0), \varphi_{\min}(\Theta)\} \leq \max\{\varphi_{\max}(\Theta^0), \varphi_{\max}(\Theta)\} < \tau_2, \tau_1, \tau_2 > 0\end{aligned}$$