Graphic Lasso: Scaled membership (Simple Case)

Jiaxin Hu

March 24, 2021

1 Simple case

Consider the model in which K categories share the same precision matrix structure with different magnitude. The optimization problem is stated below:

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

Remark 1. I just use $\|u\|_F = 1$ for simplicity. Later, we should use $\|u\|_F = K$

Lemma 1 (Precision matrix Accuracy)**.** *Let $\{u,\Theta\}$ denote the true parameters. Consider a estimation $\{\hat{u},\hat{\Theta}\}$ such that $\mathcal{L}(\hat{u},\hat{\Theta}) \geq \mathcal{L}(u,\Theta)$. With probability tends to 1 as $n \to \infty$, we have the accuracy*

$$
\sum_{k=1}^{K} \left\| \hat{\Omega}^k - \Omega^k \right\|_F = \sum_{k=1}^{K} \left\| \hat{u}_k \hat{\Theta} - u_k \Theta \right\|_F \leq 4\tau_2^2 C \sqrt{\frac{p^2 \log p}{n}}
$$

Remark 2. It is weird that there is no K in the accuracy rate.

Proof. **Step I (Decomposition):**

Considering the objective difference between the estimator and the true parameters, we define the following function.

$$
\begin{aligned}
G(\hat{u},\hat{\Theta}) &= \mathcal{L}(\hat{u},\hat{\Theta}) - \mathcal{L}(u,\Theta) \\
&= \sum_{k=1}^{K} \langle S^k, \hat{u}_k \hat{\Theta} \rangle - \langle S^k, u_k \Theta \rangle - \log\det(\hat{u}_k \hat{\Theta}) + \log\det(u_k \Theta) \\
&= G_1(\hat{u},\hat{\Theta}) + G_2(\hat{u},\hat{\Theta}),
\end{aligned}
$$

where

$$
\begin{aligned}
G_1(\hat{u},\hat{\Theta}) &= \sum_{k=1}^{K} \langle S^k, u_k(\hat{\Theta} - \Theta) \rangle - \log\det(u_k \hat{\Theta}) + \log\det(u_k \Theta), \\
G_2(\hat{u},\hat{\Theta}) &= \sum_{k=1}^{K} \langle S^k, (\hat{u}_k - u_k)\hat{\Theta} \rangle - \log \hat{u}_k / u_k.
\end{aligned}
$$

Consider G_1. Let $\Delta = \hat{\Theta} - \Theta$. By Taylor Expansion, we have

$$
\begin{aligned}
-\log\det(u_k\hat{\Theta}) + \log\det(u_k\Theta) &\geq -\langle (u_k\Theta)^{-1}, u_k\Delta\rangle + \frac{1}{2u_k^2\tau_2^2 + \|u_k\Delta\|_F^2}\|u_k\Delta\|_F^2\,, \\
&\geq -\langle u_k^{-1}\Sigma, u_k\Delta\rangle + \frac{1}{2\tau_2^2 + \|\Delta\|_F^2}\|\Delta\|_F^2\,, \qquad (1)
\end{aligned}
$$

where the first inequality uses the fact that $\varphi_{max}(u_k\Theta) \leq u_k\tau_2$. Plugging the inequality (1) into G_1, we have

$$
\begin{aligned}
G_1(\hat{u}, \hat{\Theta}) &\geq \sum_{k=1}^{K}\langle S^k - u_k^{-1}\Sigma, u_k\Delta\rangle + \frac{1}{2\tau_2^2 + \|\Delta\|_F^2}\|\Delta\|_F^2 \\
&= \langle\sum_{k=1}^{K} u_k S^k - K\Sigma, \Delta\rangle + \frac{K}{2\tau_2^2 + \|\Delta\|_F^2}\|\Delta\|_F^2\,.
\end{aligned}
$$

Now we bound the difference between the sample covariance matrix and the true covariance matrix. Let $X_1^k, ..., X_n^k \sim_{i.i.d.} \mathcal{N}_p(0, \Sigma/u_k), k = 1, ..., K$. Note that

$$
\frac{1}{K}\sum_{k=1}^{K} u_k S_{jl}^k = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{n}\sum_{i=1}^{n}\Big((\sqrt{u_k}X_{ij}^k)(\sqrt{u_k}X_{il}^k) - (\sqrt{u_k}X_{.j}^k)(\sqrt{u_k}X_{.l}^k)\Big)\,.
$$

Since $\sqrt{u_k}X_i^k \sim \mathcal{N}(0, \Sigma)$, we have

$$
\left|\frac{1}{nK}\sum_{k=1}^{K}\sum_{i=1}^{n}(\sqrt{u_k}X_{ij}^k)(\sqrt{u_k}X_{il}^k) - \Sigma_{jl}\right| \leq C\sqrt{\frac{\log p}{nK}}, \qquad (2)
$$

with hight probability. Therefore, we have a lower bound of G_1, which is

$$
\begin{aligned}
G_1(\hat{u}, \hat{\Theta}) &\geq \frac{K}{2\tau_2^2 + \|\Delta\|_F^2}\|\Delta\|_F^2 - C\sqrt{K}\sqrt{\frac{\log p}{n}}\|\Delta\|_1 \\
&\geq H_1 - H_2,
\end{aligned}
$$

where

$$
\begin{aligned}
H_1 &= \frac{K}{2\tau_2^2 + \|\Delta\|_F^2}\|\Delta\|_F^2\,, \\
H_2 &= C\sqrt{K}\sqrt{\frac{p^2\log p}{n}}\|\Delta\|_F\,. \qquad (3)
\end{aligned}
$$

Consider G_2. Note that

$$
\begin{aligned}
-\log\hat{u}_k/u_k &= -\log\det(\hat{u}_k\Theta) + \log\det(u_k\Theta) \\
&\geq -\langle u_k^{-1}\Sigma, (\hat{u}_k - u_k)\Theta\rangle + \frac{1}{2u_k^2\tau_2^2 + (\hat{u}_k - u_k)^2\|\Theta\|_F^2}(\hat{u}_k - u_k)^2\|\Theta\|_F^2\,. \qquad (4)
\end{aligned}
$$

Plugging the inequality (4) into G_2, we have

$$
G_2(\hat{u}, \hat{\Theta}) \geq H_1 + H_2 + H_3,
$$

where

$$
\begin{aligned}
H_3 &= \sum_{k=1}^{K} \langle S^k - u_k^{-1}\Sigma, (\hat{u}_k - u_k)\hat{\Theta}\rangle, \\
H_4 &= \sum_{k=1}^{K} \langle u_k^{-1}\Sigma, (\hat{u}_k - u_k)(\hat{\Theta} - \Theta)\rangle, \\
H_5 &= \sum_{k=1}^{K} \frac{1}{2u_k^2\tau_2^2 + (\hat{u}_k - u_k)^2 \left\|\Theta\right\|_F^2} (\hat{u}_k - u_k)^2 \left\|\Theta\right\|_F^2 .
\end{aligned}
$$

For H_3, similarly in G_1, we can bound the difference between the sample covariance and the true covariance matrix. With bound (2), we have

$$
|H_3| \leq \sum_{k=1}^{K} C\sqrt{\frac{p^2 \log p}{n}} \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F . \tag{5}
$$

For H_4, by the property of dual norm and inner product, we have

$$
\begin{aligned}
|H_4| &\leq \sum_{k=1}^{K} \left\|u_k^{-1}\Sigma\right\|_2 \left\|(\hat{u}_k - u_k)\Delta\right\|_* \\
&\leq \sum_{k=1}^{K} \frac{\sqrt{p}}{u_k \tau_1} |\hat{u}_k - u_k| \left\|\Delta\right\|_F .
\end{aligned}
$$

For H_5, we have

$$
H_5 \geq \sum_{k=1}^{K} \frac{1}{2u_k^2\tau_2^2 + \sum_{k=1}^{K}(\hat{u}_k - u_k)^2 \left\|\Theta\right\|_F^2} (\hat{u}_k - u_k)^2 \left\|\Theta\right\|_F^2 .
$$

Step II (Assumption):

By assumption, we have $\mathcal{L}(\hat{u}, \hat{\Theta}) \geq \mathcal{L}(u, \Theta)$. Then, we have

$$
\begin{aligned}
0 &\geq G(\hat{u}, \hat{\Theta}) = G_1(\hat{u}, \hat{\Theta}) + G_2(\hat{u}, \hat{\Theta}) \\
&\geq H_1 + H_5 - H_2 - |H_3| + H_4.
\end{aligned}
$$

Note that $H_2, |H_3| \to 0$ as $n \to \infty$. Therefore, a sufficient condition for the estimation $\{\hat{u}, \hat{\Theta}\}$ to satisfy the assumption is that

$$
\left\|\Delta\right\|_F^2 \to 0, \quad \sum_{k=1}^{K}(\hat{u}_k - u_k)^2 \to 0, \quad \text{as} \quad n \to \infty. \tag{6}
$$

Problem here! I expect that the condition (6) is a necessary condition, but I can not rigorously prove it since we do not know the sign of H_4. If H_4 always keep negative, $\left\|\Delta\right\|_F^2, \sum_{k=1}^{K}(\hat{u}_k - u_k)^2 \to \infty$ may also lead to $0 \geq G$. In the following proof, we assume the estimation satisfies the condition (6).

Suppose the estimation $\{\hat{u}, \hat{\Theta}\}$ satisfies the condition (6). Then, for n large enough, we have

$$
\begin{aligned}
H_1 &\geq \frac{K}{4\tau_2^2} \|\Delta\|_F^2, \\
H_5 &\geq \sum_{k=1}^{K} \frac{1}{4u_k^2\tau_2^2} (\hat{u}_k - u_k)^2 \left\|\hat{\Theta}\right\|_F^2.
\end{aligned}
\tag{7}
$$

Note that $\sum_{k=1}^K u_k^2 = 1$ and by Jensen's inequality we have $\sum_{k=1}^K \frac{1}{u_k^2} \geq K \times K / \sum_{k=1}^K u_k^2 = K^2 \geq K$. Similarly, we have $\sum_{k=1}^K \frac{1}{u_k} \geq \sqrt{K}$.

Two conjecture inequalities:

Since all elements $(\hat{u}_k - u_k)^2$ are near 0, we have

$$
H_5 \geq \frac{K}{4\tau_2^2} \sum_{k=1}^{K} (\hat{u}_k - u_k)^2 \left\|\hat{\Theta}\right\|_F^2, \tag{8}
$$

$$
H_4 \approx \frac{\sqrt{K}}{2\tau_2^2} \sum_{k=1}^{K} \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F \|\Delta\|_F. \tag{9}
$$

Step III(Combination):

Combine all the ingredients (7), (3), (5), (9), and (8) together. We finally have

$$
\begin{aligned}
0 &\geq H_1 + H_5 + H_4 - H_2 - |H_3| \\
&\geq \frac{1}{4\tau_2^2} \left\{ K \|\Delta\|_F^2 + K \sum_{k=1}^{K} (\hat{u}_k - u_k)^2 \left\|\hat{\Theta}\right\|_F^2 + 2\sqrt{K} \|\Delta\|_F \sum_{k=1}^{K} \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F \right\} \\
&\qquad - C\sqrt{\frac{p^2 \log p}{n}} \left[\sqrt{K} \|\Delta\|_F + \sum_{k=1}^{K} \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F \right] \\
&\geq \frac{1}{4\tau_2^2} \left[\sqrt{K} \|\Delta\|_F + \sum_{k=1}^{K} \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F \right]^2 - C\sqrt{\frac{p^2 \log p}{n}} \left[\sqrt{K} \|\Delta\|_F + \sum_{k=1}^{K} \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F \right].
\end{aligned}
$$

Hence, we obtain the accuracy rate

$$
\begin{aligned}
\sum_{k=1}^{K} \left\|\hat{\Omega}^k - \Omega^k\right\|_F &= \sum_{k=1}^{K} \left\|\hat{u}_k\hat{\Theta} - u_k\Theta\right\|_F \\
&\leq \sum_{k=1}^{K} u_k \|\Delta\|_F + \sum_{k=1}^{K} \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F \\
&\leq \sqrt{K} \|\Delta\|_F + \sum_{k=1}^{K} \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F \\
&\leq 4\tau_2^2 C\sqrt{\frac{p^2 \log p}{n}}.
\end{aligned}
$$

□