Graphic Lasso: Scaled membership

Jiaxin Hu

March 21, 2021

1 Simple case

Now we consider a simple case that K categories share the same structure of the precision matrix but with different scalar coefficients. To make the parameters identifiable, we need the assumptions mentioned in the Proposition 2 of the Note 010921. The problem is stated as following.

The original proposition 2 requires column sum to zero.
So, some u have to be negative?

$$
\begin{aligned}
\min_{\{u,\Theta\}} \quad & \mathcal{L}(u,\Theta) = \sum_{k=1}^{K} \langle S^k, \Omega^k \rangle - \log\det(\Omega^k), \\
s.t. \quad & \Omega^k = u_k \Theta, \quad k = 1, ..., K, \\
& u_k \geq 0, \|u\|_F = 1,
\end{aligned}
$$

strictly positive? Otherwise, Omega is not positive semi definite.

1.1 Trivial Accuracy

There is a trivial accuracy rate if we simply consider $u_k\Theta$ as K different precision matrix.

Lemma 1 (Trivial 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*

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

Remark 1. The Trivial Accuracy does not take the advantage of the same structure. So, the accuracy is of order $\mathcal{O}(K)$ which should not be optimal.

1.2 Sharp Accuracy

If we consider the common structure of precision matrix, my conjecture is that the accuracy should be of order $\mathcal{O}(\sqrt{K})$.

Lemma 2 (Sharp Accuracy(conjecture))**.** *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*

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

How about taking ``Omega^k \over mu_k'' as the quantity of interest?
Also, add lower boundedness constraints on ground truth and estimated mu_k.

Proof. Note that

$$\sum_{k=1}^{K} \left\| \hat{u}_k \hat{\Theta} - u_k \Theta \right\|_F^2 \leq \sum_{k=1}^{K} u_k^2 \left\| \hat{\Theta} - \Theta \right\|_F^2 + \sum_{k=1}^{K} \left\| (\hat{u}_k - u_k) \hat{\Theta} \right\|_F^2 ,$$

where $\sum_{k=1}^K u_k^2 = 1$, the first term represents the estimation error from Θ, and the second term represents the estimation error from u. Consider the 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

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

Is there any way to show \hat mu_k —> \mu_k.

Intuitively, this is similar to (continuous-version) clustering accuracy. Should be simpler than the discrete version.

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

$$\begin{aligned}
-\log\det(u_k \hat{\Theta}) + \log\det(u_k \Theta) &\geq -\langle (u_k \Theta)^{-1}, u_k \Delta \rangle + \frac{1}{4\tau^2} \left\| u_k \Delta \right\|_F^2 , \\
&\geq -\langle u_k^{-1} \Sigma, u_k \Delta \rangle + \frac{1}{4\tau^2} u_k^2 \left\| \Delta \right\|_F^2
\end{aligned}$$

where τ is the max singular value of Θ. Then, we have

$$\begin{aligned}
G_1(\hat{u}, \hat{\Theta}) &\geq \sum_{k=1}^{K} \langle S^k - u_k^{-1} \Sigma, u_k \Delta \rangle + \frac{1}{4\tau^2} u_k^2 \left\| \Delta \right\|_F^2 \\
&= \langle \sum_{k=1}^{K} u_k S^k - K\Sigma, \Delta \rangle + \frac{1}{4\tau^2} \left\| \Delta \right\|_F^2 .
\end{aligned}$$

Let $X_1^k, \ldots, X_n^k \sim_{i.i.d.} \mathcal{N}_p(0, \Sigma / u_k), k = 1, \ldots, K$. Note that

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

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

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

with hight probability. Thus, we have

$$G_1(\hat{u},\hat{\Theta}) \geq \frac{1}{4\tau^2}\|\Delta\|_F^2 - C\sqrt{K}\sqrt{\frac{\log p}{n}}\|\Delta\|_1$$
$$\geq \frac{1}{4\tau^2}\|\Delta\|_F^2 - C\sqrt{K}\sqrt{\frac{p^2\log p}{n}}\|\Delta\|_F \tag{1}$$

Consider G_2. Note that

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

Plug into the G_2. We have

$$G_2(\hat{u},\hat{\Theta}) = \sum_{k=1}^K \langle S^k - u_k^{-1}\Sigma, (\hat{u}_k - u_k)\hat{\Theta}\rangle + \langle u_k^{-1}\Sigma, (\hat{u}_k - u_k)(\hat{\Theta} - \Theta)\rangle + \frac{1}{4\tau^2}(\hat{u}_k - u_k)^2\|\Theta\|_F^2$$

Conjecture.

The conjecture might be true given the boundedness constraints on the hat mu_k and mu_k.

$$G_2(\hat{u},\hat{\Theta}) \geq \frac{1}{4\tau^2}\left(\sum_{k=1}^K \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F\right)^2 - C\sqrt{K}\sqrt{\frac{p^2\log p}{n}}\sum_{k=1}^K \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F$$
$$+ \frac{1}{2\tau^2}\|\Delta\|\sum_{k=1}^K \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F. \tag{2}$$

Combining the upper bounds (1) with (2), we have

$$0 \geq G_1(\hat{u},\hat{\Theta}) + G_2(\hat{u},\hat{\Theta})$$
$$\geq \left(\frac{1}{2\tau}\|\Delta\|_F + \frac{1}{2\tau}\sum_{k=1}^K \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F\right)^2 - C\sqrt{K}\sqrt{\frac{p^2\log p}{n}}\left(\|\Delta\|_F + \sum_{k=1}^K \left\|(\hat{u}_k - u_k)\hat{\Theta}\right\|_F\right),$$

which implies the accuracy rate. ∎