Graphic Lasso: Accuracy for precision matrices

Jiaxin Hu

March 18, 2021

1 Penalized vs Hard Constraint

<table>
<tr><th></th><th>Penalized</th><th>Hard Constraint</th></tr>
<tr><td>Problem</td><td>$$\min_{\boldsymbol{U},\Theta^l} \sum_{k=1}^{K}\langle S^k,\Omega^k\rangle - \log\det(\Omega^k) + \lambda\left\|\Omega^k\right\|_1$$ with constraint $$\Omega^k = \sum_{l=1}^{r} u_{kl}\Theta^l,\quad k=1,...,K$$ and $\lambda > 0$, $\boldsymbol{U}$ is a membership matrix.</td><td>$$\min_{\boldsymbol{U},\Theta^l} \sum_{k=1}^{K}\langle S^k,\Omega^k\rangle - \log\det(\Omega^k)$$ with constraint $$\left\|\Theta^l\right\|_0 \leq q, \Omega^k = \sum_{l=1}^{r} u_{kl}\Theta^l,\quad k=1,...,K$$ and $\lambda > 0$, $\boldsymbol{U}$ is a membership matrix., where q represents the true sparasity.</td></tr>
<tr><td>Accuracy</td><td>A1: Suppose $$\Lambda_1\sqrt{\frac{\log p}{n}} \leq \lambda \leq \Lambda_2\sqrt{\frac{1+p/q}{nK}}$$ for Λ_1, Λ_2 large enough, we have $$\sum_{k=1}^{K}\|\Delta_k\|_F \leq C\sqrt{K}\sqrt{\frac{(p+q)\log p}{n}}.$$</td><td>A2: We have $$\sum_{k=1}^{K}\|\Delta_k\|_F \leq C'\sqrt{K}\sqrt{\frac{(p+q)\log p}{n}}$$</td></tr>
</table>

2 Proof

Notations.

1. Let $\mathcal{L}(\boldsymbol{U},\Theta^l) = \sum_{k=1}^{K}\langle S^k,\Omega^k\rangle - \log\det(\Omega^k) + \lambda\left\|\Omega^k\right\|_1$, where $\Omega^k = \sum_{l=1}^{r} u_{kl}\Theta^l, k=1,...,K$.

2. Let D denote the confusion matrix between the estimation $\{\hat{\boldsymbol{U}},\hat{\Theta}^k\}$ and the true parameters, in which $D_{al} = \sum_{k=1}^{K}\boldsymbol{I}\{u_{ka} = \hat{u}_{kl} = 1\}$.

3. Let I_l denote the index set of the categories in the k-th group based on the parameter $\boldsymbol{U}$, i.e., $I_l = \{k : u_{kl} = 1\}$. Let $\hat{I}_l$ denote the estimated set by $\hat{\boldsymbol{U}}$. Note that $|\hat{I}_l \cap I_a| = D_{al}$.

4. $\tau = \max_{l\in[r]} \varphi_{\max}(\Theta^l)$ is the maximal singular value of the true precision matrices.

5. $T = \cup_{l=1}^r T_l$, where T_l is an index set such that for all $(i,j) \in T_l$ we have $\Theta^l_{i,j} \neq 0$.

6. $q = |T|$ be the upper bound of the number of non-zero entries for all Θ^l.

Proof for A_1. First, we define the function $G(\hat{\boldsymbol{U}}, \hat{\Theta}^k)$, where

$$
\begin{aligned}
G(\hat{\boldsymbol{U}}, \hat{\Theta}^l) &= \mathcal{L}(\hat{\boldsymbol{U}}, \hat{\Theta}^l) - \mathcal{L}(\boldsymbol{U}, \Theta^l) \\
&= \sum_{l=1}^r \sum_{a=1}^r \sum_{k\in \hat{I}_l\cap I_a} \langle S^k, \hat{\Theta}^l\rangle - \langle S^k, \Theta^a\rangle - \log\det(\hat{\Theta}^l) + \log\det(\Theta^a) + \lambda \left\|\hat{\Theta}^l\right\|_1 - \lambda \left\|\Theta^a\right\|_1 \\
&= \sum_{l=1}^r \sum_{a=1}^r G_{al}(\hat{\boldsymbol{U}}, \hat{\Theta}^l).
\end{aligned}
$$

Define $\Delta_{al} = \hat{\Theta}^l - \Theta^a, l \in [r], a \in [r]$. Then, the function $G(\hat{\boldsymbol{U}}, \hat{\Theta}^l)$ is a function of $\{\Delta_{al}\}$ denoted by $G(\Delta_{al})$ and $G(0) = 0$. If we take a closed convex set around 0 and show that the function G strictly positive at the boundary, then we obtain accuracy rate based on our construction of the convex set. Particularly, we set the set $\mathcal{A} = \left\{\sum_{a,l} \sqrt{D_{al}} \left\|\Delta_{al}\right\|_F \le M\sqrt{\frac{(p+q)\log p}{n}}\right\}$.

With arbitrary pair of a, l, consider the the decomposition

$$
G_{al}(\hat{\boldsymbol{U}}, \hat{\Theta}^l) = A_{al,1} + A_{al,2} + A_{al,3} + A_{al,4},
$$

where

$$
\begin{aligned}
A_{al,1} &= \sum_{k\in \hat{I}_l\cap I_a} \langle S^k - \Sigma^a, \Delta_{al}\rangle, \\
A_{al,2} &= D_{al}(\mathrm{vec}(\Delta_{al}))^T \int_0^1 (1-v)(\Theta^a + v\Delta_{al})^{-1} \otimes (\Theta^a + v\Delta_{al})^{-1} dv \mathrm{vec}(\Delta_{al}), \\
A_{al,3} &= \lambda D_{al}\left(\left\|\hat{\Theta}^l_{T^c}\right\|_1\right), \\
A_{al,4} &= \lambda D_{al}\left(\left\|\hat{\Theta}^l_{T}\right\|_1 - \left\|\Theta^a_T\right\|_1\right).
\end{aligned}
$$

For $A_{al,1}$, by Note 0115 an 0113, we have $|A_{al,1}| \le A_{al,11} + A_{al,12}$, where

$$
\begin{aligned}
A_{al,11} &\le \sqrt{D_{al}}\left(C_1\sqrt{\frac{q\log p}{n}} + C_2\sqrt{\frac{p\log p}{n}}\right)\left\|\Delta_{al}\right\|_F, \\
&\le \sqrt{D_{al}}\left(C\sqrt{\frac{(q+p)\log p}{n}}\right)\left\|\Delta_{al}\right\|_F, \\
A_{al,12} &= \sqrt{D_{al}}C_1\sqrt{\frac{\log p}{n}}\left\|\Delta_{al,T^c}\right\|_1.
\end{aligned}
$$

For $A_{al,2}$, we have

$$A_{al,2} \geq \frac{D_{al}}{2\tau^2 + \|\Delta_{al}\|_F^2} \|\Delta_{al}\|_F^2 \geq \frac{1}{4\tau^2} \|\Delta_{al}\|_F^2 ,$$

where the second inequality holds when n is large enough.

For $A_{al,4}$, we have

$$|A_{al,4}| \leq \lambda D_{al} \left(\left\| \hat{\Theta}_T^l - \Theta_T^a \right\|_1 \right) \leq \lambda D_{al} \sqrt{q} \|\Delta_{al}\|_F .$$

Combining the decomposition results for all pairs a, l, we obtain that

$$\begin{aligned}
G(\Delta_{al}) &= \sum_{a,l} G_{al}(\Delta_{al}) \\
&\geq \sum_{al} A_{al,2} - A_{al,11} - A_{al,12} + A_{al,3} - |A_{al,4}| \\
&= \frac{1}{4\tau^2} \sum_{al} D_{al} \|\Delta_{al}\|_F^2 - \sum_{al} \sqrt{D_{al}} \left(C \sqrt{\frac{(q+p)\log p}{n}} \right) \|\Delta_{al}\|_F - \sum_{al} \lambda D_{al} \sqrt{q} \|\Delta_{al}\|_F \\
&\qquad + \sum_{al} \lambda D_{al} \left(\left\| \hat{\Theta}_{T^c}^l \right\|_1 \right) - \sum_{al} \sqrt{D_{al}} C_1 \sqrt{\frac{\log p}{n}} \|\Delta_{al,T^c}\|_1 . \qquad (1)
\end{aligned}$$

Note that

$$\begin{aligned}
&\sum_{al} \lambda D_{al} \left(\left\| \hat{\Theta}_{T^c}^l \right\|_1 \right) - \sum_{al} \sqrt{D_{al}} C_1 \sqrt{\frac{\log p}{n}} \|\Delta_{al,T^c}\|_1 \\
&\qquad \geq \sum_{al} \lambda \sqrt{D_{al}} \left(\left\| \hat{\Theta}_{T^c}^l \right\|_1 \right) - \sum_{al} \sqrt{D_{al}} C_1 \sqrt{\frac{\log p}{n}} \|\Delta_{al,T^c}\|_1 \\
&\qquad = \left(\lambda - C_1 \sqrt{\frac{\log p}{n}} \right) \sum_{al} \sqrt{D_{al}} \left(\left\| \hat{\Theta}_{T^c}^l \right\|_1 \right) \geq 0,
\end{aligned}$$

where the last inequality follows by the assumption that $\lambda \geq \Lambda_1 \sqrt{\frac{\log p}{n}}$ for Λ_1 large enough.

Note that

$$\sum_{al} D_{al} \|\Delta_{al}\|_F \leq \sqrt{K} \sum_{al} \sqrt{D_{al}} \|\Delta_{al}\|_F \leq \sqrt{K} M \sqrt{\frac{(p+q)\log p}{n}},$$

and by Cauchy Schwartz we have

$$\sum_{al} D_{al} \|\Delta_{al}\|_F^2 \geq \frac{1}{r^2} (\sum_{al} \sqrt{D_{al}} \|\Delta_{al}\|_F)^2.$$

Then, the function (1) becomes

$$\begin{aligned}
G(\Delta_{al}) &\geq \frac{1}{4\tau^2} \sum_{al} D_{al} \|\Delta_{al}\|_F^2 - \sum_{al} \sqrt{D_{al}} \left(C \sqrt{\frac{(q+p)\log p}{n}} \right) \|\Delta_{al}\|_F - \sum_{al} \lambda D_{al} \sqrt{q} \|\Delta_{al}\|_F \\
&\geq \frac{1}{4\tau^2 r^2} M^2 \frac{(q+p)\log p}{n} - CM \frac{(q+p)\log p}{n} - \Lambda_2 M \frac{(q+p)\log p}{n} \\
&> 0,
\end{aligned}$$

when M is large enough. Therefore, there is a local minima inside the convex set $\mathcal{A}$ and the convex set implies that

$$\sum_{k=1}^{K} \left\| \Omega^k - \hat{\Omega}^k \right\|_F = \sum_{al} D_{al} \left\| \Delta_{al} \right\|_F \le \sqrt{K} C \sqrt{\frac{(p+q)\log p}{n}}.$$

$\square$