Supplementary Materials to "Dynamic Covariance Models"

The lower bound theory for estimating sparse nonparametric covariance matrices

We discuss the lower bound theory for estimating sparse nonparametric covariance matrices under the operator norm. A key difference of our arguments to those in Cai et al. (2010) and Cai and Zhou (2012) is that we deal with the situation where the entries of the covariance matrix vary with covariates.

Again, we assume that $Y_i \sim N(m(U_i), \Sigma(U_i))$, for $i = 1, \cdots, n$, and set $m(U_i) \equiv \mathbf{0}$ without loss of generality. All the notations, if not defined, are the same as those in the main paper. We establish the lower bound for estimating these matrices in the family of covariance matrices $\mathcal{U}(0, c_0, M_2; \Omega)$ as follows.

Let

$$g(t) = 3\exp(-\frac{t^2}{1-t^2})I(|t| \le 1),$$

and let $r_{n,p} := \frac{\sqrt{\log p}}{n^\alpha} \to 0$ as $n \to \infty$, where $\alpha > 0$. Define $\Sigma_0(u)$ as a diagonal covariance matrix, where the diagonal elements equal $M_2/2$ for a constant $M_2 > 0$, and define $\Sigma_l(u)$ as a diagonal covariance matrix whose diagonal elements are $M_2/2$ except the l-th diagonal element which is $M_2/2 + r_{n,p}g(\frac{u-u_0}{h_n})$, for $l = 1, \cdots, p$. Here $u_0 \in \Omega$, $h_n = \tau n^{-\beta}$ and $\beta > 0$.

Define $\mathcal{U}_0$ as the collection of the $p+1$ covariance matrices

$$\Sigma_0 := \{\Sigma_0(u), u \in \Omega\}, \Sigma_1 := \{\Sigma_1(u), u \in \Omega\}, \cdots, \Sigma_p := \{\Sigma_p(u), u \in \Omega\}.$$

From Theorem 1 in the paper, we have for some constant $0 < L_1 < \infty$,

$$\sup_{u \in \Omega} |\sigma_{ij}^{(2)}(u)| < L_1 \sqrt{\log p}, \text{ for } i = 1, \cdots, p; j = 1, \cdots, p.$$

It is seen then that the bias term does not affect the result in Lemma 7, if $M' h^2 \sqrt{\log p} > 3L_1 h^2 \sqrt{\log p}$ for M' sufficiently large. Therefore, if we can assume $2\beta - \alpha \leq 0$, then $\mathcal{U}_0 \subset \mathcal{U}(0, c_0, M_2; \Omega)$ with $L_1 = \frac{1}{\tau^2} \sup_t |g^{(2)}(t)| < \infty$.

For any k and l with $0 \leq k < l \leq p$, we have

$$\sup_u ||\Sigma_k(u) - \Sigma_l(u)|| = \sup_t g(t) r_{n,p} \geq 2r_{n,p}. \tag{1}$$

Thus, for any estimator $\hat{\Sigma}(u)$, it holds that

$$P_{\Sigma_l}(\sup_u ||\hat{\Sigma}(u) - \Sigma_l(u)|| \geq r_{n,p}) \geq P_{\Sigma_l}(\Upsilon^* \neq l), \ l = 0, 1, \cdots, p,$$

where $\Upsilon^* = \arg\min_{0 \leq l \leq p} \sup_u ||\hat{\Sigma}(u) - \Sigma_l(u)||$. We thus obtain

$$\begin{aligned} \inf_{\hat{\Sigma}} \sup_{\Sigma \in \mathcal{U}(0,c_0,M_2;\Omega)} E_\Sigma(r_{n,p}^{-1} \sup_u ||\hat{\Sigma}(u) - \Sigma(u)||) &\geq \inf_{\hat{\Sigma}} \sup_{\Sigma \in \mathcal{U}(0,c_0,M_2;\Omega)} P_\Sigma(r_{n,p}^{-1} \sup_u ||\hat{\Sigma}(u) - \Sigma(u)|| \geq 1) \\ &= \inf_{\hat{\Sigma}} \sup_{\Sigma \in \mathcal{U}(0,c_0,M_2;\Omega)} P_\Sigma(\sup_u ||\hat{\Sigma}(u) - \Sigma(u)|| \geq r_{n,p}) \\ &\geq \inf_{\hat{\Sigma}} \max_{\Sigma \in \mathcal{U}_0} P_\Sigma(\sup_u ||\hat{\Sigma}(u) - \Sigma(u)|| \geq r_{n,p}) \\ &\geq \inf_{\Upsilon} \max_{0 \leq l \leq p} P_{\Sigma_l}(\Upsilon \neq l), \end{aligned} \tag{2}$$

where $\inf_\Upsilon$ denotes the infimum over all tests.

Let P_l and P_0 be two probability measures having density functions p_l and p_0, respectively, where

$$
\begin{aligned}
p_l &= \prod_{i=1}^{n} p_{li} = \prod_{i=1}^{n} \frac{1}{(2\pi)^{p/2}(M_2/2 + r_{n,p}K(\frac{U_i-u_0}{h_n}))^{1/2}(M_2/2)^{(p-1)/2}} \exp\{-\frac{1}{2}Y_i^T\Sigma_l(U_i)^{-1}Y_i\} \\
&= \prod_{i=1}^{n} \Big[\{\prod_{j\neq l} \frac{1}{(2\pi)^{1/2}(M_2/2)^{1/2}} \exp\{-\frac{y_{ij}^2}{M_2}\}\} \\
&\times \frac{1}{(2\pi)^{1/2}(M_2/2 + r_{n,p}g(\frac{U_i-u_0}{h_n}))^{1/2}} \exp\{-\frac{y_{il}^2}{M_2 + 2r_{n,p}g(\frac{U_i-u_0}{h_n})}\}\Big]
\end{aligned}
$$

and

$$
p_0 = \prod_{i=1}^{n} p_{0i} = \prod_{i=1}^{n} \frac{1}{(\pi M_2)^{p/2}} \exp\{-\frac{1}{M_2}\sum_{j=1}^{p} y_{ij}^2\} = \prod_{i=1}^{n}\prod_{j=1}^{p} \frac{1}{(2\pi)^{1/2}(M_2/2)^{1/2}} \exp\{-\frac{y_{ij}^2}{M_2}\}.
$$

The Kullback-Leibler divergence of P_l and P_0 given $(U_1, \cdots, U_n)$, i.e., $K(P_l, P_0)$, defined as $\int p_l \log(p_l/p_0)dY$ (Kullback and Leibler, 1951), satisfies that, when n is large enough,

$$
K(P_l, P_0) = \sum_{i=1}^{n} \int p_{li} \log \frac{p_{li}}{p_{0i}} dY_i \leq L_1 \sum_{i=1}^{n} r_{n,p}^2 g^2(\frac{U_i - u_0}{h_n}),
$$

for some constant $L_1 > 0$. Since $K(P_1, P_0) = \cdots = K(P_p, P_0)$, by the definition of g, there exists some constant $L_2 > 0$ such that

$$
\frac{1}{p}\sum_{l=1}^{p} K(P_l, P_0) \leq nh_n \frac{1}{nh_n}\sum_{i=1}^{n} L_1 r_{n,p}^2 g^2(\frac{U_i - u_0}{h_n}) \leq L_2 nh_n r_{n,p}^2 = \tau L_2 n (\log p) n^{-2\alpha-\beta}.
$$

Setting $\tau = 1/(9L_2)$, if $2\alpha + \beta \geq 1$, we have

$$
\frac{1}{p}\sum_{l=1}^{p} K(P_l, P_0) \leq \gamma \log p, \tag{3}
$$

where $0 < \gamma < 1/8$.

Lemma 1. *(Tsybakov, 2009) If probability measures $P_0, P_1, \cdots, P_p$ satisfy*

$$
\frac{1}{p}\sum_{l=1}^{p} K(P_l, P_0) \leq \alpha^*
$$

with $0 < \alpha^* < \infty$, *then*

$$\inf_{\Upsilon} \max_{0 \le l \le p} P_l(\Upsilon \ne l) \ge \sup_{0<\phi<1} \Big\{ \frac{\phi p}{1+\phi p}(1 + \frac{\alpha^* + \sqrt{\alpha^*/2}}{\log \phi}) \Big\}.$$

By (1), (2) and (3), and the definition of P_l and P_0, applying Lemma 1, we obtain

$$\inf_{\hat{\Sigma}} \sup_{\Sigma \in \mathcal{U}(0,c_0,M_2;\Omega)} E_{\Sigma}(\sup_u ||\hat{\Sigma}(u) - \Sigma(u)||) \ge c r_{n,p},$$

where $c = \frac{1}{2}(1 - 2\gamma - \sqrt{\frac{2\gamma}{\log 2}}) > 0$. Note that the positive constants α and β satisfy $2\beta - \alpha \le 0$ and $2\alpha + \beta \ge 1$. The minimum value of α satisfying these two conditions is $2/5$. Thus, the lower bound for estimating nonparametric covariance matrices in $\mathcal{U}(0, c_0, M_2; \Omega)$ is

$$\inf_{\hat{\Sigma}} \sup_{\Sigma \in \mathcal{U}(0,c_0,M_2;\Omega)} E_{\Sigma}(\sup_u ||\hat{\Sigma}(u) - \Sigma(u)||) \ge c \frac{\sqrt{\log p}}{n^{2/5}}. \tag{4}$$

That is, for estimating the nonparametric covariance matrices in $\mathcal{U}(0, c_0, M_2; \Omega)$, our proposed estimators are optimal.

References

Cai, T.T., Zhang, C. H. and Zhou, H. H. (2010). Optimal rates of convergence for covariance matrix estimation. The Annals of Statistics, 38, 2118-2144.

Cai, T.T. and Zhou, H. H. (2012). Optimal rates of convergence for sparse covariance matrix estimation. The Annals of Statistics, 40, 2389–2420.

Kullback, S. and Leibler, R.A. (1951). On information and sufficiency. Annals of Mathematical Statistics, 22, 79–86.

Tsybakov, A. B. (2009). Introduction to nonparametric estimation. Springer, New York.