Sharp Optimal Signal Detection in Covariance and Precision Matrices

Yumou Qiu

Peking University

Collaborated with Song Xi Chen, Bin Guo and Shuyi Zhang

Outline of talk

- Sharp minimax result for covariance testing;
- Multi-level thresholding test;
- More powerful than both L_2 and $L_{\max}$ type tests;
- Testing problem for precision matrices.

Background

- Signal detection: whether a block of covariance (precision) matrix $= 0$;
- Signal detection vs. signal identification (multiple testing problem);
- Studying association (conditional association) is common in science;
- Gene co-expression network, brain connectivity studies;
- Omics data analysis: association given one type of variables.

One-sample covariance testing

- $\mathrm{X}_1, \ldots, \mathrm{X}_n \sim N(\boldsymbol{\mu}, \Sigma)$ (normality can be relaxed to sub-Gaussanity);
- $\mathrm{X}_k = (X_{k1}, \ldots, X_{kp})^{\mathrm{T}}$ for $k = 1, \ldots, n$, $p \gg n$;
- $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)^{\mathrm{T}}$ and $\Sigma = (\sigma_{j_1 j_2})_{p \times p}$;
- Testing Σ being diagonal:

$$H_0 : \sigma_{j_1 j_2} = 0 \ \text{ for all } \ j_1 \neq j_2 \quad \text{vs.} \quad H_a : \sigma_{j_1 j_2} \neq 0 \ \text{ for some } \ j_1 \neq j_2 \tag{1}$$

- Testing for sub-blocks of Σ and two-sample problems

Existing minimax rate result, Cai and Ma (2013)

- $H_0 : \Sigma = \mathrm{I}_p$ vs. $H_a : \Sigma \neq \mathrm{I}_p$ under Gaussian distribution;
- Alternative covariance class $\mathcal{U}_1(b) = \{\Sigma : \|\Sigma - \mathrm{I}_p\|_F \geq b(p/n)^{1/2}\}$;
- $\mathcal{W}_{1,\alpha}$: all α-level tests for the null hypothesis $H_0 : \Sigma = \mathrm{I}_p$;
- There exist positive constants b_0 and $\omega \in (0, 1)$ such that

$$\sup_{W \in \mathcal{W}_{1,\alpha}} \inf_{\Sigma \in \mathcal{U}_1(b_0)} P(W = 1) \leq 1 - \omega,$$

  for any $\omega \in (0, 1 - \alpha)$, as $n, p \to \infty$.

- The expression of b_0 is unknown.

Existing minimax rate result, Cai, Liu and Xia (2013)

- $H_0 : \Sigma_1 = \Sigma_2$ vs. $H_a : \Sigma_1 \neq \Sigma_2$;
- Alternative covariance class: $m_a = p^{\kappa}$ $(\kappa < 1/2)$ unequal covariances,

$$\mathcal{U}_2(c, m_a, K_0) = \Big\{(\Sigma_1, \Sigma_2) : \max_{1 \le j \le p} \{\sigma_{1,jj}, \sigma_{2,jj}\} \le K_0, \\ \|\Sigma_1 - \Sigma_2\|_F^2 \ge c m_a \log(p)/n\Big\}$$

- $\mathcal{W}_{2,\alpha}$: all α-level tests for the two-sample hypotheses;
- There exist positive constant c_0 such that, as $n, p \to \infty$,

$$\sup_{W \in \mathcal{W}_{2,\alpha}} \inf_{(\Sigma_1,\Sigma_2) \in \mathcal{U}_2(c_0, m_a, K_0)} P(W = 1) \le 1 - \omega.$$

- The expression of c_0 is unknown.

Limitation of existing minimax result

- only provide the rate of the minimum signals that can be detected;
- tight bound on the minimax power is unknown;
- Cai and Ma (2013) for non-sparse regime ($m_a \asymp p$ with strength $n^{-1/2}$);
- Cai, Liu and Xia (2013) for high sparsity regime ($m_a \ll p^{1/2}$);
- moderate sparsity regime is unknown ($p^{1/2} \ll m_a \ll p$);

Sparse and weak signals for means

- Donoho and Jin (2004)
- Test for $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)^{\mathrm{T}} = 0$ of normal distribution $N(\boldsymbol{\mu}, I)$
- Nonzero value (signal): $\mu_a = \sqrt{2r \log(p)/n}$
- Proportion of signals: $\epsilon = p^{-\beta}$
- β: signal sparsity parameter; r: signal strength parameter

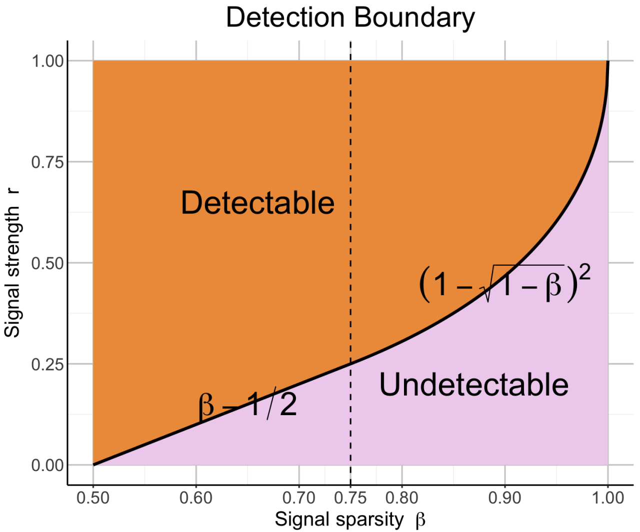
$$H_0 : \mu_j = 0 \ \text{ for all } 1 \le j \le p \ \text{ vs. } \ H_a : \mu_j \overset{i.i.d.}{\sim} (1-\epsilon)\nu_0 + \epsilon\nu_{\mu_a}$$

Detection boundary–phase transition property

$$\mathrm{DB}(\beta) = \begin{cases} \beta - 1/2 & \text{if } 1/2 < \beta \leq 3/4, \\ (1 - \sqrt{1-\beta})^2 & \text{if } 3/4 < \beta < 1. \end{cases} \quad (2)$$

- If $r < \mathrm{DB}(\beta)$, Type I error rare + Type II error rare $\to 1$ for any test;
- If $r > \mathrm{DB}(\beta)$, Type I error rare + Type II error rare $\to 0$ for some test;
- Higher Criticism test can achieve the detection boundary.

Detection boundary



Covariance class under alternative hypothesis

$$H_0 : \sigma_{j_1 j_2} = 0 \text{ for all } j_1 \neq j_2 \quad \text{vs.} \quad H_a : \sigma_{j_1 j_2} \neq 0 \text{ for some } j_1 \neq j_2$$

- Sample size n, dimension p, and $q = p(p-1)/2$
- Number of nonzero $\sigma_{j_1 j_2}$, $j_1 \neq j_2$: $m_a = \lfloor q^{(1-\beta)} \rfloor$ for $\beta \in (1/2, 1)$
- Nonzero covariance: $\sigma_{j_1 j_2} = \sqrt{2 r_{j_1 j_2} \log(q)/n}$ if $\sigma_{j_1 j_2} \neq 0$;

$$\mathcal{U}(\beta, r_0, \tau) = \big\{ \Sigma : m_a = \lfloor q^{(1-\beta)} \rfloor \text{ nonzero } \sigma_{j_1 j_2} \text{ with } r_{j_1 j_2} \geq r_0, \\ \max_{1 \leq j \leq p} \sigma_{jj} \leq \tau, \min_{1 \leq j \leq p} \sigma_{jj} \geq C^{-1} \big\}.$$

Detection boundary for covariances

Theorem (Sharp minimax result)

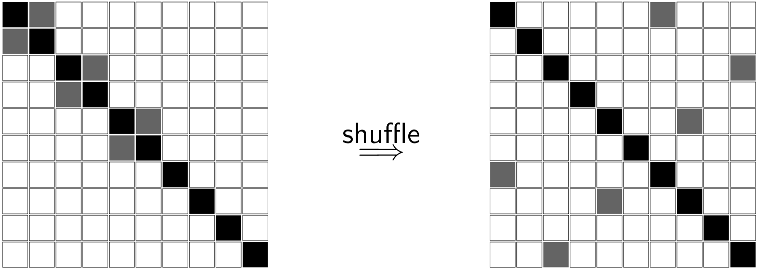
Under $\log p = o(n^{1/3})$ *and Gaussian distributed data, if* $r_0\tau^{-2} < \mathrm{DB}(\beta)$,

$$\sup_{W \in \mathcal{W}_\alpha} \inf_{\Sigma \in \mathcal{U}(\beta, r_0, \tau)} P(W = 1) \le 1 - \omega$$

for any $\omega \in (0, 1-\alpha)$, *as* $n, p \to \infty$.

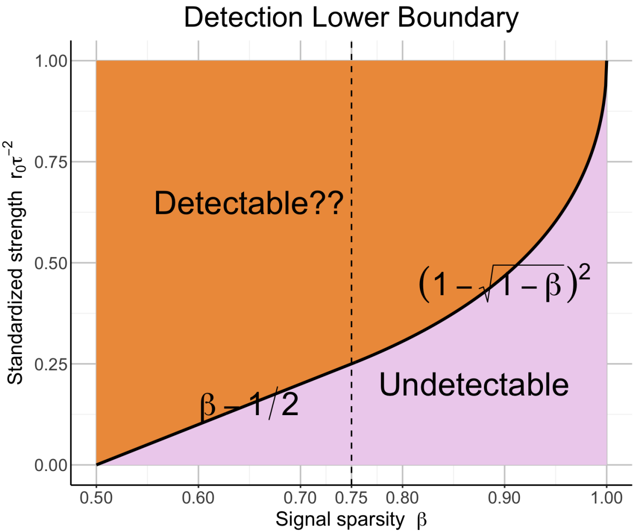
- signal sparsity parameter $\beta \in (1/2, 1)$ implies $1 \ll m_a \ll p$;
- $r_0\tau^{-2}$: minimum standardized signal strength;
- the exact minimum signal strength r_0 at the order $\{\log(p)/n\}^{1/2}$;

Proof: least favorable prior under H_a



- Least favorable H_a: uniform prior on all random shuffles
- Hellinger distance of densities under H_0 and least favorable H_a

Tight detection boundary?



Proposed test

- Sample covariance: $\widehat{\Sigma} = (\hat{\sigma}_{j_1 j_2})_{p\times p} = \frac{1}{n}\sum_{k=1}^{n}(\mathrm{X}_k - \bar{\mathrm{X}})(\mathrm{X}_k - \bar{\mathrm{X}})^{\mathrm{T}}$;
- Standardization: $G_{j_1 j_2} = \frac{\hat{\sigma}_{j_1 j_2}}{(\hat{\sigma}_{j_1 j_1}\hat{\sigma}_{j_2 j_2}/n)^{1/2}} = \sqrt{n}\hat{\rho}_{j_1 j_2},\ 1 \le j_1 < j_2 \le p$;
- Squared standardized sample covariance: $M_{j_1 j_2} = G_{j_1 j_2}^2$;
- Non-Gaussian data: $\tilde{G}_{j_1 j_2} = \sqrt{n}\hat{\sigma}_{j_1 j_2}\hat{\theta}_{j_1 j_2}^{-1/2}$ as standardization of $\hat{\sigma}_{j_1 j_2}$;
- $\theta_{j_1 j_2} = \mathrm{Var}\{(X_{kj_1} - \mu_{j_1})(X_{kj_2} - \mu_{j_2})\}$;
- $\hat{\theta}_{j_1 j_2} = \frac{1}{n}\sum_{k=1}^{n}\{(X_{kj_1} - \bar{X}_{j_1})(X_{kj_2} - \bar{X}_{j_2}) - \hat{\sigma}_{j_1 j_2}\}^2$.

Thresholding for covariance testing

Thresholding on $M_{j_1j_2}$, powerful for sparse and weak signals

$$T_n(s) = \sum_{1 \le j_1 < j_2 \le p} M_{j_1j_2} \mathbb{I}\{M_{j_1j_2} > \lambda_p(s)\}$$

$\lambda_p(s) = 4s\log(p)$ for a thresholding parameter $s \in (0,1)$.

- Cai and Jiang (2011): L_{max} statistic $M_n = \max_{1 \le j_1 < j_2 \le p} M_{j_1j_2}$
  - only use the maximal signal, powerful for strong signals
- Qiu and Chen (2012): L_2 statistic, summation of all $M_{j_1j_2}$
  - include too many uninformative entries, powerful for dense signals

Mean and variance of $T_n(s)$ under H_0

- $\mu_0(s) = \mathrm{E}\{T_n(s)|H_0\}$ and $\sigma_0^2(s) = \mathrm{Var}\{T_n(s)|H_0\}$;
- Large deviation: tail of $M_{j_1 j_2}$ can be approximated by tail of $\mathcal{X}_1^2$;
- L_p: multi-log(p) term;
- $\mu_0(s) = \tilde{\mu}_0(s)\{1 + O(L_p n^{-1/2})\}$ and $\sigma_0^2(s) = \tilde{\sigma}_0^2(s)\{1 + o(1)\}$;
- $\tilde{\mu}_0(s) = q\{2\lambda_p^{1/2}(s)\phi(\lambda_p^{1/2}(s)) + 2\bar{\Phi}(\lambda_p^{1/2}(s))\}$;
- $\tilde{\sigma}_0^2(s) = q[2\{\lambda_p^{3/2}(s) + 3\lambda_p^{1/2}(s)\}\phi(\lambda_p^{1/2}(s)) + 6\bar{\Phi}(\lambda_p^{1/2}(s))]$;
- Under Gaussian data, $\mu_0(s)$ has an exact closed form expression.

Asymptotic normality of $T_n(s)$ under H_0

Theorem (Asymptotic normality)

Under H_0, and either (i) $\log p \sim n^{\varpi}$ for $\varpi \in (0, 1/5)$ and $s > 1/2$ or (ii) $n \sim p^{\xi}$ for $\xi \in (0, 2]$ and $s > 1/2 - \xi/4$, we have

$$\sigma_0^{-1}(s)\{T_n(s) - \mu_0(s)\} \xrightarrow{d} N(0,1) \quad \text{as } n, p \to \infty.$$

- $M_{j_1 j_2}$ is dependent with $M_{j_1 j_3}$ for all j_1, j_2, j_3 (circular dependence)
- Larger s helps to alleviated the covariances among $\{M_{j_1 j_2}\}$

Selecting threshold level

- Reject H_0 of (1) if $T_n(s) > \mu_0(s) + z_\alpha \tilde{\sigma}_0(s)$
- z_α is the upper α quantile of $N(0,1)$
- Single-level thresholding test
- **How to select threshold level s?**

Selecting threshold level

- Reject H_0 of (1) if $T_n(s) > \mu_0(s) + z_\alpha \tilde{\sigma}_0(s)$
- z_α is the upper α quantile of $N(0, 1)$
- Single-level thresholding test
- **How to select threshold level s?**

Multi-level thresholding–a higher criticism approach

- Standardization of $T_n(s)$: $U_n(s) = \tilde{\sigma}_0^{-1}(s)\{T_n(s) - \mu_0(s)\}$
- P-value: $\text{pv}(s) = \mathbb{P}\{N(0,1) > U_n(s)\}$
- Choose $s \in (s_0, 1)$ that minimizes $\text{pv}(s)$: most significant threshold
- $s_0 = 1/2$ or $s_0 = 1/2 - \xi/4$ if $n \sim p^{\xi}$ for a $\xi \in (0,2)$
- Equivalently: maximize $U_n(s)$ over multiple thresholds

$$\mathcal{V}_n(s_0) = \sup_{s \in (s_0,1)} U_n(s)$$

Limiting distribution of $\mathcal{V}_n(s_0)$

Theorem

Under H_0 and the conditions for asymptotic normality of $T_n(s)$,

$$P\{a(\log(p))\mathcal{V}_n(s_0) - b(\log(p), s_0) \leq x\} \to \exp(-e^{-x}),$$

where $a(y) = (2\log(y))^{1/2}$ *and* $b(y, s_0) = 2\log(y) + 2^{-1}\log\log(y) - 2^{-1}\log(\pi) + \log(1 - s_0)$.

- $\mathcal{V}_n(s_0) = \sup_{s \in (s_0, 1)} U_n(s)$: asymptotic Gumbel
- Multi-thresholding test (MTT):

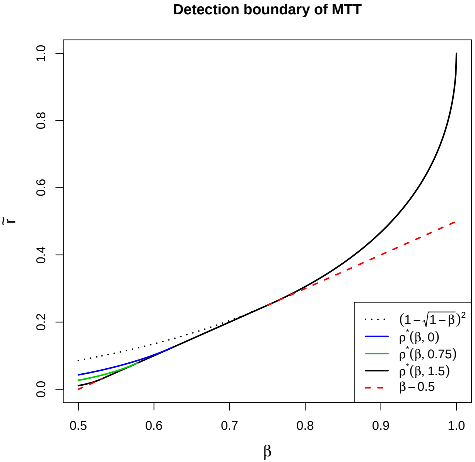
$$\mathcal{V}_n(s_0) > \{q_\alpha + b(\log(p), s_0)\}/a(\log(p))$$

Detection boundary of MTT

$$\mathrm{DB}^*(\beta,\xi) = \begin{cases} \frac{(\sqrt{4-2\xi}-\sqrt{6-8\beta-\xi})^2}{8}, & 1/2 < \beta \le 5/8 - \xi/16; \\ \beta - 1/2, & 5/8 - \xi/16 < \beta \le 3/4; \\ (1-\sqrt{1-\beta})^2, & 3/4 < \beta < 1. \end{cases} \quad (3)$$

- $n \sim p^{\xi}$ for $\xi \in (0, 2]$;
- $\xi = 0$: exponential growth of p with respect to n;
- $\bar{r} = \max_{(j_1,j_2)\in\mathcal{A}_1} \frac{r_{j_1j_2}}{\sigma_{j_1j_1}\sigma_{j_2j_2}}$ and $\underline{r} = \min_{(j_1,j_2)\in\mathcal{A}_1} \frac{r_{j_1j_2}}{\sigma_{j_1j_1}\sigma_{j_2j_2}}$;
- If $\underline{r} > \mathrm{DB}^*(\beta,\xi)$, power of MTT converges to 1;
- If $\bar{r} < \mathrm{DB}^*(\beta,\xi)$, power of MTT converges to 0,

Detection boundary of MTT



- Power $\to 1$ if $\underline{r} > \mathrm{DB}^*(\beta, \xi)$
- $\mathrm{DB}^*(\beta, \xi) = \mathrm{DB}(\beta)$ for $\beta > 5/8 - \xi/16$
- $\mathrm{DB}^*(\beta, \xi) > \mathrm{DB}(\beta)$ for $\beta < 5/8 - \xi/16$
- Due to imposing $s > s_0$
- L_{max} test: $\bar{r} > 4$
- L_2 test: powerless if $\beta > 1/2$

Sharp optimality of MTT

Theorem (Sharp optimality of MTT)

Under the Gaussian distribution and either (i) $\log p \sim n^{\varpi}$ *with* $5/8 < \beta < 1$ *or (ii)* $n \sim p^{\xi}$ *with* $5/8 - \xi/16 < \beta < 1$*, as* $n, p \to \infty$*,*

(i) if $r_0\tau^{-2} > \mathrm{DB}(\beta)$*,* $\inf_{\Sigma \in \mathcal{U}(\beta, r_0, \tau)}$ *Power of MTT* $\to 1$*;*

(ii) if $r_0\tau^{-2} < \mathrm{DB}(\beta)$*,* $\sup_{W \in \mathcal{W}_\alpha} \inf_{\Sigma \in \mathcal{U}(\beta, r_0, \tau)} \mathbb{P}(W = 1) \leq \alpha$*.*

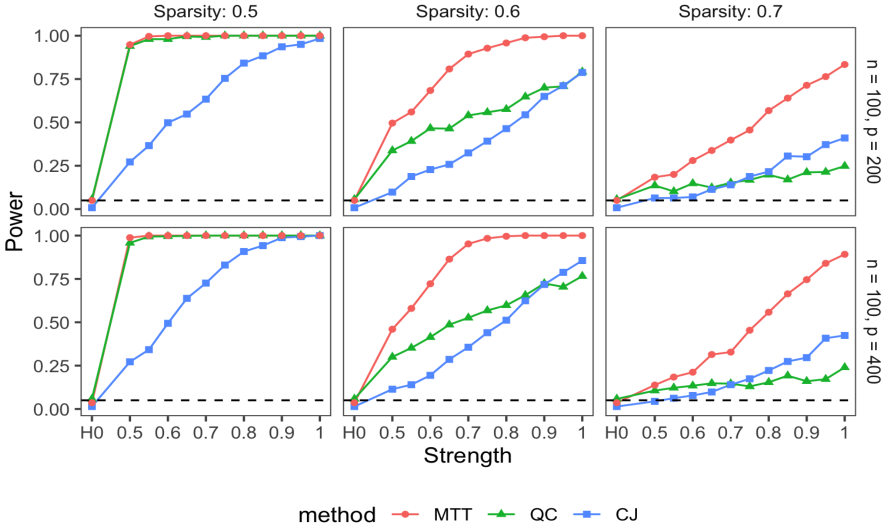
Simulation design

- $\Sigma = \sigma_0^2 \mathrm{I}$ under H_0
- $\sigma_{jj+1} = \sqrt{4r\log(p)/n}$ for $j = 1, \ldots, m_a$ under H_a
- $m_a = \lfloor q^{(1-\beta)} \rfloor$ for $q = p(p-1)/2$
- Qiu and Chen (2012) (QC): L_2 test
- Cai and Jiang (2011) (CJ): L_{max} test

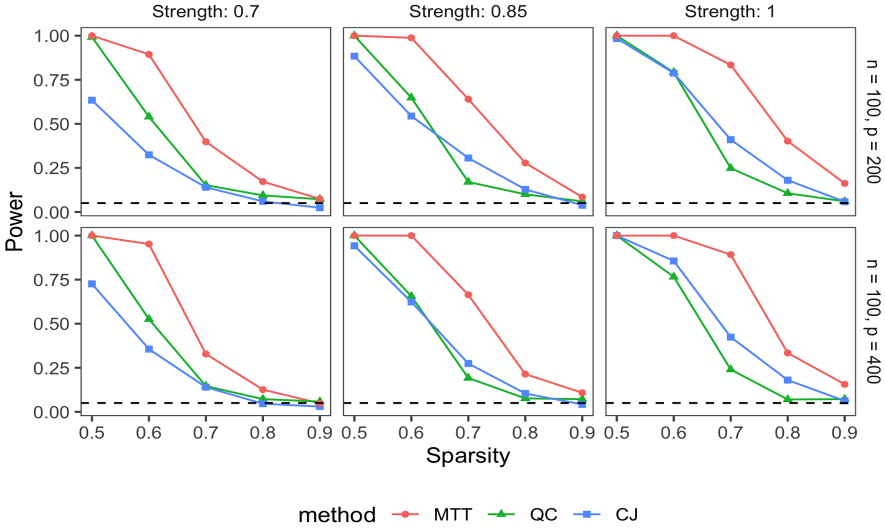
Empirical size–nominal level 5%

		$\sigma_0^2 = 1$			$\sigma_0^2 = 1.5$			$\sigma_0^2 = 2$		
n	p	MTT	CQ	CJ	MTT	CQ	CJ	MTT	CQ	CJ
60	200	0.031	0.060	0.006	0.033	0.052	0.008	0.038	0.055	0.010
	300	0.018	0.054	0.001	0.024	0.059	0.014	0.011	0.074	0.001
	400	0.015	0.067	0.002	0.012	0.062	0	0.015	0.056	0.002
80	200	0.056	0.054	0.016	0.040	0.068	0.013	0.058	0.061	0.012
	300	0.035	0.061	0.012	0.027	0.063	0.004	0.034	0.046	0.007
	400	0.032	0.065	0.008	0.037	0.042	0.017	0.027	0.034	0.004
100	200	0.054	0.074	0.011	0.056	0.062	0.023	0.062	0.046	0.018
	300	0.049	0.044	0.015	0.046	0.040	0.012	0.016	0.061	0.006
	400	0.031	0.058	0.010	0.039	0.071	0.013	0.023	0.058	0.009

Power versus signal strength



Power versus signal sparsity



Two sample covariances testing

- $\mathrm{X}_1, \ldots, \mathrm{X}_{n_1} \overset{i.i.d.}{\sim} \mathrm{F}_1$ and $\mathrm{Y}_1, \ldots, \mathrm{Y}_{n_2} \overset{i.i.d.}{\sim} \mathrm{F}_2$
- $\mathrm{X}_k = (X_{k1}, \ldots, X_{kp})^{\mathrm{T}}$ and $\mathrm{Y}_k = (Y_{k1}, \ldots, Y_{kp})^{\mathrm{T}}$
- Covariances: $\Sigma_1 = (\sigma_{ij1})_{p \times p}$ and $\Sigma_2 = (\sigma_{ij2})_{p \times p}$

$$H_0 : \Sigma_1 = \Sigma_2 \quad \text{vs.} \quad H_a : \Sigma_1 \neq \Sigma_2$$

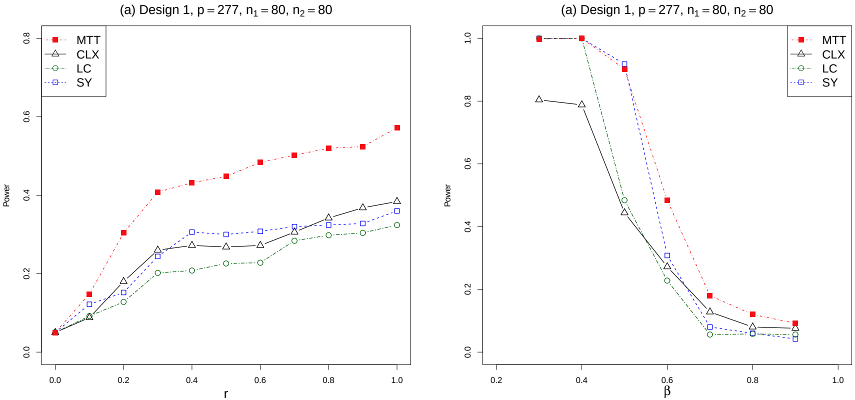
Standardization of Sample Covariances

- $\mathrm{S}_{n1} = (s_{ij1})$, $\mathrm{S}_{n2} = (s_{ij2})$: sample covariances from $\{\mathrm{X}_k\}$ and $\{\mathrm{Y}_k\}$
- $\theta_{ij1} = \mathrm{Var}\{(X_{ki} - \mu_{1i})(X_{kj} - \mu_{1j})\}$ and $\theta_{ij2} = \mathrm{Var}\{(Y_{ki} - \mu_{2i})(Y_{kj} - \mu_{2j})\}$
- $\hat{\theta}_{ij1} = \frac{1}{n_1}\sum_{k=1}^{n_1}\{(X_{ki} - \bar{X}_i)(X_{kj} - \bar{X}_j) - s_{ij1}\}^2 \to \theta_{ij1}$
- $\hat{\theta}_{ij2} = \frac{1}{n_2}\sum_{k=1}^{n_2}\{(Y_{ki} - \bar{Y}_i)(Y_{kj} - \bar{Y}_j) - s_{ij2}\}^2 \to \theta_{ij2}$
- Apply multi-level thresholding procedure on

$$M_{ij} = \frac{(s_{ij1} - s_{ij2})^2}{\hat{\theta}_{ij1}/n_1 + \hat{\theta}_{ij2}/n_2}, \quad 1 \le i \le j \le p.$$

Simulation Results

Left: $\beta = 0.6$; Right: $r = 0.6$



Signal detection in precision matrices

- Precision matrix: $\Omega = \Sigma^{-1} = (\omega_{j_1 j_2})$
- Gaussian graphical model (GGM)
- Target variables: $\mathcal{A}_1 \subset \{1, \ldots, p\}$

$$\begin{aligned} &H_0 : \omega_{j_1 j_2} = 0 \text{ for all } j_1, j_2 \in \mathcal{A}_1 \text{ and } j_1 \neq j_2 \quad \text{vs.} \\ &H_a : \omega_{j_1 j_2} \neq 0 \text{ for some } j_1, j_2 \in \mathcal{A}_1 \text{ and } j_1 \neq j_2. \end{aligned} \tag{4}$$

- Applied to evaluate spatial dependence structure using "concliques"

Precision matrix class under alternative hypothesis

- Sample size n, $p_1 = |\mathcal{A}|_1$, and $q_1 = p_1(p_1 - 1)/2$
- Number of signals: $m_a = \sum_{(j_1,j_2)\in\mathcal{A}_1, j_1<j_2} \mathbb{I}(\omega_{j_1j_2} \neq 0) = \lfloor q_1^{(1-\beta)} \rfloor$
- $\beta \in (1/2, 1)$
- Nonzero $\omega_{j_1j_2}$: $\omega_{j_1j_2} = \sqrt{2 r_{j_1j_2} \log(q_1)/n}$ if $\omega_{j_1j_2} \neq 0$;

$$\mathcal{U}(\beta, r_0, \tau) = \Big\{\Omega: \; m_a \geq \lfloor q_1^{(1-\beta)} \rfloor \text{ nonzero } \omega_{j_1j_2} \text{ with } r_{j_1j_2} \geq r_0, \\ \max_{j\in\mathcal{A}_1} \omega_{jj} \leq \tau, \; \min_{j\in\mathcal{A}_1} \omega_{jj} \geq C^{-1}\Big\}.$$

Detection boundary for precision matrices

Theorem (Sharp minimax result)

Under $n \asymp p^{\xi}$ for $\xi \in (0,1]$ and Gaussian distributed data, if $r_0\tau^{-2} < \mathrm{DB}(\beta)$ and $3/4 - \xi/4 < \beta < 1$, we have, as $n, p \to \infty$,

$$\sup_{W \in \mathcal{W}_\alpha} \inf_{\Omega \in \mathcal{U}(\beta, r_0, \tau)} \mathbb{P}(W = 1) \le \alpha$$

- $\mathcal{W}_{1,\alpha}$: all α-level tests;
- requires $3/4 - \xi/4 < \beta < 1$;
- if $n \asymp p$, this covers the entire sparse regime $1/2 < \beta < 1$;
- $r_0\tau^{-2}$: minimum standardized signal strength

Estimation of Ω–node-wise regression

$$X_{ij_1} - \mu_{j_1} = \sum_{j_2 \neq j_1} \beta_{j_1 j_2}(X_{ij_2} - \mu_{j_2}) + \epsilon_{ij_1}$$

- $\epsilon_{ij_1} \perp X_{ij_2}$ for $j_2 \neq j_1$ if and only if $\beta_{j_1 j_2} = -\omega_{j_1 j_2}/\omega_{j_1 j_1}$;
- error covariance satisfies $\mathrm{Cov}(\epsilon_{ij_1}, \epsilon_{ij_2}) = \omega_{j_1 j_2}/(\omega_{j_1 j_1}\omega_{j_2 j_2})$;
- $\omega_{j_1 j_2}$ can be estimated by the residuals from node-wise regressions

Estimation of Ω–Liu (2013)

$$X_{ij_1} - \mu_{j_1} = \sum_{j_2 \neq j_1} \beta_{j_1 j_2}(X_{ij_2} - \mu_{j_2}) + \epsilon_{ij_1}$$

- Lasso estimates: $\widehat{\boldsymbol{\beta}}_j = \underset{\boldsymbol{\beta} \in \mathbb{R}^p, \beta_{jj} = -1}{\operatorname{argmin}} \left[\frac{1}{n} \sum_{i=1}^{n} \{\boldsymbol{\beta}^{\mathrm{T}}(\mathrm{X}_i - \bar{\mathrm{X}})\}^2 + 2\lambda_j |\boldsymbol{\beta}|_1 \right]$;
- set $\widehat{\boldsymbol{\beta}}_{jj} = -1$ for notation simplicity;
- residuals: $\widehat{\epsilon}_{ij} = -\widehat{\boldsymbol{\beta}}_j^{\mathrm{T}}(\mathrm{X}_i - \bar{\mathrm{X}})$;

$$\widehat{\omega}_{j_1 j_2} = \frac{\widehat{v}_{j_1 j_2}}{\widehat{v}_{j_1 j_1} \widehat{v}_{j_2 j_2}} \quad \text{where} \quad \widehat{v}_{j_1 j_2} = -\frac{1}{n} \sum_{i=1}^{n} (\widehat{\epsilon}_{ij_1} \widehat{\epsilon}_{ij_2} + \widehat{\beta}_{j_1 j_2} \widehat{\epsilon}_{ij_2}^2 + \widehat{\beta}_{j_2 j_1} \widehat{\epsilon}_{ij_1}^2).$$

MTT for precision matrices

- $\widehat{\psi}_{j_1 j_2} = \widehat{v}_{j_1 j_2}(\widehat{v}_{j_1 j_1}\widehat{v}_{j_2 j_2})^{-1/2}$: $-\widehat{\psi}_{j_1 j_2}$ is the estimated partial correlation;
- Standardized statistic for testing $\omega_{j_1 j_2} = 0$: $V_{j_1 j_2} = n\widehat{\psi}^2_{j_1 j_2}$;
- Thresholding statistic: $T(t) = \sum_{j_1, j_2 \in \mathcal{A}_1, j_1 < j_2} V_{j_1 j_2} \mathbb{I}\{V_{j_1 j_2} \geq \lambda_q(t)\}$;
- Derive large deviation results for $\widehat{\psi}_{j_1 j_2}$;
- MTT procedure on $T(t)$.

Detection boundary of MTT for precision matrices

$$\mathrm{DB}(\beta, \xi/\gamma, \nu) = \begin{cases} \frac{\left\{\sqrt{8-8c(\xi/\gamma,\nu)}-\sqrt{8-8\beta-4c(\xi/\gamma,\nu)}\right\}^2}{8}, & 1/2 < \beta \le 3/4 - c(\xi/\gamma,\nu)/4, \\ \beta - 1/2, & 3/4 - c(\xi/\gamma,\nu)/4 < \beta \le 3/4, \\ (1-\sqrt{1-\beta})^2, & 3/4 < \beta < 1, \end{cases}$$

- $c(\xi/\gamma, \nu) = \min\{(\xi/\gamma)(1/2 - \nu), 1\}$;
- $\xi \in (0, 1]$: $n \asymp p^{\xi}$;
- $\gamma \in (0, 1]$: $p_1 = |\mathcal{A}_1| \asymp p^{\gamma}$;
- $\nu \in [0, 1/2)$: $\max_{1 \le j_1 \le p} \sum_{j_2=1}^{p} \mathbb{I}(\omega_{j_1 j_2} \ne 0) \le C n^{\nu}$.

Detection boundary of MTT for precision matrices

$$\mathrm{DB}(\beta,\xi/\gamma,\nu)=\begin{cases}\frac{\left\{\sqrt{8-8c(\xi/\gamma,\nu)}-\sqrt{8-8\beta-4c(\xi/\gamma,\nu)}\right\}^2}{8}, & 1/2<\beta\le 3/4-c(\xi/\gamma,\nu)/4,\\ \beta-1/2, & 3/4-c(\xi/\gamma,\nu)/4<\beta\le 3/4,\\ (1-\sqrt{1-\beta})^2, & 3/4<\beta<1,\end{cases}$$

- Power of MTT$\to 1$ if $r_0\tau^{-2}>\mathrm{DB}(\beta,\xi/\gamma,\nu)$;
- $\mathrm{DB}(\beta,\xi/\gamma,\nu)=\mathrm{DB}(\beta)$ if $\beta>3/4-c(\xi/\gamma,\nu)/4$;
- $\mathrm{DB}(\beta,\xi/\gamma,\nu)>\mathrm{DB}(\beta)$ if $\beta<3/4-c(\xi/\gamma,\nu)/4$.

Conclusion

- Tight minimax result for testing covariance and precision matrices.
- Sharp optimal multi-level thresholding test.
- More powerful tests for sparse and weak signals.

Reference

Chen, S. X., Guo, B. and Qiu, Y. (2023). Global Test and Signal Identification for Two-sample Covariances via Multi-level Thresholding. *Journal of Econometrics*.

Chen, S. X., Qiu, Y. and Zhang, S. (2023). Sharp Optimality for High Dimensional Covariance Testing under Sparse Signals. *The Annals of Statistics*.

Qiu, Y. and Guo, B. (2023). Minimax Detection Boundary and Sharp Optimal Test for Gaussian Graphical Models. *Manuscript*