Hypothesis Testing on Linear Structures of High Dimensional Covariance Matrix

Runze Li

The Pennsylvania State University

Joint work with Shurong Zheng, Zhao Chen and Hengjian Cui

Introduction

Impact of dimensionality on test for high-dimensional data

Bai and Saranadasa (1996) demonstrated the effect of dimensionality for test of two sample high-dimensional (HD) normal means. Hotelling's T^2 has very low power when $y_n = p/n \to y \in (0, 1)$.

Replacing inverse of sample covariance matrix in Hotelling T^2 test by the identify matrix may yield a more powerful test than T^2.

Test for one sample mean problem with $p > n$ goes back to Dempster (1958, 1960);

Test of HD one/two sample mean problems becomes very active now.

Test of hypotheses in HD data analysis

Since Bai and Saranadasa (1996), there are many publications on test on HD one/two sample mean problems:

Srivastava & Du (2008); Srivastava (2009); Chen & Qin (2010); Lee, et al (2012); Thulin (2014); Wang, Peng & Li (2015); Gregory, Carroll, Baladandayuthapani and Lahiri (2015)

Chen, Paul, Prentice & Wang (2011);

Lauter (1996); Lauter, et al. (1998); Lopes, Jacob & Wainwright (2011, 2012); Wei, Lee, Wichers, Li & Marron (2016); Srivastava, Li & Ruppert (2016).

Test of covariance structure

Test of covariance structure is of great interest in classical multivariate data analysis:

◦ Chapters 8 and 11 of Muirhead (1982)

◦ Chapters 9 and 10 of Anderson (2003)

Test on HD covariance structure

Early works include Ledoit & Wolf (2002), Srivastava (2005), Birke and Dette (2005), Schott (2007)

Large dimensional random matrix theory were used to construct tests of covariance structure:

(a). $H_0 : \Sigma = \Sigma_0$ (known) was studied in Bai, Jiang, Yao and Zheng (2009) proposed correcting LRT with $p/n \to y \in (0, 1)$

(b). Test of sphericity: $H_0 : \Sigma = \sigma^2 I_p$ was studied in Chen, Zhang and Zhong (2010), Wang and Yao (2013), Jiang and Yang (2013)

(c). Test of bandedness was studied in Qiu and Chen (2012).

Under **normality assumption**, Jiang and Yang (2013) and Jiang and Qi (2015) obtained the limiting distribution of LRT for tests studied in Ch. 8 of Muirhead (1982) with $p < n$ and $p/n \to y \in (0, 1]$. Cai and Ma (2013) proposed a further corrected LRT that allows $p/n \to \infty$.

Today's goal

Develop two tests on linear covariance structure for HD data.

For a set of pre-specified symmetric $p \times p$ matrices $(\mathbf{A}_1, \ldots, \mathbf{A}_K)$ with fixed and finite K,

$$H_0 : \mathbf{\Sigma} = \theta_1 \mathbf{A}_1 + \theta_2 \mathbf{A}_2 + \ldots + \theta_K \mathbf{A}_K,$$

where $\{\theta_j, j = 1, \ldots, K\}$ are unknown parameters

Statistical Setting

- *Assumption A.* Assume that the population $\mathbf{X}$ can be represented as $\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{W}$, where $\mathbf{W} = (w_1, \cdots, w_p)^T$, and $w_1, \cdots, w_p$ being IID and $E(w_j) = 0$, $E(w_j^2) = 1$ and $E(w_j^4) = \kappa < \infty$.
- *Assumption B.* Denote $y_n = p/n$. Assume that $y_n \to y \in (0, \infty)$.

Assumption A relaxes the normality assumption by imposing the moment condition and is referred to as **independent component model** (ICM). This is a typical assumption in random matrix theories.

Regarding to representation in Assumption A, it is natural to assume that w_j is standardized so that $E(w_j) = 0$ and $E(w_j^2) = 1$. If w_j has finite kurtosis, then Assumption A is satisfied.

Many distributions including normal satisfy Assumption A. Multinormal is only class belonging to both ICM and elliptical distribution.

Parameter estimation

Without a likelihood, we propose estimating $\boldsymbol{\theta}$ by minimizing the following squared loss function

$$\min_{\boldsymbol{\theta}} \mathrm{tr}(\mathbf{S}_n - \theta_1 \mathbf{A}_1 - \ldots - \theta_K \mathbf{A}_K)^2, \tag{1}$$

where $\mathbf{S}_n$ is the sample covariance matrix.
Let $\mathbf{C}$ be a $K \times K$ matrix with (i, j)-element being $\mathrm{tr}\mathbf{A}_i\mathbf{A}_j$ and $\mathbf{a}$ be a $K \times 1$ vector with j-th element being $\mathrm{tr}\mathbf{S}_n\mathbf{A}_j$. Minimizing (1) yields a least squares type estimate for $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}} = \mathbf{C}^{-1}\mathbf{a}. \tag{2}$$

We can show that under Assumptions A and B, $\hat{\theta}_k = \theta_k + O_p(n^{-1})$, $k = 1, \cdots, K$.

Tests

Denote $\mathbf{\Sigma}_0 = \theta_1 \mathbf{A}_1 + \theta_2 \mathbf{A}_2 + \ldots + \theta_K \mathbf{A}_K$.

$$H_0 : \mathbf{\Sigma} = \mathbf{\Sigma}_0.$$

Under H_0, we estimate $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}$ given in (2), and then an estimator of $\mathbf{\Sigma}$ is $\widehat{\mathbf{\Sigma}}_0 = \hat{\theta}_1 \mathbf{A}_1 + \ldots + \hat{\theta}_K \mathbf{A}_K$.

Without the linear structure assumption, a natural estimator for $\mathbf{\Sigma}$ is the sample covariance matrix $\mathbf{S}_n$.

Entropy loss based test

Motivated by the entropy loss (EL) used for covariance matrix estimation (James and Stein, 1961; Muirhead, 1982), we propose our first test for H_0. For $p < n-1$,

$$T_{n1} = \mathrm{tr}\mathbf{S}_n\widehat{\boldsymbol{\Sigma}}_0^{-1} - \log(|\mathbf{S}_n\widehat{\boldsymbol{\Sigma}}_0^{-1}|).$$

Denote $\{\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p\}$ to be the eigenvalues of $\widehat{\boldsymbol{\Sigma}}_0^{-1/2}\mathbf{S}_n\widehat{\boldsymbol{\Sigma}}_0^{-1/2}$. Then we can write T_{n1} as

$$T_{n1} = p\left(p^{-1}\sum_{j=1}^{p}\lambda_j - p^{-1}\sum_{j=1}^{p}\log\lambda_j\right).$$

This motivates us to further extend the test to the situation with $p \geq n-1$ by defining

$$T_{n1} = (n-1)\left(p^{-1}\sum_{j=1}^{n-1}\lambda_j - (n-1)^{-1}\sum_{j=1}^{n-1}\log\lambda_j\right).$$

Entropy loss based test

Define $q = \min\{p, n-1\}$. T_{n1} can be written in a unified form for $p < n-1$ and $p \geq n-1$:

$$T_{n1} = q\left(p^{-1}\sum_{j=1}^{q}\lambda_j - q^{-1}\sum_{j=1}^{q}\log\lambda_j\right). \tag{3}$$

Since this test is motivated by the entropy loss, we refer this test to be as **EL-test**.

Quadratic loss based test

Motivated by the quadratic loss (QL), another popular loss function in covariance matrix estimation (Olkin and Selliah, 1977; Haff, 1980; Muirhead, 1982), we propose the second test

$$T_{n2} = \mathrm{tr}(\mathbf{S}_n \widehat{\boldsymbol{\Sigma}}_0^{-1} - \mathbf{I}_p)^2, \tag{4}$$

and refer this test to be as **QL-test**.

Example 1. Test of sphericity: $H_{10} : \Sigma = \sigma^2 I_p$

Chen, Zhang & Zhong (2010) demonstrated that the classical LRT may become invalid for HD data and proposed a test based on U-statistics with $p, n \to \infty$.

Jiang & Yang (2013) derived the asymptotic distribution of the LRT with $p/n \to (0, 1]$ under normality assumption on $\mathbf{W}$.

Wang & Yao (2013) proposed the corrected LRT with $p/n \to (0, 1)$ and the corrected John's (CJ) test with $p/n \to y \in (0, \infty)$

The CJ test behaves similarly with the proposal by Chen, Zhang & Zhong (2010) on powers as $p/n \to y \in (0, \infty)$ and the corrected LRT had greater powers than the CJ test and Chen, Zhang & Zhong (2010)'s test when the dimension p is not large relative to the sample size n. But when p is large relative to n ($p < n$), the corrected LRT had less powers than the CJ test and the test proposed in Chen, Zhang & Zhong (2010).

Example 1. Test of sphericity: $H_{10} : \Sigma = \sigma^2 I_p$

Under H_{10}, σ^2 can be estimated by $\hat{\sigma}^2 = p^{-1}\mathrm{tr}\mathbf{S}_n$. It is easy to see that $\hat{\sigma}^2 = \sigma^2 + O_p(n^{-1})$ under the condition that $p/n \to y \in (0, \infty)$. Let $\{\lambda_1 \geq \lambda_2 \ldots \geq \lambda_p\}$ be the eigenvalues of $\mathbf{S}_n/(p^{-1}\mathrm{tr}\mathbf{S}_n)$. The EL and QL tests have the following expressive forms:

$$T_{n1} = q\left(p^{-1}\sum_{j=1}^{q}\lambda_j - q^{-1}\sum_{j=1}^{q}\log\lambda_j\right), \tag{5}$$

where $q = \min\{p, n-1\}$, and

$$T_{n2} = \mathrm{tr}[\mathbf{S}_n/(p^{-1}\mathrm{tr}\mathbf{S}_n) - \mathbf{I}_p]^2. \tag{6}$$

T_{n1} is equivalent to the LRT under normality assumption and $p < n-1$. T_{n2} in (6) coincides with the CJ test proposed by Wang & Yao (2013).

Example 2: Compound symmetric structure

Let $\mathbf{A}_1 = \mathbf{I}_p$ and $\mathbf{A}_2 = \mathbf{1}_p\mathbf{1}_p^T$, where $\mathbf{1}_p$ stands for a p-dimensional column vectors with all elements being 1. Testing compound symmetric structure is to test

$$H_{20} : \mathbf{\Sigma} = \theta_1 \mathbf{A}_1 + \theta_2 \mathbf{A}_2,$$

where $\theta_1 > 0$ and $-1/(p-1) < \theta_2/(\theta_1 + \theta_2) < 1$.

Under normality assumption, Kato, Yamada and Fujikoshi (2010) studied the asymptotic behavior of the corresponding likelihood ratio test when $p < n$, and Srivastava and Reid (2012) proposed a new test statistic for H_{20} when $p \geq n$.

Example 2: Compound symmetric structure

Under H_{20},

$$\hat{\theta}_1 = p^{-1}(p-1)^{-1}(p\,\mathrm{tr}\mathbf{S}_n - \mathbf{1}_p^T\mathbf{S}_n\mathbf{1}_p), \quad \hat{\theta}_2 = p^{-1}(p-1)^{-1}(\mathbf{1}_p^T\mathbf{S}_n\mathbf{1}_p - \mathrm{tr}\mathbf{S}_n).$$

respectively. It can be shown that $\hat{\theta}_k = \theta_k + O_p(n^{-1})$, $k = 1, 2$

T_{n1} is equivalent to the LRT in Kato, Yamada and Fujikoshi (2010) under normality assumption when $p < n - 1$.

Srivastava and Reid (2012) recast this test problem to test of independence under normality. So it is different from the LRT, T_{n1} and T_{n2}.

Example 3: Testing bandedness structure

Let $\mathbf{A}_1 = \mathbf{I}_p$ and for $2 \le k \le K$, $\mathbf{A}_k$ to be a $p \times p$ matrix with (i, j)-element being 1 if $|i - j| = k - 1$ and 0 otherwise. Testing the $(K-1)$-banded covariance matrix is to test

$$H_{30} : \mathbf{\Sigma} = \theta_1 \mathbf{A}_1 + \cdots + \theta_K \mathbf{A}_K,$$

where θ_k's are unknown. By (2), we have

$$\begin{aligned}\hat{\theta}_1 &= p^{-1} \mathrm{tr} \mathbf{S}_n \\ \hat{\theta}_k &= \frac{1}{2}(p - k + 1)^{-1} \mathrm{tr} \mathbf{S}_n \mathbf{A}_k, \quad \text{for} \quad 2 \le k \le K.\end{aligned}$$

Under Assumptions A and C, $\hat{\theta}_k = \theta_k + O_p(n^{-1})$ if $p/n \to y \in (0, \infty)$ when K is a finite.

Qiu and Chen (2012) proposed a U-statistic test for the banded structure and is different from T_{n1} and T_{n2}.

Example 4: Factor model

A factor model assumes $\mathbf{X} = v_1\mathbf{U}_1 + \cdots + v_{K-1}\mathbf{U}_{K-1} + \epsilon$, where $v_1, \cdots, v_{K-1}$ are random variables and $\mathbf{U}_k$, $k = 1, \cdots, K-1$ are random vectors.

Suppose that $v_1, \cdots, v_{K-1}$, $\mathbf{U}_1, \ldots, \mathbf{U}_{K-1}$ and ϵ are mutually independent. Suppose that $\mathrm{Cov}(\epsilon) = \theta_1\mathbf{I}_p$ and conditioning on $\mathbf{U}_k$, $k = 1, \cdots, K-1$, the factor model has a covariance structure $\mathbf{\Sigma} = \theta_1\mathbf{I}_p + \theta_2\mathbf{U}_1\mathbf{U}_1^T + \cdots + \theta_K\mathbf{U}_{K-1}\mathbf{U}_{K-1}^T$, where $\theta_{k+1} = \mathrm{Var}(v_k)$.

Let $\mathbf{A}_1 = I_p$, and $\mathbf{A}_{k+1} = \mathbf{U}_k\mathbf{U}_k^T$ for $k = 1, \cdots, K-1$. Thus, it is of interest to test

$$H_{40} : \mathbf{\Sigma} = \theta_1\mathbf{A}_1 + \cdots + \theta_K\mathbf{A}_K,$$

where θ_k's are unknown parameters.

Example 4: Factor model

In practice one typically sets $\mathbf{U}_k$ to be orthogonal to each other and have been standardized so that $\mathbf{U}_s^T\mathbf{U}_t = p$ for $s = t$ and 0 for $s \neq t$. The parameters θ_k can be estimated by

$$\hat{\theta}_1 = (p - K + 1)^{-1}\left(\mathrm{tr}\mathbf{S}_n - p^{-1}\sum_{k=1}^{K-1}\mathbf{U}_k^T\mathbf{S}_n\mathbf{U}_k\right)$$

$$\hat{\theta}_{k+1} = p^{-2}(\mathbf{U}_k^T\mathbf{S}_n\mathbf{U}_k - p\hat{\theta}_1)$$

for $k = 1, \cdots, K-1$. Thus, when K is finite and p/n has a finite positive limit, then $\hat{\theta}_k = \theta_k + O_p(n^{-1})$ under H_{40}. Then testing H_{40} can be carried out by using the EL and QL tests.

Example 5: Test a particular pattern (McDonald, 1974)

For even p which is finite and fixed, McDonald (1974) considered

$$H_{50}: \mathbf{\Sigma} = \begin{pmatrix} \theta_1 \mathbf{I}_{p/2} + \theta_2 \mathbf{1}_{p/2}\mathbf{1}_{p/2}^T & \theta_3 \mathbf{I}_{p/2} \\ \theta_3 \mathbf{I}_{p/2} & \theta_1 \mathbf{I}_{p/2} + \theta_2 \mathbf{1}_{p/2}\mathbf{1}_{p/2}^T \end{pmatrix}.$$

Let $\mathbf{A}_1 = \mathbf{I}_p$,

$$\mathbf{A}_2 = \begin{pmatrix} \mathbf{1}_{p/2}\mathbf{1}_{p/2}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{p/2}\mathbf{1}_{p/2}^T \end{pmatrix} \quad \text{and} \quad \mathbf{A}_3 = \begin{pmatrix} \mathbf{0} & \mathbf{I}_{p/2} \\ \mathbf{I}_{p/2} & \mathbf{0} \end{pmatrix}.$$

Then H_{50} can be written as $H_{50}: \mathbf{\Sigma} = \theta_1 \mathbf{A}_1 + \theta_2 \mathbf{A}_2 + \theta_3 \mathbf{A}_3$. Thus, the proposed EL and QL test procedure can be used to test H_{50} with high-dimensional data.

Limiting distributions of tests

To derive the limiting distribution of T_{n1} and T_{n2}, we first derive some theoretical results on linear functionals of $\mathbf{S}_n$.

Theorem 1.

Under H_0 and Assumptions A and B, it follows that

(a) For $p < n-1$,

$$\frac{T_{n1} + (p-n+1)\log(1-y_{n-1}) - 2p + \alpha_1(y_{n-1})}{\sqrt{\alpha_2(y_{n-1})}} \to N(0,1), \qquad (7)$$

where $\alpha_1(y) = 0.5\log(1-y) - 0.5(\kappa-3)y$ and $\alpha_2(y) = -2y - 2\log(1-y)$;

Theorem 1

(b) For $p \geq (n-1)$,

$$\frac{T_{n1} - p[y_{n-1}^{-1} - \alpha_3(y_{n-1})] - y_{n-1}^{-1} m_1(y_{n-1}) + m_2(y_{n-1})}{\sqrt{y^{-2} v_{11}(y_{n-1}) + v_{22}(y_{n-1}) - 2y_{n-1}^{-1} v_{12}(y_{n-1})}} \to N(0,1) \quad (8)$$

where $\alpha_3(y)$, $m_1(y)$, $m_2(y)$, $v_{11}(y)$, $v_{22}(y)$, $v_{12}(y)$ are given in the paper.
(c) $\frac{1}{2}[T_{n2} - py_{n-1} - (\kappa-2)y]/\sigma \to N(0,1)$, where

$$\sigma^2 = y_{n-1}^2 + 2y_{n-1}^3 p^{-1} \mathrm{tr}(\mathbf{\Sigma}_0 \mathbf{B})^2 + (\kappa-3) y_{n-1}^3 p^{-1} \sum_{i=1}^{p} (\mathbf{e}_i^T \mathbf{\Gamma}^T \mathbf{B} \mathbf{\Gamma} \mathbf{e}_i)^2 - (\kappa-1) y_{n-}^3$$

with $\mathbf{B} = \sum_{k=1}^K d_k \mathbf{A}_k$, $\mathbf{d} = (d_1, \cdots, d_K)^T = \mathbf{C}^{-1}\mathbf{c}$ and $\mathbf{c}$ being a K-dimensional column vector with k-th element being $\mathrm{tr}\mathbf{A}_k \mathbf{\Sigma}_0^{-1}$.

The limiting null distribution can be used to construct the rejection regions for T_{n1} and T_{n2}.

We next study the asymptotic power of these two tests. Define

$$\begin{pmatrix} \theta_1^* \\ \theta_2^* \\ \dots \\ \theta_K^* \end{pmatrix} = \begin{pmatrix} \mathrm{tr}\mathbf{A}_1^2 & \mathrm{tr}\mathbf{A}_2\mathbf{A}_1 & \dots & \mathrm{tr}\mathbf{A}_K\mathbf{A}_1 \\ \mathrm{tr}\mathbf{A}_1\mathbf{A}_2 & \mathrm{tr}\mathbf{A}_2^2 & \dots & \mathrm{tr}\mathbf{A}_K\mathbf{A}_2 \\ \dots & \dots & \dots & \dots \\ \mathrm{tr}\mathbf{A}_1\mathbf{A}_K & \mathrm{tr}\mathbf{A}_2\mathbf{A}_K & \dots & \mathrm{tr}\mathbf{A}_K^2 \end{pmatrix}^{-1} \begin{pmatrix} \mathrm{tr}\boldsymbol{\Sigma}\mathbf{A}_1 \\ \mathrm{tr}\boldsymbol{\Sigma}\mathbf{A}_2 \\ \dots \\ \mathrm{tr}\boldsymbol{\Sigma}\mathbf{A}_K \end{pmatrix}$$

and $\boldsymbol{\Sigma}_{0*} = \theta_1^*\mathbf{A}_1 + \ldots + \theta_K^*\mathbf{A}_K$. Under $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, it follows that $\boldsymbol{\Sigma}_{0*} = \boldsymbol{\Sigma}_0$. Under Assumptions A and B, it can be shown that $\hat{\theta}_k = \theta_k^* + O_p(n^{-1}), k = 1, \ldots, K$.

Let $G_p(t) = p^{-1}\sum_{j=1}^p I(\lambda_j \le t)$ be the ESD of $\boldsymbol{\Sigma}_{0*}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}_{0*}^{-1/2}$.

If $G_p(t) \to G(t)$, in which $G(t)$ is assumed not to be degenerated to a single point distribution. That is, $\boldsymbol{\Sigma} \neq \tau\boldsymbol{\Sigma}_{0*}$ for some constant $\tau > 0$. Under this condition, we have the limiting distributions of T_{n1} and T_{n2}.

Limiting distributions under alternative

Theorem 2

Under Assumptions A and B, and under $H_1 : G_p(t) \to G(t)$, a non-degenerated distribution, it follows that

$$\frac{T_{n1} - pF_2^{y_{n-1},G} - \mu_2^{(1)}}{\sigma_{2n}^{(1)}} \to N(0,1), \quad \text{for} \quad p < n-1$$

$$\frac{T_{n1} - pF_3^{y_{n-1},G} - \mu_3^{(1)}}{\sigma_{3n}^{(1)}} \to N(0,1), \quad \text{for} \quad p \geq n-1$$

$$\frac{T_{n2} - \mu_1^{(1)}}{\sigma_{1n}^{(1)}} \to N(0,1),$$

where $\mu_j^{(1)}, j = 1,2,3$, $F_j^{y_{n-1},G}, j = 2,3$ and $\sigma_{jn}^{(1)}, j = 1,2,3$ are given in the proof of Theorem .

Power function of T_{n2} and unbiased test

For a level α, the power function of T_{n2} is
$1-\Phi((\mu_0-\mu_1^{(1)})/\sigma_{1n}^{(1)}-2q_{\alpha/2}\sigma/\sigma_{1n}^{(1)})+\Phi((\mu_0-\mu_1^{(1)})/\sigma_{1n}^{(1)}+2q_{\alpha/2}\sigma/\sigma_{1n}^{(1)})$,
where $q_{\alpha/2}$ is the $\alpha/2$ quantile of $N(0,1)$ and $\mu_0 = py_{n-1}+(\kappa-2)y$.

Theorem 3

Suppose that Assumptions A and B are satisfied and the limit of $\sigma_{1n}^{(1)}$ exists. If $\boldsymbol{\Gamma}^T\boldsymbol{\Sigma}_{0*}^{-1}\boldsymbol{\Gamma} = \mathbf{I}_p + \mathbf{A}$ whose ESD weakly converges and $\mathrm{tr}\mathbf{A}^2 > \delta > 0$, then we have

$$\beta_{T_{n2}} > \alpha$$

when n is large enough and δ is any given small constant.

Thus, T_{n2} is an asymptotically unbiased test. Furthermore, if $p^{-1}\mathrm{tr}\mathbf{A} \to c_1 \neq 0$, then $\beta_{T_{n2}} \to 1$ when $n \to \infty$.

Estimation of κ: Under Assumption A, we construct an consistent estimate for κ by using

$$E\{\mathbf{W}^T\Sigma\mathbf{W} - \mathrm{tr}(\Sigma)\}^2 = 2\mathrm{tr}(\Sigma^2) + (\kappa - 3)\sum_{j=1}^{p}\sigma_{jj}^2.$$

Our simulation shows that we may get more stable estimate for κ by using the linear structure under H_0.

Numerical Examples

In our simulation, $\mathbf{X} = \mathbf{\Sigma}^{1/2}\mathbf{W}$, where $\mathbf{\Sigma}$ will be set according to the hypothesis to be tested.

In order to examine the performances of the proposed tests under different distributions, we consider the elements of $\mathbf{W}$ being independent and identically distributed as (a) $N(0, 1)$ or (b) Gamma(4,2)-2. The two distributions are both normalized so that their means 0 and variances 1.

For each setting, we conduct 1000 Monte Carlo simulations. The Monte Carlo simulation error rate is $1.96\sqrt{0.05 \times 0.95/1000} \approx 0.0135$ at level 0.05.

Test of compound symmetry $H_{20} : \mathbf{\Sigma} = \theta_1 \mathbf{I}_p + \theta_2 \mathbf{1}_p \mathbf{1}_p^T$

We set $\mathbf{\Sigma} = \theta_1 \mathbf{I}_p + \theta_2 \mathbf{1}_p \mathbf{1}_p^T + \theta_3 \mathbf{u}_p \mathbf{u}_p^T$, where $\mathbf{u}_p$ is a p-dimensional random vector following uniform distribution over $[-1, 1]$.

The third term is to examine the empirical powers when $\theta_3 \neq 0$. In our simulation, we set $(\theta_1, \theta_2) = (6, 1)$ and $\theta_3 = 0, 0.5, 1$, respectively.

We set $\theta_3 = 0$ to examine Type I error rate and $\theta_3 = 0.5, 1$ to study the empirical power of the proposed tests. The sample size is set as $n = 100, 200$ and the dimension is taken to be $p = 50, 100, 500, 1000$.

Compare the performances of proposed testing procedures and the test proposed in Srivastava and Reid (2012)

Table: Empirical power for H_{20} (in percentage) with $n = 100$

		$W_j \sim N(0,1)$				$W_j \sim$ Gamma(4,2)−2			
θ_3	Test	p =50	100	500	1000	50	100	500	1000
0	QL	5.23	5.40	5.12	5.19	6.48	5.99	5.64	5.41
	EL	5.32	6.51	5.12	5.18	5.77	6.35	5.46	5.54
	SR	4.90	5.01	4.91	4.98	9.60	8.96	8.15	8.04
0.5	QL	40.25	80.58	100.00	100.00	41.01	80.70	100.00	100.00
	EL	13.42	11.38	99.78	100.00	13.74	11.42	99.74	100.00
	SR	24.46	59.04	99.99	100.00	41.22	73.86	100.00	100.00
1.0	QL	95.88	99.97	100.00	100.00	95.98	99.97	100.00	100.00
	EL	53.53	29.97	100.00	100.00	53.71	30.18	100.00	100.00
	SR	87.90	99.67	100.00	100.00	93.61	99.87	100.00	100.00

Table: Empirical power for H_{20} (in percentage) with $n = 200$

		$W_j \sim N(0,1)$				$W_j \sim$ Gamma(4,2)−2			
θ_3	Test	p =50	100	500	1000	50	100	500	1000
0	QL	5.22	5.14	5.12	5.19	6.32	5.78	5.31	5.34
	EL	5.18	5.12	5.05	5.13	5.94	5.42	5.23	5.31
	SR	4.98	4.93	4.93	5.03	9.95	9.23	8.44	8.43
0.5	QL	79.86	99.32	100.00	100.00	78.56	99.28	100.00	100.00
	EL	42.00	58.62	100.00	100.00	41.22	58.62	100.00	100.00
	SR	61.74	95.79	100.00	100.00	75.78	98.24	100.00	100.00
1.0	QL	99.98	100.00	100.00	100.00	99.96	100.00	100.00	100.00
	EL	97.23	99.53	100.00	100.00	96.86	99.55	100.00	100.00
	SR	99.81	100.00	100.00	100.00	99.91	100.00	100.00	100.00

Test covariance matrix structure in H_{30}

We construct a banded matrix defined in Example 3 with width of band $K = 3$. Therefore, the null hypothesis H_{30} has the linear decomposition $\mathbf{\Sigma} = \theta_1 \mathbf{I}_p + \theta_2 \mathbf{A}_2 + \theta_3 \mathbf{A}_3 + \theta_4 \mathbf{u}_p \mathbf{u_p}^T$, where $\mathbf{A}_2$ and $\mathbf{A}_3$ are defined in Example 3 and $\mathbf{u}_p$ is generated by the same way in the last example.

We take $(\theta_1, \theta_2, \theta_3) = (6, 1, 0.5)$ and $\theta_4 = 0$ to examine Type I error rates and take $\theta_4 = 0.5, 1$ to examine powers. In the simulation studies, we still set sample size $n = 100, 200$ and dimension $p = 50, 100, 500, 1000$.

In this example, we compare the test proposed in Qiu and Chen (2012) for the banded covariance matrix with our proposed tests, and referred their test as "QC" test hereinafter.

Table: Simulation Results for H_{30} (in percentage)

			$W_j \sim N(0,1)$				$W_j \sim$ Gamma(4,2)−2			
θ_4	n	Test	p =50	100	500	1000	50	100	500	1000
0	100	QL	5.30	5.19	5.29	5.34	6.61	6.19	5.83	5.89
		EL	5.31	6.34	5.20	5.36	5.83	6.32	5.53	5.81
		QC	5.00	5.50	5.50	5.90	5.00	5.80	6.10	5.20
0.5	100	QL	46.01	84.91	100.00	100.00	45.13	83.62	100.00	100.00
		EL	15.20	12.23	99.90	100.00	15.35	12.13	99.90	100.00
		QC	17.00	70.00	100.00	100.00	24.00	85.30	100.00	100.00
1.0	100	QL	97.49	99.98	100.00	100.00	96.96	99.98	100.00	100.00
		EL	60.45	33.68	100.00	100.00	59.80	33.71	100.00	100.00
		QC	83.00	100.00	100.00	100.00	78.00	99.90	100.00	100.00

Table: Simulation Results for H_{30} (in percentage)

			$W_j \sim N(0,1)$				$W_j \sim$ Gamma(4,2)-2			
θ_4	n	Test	p =50	100	500	1000	50	100	500	1000
0	200	QL	5.25	5.19	5.16	5.10	6.44	5.84	5.56	5.55
		EL	5.24	5.15	5.04	4.94	6.01	5.46	5.25	5.26
		QC	6.00	5.40	5.50	4.50	4.00	4.50	5.80	5.10
0.5	200	QL	85.30	99.66	100.00	100.00	84.09	99.64	100.00	100.00
		EL	48.46	65.39	100.00	100.00	47.62	65.20	100.00	100.00
		QC	49.00	100.00	100.00	100.00	49.00	99.70	100.00	100.00
1.0	200	QL	99.99	100.00	100.00	100.00	99.98	100.00	100.00	100.00
		EL	98.27	99.80	100.00	100.00	98.00	99.76	100.00	100.00
		QC	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Test of factor model H_{40}

In this example, we examine Type I error rate and empirical power of the proposed tests for factor model.

We first generate several mutually orthogonal factors. Suppose that $\mathbf{u}_k^*$, $k = 1, \cdots, K$ are IID random vectors following $N_p(\mathbf{0}, \mathbf{I}_p)$. Let $\mathbf{u}_1 = \mathbf{u}_1^*$ and $\mathbf{u}_k = (\mathbf{I}_p - \mathbf{P}_k)\mathbf{u}_k^*$, where $\mathbf{P}_k$ is the projection matrix on $\mathbf{u}_1, \cdots, \mathbf{u}_{k-1}$ for $k = 2, \cdots, K$.

Given $\mathbf{u}_k$'s, we have the covariance matrix structure

$$\boldsymbol{\Sigma} = \theta_0 \mathbf{I}_p + \sum_{k=1}^{K} \theta_k \mathbf{u}_k \mathbf{u}_k^T$$

for the factor model.

We set $K = 4$ and the coefficient vector $(\theta_0, \theta_1, \theta_2, \theta_3)^T = (4, 3, 2, 1)^T$.

Similarly, $\theta_4 = 0$ is for Type I error rates and $\theta_4 = 0.5, 1$ is for powers.

Table: Empirical powers for H_{40} (in percentage)

			$W_j \sim N(0,1)$				$W_j \sim$ Gamma(4,2)-2			
θ_5	n	Test	p =50	100	500	1000	50	100	500	1000
0	100	QL	5.46	5.46	6.02	6.27	6.89	6.53	6.58	6.94
		EL	5.40	6.40	5.79	6.20	6.03	6.42	6.30	6.52
0.5	100	QL	99.99	100.00	100.00	100.00	99.96	100.00	100.00	100.00
		EL	97.58	86.91	100.00	100.00	97.38	87.25	100.00	100.00
1.0	100	QL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		EL	100.00	99.93	100.00	100.00	99.99	99.93	100.00	100.00
0	200	QL	5.32	5.28	5.42	5.52	6.57	6.05	5.89	6.05
		EL	5.30	5.22	5.33	5.65	6.15	5.61	5.60	5.65
0.5	200	QL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		EL	99.99	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1.0	200	QL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		EL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Test for special pattern H_{50}

To investigate the performance of QL and EL tests for H_{50}, represent $\boldsymbol{\Sigma}$ as a linear combination

$$\boldsymbol{\Sigma} = \sum_{k=1}^{4} \theta_k \mathbf{A}_k,$$

where $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$ are defined in Example 5 and $\mathbf{A}_4 = \mathbf{u}_p \mathbf{u}_p^T$ with $\mathbf{u}_p \sim N_p(\mathbf{0}, \mathbf{I}_p)$. We set the first three coefficients $(\theta_1, \theta_2, \theta_3) = (6, 0.5, 0.1)$ and $\theta_4 = 0$, 0.5 and 1 for examining Type I error rates and powers, respectively.

Table: Empirical power for H_{50} (in percentage)

			$W_j \sim N(0,1)$				$W_j \sim$ Gamma(4,2)−2			
θ_5	n	Test	p =50	100	500	1000	50	100	500	1000
0	100	QL	5.28	5.19	5.23	5.41	6.59	6.15	5.84	6.16
		EL	5.27	6.33	5.17	5.48	5.85	6.35	5.61	5.73
0.5	100	QL	99.64	100.00	100.00	100.00	99.61	100.00	100.00	100.00
		EL	85.90	57.86	100.00	100.00	85.89	59.17	100.00	100.00
1.0	100	QL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		EL	99.87	97.67	100.00	100.00	99.81	97.86	100.00	100.00
0	200	QL	5.25	5.11	5.08	5.18	6.40	5.84	5.58	5.62
		EL	5.25	5.15	5.06	5.05	6.01	5.44	5.29	5.30
0.5	200	QL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		EL	99.75	99.99	100.00	100.00	99.72	99.99	100.00	100.00
1.0	200	QL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		EL	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Real data example

The data are extracted from a commercial database containing weekly returns for some stocks traded on the Chinese stock market during the period from September 1, 2006 to December 25, 2009.

After data cleaning including excluding observations with missing values, we obtain a data subset containing 132 weekly returns for 97 stocks.

It is of great interest to examine whether the Chinese stock market returns follow two classical pricing models:
capital asset pricing model (CAPM; Markowitz, 1952)
Fama and French's three-factor model (TFM; Fama and French, 1993).

To this end, we formulate this problem as testing the covariance matrix structure on the residual vector.

Consider three factors:

z_1: weekly return of the Shanghai Composite Index (i.e., the market index);
z_2: the difference in returns between portfolios of small capitalization firms and large capitalization firms
z_3: the difference in returns between portfolios of high book-to-market ratio firms and low book-to-market ratio firms.

The CAPM includes only the factor z_1 under the efficient market assumption. Let $\mathbf{y}_i$ be a 97-dimensional row vector for the weekly returns of 97 stocks at the ith week.

The CAPM follows the multi-response regression model

$$\mathbf{y}_i = \mathbf{b}_{10} + \mathbf{b}_{11} z_{1i} + \mathbf{e}_{1i}, \tag{9}$$

Let $\mathbf{z}_i = (z_{1i}, z_{2i}, z_{3i})^T$. The TFM follows

$$\mathbf{y}_i = \mathbf{b}_{20} + \mathbf{b}_{21} \mathbf{z}_i + \mathbf{e}_{2i}. \tag{10}$$

We first examine whether the CAPM is sufficient for deciding the stock returns. This can be formulated to test the correlation matrix of the corresponding residual vector having the structure: $\mathbf{I}_p + \theta(\mathbf{1}_p\mathbf{1}_p^T - \mathbf{I})$.

We apply the proposed methods to test

$$\mathbf{H}_{20a}: \ \mathrm{Corr}(\mathbf{e}_{1i}) = \mathbf{I}_p + \theta_1(\mathbf{1}_p\mathbf{1}_p^T - \mathbf{I}_p)$$

The estimator is $\hat{\theta}_1 = 0.099$. The EL-test and QL-test statistics are 10.9834 and 23.5712, respectively.

Both EL-test and QL-test reject the null hypothesis with the p-values 0.00 and 0.00, respectively. This implies that the residuals are not equally correlated and $\theta_1 \neq 0$.

The CAPM model was proposed based on the assumption of efficient market and all stock returns should be on the efficient frontier. Both EL-test and QL-test implies that Chinese market is not efficient

We next examine whether TFM can describe the stock pricing well by testing

$$\mathbf{H}_{20b}: \ \mathrm{Corr}(\mathbf{e}_{2i}) = \mathbf{I}_p + \theta_2(\mathbf{1}_p\mathbf{1}_p^T - \mathbf{I}_p)$$

The estimator is $\hat{\theta}_2 = 0.031$ which is less than $\hat{\theta}_1$. This implies that the TFM indeed improves the results. The EL-test and QL-test statistics are 11.3252 and 21.7684, respectively.

Both EL-test and QL-test also reject the null hypothesis H_{20b} with the p-values 0.00 and 0.00, respectively.

This empirical analysis seems to imply that the two classical pricing models do not describe the Chinese stocks returns sufficiently.

THANK YOU