Adaptive minimax tests for high dimensional covariance matrices with incomplete data

Cristina Butucea

joint work with Rania Zgheib

CREST, ENSAE, Université Paris-Saclay

Inverse problems - Heidelberg - 29 October 2016

Overview

1. Introduction

Statistical model

$X_1, \dots, X_n$ i.i.d p-dimensional vectors with Gaussian $\mathcal{N}_p(0, \Sigma)$ law, $\Sigma_{ii} = 1$.
We denote by $X_k = (X_{k,1}, \dots, X_{k,p})^\top$, for all $k = 1, \dots, n$.

We observe $Y_1, \dots, Y_n$ i.i.d p-dimensional vectors such that

$$Y_k = (\varepsilon_{k,1} \cdot X_{k,1}, \dots, \varepsilon_{k,p} \cdot X_{k,p})^\top \quad \text{for all } k = 1, \dots, n$$

- $\{\varepsilon_{k,j}\}_{1 \le k \le n, 1 \le j \le p}$ i.i.d. Bernoulli random variables $\mathcal{B}(a)$; $a \in (0,1)$
- $\{\varepsilon_{k,j}\}_{k,l}$ independent from $X_1, \dots, X_n$

Σ unknown but belongs to an ellipsoid $\mathcal{F}_1(\alpha, L)$, for $\alpha, L > 0$.

We assume that n and $p \to \infty$ and $a \to 0$.

Problem

- Testing problem:

$$\begin{aligned} H_0 &: \quad \Sigma = I \\ H_1 &: \quad \Sigma \in \mathcal{F}_1(\alpha, L), \quad \text{such that } \frac{1}{2p}\|\Sigma - I\|_F^2 \geq \varphi^2 \end{aligned}$$

where $\varphi = \varphi(n, p)$ tends to 0 as $n, p \to +\infty$.

Denote by $Q(\alpha, L, \varphi) = \left\{ \Sigma : \Sigma \in \mathcal{F}_1(\alpha, L); \frac{1}{2p}\|\Sigma - I\|_F^2 \geq \varphi^2 \right\}$

Problem

- Testing problem:

$$
\begin{aligned}
H_0 &: \quad \Sigma = I \\
H_1 &: \quad \Sigma \in \mathcal{F}_1(\alpha, L), \quad \text{such that } \frac{1}{2p}\|\Sigma - I\|_F^2 \geq \varphi^2
\end{aligned}
$$

  where $\varphi = \varphi(n, p)$ tends to 0 as $n, p \to +\infty$.

  Denote by $Q(\alpha, L, \varphi) = \left\{\Sigma : \Sigma \in \mathcal{F}_1(\alpha, L); \frac{1}{2p}\|\Sigma - I\|_F^2 \geq \varphi^2\right\}$

- To make a decision if either H_0 or H_1 is true, we need to construct a test procedure Δ: measurable function with respect to the observations taking values 0 or 1.

Quality of a test procedure and optimality

- For a given test Δ,
  - Type I error probability:

$$\eta(\Delta) := \eta(\Delta, I) = \mathbb{P}_I(\Delta = 1)$$

  - Maximal type II error probability:

$$\beta(\Delta, \mathcal{F}_1, \varphi) := \sup_{\Sigma \in Q(\alpha, L, \varphi)} \mathbb{P}_\Sigma(\Delta = 0)$$

  - Total error probability:

$$\gamma(\Delta, \mathcal{F}_1, \varphi) := \eta(\Delta) + \beta(\Delta, \mathcal{F}_1, \varphi)$$

Quality of a test procedure and optimality

- For a given test Δ,
  - Type I error probability:

$$\eta(\Delta) := \eta(\Delta, I) = \mathbb{P}_I(\Delta = 1)$$

  - Maximal type II error probability:

$$\beta(\Delta, \mathcal{F}_1, \varphi) := \sup_{\Sigma \in Q(\alpha, L, \varphi)} \mathbb{P}_\Sigma(\Delta = 0)$$

  - Total error probability:

$$\gamma(\Delta, \mathcal{F}_1, \varphi) := \eta(\Delta) + \beta(\Delta, \mathcal{F}_1, \varphi)$$

- Minimax optimality
  - Minimax total error probability:

$$\gamma(\varphi) := \inf_{\Delta} \gamma(\Delta, \mathcal{F}_1, \varphi)$$

where the infimum is taken over all possible tests Δ.

Minimax rate and sharp asymptotic optimality

- $\tilde{\varphi}$ is called minimax separation rate if:
  - Upper bound: there exists a test procedure Δ^* allowing us to distinguish between the two hypotheses i.e

$$\gamma(\Delta^*, \mathcal{F}_1, \varphi) \to 0 \quad \text{when } \varphi/\tilde{\varphi} \to +\infty.$$

  - Lower bound: there is no test able to distinguish between the two hypotheses i.e

$$\gamma(\varphi) = \inf_{\Delta} \gamma(\Delta, \mathcal{F}_1, \varphi) \to 1 \quad \text{when } \varphi/\tilde{\varphi} \to 0$$

Minimax rate and sharp asymptotic optimality

- $\tilde{\varphi}$ is called minimax separation rate if:
  - Upper bound: there exists a test procedure Δ^* allowing us to distinguish between the two hypotheses i.e

$$\gamma(\Delta^*, \mathcal{F}_1, \varphi) \to 0 \quad \text{when } \varphi/\tilde{\varphi} \to +\infty.$$

  - Lower bound: there is no test able to distinguish between the two hypotheses i.e

$$\gamma(\varphi) = \inf_{\Delta} \gamma(\Delta, \mathcal{F}_1, \varphi) \to 1 \quad \text{when } \varphi/\tilde{\varphi} \to 0$$

- $\gamma(\varphi)$ has an asymptotic Gaussian shape if, for $\varphi \asymp \tilde{\varphi}$:
  - there exists a continuous function $u \in]-\infty, 0]$ such that:

$$u(\tilde{\varphi}) = 1 \quad \text{and} \quad \gamma(\varphi) = \inf_{\Delta} \gamma(\Delta, \mathcal{F}_1, \varphi) = 2\Phi(-u(\varphi)) + o(1)$$

where Φ is the cumulative distribution function of a standard Gaussian random variable.

Bibliography

Testing the null hypothesis $\Sigma = I$ has been considered first for p fixed.

- A test based on the likelihood ratio was proposed by Mauchly(1940):

$$LR = \frac{1}{(\det S_n)^{\frac{n}{2}}} \cdot \exp\Big(- \frac{n}{2}(tr(S_n - I))\Big)$$

where $S_n = (1/n) \sum_{k=1}^n X_k X_k^\top$ is the sample covariance matrix.

Bibliography

Testing the null hypothesis $\Sigma = I$ has been considered first for p fixed.

- A test based on the likelihood ratio was proposed by Mauchly(1940):

$$LR = \frac{1}{(\det S_n)^{\frac{n}{2}}} \cdot \exp\Big(- \frac{n}{2}(tr(S_n - I))\Big)$$

  where $S_n = (1/n)\sum_{k=1}^n X_k X_k^\top$ is the sample covariance matrix.

- Another test based on the quadratic form was proposed by Nagao(1973):

$$QF = \frac{1}{2} tr(S_n - I)^2.$$

  It is shown that $-2\log(LR)$ and nQF converge in law to $\chi^2_{p(p+1)/2}$ under H_0.

Bibliography

- To cover the case $p \to +\infty$ several modifications of LR and FQ were suggested.
  - Ledoit and Wolf(2002), Srivastava(2005), Chen *et al.*(2010), Srivastava *et al.*(2014) introduced new test statistics based on modifications of FQ.
  - Srivastava(2006), Bai *et al.*(2009), Jiang *et al.*(2012) proposed corrections to the LR.

Bibliography

- To cover the case $p \to +\infty$ several modifications of LR and FQ were suggested.
  - Ledoit and Wolf(2002), Srivastava(2005), Chen *et al.*(2010), Srivastava *et al.*(2014) introduced new test statistics based on modifications of FQ.
  - Srivastava(2006), Bai *et al.*(2009), Jiang *et al.*(2012) proposed corrections to the LR.
- Previous papers show control of the type I error probability. However, Srivastava(2005) and Chen *et al.*(2010) also studied the test power $\mathbb{P}_\Sigma(\Delta = 1)$, for an arbitrary fixed Σ under the alternative.

Bibliography

- To cover the case $p \to +\infty$ several modifications of LR and FQ were suggested.
  - Ledoit and Wolf(2002), Srivastava(2005), Chen *et al.*(2010), Srivastava *et al.*(2014) introduced new test statistics based on modifications of FQ.
  - Srivastava(2006), Bai *et al.*(2009), Jiang *et al.*(2012) proposed corrections to the LR.
- Previous papers show control of the type I error probability. However, Srivastava(2005) and Chen *et al.*(2010) also studied the test power $\mathbb{P}_\Sigma(\Delta = 1)$, for an arbitrary fixed Σ under the alternative.
- Cai and Ma(2013) are first to give minimax separation rates:
  - H_1: $\Sigma > 0$ such that $\|\Sigma - I\|_F^2 \geq \varphi^2$
  - U-statistic of order 2: $U_n = \frac{1}{n(n-1)} \sum_{\substack{l,k=1 \\ i \neq j}}^{n} (X_l^T X_k)^2 - \frac{2}{n} \sum_{i=1}^{n} X_i^T X_i + 1$
  - Minimax separation rate: $\tilde{\varphi} = b\sqrt{p/n}$.

Bibliography

- B. and Zgheib (2016, EJS)
  - H_1: $\Sigma > 0$ such that $\Sigma \in \mathcal{F}_1(\alpha, L)$; $\|\Sigma - I\|_F^2 \geq \varphi^2$
  - U-statistic of order 2:
    $\widehat{\mathcal{D}}_n = \frac{1}{n(n-1)p} \sum_{\substack{l,k=1 \\ l \neq k}}^{n} \sum_{\substack{i,j=1 \\ i<j}}^{p} w_{ij}^* X_{k,i} X_{k,j} X_{l,i} X_{l,j}$, with optimal weights w_{ij}^*.
  - Minimax sharp separation rate: $\tilde{\varphi}_{n,p} = (C^{1/2}(\alpha, L) n \sqrt{p})^{-\frac{2\alpha}{4\alpha+1}}$.

Bibliography

- B. and Zgheib (2016, EJS)
  - *H_1*: $\Sigma > 0$ such that $\Sigma \in \mathcal{F}_1(\alpha, L)$; $\|\Sigma - I\|_F^2 \geq \varphi^2$
  - U-statistic of order 2: $\widehat{\mathcal{D}}_n = \frac{1}{n(n-1)p} \sum_{\substack{l,k=1 \\ l \neq k}}^{n} \sum_{\substack{i,j=1 \\ i<j}}^{p} w_{ij}^* X_{k,i} X_{k,j} X_{l,i} X_{l,j}$, with optimal weights w_{ij}^*.
  - Minimax sharp separation rate: $\tilde{\varphi}_{n,p} = (C^{1/2}(\alpha, L) n\sqrt{p})^{-\frac{2\alpha}{4\alpha+1}}$.
- B. Zgheib (2016, JMVA) - If moreover Σ is Toeplitz,
  - Minimax sharp separation rate: $\varphi_{n,p} = (C^{1/2}(\alpha, L) np)^{-\frac{2\alpha}{4\alpha+1}}$.

Bibliography

- B. and Zgheib (2016, EJS)
  - H_1: $\Sigma > 0$ such that $\Sigma \in \mathcal{F}_1(\alpha, L)$; $\|\Sigma - I\|_F^2 \geq \varphi^2$
  - U-statistic of order 2:
    $\widehat{\mathcal{D}}_n = \frac{1}{n(n-1)p} \sum_{\substack{l,k=1 \\ l \neq k}}^{n} \sum_{\substack{i,j=1 \\ i<j}}^{p} w_{ij}^* X_{k,i} X_{k,j} X_{l,i} X_{l,j}$, with optimal weights w_{ij}^*.
  - Minimax sharp separation rate: $\tilde{\varphi}_{n,p} = (C^{1/2}(\alpha, L) n\sqrt{p})^{-\frac{2\alpha}{4\alpha+1}}$.
- B. Zgheib (2016, JMVA) - If moreover Σ is Toeplitz,
  - Minimax sharp separation rate: $\varphi_{n,p} = (C^{1/2}(\alpha, L) np)^{-\frac{2\alpha}{4\alpha+1}}$.
- Lounici (2014, Bernoulli): Estimation of the covariance matrix with missing data;

Bibliography

- B. and Zgheib (2016, EJS)
  - H_1: $\Sigma > 0$ such that $\Sigma \in \mathcal{F}_1(\alpha, L)$; $\|\Sigma - I\|_F^2 \geq \varphi^2$
  - U-statistic of order 2: $\widehat{\mathcal{D}}_n = \frac{1}{n(n-1)p} \sum_{\substack{l,k=1 \\ l \neq k}}^{n} \sum_{\substack{i,j=1 \\ i<j}}^{p} w_{ij}^* X_{k,i} X_{k,j} X_{l,i} X_{l,j}$, with optimal weights w_{ij}^*.
  - Minimax sharp separation rate: $\tilde{\varphi}_{n,p} = (C^{1/2}(\alpha, L) n \sqrt{p})^{-\frac{2\alpha}{4\alpha+1}}$.
- B. Zgheib (2016, JMVA) - If moreover Σ is Toeplitz,
  - Minimax sharp separation rate: $\varphi_{n,p} = (C^{1/2}(\alpha, L) np)^{-\frac{2\alpha}{4\alpha+1}}$.
- Lounici (2014, Bernoulli): Estimation of the covariance matrix with missing data;
- Jurczak, Rohde (2015): Asymptotic distribution of the spectrum of the empirical covariance matrix with missing data.

2. Procedure and results

Test statistic

- Ellipsoid: for $\alpha, L > 0$,

$$\mathcal{F}_1(\alpha, L) = \Big\{\Sigma > 0 \; ; \frac{1}{p} \sum_{1 \le i < j \le p} \sigma_{ij}^2 |i-j|^{2\alpha} \le L \text{ for all } p \text{ and } \sigma_{ii} = 1 \text{ for all } i = 1, \dots, p\Big\}$$

Test statistic

- Ellipsoid: for $\alpha, L > 0$,

$$\mathcal{F}_1(\alpha, L) \quad = \quad \Big\{ \Sigma > 0 \; ; \frac{1}{p} \sum_{1 \le i < j \le p} \sigma_{ij}^2 |i-j|^{2\alpha} \le L \text{ for all } p \\ \text{and } \sigma_{ii} = 1 \text{ for all } i = 1, \ldots, p \Big\}$$

- Test statistic with constant weights:

$$\widehat{\mathcal{D}}_{n,p,m} = \frac{1}{n(n-1)p} \cdot \frac{1}{\sqrt{2m}} \sum_{1 \le k \ne l \le n} \sum_{\substack{1 \le i < j \le p \\ |i-j| < m}} Y_{k,i} Y_{k,j} Y_{l,i} Y_{l,j}$$

where $m \asymp \varphi^{-\frac{1}{\alpha}}$ is a large enough integer.

Test statistic

- Ellipsoid: for $\alpha, L > 0$,

$$\mathcal{F}_1(\alpha, L) = \Big\{ \Sigma > 0 \; ; \frac{1}{p} \sum_{1 \le i < j \le p} \sigma_{ij}^2 |i-j|^{2\alpha} \le L \text{ for all } p \text{ and } \sigma_{ii} = 1 \text{ for all } i = 1, \ldots, p \Big\}$$

- Test statistic with constant weights:

$$\widehat{\mathcal{D}}_{n,p,m} = \frac{1}{n(n-1)p} \cdot \frac{1}{\sqrt{2m}} \sum_{1 \le k \ne l \le n} \sum_{\substack{1 \le i < j \le p \\ |i-j| < m}} Y_{k,i} Y_{k,j} Y_{l,i} Y_{l,j}$$

where $m \asymp \varphi^{-\frac{1}{\alpha}}$ is a large enough integer.

- Results are obtained as $a \to 0$ such that $a^2 n \sqrt{p} \to +\infty$. Note that $a \to 0$ turns the problem into an ill-posed inverse problem.

Test statistic properties

- Under the null hypothesis:

$$\mathbb{E}_I(\widehat{\mathcal{D}}_n) = 0; \ \mathsf{Var}_I(\widehat{\mathcal{D}}_n) = \frac{a^4}{n(n-1)p} \quad \text{and} \quad \frac{n\sqrt{p}}{a^2}\,\widehat{\mathcal{D}}_{n,p,m} \xrightarrow{d} \mathcal{N}(0,1).$$

Test statistic properties

- Under the null hypothesis:

$$\mathbb{E}_I(\widehat{\mathcal{D}}_n) = 0; \; \mathrm{Var}_I(\widehat{\mathcal{D}}_n) = \frac{a^4}{n(n-1)p} \quad \text{and} \quad \frac{n\sqrt{p}}{a^2}\,\widehat{\mathcal{D}}_{n,p,m} \xrightarrow{d} \mathcal{N}(0,1).$$

- Under the alternative hypothesis, for all $\Sigma \in Q(\alpha, L, \varphi)$, if $\alpha > 1/2$ and $m \to \infty, m/p \to 0$ we have:

$$\mathbb{E}_\Sigma(\widehat{\mathcal{D}}_n) = \frac{a^4}{p\sqrt{2m}} \sum_{i<j} \sigma_{ij}^2 \text{ and } \mathrm{Var}_\Sigma(\widehat{\mathcal{D}}_n) = \frac{T_1}{n(n-1)p} + \frac{T_2}{np}$$

where,

$$\begin{aligned} T_1 &\leq a^4(1+o(1)) + \mathbb{E}_\Sigma(\widehat{\mathcal{D}}_{n,p,m}) \cdot O(a^2 m\sqrt{m}) \\ T_2 &\leq \mathbb{E}_\Sigma(\widehat{\mathcal{D}}_{n,p,m}) \cdot O(a^2\sqrt{m}) \\ &\quad + \sqrt{p} \cdot \left(\mathbb{E}_\Sigma^{3/2}(\widehat{\mathcal{D}}_{n,p,m}) \cdot O(a^2 m^{3/4}) + \mathbb{E}_\Sigma(\widehat{\mathcal{D}}_{n,p,m}) \cdot o(a^4) \right) \end{aligned}$$

Upper bound theorem

The test procedure $\Delta^* := \Delta^*(t) = \mathbb{1}(\widehat{\mathcal{D}}_{n,p,m} > t), \quad t > 0$ is s.t.:

Type I error probability : if $a^2 n\sqrt{p} \cdot t \to +\infty$ then $\eta(\Delta^*) \to 0$.

Type II error probability : if $\alpha > 1/2$ and if

$$m \to \infty, \quad m/p \to 0 \quad \text{and} \quad a^2 n\sqrt{p}\,\varphi^{2+\frac{1}{2\alpha}} \to +\infty$$

then, uniformly over t such that $t \leq c \cdot a^4 \varphi^{2+\frac{1}{2\alpha}}$, for some constant c small enough, we have

$$\beta(\Delta^*(t), \mathcal{F}_1, \varphi) \to 0$$

If all previous assumptions are satisfied, then $\Delta^*(t)$ is asymptotically minimax consistent:

$$\gamma(\Delta^*(t), \mathcal{F}_1, \varphi) \to 0$$

Lower bound theorem

Assume that, $\alpha \geq 1/2$. If $m \to \infty$, $m/p \to 0$,

$$a^2 n \to \infty\,, \quad (a^2 n)^{4\alpha - 1}/p \to \infty \quad \text{and} \quad a^2 n \sqrt{p}\, \varphi^{2 + \frac{1}{2\alpha}} \to 0$$

then

$$\gamma(\varphi) = \inf_{\Delta} \gamma(\Delta, \mathcal{F}_1, \varphi) \to 1$$

where the infimum is taken over all test statistics Δ.

Lower bound theorem

Assume that, $\alpha \geq 1/2$. If $m \to \infty$, $m/p \to 0$,

$$a^2 n \to \infty\,, \quad (a^2 n)^{4\alpha-1}/p \to \infty \quad \text{and} \quad a^2 n \sqrt{p}\, \varphi^{2+\frac{1}{2\alpha}} \to 0$$

then

$$\gamma(\varphi) = \inf_{\Delta} \gamma(\Delta, \mathcal{F}_1, \varphi) \to 1$$

where the infimum is taken over all test statistics Δ.

- As a consequence of the upper and lower bound theorems, the minimax separation rate is:

$$\tilde{\varphi} \asymp \left(a^2 n \sqrt{p}\right)^{-\frac{2\alpha}{(4\alpha+1)}}$$

Proof of lower bounds

- Reduce the set $Q(\alpha, L, \varphi)$

$$Q^* = \{\Sigma_U : [\Sigma_U]_{ij} = I(i=j) + u_{ij}\sigma I(i \neq j),\ U = [u_{ij}]_{1 \leq i,\, j \leq p} \in \mathcal{U}\}$$

where $\sigma = \varphi^{1+\frac{1}{2\alpha}}$, $T \asymp \varphi^{-\frac{1}{\alpha}}$

$$\mathcal{U} = \{U = [u_{ij}]_{i,\,j} : u_{ii} = 0, \forall i \text{ and } u_{ij} = u_{ji} = \pm 1 \cdot I(1 \leq |i-j| < T)\}.$$

If $\alpha > 1/2$, Σ_U is positive definite, for all $U \in \mathcal{U}$.

$\mathbb{P}_U$ the probability measure when $X_1, \ldots, X_n \overset{i.i.d}{\sim} \mathcal{N}_p(0, \Sigma_U)$.

- The average measure over Q^*:

$$\mathbb{P}_\pi = \frac{1}{2^{(p-\frac{T}{2})(T-1)}} \sum_{U \in \mathcal{U}} \mathbb{P}_U$$

$\mathbb{P}_I$ the probability measure when $X_1, \ldots, X_n \overset{i.i.d}{\sim} \mathcal{N}_p(0, I)$

Proof of lower bounds

- Minimax total error probability

$$\gamma(\varphi) \quad \geq \quad \underbrace{\inf_{\Delta}\{\mathbb{P}_I(\Delta = 1) + \mathbb{P}_\pi(\Delta = 0)\}}_{\gamma_1} = \underbrace{1 - \frac{1}{2}\|\mathbb{P}_I - \mathbb{P}_\pi\|_1}_{\gamma_2}$$

Proof of lower bounds

- Minimax total error probability

$$\gamma(\varphi) \quad \geq \quad \underbrace{\inf_{\Delta}\{\mathbb{P}_I(\Delta = 1) + \mathbb{P}_\pi(\Delta = 0)\}}_{\gamma_1} = \underbrace{1 - \frac{1}{2}\|\mathbb{P}_I - \mathbb{P}_\pi\|_1}_{\gamma_2}$$

  - $u_n := n\sqrt{p}b(\varphi) \to 0$

$$\gamma(\varphi) \geq \gamma_2 \quad \text{and show that} \quad \|\mathbb{P}_I - \mathbb{P}_\pi\|_1^2 \leq -\frac{1}{2}\mathbb{E}_I \log\left(\frac{d\mathbb{P}_\pi}{d\mathbb{P}_I}\right) \longrightarrow 0$$

Proof of lower bounds

- Minimax total error probability

$$\gamma(\varphi) \quad \geq \quad \underbrace{\inf_{\Delta}\{\mathbb{P}_I(\Delta = 1) + \mathbb{P}_\pi(\Delta = 0)\}}_{\gamma_1} = \underbrace{1 - \frac{1}{2}\|\mathbb{P}_I - \mathbb{P}_\pi\|_1}_{\gamma_2}$$

  - $u_n := n\sqrt{p}b(\varphi) \to 0$

    $$\gamma(\varphi) \geq \gamma_2 \quad \text{and show that} \quad \|\mathbb{P}_I - \mathbb{P}_\pi\|_1^2 \leq -\frac{1}{2}\mathbb{E}_I \log\left(\frac{d\mathbb{P}_\pi}{d\mathbb{P}_I}\right) \longrightarrow 0$$

  - $u_n \asymp 1$

    $$\gamma(\varphi) \geq \gamma_1 =: \inf_{\Delta} \gamma(\Delta, \mathbb{P}_I, \mathbb{P}_\pi) \geq 2\Phi(-n\sqrt{p}\,\frac{b(\varphi)}{2}) + o(1)$$

    (Ingster and Suslina(2003)) if we show that

    $$L_{n,p} := \log \frac{d\mathbb{P}_\pi}{d\mathbb{P}_I}(X_1, ..., X_n) = u_n Z_n - \frac{u_n^2}{2} + \xi \quad \text{in } \mathbb{P}_I\text{-probability}$$

    where $u_n = n\sqrt{p}b(\varphi)$, $Z_n \xrightarrow{d} \mathcal{N}(0,1)$ and $\xi \xrightarrow{\mathbb{P}_I} 0$.

Likelihood ratio

- Conditionally on the random variables ε_k:

$$\widetilde{X}_k := Y_k/\varepsilon_k \sim \mathcal{N}(0, \Sigma_U \star (\varepsilon_k \varepsilon_k^\top))$$

are independent, degenerate Gaussian.

The Kullback-Leibler distance writes

$$-\mathbb{E}_I \log\left(\frac{d\mathbb{P}_\pi}{d\mathbb{P}_I}\right) = -\mathbb{E}_\varepsilon \mathbb{E}_I^{Y/\varepsilon} \log \frac{d\mathbb{P}_\pi^{Y/\varepsilon}}{d\mathbb{P}_I^{Y/\varepsilon}}.$$

Likelihood ratio

- Conditionally on the random variables ε_k:

$$\widetilde{X}_k := Y_k/\varepsilon_k \sim \mathcal{N}(0, \Sigma_U \star (\varepsilon_k \varepsilon_k^\top))$$

are independent, degenerate Gaussian.

The Kullback-Leibler distance writes

$$-\mathbb{E}_I \log\left(\frac{d\mathbb{P}_\pi}{d\mathbb{P}_I}\right) = -\mathbb{E}_\varepsilon \mathbb{E}_I^{Y/\varepsilon} \log \frac{d\mathbb{P}_\pi^{Y/\varepsilon}}{d\mathbb{P}_I^{Y/\varepsilon}}.$$

- Moreover,

$$\begin{aligned} L_{n,p} &:= \log \frac{d\mathbb{P}_\pi^{Y/\varepsilon}}{d\mathbb{P}_I^{Y/\varepsilon}} = \log \mathbb{E}_U \frac{d\mathbb{P}_{\Sigma_U}^{Y/\varepsilon}}{d\mathbb{P}_I^{Y/\varepsilon}} \\ &= \log \mathbb{E}_U \exp\left(-\frac{1}{2}\sum_{k=1}^{n}(\widetilde{X}_k^\top((\Sigma_U^{\varepsilon_k})^{-1} - I^{\varepsilon_k})\widetilde{X}_k + \log\det(\Sigma_U^{\varepsilon_k}))\right). \end{aligned}$$

Likelihood ratio

- Conditionally on the random variables ε_k:

$$\widetilde{X}_k := Y_k/\varepsilon_k \sim \mathcal{N}(0, \Sigma_U \star (\varepsilon_k \varepsilon_k^\top))$$

  are independent, degenerate Gaussian.

  The Kullback-Leibler distance writes

$$-\mathbb{E}_I \log\left(\frac{d\mathbb{P}_\pi}{d\mathbb{P}_I}\right) = -\mathbb{E}_\varepsilon \mathbb{E}_I^{Y/\varepsilon} \log \frac{d\mathbb{P}_\pi^{Y/\varepsilon}}{d\mathbb{P}_I^{Y/\varepsilon}}.$$

- Moreover,

$$\begin{aligned} L_{n,p} &:= \log \frac{d\mathbb{P}_\pi^{Y/\varepsilon}}{d\mathbb{P}_I^{Y/\varepsilon}} = \log \mathbb{E}_U \frac{d\mathbb{P}_{\Sigma_U}^{Y/\varepsilon}}{d\mathbb{P}_I^{Y/\varepsilon}} \\ &= \log \mathbb{E}_U \exp\left(-\frac{1}{2}\sum_{k=1}^n (\widetilde{X}_k^\top ((\Sigma_U^{\varepsilon_k})^{-1} - I^{\varepsilon_k})\widetilde{X}_k + \log\det(\Sigma_U^{\varepsilon_k}))\right). \end{aligned}$$

- Note that

$$\Sigma_U^{\varepsilon_k} = I^{\varepsilon_k} + \Delta_U^{\varepsilon_k}, \text{ with } \|\Delta_U^{\varepsilon_k}\| \le 2T\sigma = O(\varphi^{1-\frac{1}{2\alpha}}).$$

Likelihood ratio

We use

$$\begin{aligned} L_1(\Delta_U^{\varepsilon_k}) \leq &\quad -(\Sigma_U^{\varepsilon_k})^{-1} - I^{\varepsilon_k} &\leq U_1(\Delta_U^{\varepsilon_k}) \\ L_2(\Delta_U^{\varepsilon_k}) \leq &\quad -\log\det(\Sigma_U^{\varepsilon_k}) &\leq U_2(\Delta_U^{\varepsilon_k}) \end{aligned}$$

for φ small enough, such that $\|\Delta_U^{\varepsilon_k}\| \leq 1/2$ where

$$U_1(\Delta) = \Delta - \Delta^2 + \Delta^3, \qquad L_1(\Delta) = U_1(\Delta) - 2\Delta^4$$

$$L_2(\Delta) = \frac{1}{2}tr(\Delta^2) - \frac{1}{3}tr(\Delta^3), \quad U_2(\Delta) = L_2(\Delta) + \frac{1}{2}tr(\Delta^4).$$

We bound $L_{n,p}$ from above and from below by terms that tend to 0.

In particular,

$$\log \mathbb{E}_U \exp\left(\sum_{i<j:\, 1<j-i<T} \sigma u_{ij} \sum_{k=1}^{n} \varepsilon_{k,i}\varepsilon_{k,j} X_{k,i} X_{k,j} \right)$$
$$= \sum_{i<j:\, 1<j-i<T} \log\cosh\left(\sigma \sum_{k=1}^{n} \varepsilon_{k,i}\varepsilon_{k,j} X_{k,i} X_{k,j} \right)$$

and also use that, for all real number u,

$$\frac{u^2}{2} - \frac{u^4}{12} \leq \log\cosh(u) \leq \frac{u^2}{2}.$$

Moreover, u_{ij}, $u_{ij}u_{jk}$, $u_{ij}u_{jk}u_{kl}$ are i.i.d. Rademacher distributed for all $i \neq j \neq k \neq l \neq i$.

Toeplitz covariance matrices

- $X_1, \ldots, X_n \overset{i.i.d}{\sim} \mathcal{N}_p(0, \Sigma)$, with Σ Toeplitz and belongs to

$$\mathcal{T}(\alpha, L) = \{\Sigma > 0, \Sigma \text{ is Toeplitz } ; \sum_{j \geq 1} \sigma_j^2 j^{2\alpha} \leq L \text{ and } \sigma_0 = 1\},\ \alpha > 0,\ L > 0.$$

- Alternative hypothesis

$$H_1 : \Sigma \in \mathcal{T}(\alpha, L) \text{ such that } \sum_{j \geq 1} \sigma_j^2 \geq \psi^2$$

- Test statistic: for m large such that $m \asymp \psi^{-\frac{1}{\alpha}}$

$$\widehat{\mathcal{A}}_n = \frac{1}{n(n-1)(p-m)^2} \cdot \frac{1}{\sqrt{2m}} \sum_{1 \leq k \neq l \leq n} \sum_{j=1}^{m} \sum_{m+1 \leq i_1, i_2 \leq p} Y_{k,i_1} Y_{k,i_1-j} Y_{l,i_2} Y_{l,i_2-j}$$

- Test procedure

$$\chi^* := \chi^*(t) = \mathbb{1}(\widehat{\mathcal{A}}_n > t), \quad t > 0$$

Toeplitz matrices v.s non-Toeplitz matrices

Σ	Toeplitz	non-Toeplitz
Minimax separation rate	$\left(a^2 \cdot np\right)^{-\frac{2\alpha}{4\alpha+1}}$	$\left(a^2 \cdot n\sqrt{p}\right)^{-\frac{2\alpha}{4\alpha+1}}$
Upper bound(UB)	$\alpha > 1/4$, $a^2 np\psi^{2+\frac{1}{2\alpha}} \to +\infty$	$\alpha > 1/2$, $a^2 n\sqrt{p}\varphi^{2+\frac{1}{2\alpha}} \to +\infty$
Exact UB: $a = 1$	$\alpha > 1/4$, $np\psi^{2+\frac{1}{2\alpha}} \asymp 1$	$\alpha > 1/2$, $n\sqrt{p}\varphi^{2+\frac{1}{2\alpha}} \asymp 1$
Lower bound(LB)	$\alpha > 1/2$, $a^2 np \to \infty$ and $a^2 np\psi^{2+\frac{1}{2\alpha}} \to 0$	$\alpha > 1/2$, $(a^2 n)^{4\alpha-1}/p \to \infty$ and $a^2 n\sqrt{p}\varphi^{2+\frac{1}{2\alpha}} \to 0$
Exact LB: $a = 1$	$\alpha > 1$, $np\psi^{2+\frac{1}{2\alpha}} \asymp 1$	$\alpha > 1$, $np\varphi^4 \to 0$ and $n\sqrt{p}\varphi^{2+\frac{1}{2\alpha}} \asymp 1$

3. Adaptation

Adaptive tests

- Assume that α is unknown and belongs to an interval A.

$$H_0 : \Sigma = I \text{ vs. } H_1 : \Sigma \in \bigcup_{\alpha \in A} \left\{ \mathcal{F}_1(\alpha, 1) \ ; \ \frac{1}{2p} \sum_{i<j} \sigma_{ij}^2 \geq (\mathcal{C}\Phi_\alpha)^2 \right\}$$

where $\mathcal{C} > 0$ is some constant and $\Phi_\alpha = (\rho_{n,p}/a^2 n\sqrt{p})^{\frac{2\alpha}{4\alpha+1}}$

Adaptive tests

- Assume that α is unknown and belongs to an interval A.

$$H_0 : \Sigma = I \text{ vs. } H_1 : \Sigma \in \bigcup_{\alpha \in A} \left\{ \mathcal{F}_1(\alpha, 1) \; ; \; \frac{1}{2p} \sum_{i<j} \sigma_{ij}^2 \geq (\mathcal{C}\Phi_\alpha)^2 \right\}$$

where $\mathcal{C} > 0$ is some constant and $\Phi_\alpha = (\rho_{n,p}/a^2 n\sqrt{p})^{\frac{2\alpha}{4\alpha+1}}$

- The goal is to construct a test procedure Δ_{ad} free of α and find the loss $\rho_{n,p}$ and the constant $\mathcal{C}_0$ such that:

$$\eta(\Delta_{ad}) + \sup_{\alpha \in A} \beta(\Delta_{ad}, \mathcal{F}_1(\alpha, 1), \mathcal{C}\Phi_\alpha) \to 0 \quad \text{when } \mathcal{C} > \mathcal{C}_0$$

Adaptive test procedure

- $A := [\alpha_*, \alpha^*_{n,p}] \subset]1/2, +\infty[$, with $\alpha^*_{n,p} \to +\infty$ and $\alpha^*_{n,p} = o(\ln(a^2 n \sqrt{p}))$

Adaptive test procedure

- $A := [\alpha_*, \alpha^*_{n,p}] \subset]1/2, +\infty[$, with $\alpha^*_{n,p} \to +\infty$ and $\alpha^*_{n,p} = o(\ln(a^2 n\sqrt{p}))$
- For each $\alpha \in [\alpha_*, \alpha^*_{n,p}]$, there exists $l \in \mathbb{N}^*$ ($m = 2^l$) such that

$$2^{l-1} \leq (\Phi_\alpha)^{-\frac{1}{\alpha}} < 2^l, \quad \text{it suffices to take } l \sim \frac{\frac{2}{4\alpha+1}\ln(a^2 n\sqrt{p})}{\ln(2)}$$

- $L_*, L^* \in \mathbb{N}^*$ such that

$$L_* = \left(\frac{2}{(4\alpha^*_{n,p}+1)\ln 2}\right)\ln(a^2 n\sqrt{p}) \quad \text{and} \quad L^* = \left(\frac{2}{(4\alpha_*+1)\ln 2}\right)\ln(a^2 n\sqrt{p})$$

$l \in \{L_*, \dots, L^*\}$:

$$\Delta_{2^l}(t_l) = \mathbb{1}(\widehat{\mathcal{D}}_{n,p,2^l} > t_l), \quad t_l > 0.$$

Adaptive test procedure

- $A := [\alpha_*, \alpha^*_{n,p}] \subset]1/2, +\infty[$, with $\alpha^*_{n,p} \to +\infty$ and $\alpha^*_{n,p} = o(\ln(a^2 n\sqrt{p}))$
- For each $\alpha \in [\alpha_*, \alpha^*_{n,p}]$, there exists $l \in \mathbb{N}^*$ $(m = 2^l)$ such that

$$2^{l-1} \leq (\Phi_\alpha)^{-\frac{1}{\alpha}} < 2^l, \quad \text{it suffices to take } l \sim \frac{\frac{2}{4\alpha+1}\ln(a^2 n\sqrt{p})}{\ln(2)}$$

- $L_*, L^* \in \mathbb{N}^*$ such that

$$L_* = \left(\frac{2}{(4\alpha^*_{n,p}+1)\ln 2}\right)\ln(a^2 n\sqrt{p}) \quad \text{and} \quad L^* = \left(\frac{2}{(4\alpha_*+1)\ln 2}\right)\ln(a^2 n\sqrt{p})$$

$l \in \{L_*, \ldots, L^*\}$:

$$\Delta_{2^l}(t_l) = \mathbb{1}(\widehat{\mathcal{D}}_{n,p,2^l} > t_l), \quad t_l > 0.$$

- Adaptive test

$$\Delta_{ad} = \max_{L_* \leq l \leq L^*} \Delta_{2^l}(t_l)$$

Adaptivity results

non-Toeplitz	Toeplitz
$t_l = a^2 \frac{\sqrt{\mathcal{C}^* \ln l}}{n\sqrt{p}}$	$t_l = a^2 \frac{\sqrt{\mathcal{C}^* \ln l}}{np}$
$\rho_{n,p} = \sqrt{\ln\ln(a^2 n\sqrt{p})}$	$\rho_{n,p} = \sqrt{\ln\ln(a^2 np)}$
if $\mathcal{C}^* > 4$ then $\eta(\Delta_{ad}) \to 0$	if $\mathcal{C}^* > 4$ then $\eta(\chi_{ad}) \to 0$
if $\mathcal{C}^2 > 1 + 4\sqrt{\mathcal{C}^*}$ and if	if $\mathcal{C}^2 > 1 + 4\sqrt{\mathcal{C}^*}$ and if
$a^2 n\sqrt{p} \to +\infty\,, \quad 2^{L^*}/p \to 0$	$a^2 np \to +\infty\,, \quad 2^{L^*}/p \to 0$
and $\ln(a^2 n\sqrt{p})/n \to 0$ then	and $\ln(a^2 np)/n \to 0$ then
$\sup_{\alpha\in A} \beta(\Delta_{ad}\,, \mathcal{F}_1(\alpha,1), \mathcal{C}\Phi_\alpha) \to 0$	$\sup_{\alpha\in A} \beta(\Delta_{ad}\,, \mathcal{F}_1(\alpha,1), \mathcal{C}\Psi_\alpha) \to 0$

4. Numerical behavior

Simulation

Example 1 : $\Sigma(M) = [\sigma_{ij}(M)]_{1 \le i,j \le p}$; $\sigma_{ij}(M) = \mathbb{1}_{(i=j)} + \dfrac{u_{ij} \cdot |i-j|^{-2}}{M} \cdot \mathbb{1}_{(i \neq j)}$

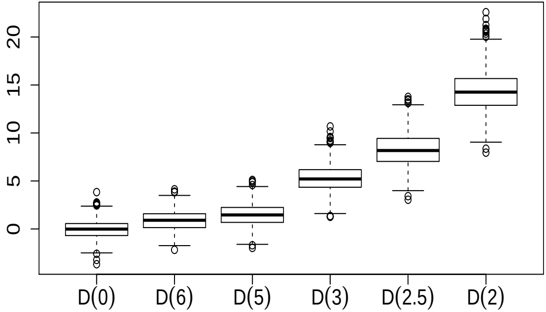

Figure : Distribution of $D(M) = n\sqrt{p}\widehat{\mathcal{D}}_n$, when $\Sigma = \Sigma(M)$ and $\Sigma(0) = I$, for $n = 50$ and $p = 80$, from 1000 repetitions

Simulation

$$\varphi^2(M) = (\sum_{i \neq j} \sigma_{ij}^4(M))/2p$$

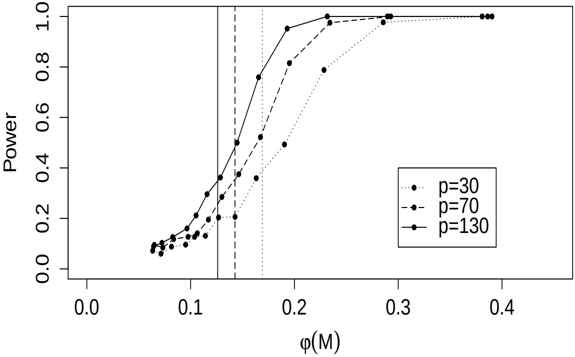

Figure : Power curves of the Δ-test as function of $\varphi(M)$ for $n = 30$ and $p \in \{30, 70, 130\}$

Simulation

Example 2 : $\Sigma(\rho) = [\rho I_{\{|i-j|=1\}}]_{1 \le i,j \le p}$

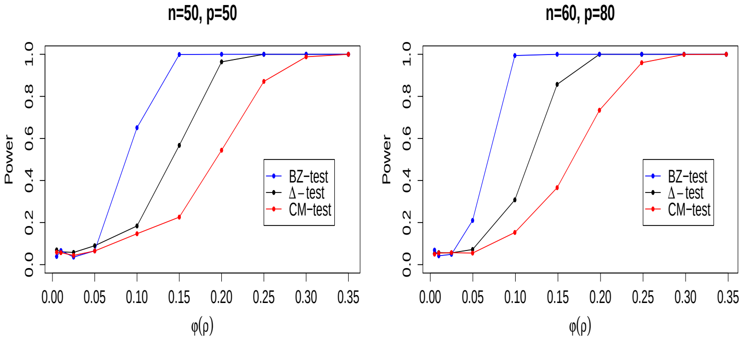

Figure : Power curves of the BZ-test, Δ-test and CM-test as function of $\varphi(\rho)$ for MA(1) Gaussian processes

Simulation

Example 3 : $\Sigma(\rho) = [\rho^{|i-j|}]_{1 \le i,j \le p}$

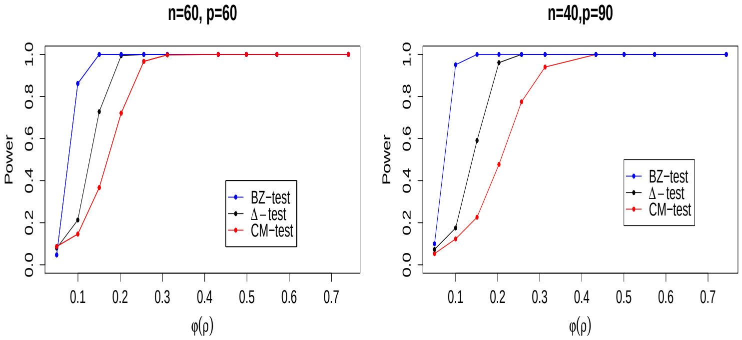

Figure : Power curves of the BZ-test, Δ-test and CM-test as functions of $\varphi(\rho)$, for AR(1) Gaussian processes

Simulation

Example 4: $\Sigma(M) = [\sigma_{ij}(M)]_{1\leq i,j\leq p}$; $\sigma_{ij}(M) = \mathbb{1}_{(i=j)} + \dfrac{|i-j|^{-2}}{M} \cdot \mathbb{1}_{(i\neq j)}$

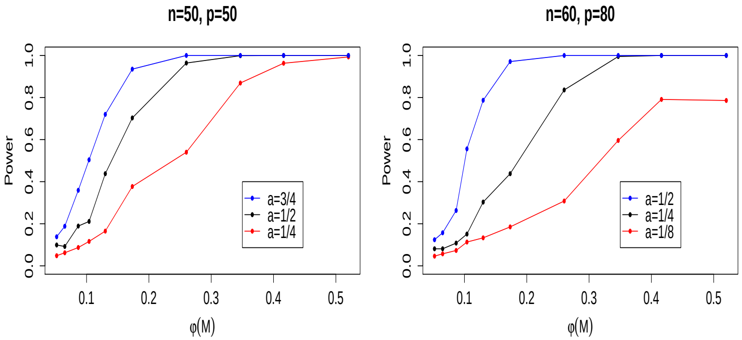

Figure : Power curves of the test based on the test statistic with constant weights in presence of missing data as function of $\varphi(M)$