Determinants of Correlation Matrices with Applications

Tiefeng Jiang

School of Statistics, University of Minnesota

July 14, 2016

Outline

- Correlation matrix

Outline

- Correlation matrix
- What we know about it?

Outline

- Correlation matrix
- What we know about it?
- Why study determinant?

Outline

- Correlation matrix
- What we know about it?
- Why study determinant?
- Our result

Outline

- Correlation matrix
- What we know about it?
- Why study determinant?
- Our result
- Application

Outline

- Correlation matrix
- What we know about it?
- Why study determinant?
- Our result
- Application
- Proof

Correlation Matrix

Pearson's correlation coefficient:

$$\mathbf{a} = (a_1, \cdots, a_n); \quad \mathbf{b} = (b_1, \cdots, b_n)$$

Correlation Matrix

Pearson's correlation coefficient:

$$\mathbf{a} = (a_1, \cdots, a_n); \quad \mathbf{b} = (b_1, \cdots, b_n)$$

$$\bar{\mathbf{a}} = \frac{1}{n}\sum_{i=1}^{n} a_i; \quad \bar{\mathbf{b}} = \frac{1}{n}\sum_{i=1}^{n} b_i$$

Correlation Matrix

Pearson's correlation coefficient:

$\mathbf{a} = (a_1, \cdots, a_n); \quad \mathbf{b} = (b_1, \cdots, b_n)$

$\bar{\mathbf{a}} = \frac{1}{n}\sum_{i=1}^{n} a_i; \quad \bar{\mathbf{b}} = \frac{1}{n}\sum_{i=1}^{n} b_i$

$$r_{\mathbf{a},\mathbf{b}} = \frac{\sum (a_i - \bar{\mathbf{a}})(b_i - \bar{\mathbf{b}})}{\sqrt{\sum (a_i - \bar{\mathbf{a}})^2}\sqrt{\sum (b_i - \bar{\mathbf{b}})^2}}$$

Correlation Matrix

Pearson's correlation coefficient:

$\mathbf{a} = (a_1, \cdots, a_n); \quad \mathbf{b} = (b_1, \cdots, b_n)$

$\bar{\mathbf{a}} = \frac{1}{n}\sum_{i=1}^{n} a_i; \quad \bar{\mathbf{b}} = \frac{1}{n}\sum_{i=1}^{n} b_i$

$$r_{\mathbf{a},\mathbf{b}} = \frac{\sum(a_i - \bar{\mathbf{a}})(b_i - \bar{\mathbf{b}})}{\sqrt{\sum(a_i - \bar{\mathbf{a}})^2}\sqrt{\sum(b_i - \bar{\mathbf{b}})^2}}$$

$r_{\mathbf{a},\mathbf{b}}$ is cosine of angle between
$(a_1 - \bar{\mathbf{a}}, \cdots, a_n - \bar{\mathbf{a}})$ and $(b_1 - \bar{\mathbf{b}}, \cdots, b_n - \bar{\mathbf{b}})$

$\mathbf{x}_1, \cdots, \mathbf{x}_n$ i.i.d from $N_p(\mu, \Sigma)$, $\quad \Sigma = (\sigma_{ij})_{p \times p}$.

$\mathbf{x}_1, \cdots, \mathbf{x}_n$ i.i.d from $N_p(\mu, \Sigma)$, $\quad \Sigma = (\sigma_{ij})_{p \times p}$.

• Correlation matrix $\mathbf{R}_n = (r_{ij})_{p \times p}$, where

$$r_{ii} = 1 \quad \text{and} \quad r_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

$\mathbf{x}_1, \cdots, \mathbf{x}_n$ i.i.d from $N_p(\mu, \Sigma)$, $\quad \Sigma = (\sigma_{ij})_{p\times p}$.

- Correlation matrix $\mathbf{R}_n = (r_{ij})_{p\times p}$, where

$$r_{ii} = 1 \quad \text{and} \quad r_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

- Sample correlation matrix $\hat{\mathbf{R}}_n := (\hat{r}_{ij})_{p\times p}$, where

$\mathbf{x}_1, \cdots, \mathbf{x}_n$ i.i.d from $N_p(\mu, \Sigma)$, $\quad \Sigma = (\sigma_{ij})_{p\times p}$.

• Correlation matrix $\mathbf{R}_n = (r_{ij})_{p\times p}$, where

$$r_{ii} = 1 \quad \text{and} \quad r_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

• Sample correlation matrix $\hat{\mathbf{R}}_n := (\hat{r}_{ij})_{p\times p}$, where

$\hat{r}_{ij}$ is Pearson's correlation coefficient of ith row and jth row of $(\mathbf{x}_1, \cdots, \mathbf{x}_n)_{p\times n}$

$\mathbf{x}_1, \cdots, \mathbf{x}_n$ i.i.d from $N_p(\mu, \Sigma)$, $\Sigma = (\sigma_{ij})_{p\times p}$.

- Correlation matrix $\mathbf{R}_n = (r_{ij})_{p\times p}$, where

$$r_{ii} = 1 \quad \text{and} \quad r_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

- Sample correlation matrix $\hat{\mathbf{R}}_n := (\hat{r}_{ij})_{p\times p}$, where

$\hat{r}_{ij}$ is Pearson's correlation coefficient of ith row and jth row of $(\mathbf{x}_1, \cdots, \mathbf{x}_n)_{p\times n}$

Remember difference between $\mathbf{R}_n$ and $\hat{\mathbf{R}}_n$

Connections

- Largest entry of $\hat{\mathbf{R}}_n$ test independence of entries of $N(\mu, \Sigma)$
  Cai and J. (2011)

Connections

- Largest entry of $\hat{\mathbf{R}}_n$ test independence of entries of $N(\mu, \Sigma)$
  Cai and J. (2011)

- Largest entry of $\hat{\mathbf{R}}_n$ is a statistic in Compressed Sensing; MIP
  Cai and J. (2011)

Connections

- Largest entry of $\hat{\mathbf{R}}_n$ test independence of entries of $N(\mu, \Sigma)$
  Cai and J. (2011)

- Largest entry of $\hat{\mathbf{R}}_n$ is a statistic in Compressed Sensing; MIP
  Cai and J. (2011)

- Random packing on sphere $\mathbb{S}^{p-1}$
  Cai, Fan and J. (2013).

What We Know about Sample Correlation Matrix?

Assume $\Sigma = \mathbf{I}$.

- Joint pdf of eigenvalues unknown, no invariance

What We Know about Sample Correlation Matrix?

Assume $\Sigma = \mathbf{I}$.

- Joint pdf of eigenvalues unknown, no invariance
- Empirical dist. of eigenvalues of $\hat{\mathbf{R}}_n$ is asymptotically Marchenko-Pastur law. J. (2004)

What We Know about Sample Correlation Matrix?

Assume $\Sigma = \mathbf{I}$.

- Joint pdf of eigenvalues unknown, no invariance
- Empirical dist. of eigenvalues of $\hat{\mathbf{R}}_n$ is asymptotically Marchenko-Pastur law. J. (2004)
- Largest eigenvalue of $\hat{\mathbf{R}}_n \to$ Tracy-Widom law Bao, Pan & Zhou. (2012)

What We Know about Sample Correlation Matrix?

Assume $\Sigma = \mathbf{I}$.

- Joint pdf of eigenvalues unknown, no invariance
- Empirical dist. of eigenvalues of $\hat{\mathbf{R}}_n$ is asymptotically Marchenko-Pastur law. J. (2004)
- Largest eigenvalue of $\hat{\mathbf{R}}_n \to$ Tracy-Widom law
  Bao, Pan & Zhou. (2012)
- $\log|\hat{\mathbf{R}}_n|$ satisfies CLT
  Jiang and Yang (2013) and Jiang and Qi (2015)

In particular, Jiang, Yang & Qi proved

Theorem

Assume $p := p_n$ *satisfy that* $n > p + 4$ *and* $p \to \infty$. *Set*

$$\mu_n = (p - n + \frac{3}{2}) \log \left(1 - \frac{p}{n-1}\right) - \frac{n-2}{n-1} p \, ;$$

$$\sigma_n^2 = -2 \left[\frac{p}{n-1} + \log \left(1 - \frac{p}{n-1}\right) \right].$$

Then, $(\log |\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n \to N(0, 1)$.

Little is known when $\mathbf{R}_n \neq \mathbf{I}$.

Why study $|\hat{\mathbf{R}}_n|$ under $\mathbf{R}_n \neq \mathbf{I}$?

- New understanding of sample correlation matrix $|\hat{\mathbf{R}}_n|$

Why study $|\hat{\mathbf{R}}_n|$ under $\mathbf{R}_n \neq \mathbf{I}$?

- New understanding of sample correlation matrix $|\hat{\mathbf{R}}_n|$
- Recently, Tao and Vu (2012); Nguyen & Vu (2014): CLT for determinant of Wigner matrix

Why study $|\hat{\mathbf{R}}_n|$ under $\mathbf{R}_n \neq \mathbf{I}$?

- New understanding of sample correlation matrix $|\hat{\mathbf{R}}_n|$
- Recently, Tao and Vu (2012); Nguyen & Vu (2014): CLT for determinant of Wigner matrix
- Cai, Liang, Zhou (2015) study CLT for determinant of Wishart matrix

Why study $|\hat{\mathbf{R}}_n|$ under $\mathbf{R}_n \neq \mathbf{I}$?

- New understanding of sample correlation matrix $|\hat{\mathbf{R}}_n|$
- Recently, Tao and Vu (2012); Nguyen & Vu (2014): CLT for determinant of Wigner matrix
- Cai, Liang, Zhou (2015) study CLT for determinant of Wishart matrix
- We have a problem from high-dimensional statistics on $|\hat{\mathbf{R}}_n|$

High-dimensional statistics + Machine Learning = Big Data

What kind of $\mathbf{R}_n$ are interesting?

What kind of $\mathbf{R}_n$ are interesting?

- Compound symmetry structure

$$\begin{pmatrix} 1 & a & a & \cdots & a \\ a & 1 & a & \cdots & a \\ a & a & 1 & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & 1 \end{pmatrix}$$

What kind of $\mathbf{R}_n$ are interesting?

- Compound symmetry structure

$$\begin{pmatrix} 1 & a & a & \cdots & a \\ a & 1 & a & \cdots & a \\ a & a & 1 & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & 1 \end{pmatrix}$$

- Autoregressive process of order 1

$$\begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{p-1} \\ \rho & 1 & \rho & \cdots & \rho^{p-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{p-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \cdots & 1 \end{pmatrix}$$

Banded matrix

$$\begin{pmatrix} 1 & a_{12} & 0 & \cdots & 0 & 0 \\ a_{21} & 1 & a_{23} & \cdots & 0 & 0 \\ 0 & a_{32} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{p-1\,p} \\ 0 & 0 & 0 & \cdots & a_{p\,p-1} & 1 \end{pmatrix}$$

Banded matrix

$$\begin{pmatrix} 1 & a_{12} & 0 & \cdots & 0 & 0 \\ a_{21} & 1 & a_{23} & \cdots & 0 & 0 \\ 0 & a_{32} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{p-1\,p} \\ 0 & 0 & 0 & \cdots & a_{p\,p-1} & 1 \end{pmatrix}$$

Other important matrices:

Toeplitz, Hankel, symmetric circulant matrices

Brockwell and Davis (2002)

Our Result

$\mathbf{x}_1, \cdots, \mathbf{x}_n$ i.i.d from $N_p(\mu, \Sigma)$, $\Sigma = (\sigma_{ij})_{p\times p}$.

- Correlation matrix $\mathbf{R}_n = (r_{ij})_{p\times p}$, where

$$r_{ii} = 1 \quad \text{and} \quad r_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

- Sample correlation matrix $\hat{\mathbf{R}}_n := (\hat{r}_{ij})_{p\times p}$, where

$\hat{r}_{ij}$ is Pearson's correlation coefficient of ith row and jth row of $(\mathbf{x}_1, \cdots, \mathbf{x}_n)_{p\times n}$

Theorem

Assume $p := p_n$*:*
$n > p + 4 \to \infty$ *and* $\inf_{n \geq 6} \lambda_{min}(\mathbf{R}_n) > \frac{1}{2}$*. Set*

$$\mu_n = (p - n + \frac{3}{2}) \log \left(1 - \frac{p}{n-1}\right) - \frac{n-2}{n-1} p + \log |\mathbf{R}_n| ;$$

$$\sigma_n^2 = -2\left[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\right] + \frac{2}{n-1} tr\left[(\mathbf{R}_n - \mathbf{I})^2\right].$$

Then, $(\log |\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n \to N(0,1)$ *if*

$\frac{p}{n} \to c > 0 \quad or \quad \sup_{n \geq 6} \frac{p_n \|\mathbf{R}_n - \mathbf{I}\|_\infty}{\|\mathbf{R}_n - \mathbf{I}\|_2} < \infty$

$\frac{0}{0} := 1$

Why we need $\inf_{n\geq 6} \lambda_{min}(\mathbf{R}_n) > \frac{1}{2}$?

- Essentially, it is from $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$
- Technically, it will be clear latter
- If $\lambda_{min}(\mathbf{R}_n) < \frac{1}{2}$, a different limit dist. is possible

Corollaries

- Compound symmetry structure

$$\mathbf{R}_n = \begin{pmatrix} 1 & a & a & \cdots & a \\ a & 1 & a & \cdots & a \\ a & a & 1 & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & 1 \end{pmatrix}$$

Smallest eigenvalue is $1 - a$.

$\|\mathbf{R}_n - \mathbf{I}\|_\infty = a$ and $\|\mathbf{R}_n - \mathbf{I}\|_2 = a\sqrt{p(p-1)}$. Then

$\|\mathbf{R}_n - \mathbf{I}\|_\infty = a$ and $\|\mathbf{R}_n - \mathbf{I}\|_2 = a\sqrt{p(p-1)}$. Then

$$\frac{p\|\mathbf{R}_n - \mathbf{I}\|_\infty}{\|\mathbf{R}_n - \mathbf{I}\|_2} = \sqrt{\frac{p}{p-1}} < 2$$

$\|\mathbf{R}_n - \mathbf{I}\|_\infty = a$ and $\|\mathbf{R}_n - \mathbf{I}\|_2 = a\sqrt{p(p-1)}$. Then

$$\frac{p\|\mathbf{R}_n - \mathbf{I}\|_\infty}{\|\mathbf{R}_n - \mathbf{I}\|_2} = \sqrt{\frac{p}{p-1}} < 2$$

Corollary

Assume $p := p_n$ *satisfy* $n > p + 4 \to \infty$. *Let* $a \in (0, 1/2)$. *Then,* $(\log|\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n \to N(0,1)$, *where*

$$|\mathbf{R}_n| = (1 + a(p-1))(1-a)^{p-1} \quad \textit{and} \quad tr\left[(\mathbf{R}_n - \mathbf{I})^2\right] = p(p-1)a^2$$

- Autoregressive process of order 1
- Banded matrix

- Autoregressive process of order 1
- Banded matrix

Corollary

Assume $\frac{p}{n} \to c > 0$. *Then*
(i) If $\mathbf{R}_n = (r_{ij})_{p\times p}$ *with* $r_{ij} = \rho^{|i-j|}$ *and* $|\rho| < \frac{1}{5}$ *then CLT holds*

- Autoregressive process of order 1
- Banded matrix

Corollary

Assume $\frac{p}{n} \to c > 0$. *Then*
(i) If $\mathbf{R}_n = (r_{ij})_{p\times p}$ *with* $r_{ij} = \rho^{|i-j|}$ *and* $|\rho| < \frac{1}{5}$ *then CLT holds*

(ii) Let $\mathbf{R}_n = (r_{ij})_{p\times p}$ *satisfy* $r_{ij} = 0$ *for* $|j-i| > k$ *and* $\max_{i\neq j} |r_{ij}| < \frac{1}{4k}$. *Then, CLT holds*

Application

$\mathbf{x}_1, \ldots, \mathbf{x}_n$ are i.i.d. from $N_p(\mu, \Sigma)$ with large p. Consider

$$H_0 : \mathbf{R} = \mathbf{I}_p \ \text{ vs } \ H_a : \mathbf{R} \neq \mathbf{I}_p.$$

Entries of $\mathbf{x}_1$ are independent.

Application

$\mathbf{x}_1, \ldots, \mathbf{x}_n$ are i.i.d. from $N_p(\mu, \Sigma)$ with large p. Consider

$$H_0 : \mathbf{R} = \mathbf{I}_p \ \text{ vs } \ H_a : \mathbf{R} \neq \mathbf{I}_p.$$

Entries of $\mathbf{x}_1$ are independent.

Likelihood Ratio Test: rejection region $\{|\hat{\mathbf{R}}_n| \leq c_\alpha\}$, where $c_\alpha = \mu_{n,0} + \sigma_{n,0}\Phi^{-1}(\alpha)$ and

$$\mu_{n,0} = (p - n + \frac{3}{2}) \log(1 - \frac{p}{n-1}) - \frac{n-2}{n-1}p \, ;$$
$$\sigma_{n,0}^2 = -2\Big[\frac{p}{n-1} + \log\Big(1 - \frac{p}{n-1}\Big)\Big].$$

Jiang, Qi and Yang (2013, 2015).

To evaluate test, need to compute power:
$P(\text{reject} \mid \text{null hypothesis is wrong})$. That is,

$$\beta(\mathbf{R}) \quad = \quad P(\log|\hat{\mathbf{R}}_n| \leq c_\alpha | \mathbf{R})$$

To evaluate test, need to compute power:
$P(\text{reject} \mid \text{null hypothesis is wrong})$. That is,

$$\begin{aligned} \beta(\mathbf{R}) &= P(\log|\hat{\mathbf{R}}_n| \leq c_\alpha | \mathbf{R}) \\ &\sim \Phi\Big(\frac{c_\alpha - \mu_n}{\sigma_n}\Big) \end{aligned}$$

where

To evaluate test, need to compute power:
$P(\text{reject} \mid \text{null hypothesis is wrong})$. That is,

$$\begin{aligned} \beta(\mathbf{R}) &= P(\log|\hat{\mathbf{R}}_n| \leq c_\alpha|\mathbf{R}) \\ &\sim \Phi\Big(\frac{c_\alpha - \mu_n}{\sigma_n}\Big) \end{aligned}$$

where

$$\mu_n = \big(p - n + \frac{3}{2}\big)\log\left(1 - \frac{p}{n-1}\right) - \frac{n-2}{n-1}p + \log|\mathbf{R}_n|\,;$$

$$\sigma_n^2 = -2\Big[\frac{p}{n-1} + \log\left(1 - \frac{p}{n-1}\right)\Big] + \frac{2}{n-1}\,\mathrm{tr}\,[(\mathbf{R}_n - \mathbf{I})^2].$$

Sketch of Proof

- $Z_n \to N(0,1)$ if $Ee^{tZ_n} \to e^{t^2/2}$ for $|t| \leq 1$

The above may fail if $Ee^{tZ_n} \to e^{t^2/2}$ for $0 < t \leq 1$ only.

Sketch of Proof

- $Z_n \to N(0,1)$ if $Ee^{tZ_n} \to e^{t^2/2}$ for $|t| \le 1$

The above may fail if $Ee^{tZ_n} \to e^{t^2/2}$ for $0 < t \le 1$ only.

- So to prove $(\log|\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n \to N(0,1)$, need to evaluate $E(|\hat{\mathbf{R}}_n|^s)$

Sketch of Proof

- $Z_n \to N(0,1)$ if $Ee^{tZ_n} \to e^{t^2/2}$ for $|t| \le 1$

The above may fail if $Ee^{tZ_n} \to e^{t^2/2}$ for $0 < t \le 1$ only.

- So to prove $(\log|\hat{\mathbf{R}}_n| - \mu_n)/\sigma_n \to N(0,1)$, need to evaluate $E(|\hat{\mathbf{R}}_n|^s)$

- Generalized Gamma function:

$$\Gamma_p(z) := \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\Big(z - \frac{1}{2}(i-1)\Big)$$

for z with $\mathrm{Re}(z) > \frac{1}{2}(p-1)$.

Proposition

$\mathbf{x}_1, \cdots, \mathbf{x}_n$: *i.i.d. with* $N_p(\mu, \Sigma)$ *and* $n = m + 1 > p$.
Set $\Delta_n = \mathbf{R}_n - \mathbf{I}$. *Then*

$$E[|\hat{\mathbf{R}}_n|^t] = \left(\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m}{2}+t)}\right)^p \cdot \frac{\Gamma_p(\frac{m}{2}+t)}{\Gamma_p(\frac{m}{2})} \cdot |\mathbf{R}_n|^t \cdot E\left[|\mathbf{I} + \Delta_n \cdot diag(V_1, \cdots, V_p)|^{-(m/2)-t}\right]$$

for $t > 0$, *where* $V_1, \cdots, V_p$: *i.i.d.* $Beta(t, \frac{m}{2})$*-dist.*

Proposition

$\mathbf{x}_1, \cdots, \mathbf{x}_n$: *i.i.d. with* $N_p(\mu, \Sigma)$ *and* $n = m + 1 > p$. *Set* $\Delta_n = \mathbf{R}_n - \mathbf{I}$. *Then*

$$E[|\hat{\mathbf{R}}_n|^t] = \left(\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m}{2}+t)}\right)^p \cdot \frac{\Gamma_p(\frac{m}{2}+t)}{\Gamma_p(\frac{m}{2})} \cdot |\mathbf{R}_n|^t \cdot E\left[|\mathbf{I} + \Delta_n \cdot diag(V_1, \cdots, V_p)|^{-(m/2)-t}\right]$$

for $t > 0$, *where* $V_1, \cdots, V_p$: *i.i.d.* $Beta(t, \frac{m}{2})$*-dist.*

• Drawback: Beta$(t, \frac{m}{2})$-dist forces $t > 0$. Traditional method not work; Moments not explicit

Proposition

$\mathbf{x}_1, \cdots, \mathbf{x}_n$: *i.i.d. with* $N_p(\mu, \Sigma)$ *and* $n = m + 1 > p$.
Set $\Delta_n = \mathbf{R}_n - \mathbf{I}$. *Then*

$$E[|\hat{\mathbf{R}}_n|^t] = \Big(\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m}{2}+t)}\Big)^p \cdot \frac{\Gamma_p(\frac{m}{2}+t)}{\Gamma_p(\frac{m}{2})} \cdot |\mathbf{R}_n|^t \cdot E\Big[|\mathbf{I} + \Delta_n \cdot diag(V_1, \cdots, V_p)|^{-(m/2)-t}\Big]$$

for $t > 0$, *where* $V_1, \cdots, V_p$: *i.i.d.* $Beta(t, \frac{m}{2})$*-dist.*

- Drawback: Beta$(t, \frac{m}{2})$-dist forces $t > 0$. Traditional method not work; Moments not explicit

- Chance: i.i.d.!

If $\mathbf{R}_n = (r_{ij})$ with $r_{ij} = a$ for $i \neq j$. Then

$$\begin{aligned} & |\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \cdots, V_p)| \\ = \; & \Big[\prod_{i=1}^{p}(1 - aV_i)\Big] \cdot \Big(1 + \sum_{i=1}^{p} \frac{aV_i}{1 - aV_i}\Big) \end{aligned}$$

where $V_1, \cdots, V_p$: i.i.d. Beta$(t, \frac{m}{2})$-dist.

Proposition

$\{Z_n;\, n \geq 1\}$: $\sup_{n\geq 0} E(|Z_n|^p) < \infty$ *for* $p \geq 1$ *and* $\lim_{n\to\infty} Ee^{tZ_n} = Ee^{tZ_0}$ *for* $t \in [0, \delta]$.

Proposition

$\{Z_n;\ n \geq 1\}$: $\sup_{n\geq 0} E(|Z_n|^p) < \infty$ *for* $p \geq 1$ *and* $\lim_{n\to\infty} Ee^{tZ_n} = Ee^{tZ_0}$ *for* $t \in [0, \delta]$.

If dist. of Z_0 *can be determined uniquely by moments* $\{E(Z_0^p);\ p = 1, 2, \cdots\}$, *then* $Z_n \to Z_0$

- Step 1

$$|\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \cdots, V_p)|^{-(m/2)-t} \sim e^{-\frac{t^2}{m} \cdot \mathrm{tr}\,(\Delta_n^2)}$$

in probability as $n \to \infty$

• Step 1

$$|\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \cdots, V_p)|^{-(m/2)-t} \sim e^{-\frac{t^2}{m} \cdot \mathrm{tr}\,(\Delta_n^2)}$$

in probability as $n \to \infty$

• Step 2 (most efforts)

$$\left\{\frac{LHS}{RHS};\, n \geq 6\right\} \text{ is uniformly integrable}$$

• Step 1

$$|\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \cdots, V_p)|^{-(m/2)-t} \sim e^{-\frac{t^2}{m} \cdot \mathbf{tr}\,(\Delta_n^2)}$$

in probability as $n \to \infty$

• Step 2 (most efforts)

$$\left\{\frac{LHS}{RHS};\, n \geq 6\right\} \text{ is uniformly integrable}$$

Then

$$E\left[|\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \cdots, V_p)|^{-(m/2)-t}\right] \sim e^{-\frac{t^2}{m} \cdot \mathbf{tr}\,(\Delta_n^2)}$$

Write

$$
\begin{aligned}
& E[|\hat{\mathbf{R}}_n|^t] \\
= \quad & E[|\hat{\mathbf{R}}_{n,0}|^t] \cdot |\mathbf{R}_n|^t \cdot E\left[|\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \cdots, V_p)|^{-\frac{n-1}{2}-t}\right]
\end{aligned}
$$

Write

$$
\begin{aligned}
& E[|\hat{\mathbf{R}}_n|^t] \\
= \quad & E[|\hat{\mathbf{R}}_{n,0}|^t] \cdot |\mathbf{R}_n|^t \cdot E\left[|\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \cdots, V_p)|^{-\frac{n-1}{2}-t}\right]
\end{aligned}
$$

where $E[|\hat{\mathbf{R}}_{n,0}|^t]$ is the case when $\mathbf{R}_n = \mathbf{I}$. By earlier result, we know behavior of $E[|\hat{\mathbf{R}}_{n,0}|^t]$.

Write

$$
\begin{aligned}
& E[|\hat{\mathbf{R}}_n|^t] \\
= \quad & E[|\hat{\mathbf{R}}_{n,0}|^t] \cdot |\mathbf{R}_n|^t \cdot E\big[|\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \cdots, V_p)|^{-\frac{n-1}{2}-t}\big]
\end{aligned}
$$

where $E[|\hat{\mathbf{R}}_{n,0}|^t]$ is the case when $\mathbf{R}_n = \mathbf{I}$. By earlier result, we know behavior of $E[|\hat{\mathbf{R}}_{n,0}|^t]$.

Combine them to have

$$
E \exp\left(\frac{\log|\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n} s\right) \to e^{s^2/2}
$$

for $0 \le s \le \delta$.

- Step 3

$$\sup_{n\geq 6} E\Big[\Big(\frac{\log|\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n}\Big)^{2k}\Big] < \infty$$

for $k = 1, 2, \cdots$

- Step 3

$$\sup_{n\geq 6} E\Big[\Big(\frac{\log|\hat{\mathbf{R}}_n| - \mu_n}{\sigma_n}\Big)^{2k}\Big] < \infty$$

for $k = 1, 2, \cdots$

- $N(0, 1)$ is uniquely determined by its moments.

The proof is complete.

Two Concentration Inequalities

Lemma

Given $s > 0$, *define* $t = t_n = \frac{s}{\sigma_n}$.
$V_1, \cdots, V_p$: *i.i.d. Beta*$(t, \frac{m}{2})$*-dist.*
Then, for $\rho \in (0, \frac{1}{2})$, *there exist* $M > 0$ *and* $n_0 \geq 1$ *s.t.*

$$P\Big(\sum_{i=1}^{p} V_i > y\Big) \leq e^{-\rho m y}$$

as $y \geq M\frac{pt}{m}$ *and* $n \geq n_0$

Not tandard Chernoff bound

Lemma

Given $s > 0$, *define* $t = \frac{s}{\sigma_n}$. *Recall* $\mathbf{R}_n = (r_{ij})$.
$V_1, \cdots, V_p$: *i.i.d. Beta*$(t, \frac{m}{2})$*-dist.*
Assume $\inf_{n \geq 6} \frac{p_n}{n} > 0$. *Then, there exists* $\delta > 0$ *s.t.*

$$P\Big(\sum_{i \neq j} r_{ij}^2 V_i V_j \geq y\Big) \leq \exp\Big(-\frac{1}{256} \cdot \frac{m^2 y}{pt + m\sqrt{y}}\Big)$$

for all $y > \frac{1}{m}$, $s \in (0, \delta]$ *and* $n \geq 6$

The proof is based on Yurinskii's ineq. + matrix tricks

Hanson-Wright ineq. is not enough (Rudelson, Vershynin, 2013)

Let $Q_n = |\mathbf{I} + \Delta_n \cdot \text{diag}(V_1, \cdots, V_p)|^{-\frac{m}{2}-t}$. Then

$$Q_n \preceq \exp\left(m\frac{1-\lambda_{min}}{2\lambda_{min}}\sum_{i=1}^{p} V_i\right).$$

$\lambda_{min} = \lambda_{min}(\mathbf{R}_n)$. To bound EQ_n, use formula

Let $Q_n = |\mathbf{I} + \Delta_n \cdot \mathrm{diag}(V_1, \cdots, V_p)|^{-\frac{m}{2}-t}$. Then

$$Q_n \preceq \exp\left(m\frac{1-\lambda_{min}}{2\lambda_{min}}\sum_{i=1}^{p} V_i\right).$$

$\lambda_{min} = \lambda_{min}(\mathbf{R}_n)$. To bound EQ_n, use formula

$$Ee^{\gamma H} \leq c + \int_0^\infty e^{\gamma x} P(H > x)\, dx$$

where H is r.v. and $\gamma > 0$ is const.

By earlier concentration ineq., this forces $\lambda_{min} > \frac{1}{2}$.

The End!