Statistical Methods in Portfolio Theory

Stanislas Muhinyuza

Department of Mathematics, University of Rwanda

Department of Mathematics, Stockholm University

Second Network Meeting for Sida- and ISP-funded PhD Students in Mathematics

Stockholm 26–27 February 2018

My Advisors

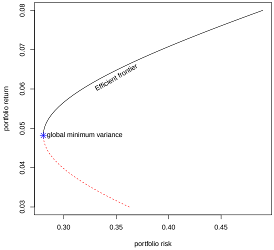
Taras Bodnar	Marcel R.Ndengo
Main advisor	*Assistant advisor*
Stockholm University	*University of Rwanda*

Outline

- Introduction
- Paper I(A test on the location of the tangency portfolio on the set of feasible portfolios)
  - Statistical test theory on the tangency portfolio
  - Out-of-sample performance
  - Empirical study
- Paper II(On the product of a singular Wishart matrix and a singular Gaussian vector in high dimension)
  - Assumption
  - Main results
  - Finite sample performance

Introduction

- The question of wealth allocation is relevant for both individuals (e.g. retirement savings), as well as for banks and other institutional investors.
- The mean variance analysis of [Markowitz, 1952] and later extended by [Merton, 1972]

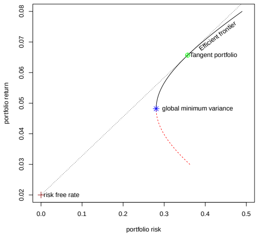


(a) Efficient frontier(EF)

- [Bodnar and Schmid, 2008, Kan and Zhou, 2008, Bodnar and Schmid, 2009] give properties and distribution of sample efficient frontier,
- [Bodnar et al., 2017] gives the properties of GMV.

Introduction(cont...)

- If there is a possibility to invest into a risk-free asset, the efficient frontier degenerates to straight line in the mean-variance space drawn from the return of the risk-free asset to the parabola.The tangent point is known as the tangent portfolio (TP) [Ingersoll, 1987].



(b) EF(Risk free)

- The TP is the only portfolio that maximizes the Sharpe ratio,
- The TP properties can be found in [Britten-Jones, 1999, Okhrin and Schmid, 2006, Bodnar and Zabolotskyy, 2017]
.
- The location of TP on EF is not treated up to know!!

Introduction(cont...)

- Asset allocation problem in presence of risk free asset

$$\max \mathbf{w}'(\boldsymbol{\mu} - r_f\mathbf{1}) + r_f - \frac{\gamma}{2}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} \text{ s.t. } \mathbf{w}'\mathbf{1} + w_0 = 1$$

- The optimal allocation $\mathbf{w}^* = \gamma^{-1}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f)\mathbf{1}, w_0 = 1 - \mathbf{1}'\mathbf{w}^*$
- The weights, expected return and variance of TP are given by

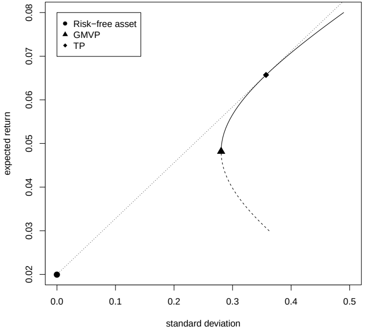
$$\mathbf{w}_T = \frac{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1})}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1})}, R_T = r_f + \frac{(\boldsymbol{\mu} - r_f\mathbf{1})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1})}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1})}$$

$$\text{and } V_T = \frac{(\boldsymbol{\mu} - r_f\mathbf{1})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1})}{(\mathbf{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1}))^2}$$

- **Problem:** $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$ unknown $\rightsquigarrow$ construct estimator $\hat{\boldsymbol{\Sigma}}$ and $\hat{\boldsymbol{\mu}}$
- [Muirhead, 1982] showed that $\mathbf{A} = (n-1)\hat{\boldsymbol{\Sigma}} \sim W_k(n-1, \boldsymbol{\Sigma})$ and $\mathbf{z} = \hat{\boldsymbol{\mu}} \sim N_k(\boldsymbol{\mu}, 1/n\boldsymbol{\Sigma})$, $\mathbf{A}, \mathbf{z}$-independent
- Properties& distribution of **Az** [Bodnar and Okhrin, 2011, Bodnar et al., 2013, Bodnar et al., 2015]. **Not all cases!!**

Paper I:Problem statement

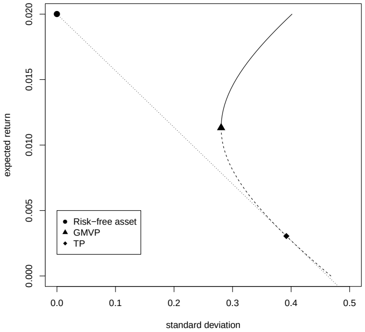
- The location of the TP portfolio on the set of feasible portfolio depends crucially on the relation between the expected return of the GMVP and the return of the risk-free asset.



(c) $R_{GMV} \geq r_f$

- The TP belongs to the upper part of the efficient frontier [Ingersoll, 1987, chapter 4, page 57]
- The TP is mean variance efficient.

Problem statement(cont...)



(d) $R_{GMV} < r_f$

- The TP belongs to the set of the feasible portfolios which are located on the lower part of the parabola
- The TP is not mean-variance efficient.

The investor would then prefer to invest into the risk-free asset or in the GMVP which lies in the vertex of the efficient frontier.

Statistical test theory on the tangency portfolio

- If the investor want to be sure in the investment into the TP, (s)he has to check if $R_{GMV} > r_f$.

$$H_0: \; R_{GMV} \leq r_f \quad \text{against} \quad H_1: \; R_{GMV} > r_f. \tag{1}$$

- The rejection of the null hypothesis means that:
  - TP lies on the upper part of the parabola(mean-variance efficient): Figure 1(c)
- if the null hypothesis in (1) cannot be rejected,
  - TP is not mean-variance efficient(Figure 1(d)),
  - the investor could not be sure in investing money into the TP and the allocation of the whole wealth into the risk-free asset could be considered as a suitable alternative.

Test statistic

- Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be an independent k-dimensional sample of the asset returns. Following the derivation in [Bodnar and Schmid, 2009] we obtain the test statistic given by

$$T = \frac{\sqrt{n-k}}{\sqrt{n-1}} \frac{\hat{R}_{GMV} - r_f}{\sqrt{1 + \frac{n}{n-1}\hat{s}}\sqrt{\frac{\hat{V}_{GMV}}{n}}}, \quad (2)$$

where $\hat{R}_{GMV} = \frac{\mathbf{1}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}$, $\hat{V}_{GMV} = \frac{1}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}$, and $\hat{s} = \hat{\boldsymbol{\mu}}'\hat{\mathbf{R}}\hat{\boldsymbol{\mu}}$, with $\hat{\mathbf{R}} = \hat{\boldsymbol{\Sigma}}^{-1} - \frac{\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\mathbf{1}'\hat{\boldsymbol{\Sigma}}^{-1}}{\mathbf{1}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}$ are the sample estimators for R_{GMV}, V_{GMV}, and s. with

$$\hat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i \quad \text{and} \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i - \hat{\boldsymbol{\mu}})(\mathbf{X}_i - \hat{\boldsymbol{\mu}})'$$

the sample mean vector and the sample covariance matrix, respectively

Distribution of T

Proposition 1

- Let $\mathbf{X}_t \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ for $t = 1, \ldots, n$ be independent random vectors of asset return,
- Assume that $\boldsymbol{\Sigma}$ is positive definite and $n > k$.
- Then the density of T is given by

$$f_T(x) = \frac{n(n-k+1)}{(k-1)(n-1)} \int_0^\infty f_{t_{n-k,\delta(y)}}(x) f_{F_{k-1,n-k+1,ns}}\left(\frac{n(n-k+1)}{(k-1)(n-1)} y\right) dy \quad (3)$$

where

$$\delta(y) = \sqrt{\frac{n}{1 + y(n/(n-1))}} S_{GMV} \quad \text{with} \quad S_{GMV} = \frac{R_{GMV} - r_f}{\sqrt{V_{GMV}}}$$

is the Sharpe ratio of the GMVP and the slope parameter $_s$ is defined above.

The critical for the test

Proposition 2

Under the conditions of Proposition (1), it holds that

$$\sup_{V_{GMV}>0, s\geq 0, R_{GMV}\leq r_f} G_{T,\alpha}(S_{GMV}, s) \leq \mathbb{P}_{H_0: R_{GMV}=r_f}(T > t_{n-k,1-\alpha}) = \alpha,$$

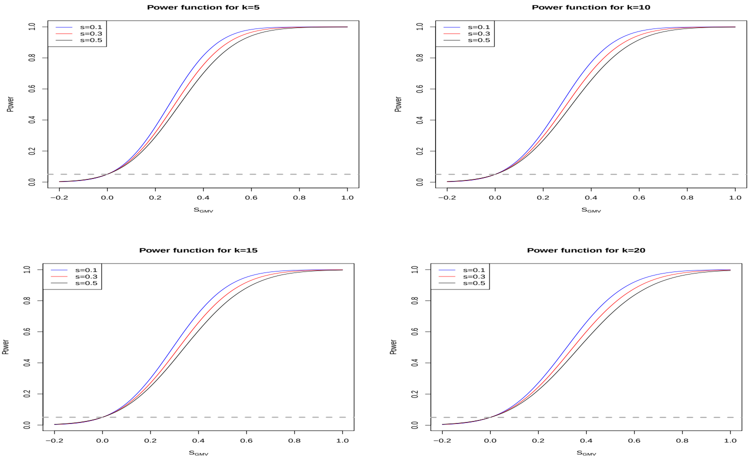
where

$$G_{T,\alpha}(S_{GMV}, s) = \mathbb{P}(T > c) = \int_c^{\infty} f_T(x) dx\,.$$

- Following Proposition (2) the test (1) rejects H_0 in favour to H_1 as soon as $T \geq t_{n-k,1-\alpha}$.
- The power function depends on $\boldsymbol{\mu}$ & $\boldsymbol{\Sigma}$ through S_{GMV} & s,

$$G_{T,\alpha}(S_{GMV}, s) = \mathbb{P}(T > t_{n-k,1-\alpha})$$
$$= \frac{n(n-k+1)}{(k-1)(n-1)} \int_0^{\infty} \left(1 - F_{t_{n-k,\delta(y)}}(t_{n-k,1-\alpha})\right) f_{F_{k-1,n-k+1,ns}}\left(\frac{n(n-k+1)}{(k-1)(n-1)} y\right) dy.$$

Power function for $k \in \{5, 10, 15, 20\}$ and $n = 50$



It means that smaller slope coefficient s yields smaller expected return of the optimal portfolio for a higher level of risk and it is evidence of a crisis of the market.

Out-of-sample performance

- The aim is to provide the statements about the following two conditional probabilities:

$$P_1 = \mathbb{P}\left(\hat{R}_{GMV,n+1} > r_f | \hat{R}_{GMV} > r_f\right) \quad (4)$$

and

$$P_2 = \mathbb{P}\left(\hat{R}_{GMV,n+1} > r_f | T > t_{n-k,1-\alpha}\right) \quad (5)$$

where $\hat{R}_{GMV,n+1} = \hat{\mathbf{w}}'_{GMV}\mathbf{X}_{n+1}$ is the realized expected return of the GMVP at the time point $n+1$, $\mathbf{X}_{n+1}$ is the vector of asset returns at time point $n+1$ and $\hat{\mathbf{w}}_{GMV} = \hat{\mathbf{\Sigma}}^{-1}\mathbf{1}/(\mathbf{1}'\hat{\mathbf{\Sigma}}^{-1}\mathbf{1})$ are the estimated weights of the GMVP by using asset return data $\mathbf{X}_1, ..., \mathbf{X}_n$.

Stochastic representation

Under the conditions of Proposition (1). Then, it holds that:

(a) the stochastic representation for $(\hat{R}_{GMV}, \hat{R}_{GMV,n+1})$ is given by

$$\hat{R}_{GMV} \stackrel{d}{=} R_{GMV} + \frac{\sqrt{V_{GMV}}}{\sqrt{n}} z_4 + \sqrt{\frac{1}{n}\xi_3 + \frac{1}{n}\left(\sqrt{ns} + z_5\right)^2} \sqrt{V_{GMV}} \frac{z_1}{\sqrt{\xi_1}} \tag{6}$$

and

$$\hat{R}_{GMV,n+1} \stackrel{d}{=} R_{GMV} + \sqrt{V_{GMV}} z_6 + \sqrt{V_{GMV}} \left(\frac{\sqrt{s}(\sqrt{ns} + z_5)}{\sqrt{\xi_3 + (\sqrt{ns} + z_5)^2}} + z_7 \right) \frac{z_1}{\sqrt{\xi_1}} + \sqrt{V_{GMV}} \sqrt{\xi_4} \left(\frac{z_3}{\sqrt{\xi_2}} \frac{z_1}{\sqrt{\xi_1}} + \frac{z_2}{\sqrt{\xi_2}} \right) \tag{7}$$

where $z_1, z_2, z_3, z_4, z_5, z_6, z_7 \sim \mathcal{N}(0,1)$, $\xi_1 \sim \chi^2_{n-k+1}$, $\xi_2 \sim \chi^2_{n-k+2}$, $\xi_3 \sim \chi^2_{k-2}$, $\xi_4 | z_5, \xi_3 \sim \chi^2_{k-2;\delta^2(s,\xi_3,z_5)}$ with $\delta^2(s, \xi_3, z_5) = \frac{s\xi_3}{\xi_3 + (\sqrt{ns} + z_5)^2}$; $z_1, z_2, z_3, z_4, z_6, z_7, \xi_1, \xi_2, (z_5, \xi_3, \xi_4)$ are mutually independent.

(b) the stochastic representation for $(T, \hat{R}_{GMV,n+1})$ is given by (7) and

$$T \stackrel{d}{=} \frac{\sqrt{n-k}}{\sqrt{\xi_5}} \frac{1}{\sqrt{1 + \frac{\xi_3 + (\sqrt{ns} + z_5)^2}{\xi_1}}} \left(\sqrt{n} \frac{R_{GMV} - r_f}{\sqrt{V_{GMV}}} + z_4 + \sqrt{\frac{\xi_3 + \left(\sqrt{ns} + z_5\right)^2}{\xi_1}} z_1 \right) \tag{8}$$

where $\xi_5 \sim \chi^2_{n-k}$ independent of $z_1, z_2, z_3, z_4, z_6, z_7, \xi_1, \xi_2, (z_5, \xi_3, \xi_4)$.

Algorithm for $\hat{P}_1$

- In the case of $(\hat{R}_{GMV}, \hat{R}_{GMV,n+1})$, the following algorithm can be used to evaluate P_1:
  - (i) fix the values of r_f and (R_{GMV}, V_{GMV}, s);
  - (ii) generate independently $z_1^b, z_2^b, z_3^b, z_4^b, z_5^b, z_6^b, z_7^b \sim \mathcal{N}(0,1)$, $\xi_1^b \sim \chi^2_{n-k+1}$, $\xi_2^b \sim \chi^2_{n-k+2}$, $\xi_3^b \sim \chi^2_{k-2}$;
  - (iii) generate $\xi_4 \sim \chi^2_{k-2;\delta^2(s,\xi_3^b,z_5^b)}$ with $\delta^2(s,\xi_3^b,z_5^b) = \frac{s\xi_3^b}{\xi_3^b + \left(\sqrt{ns} + z_5^b\right)^2}$;
  - (iv) compute $(\hat{R}^b_{GMV}, \hat{R}^b_{GMV,n+1})$ as in (6) and (7) by using $z_1^b, z_2^b, z_3^b, z_4^b, z_5^b, z_6^b, z_7^b$, $\xi_1^b, \xi_2^b, \xi_3^b, \xi_4^b$;
  - (v) determine

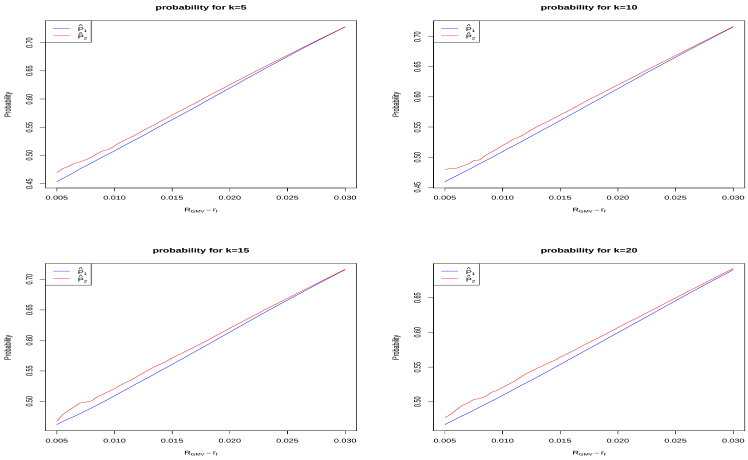
    $$c_1^b = \mathbb{1}_{\{\hat{R}^b_{GMV} > r_f, \hat{R}^b_{GMV,n+1} > r_f\}} \quad \text{and} \quad c_2^b = \mathbb{1}_{\{\hat{R}^b_{GMV} > r_f\}},$$

    where $\mathbb{1}_{\{\mathcal{A}\}}$ is the indicator function of set $\mathcal{A}$;
  - (vi) repeat steps (i)-(v) for $b = 1, ..., B$ and approximate P_1 by

    $$\hat{P}_1 = \frac{\sum_{b=1}^B c_1^b}{\sum_{b=1}^B c_2^b}$$

$\hat{P}_1$ and $\hat{P}_2$

Using data sets of the empirical illustration of Section 5.1 in [Bodnar and Schmid, 2009].



Simulation Study

- fix $\boldsymbol{\Sigma} = \boldsymbol{I}_k$,
- the values of mean vector $\boldsymbol{\mu}$ are fixed as follow:
  - $\boldsymbol{\mu}_1 = (0.1, 0, \ldots, 0)'$
  - $\boldsymbol{\mu}_2 = (0.1, 0.1, 0, \ldots, 0)'$
  - $\boldsymbol{\mu}_3 = (0.1, 0.1, 0.1, 0 \ldots, 0)'$
  - $\boldsymbol{\mu}_4 = (0.1, 0.1, 0.1, 0.1, 0, \ldots, 0)'$
  - $\boldsymbol{\mu}_5 = (0.1, 0.1, 0.1, 0.1, 0.1, 0, \ldots, 0)'$
- The dimension of mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ will depend on the size of the portfolio $k \in \{5, 10, 15, 20\}$ and $n = 50$
- $r_f = 0.01$ and $\alpha = 0.05$
- Data from normal distribution and t_5 and t_{10} are generated,
- compute the power function of the test (1),
- $B = 10^6$

Power function

k	Distribution	μ_1	μ_2	μ_3	μ_4	μ_5
	Normal	0.0669	0.1151	0.1833	0.2731	0.3827
5	t_5	0.0638	0.1061	0.1670	0.2473	0.3438
	t_{10}	0.0660	0.1110	0.1767	0.2622	0.3655
	Normal	0.0497	0.0738	0.1055	0.1456	0.1952
10	t_5	0.0464	0.0683	0.0952	0.1316	0.1752
	t_{10}	0.0485	0.0710	0.1012	0.1402	0.1866
	Normal	0.0426	0.0582	0.0781	0.1020	0.1310
15	t_5	0.0392	0.0526	0.0699	0.0912	0.1163
	t_{10}	0.0414	0.0561	0.0745	0.0969	0.1244
	Normal	0.0389	0.0497	0.0634	0.0801	0.0991
20	t_5	0.0348	0.0447	0.0567	0.0707	0.0873
	t_{10}	0.0371	0.0480	0.0607	0.0762	0.0941

Empirical study

- Weekly returns of 29 assets listed DJI index and traded on the period from 0.1.01.2006 to 31.12.2015,
- the risk free asset, we use the 13 week treasury bill covering the aforementioned period,
- the most recent data of one year for each portfolio $n = 50$, i.e. approximately year 2015 is considered for different portfolios,
- the p-value of the hypothesis test problem in (1) is computed for different portfolio size $k \in \{5, 10, 15, 20\}$ selected randomly without repetitions,
- generate 10^3 different portfolio and plot a histogram of computed p-values

Number of rejections

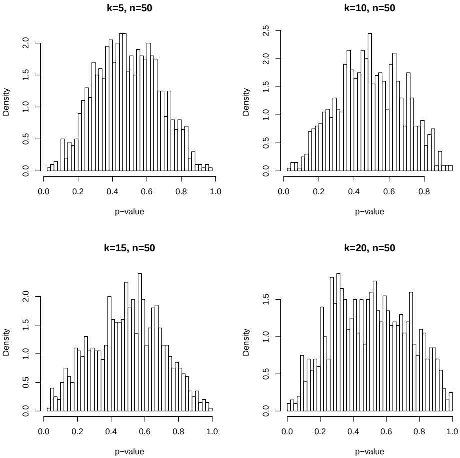
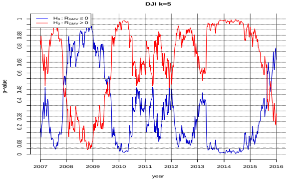


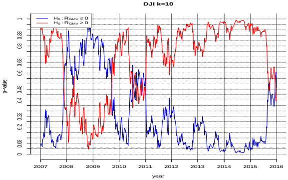
Figure: Histograms of p-values for different portfolios consisting of $k = 5$(top left), $k = 10$(top right), $k = 15$(bottom left) and $k = 20$(bottom right) randomly sampled DJI markets for $n = 50$.

Rolling window

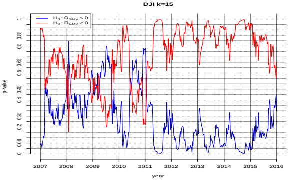
- weekly returns 29 assets listed DJI index and traded on the period from 0.1.01.2006 to 31.12.2015,
- a portfolio of size k is formed by choosing k stocks in the alphabetic order, for example for $k = 5$ we choose $\{AAPL, AXP, BA, CAT, CSCO\}$,
- the risk free asset, we use the 13 week treasury bill covering the aforementioned period,
- window size of $n = 50$,
- we compute the p-value over the years



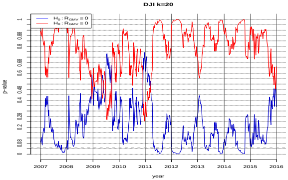
(a) Rolling window of DJI for $k = 5$



(b) Rolling window of DJI for $k = 10$



(c) Rolling window of DJI for $k = 15$



(d) Rolling window of DJI for $k = 20$

Figure: P-values of fixed number of stocks $k \in \{5, 10, 15, 20\}$ calculated by using the weekly asset returns of 30 stocks included into DJI index

Paper II

Recall

Asymptotics

- standard asymptotics
  - fixed dimension k and the sample size n goes to infinity;
  - classical limit theorems hold
- Large dimensional asymptotics (Asymptotic distribution under double asymptotic regime)
  - both the dimension k and the sample size n goes to infinity;
  - the ratio k/n tends to a positive constant $c > 0$;
  - classical limit theorems do not hold anymore.

Recall(cont...)

- Let $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ be a sample of size n from k-variate normal distribution, That is $\mathbf{X}_i \sim \mathcal{N}_k(\mathbf{0}, \mathbf{\Sigma})$ for $i = 1, \ldots, n$.
- $\mathbf{A} = \mathbf{X}\mathbf{X}' \sim \mathcal{W}_k(n, \mathbf{\Sigma}); k \leq n$, **Wishart distribution** . The properties of $\mathbf{A}$ are detailed in [Muirhead, 1982].
- [Srivastava, 2003] defined $\mathbf{A} = \mathbf{X}\mathbf{X}' \sim \mathcal{W}_k(n, \mathbf{\Sigma})$, $k > n$ as a **singular Wishart matrix.**
- **Singular Gaussian vector** is the normal distributed vector with a singular covariance matrix.

Assumptions

- $\mathbf{z} \sim \mathcal{N}_k(\boldsymbol{\mu}, \kappa\boldsymbol{\Sigma})$, $\kappa > 0$ with $rank(\boldsymbol{\Sigma}) = r < k$;
- $\mathbf{A} \sim \mathcal{W}_k(n, \boldsymbol{\Sigma})$ with $rank(\boldsymbol{\Sigma}) = r < k$;
- $\mathbf{M}$: $p \times k$ matrix of constants with $rank(\mathbf{M}) = p \leq r \leq \min\{n, k\}$ such that $\mathbf{M}\boldsymbol{\Sigma} \neq \mathbf{0}$;
- $\mathbf{A}$ and $\mathbf{z}$ are independent.

Aim: Distribution of $\mathbf{MAz}$

Stochastic representation

Theorem 1[Bodnar et al., 2016]

1. Let $P(rank((\mathbf{M}^T, \mathbf{z})^T \boldsymbol{\Sigma}) = p + 1 \leq r) = 1$, and let $\mathbf{Q} = \mathbf{P}^T\mathbf{P}$ with $\mathbf{P} = (\mathbf{M}\boldsymbol{\Sigma}\mathbf{M}^T)^{-1/2}\mathbf{M}\boldsymbol{\Sigma}^{1/2}$. Then

$$\mathbf{MAz} \stackrel{d}{=} \zeta\mathbf{M}\boldsymbol{\Sigma}^{1/2}\mathbf{t} + \sqrt{\zeta}(\mathbf{M}\boldsymbol{\Sigma}\mathbf{M}^T)^{1/2} \times \left[\sqrt{\mathbf{t}^T\mathbf{t}}\mathbf{I}_p - \frac{\sqrt{\mathbf{t}^T\mathbf{t}} - \sqrt{\mathbf{t}^T(\mathbf{I}_k - \mathbf{Q})\mathbf{t}}}{\mathbf{t}^T\mathbf{Q}\mathbf{t}}\mathbf{P}\mathbf{t}\mathbf{t}^T\mathbf{P}^T\right]\mathbf{z}_0,$$

where $\zeta \sim \chi^2_n$, $\mathbf{t} \sim \mathcal{N}_k(\boldsymbol{\Sigma}^{1/2}\boldsymbol{\mu}, \kappa\boldsymbol{\Sigma}^2)$, and $\mathbf{z}_0 \sim \mathcal{N}_p(\mathbf{0}, \mathbf{I}_p)$; ζ, $\mathbf{t}$, and $\mathbf{z}_0$ are mutually independent.

2. Let $\mathbf{m}$ be a k-dimensional vector of constants such that $\mathbf{m}^T\boldsymbol{\Sigma}\mathbf{m} > 0$ and $P(\mathbf{z}^T\boldsymbol{\Sigma}\mathbf{z} = 0) = 0$. Then

$$\mathbf{m}^T\mathbf{Az} \stackrel{d}{=} \zeta\mathbf{m}^T\boldsymbol{\Sigma}\mathbf{z} + \sqrt{\zeta}\left[\mathbf{z}^T\boldsymbol{\Sigma}\mathbf{z}\cdot\mathbf{m}^T\boldsymbol{\Sigma}\mathbf{m} - (\mathbf{m}^T\boldsymbol{\Sigma}\mathbf{z})^2\right]^{1/2} z_0,$$

where $\zeta \sim \chi^2_n$ and $z_0 \sim \mathcal{N}(0,1)$; ζ, z_0, and $\mathbf{z}$ are mutually indep.

Characteristic function

Theorem 2[Bodnar et al., 2016]

The characteristic function of $\mathbf{Az}$ is given by

$$\begin{aligned}
\varphi_{\mathbf{Az}}(\mathbf{u}) &= \frac{\exp\left(-\frac{\kappa^{-1}}{2}\boldsymbol{\mu}^T\mathbf{R}\boldsymbol{\Lambda}^{-1}\mathbf{R}^T\boldsymbol{\mu}\right)}{\kappa^{r/2}|\boldsymbol{\Lambda}|^{1/2}}\int_0^\infty |\boldsymbol{\Omega}|^{-1/2} f_{\chi^2_n}(\zeta) \\
&\times \exp\left(i\zeta\boldsymbol{\nu}^T\boldsymbol{\Lambda}\mathbf{R}^T\mathbf{u} - \frac{\zeta^2}{2}\mathbf{u}^T\mathbf{R}\boldsymbol{\Lambda}\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}\mathbf{R}^T\mathbf{u} + \frac{1}{2}\boldsymbol{\nu}^T\boldsymbol{\Omega}\boldsymbol{\nu}\right)\mathrm{d}\zeta,
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\Omega} &= \boldsymbol{\Omega}(\zeta) = \kappa^{-1}\boldsymbol{\Lambda}^{-1} + \zeta\left[\boldsymbol{\Lambda}\cdot\mathbf{u}^T\boldsymbol{\Sigma}\mathbf{u} - \boldsymbol{\Lambda}\mathbf{R}^T\mathbf{u}\mathbf{u}^T\mathbf{R}\boldsymbol{\Lambda}\right], \\
\boldsymbol{\nu} &= \boldsymbol{\nu}(\zeta) = \kappa^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{R}^T\boldsymbol{\mu},
\end{aligned}$$

$\boldsymbol{\Sigma} = \mathbf{R}\boldsymbol{\Lambda}\mathbf{R}^T$ singular value decomposition of $\boldsymbol{\Sigma}$, with matrix $\boldsymbol{\Lambda}$ consisting of all r non-zero eigenvalues of $\boldsymbol{\Sigma}$ and $\mathbf{R}$ is the $k \times r$ matrix of the corresponding eigenvectors.

Asymptotic distribution

(A1) $(\lambda_i, \boldsymbol{u}_i)$ denote the set of non-zero eigenvalues and the corresponding eigenvectors of $\boldsymbol{\Sigma}$. There exist l_1 and L_1 such that $0 < l_1 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_r \leq L_1 < \infty$ uniformly on k.

(A2) There exists L_2 such that $|\boldsymbol{u}_i^T \boldsymbol{\mu}| \leq L_2$ for all $i = 1, \ldots, r$ uniformly on k.

Asymptotic distribution(cont...)

Theorem 3[Bodnar et al., 2016]

Assume $\frac{r}{n} = c + o(n^{-1/2}), c \in [0, 1)$ and $\kappa r = O(1)$ as $n \to \infty$. Let $\mathbf{m}$ be a k-dimensional vector of constants s.t $\mathbf{m}^T \boldsymbol{\Sigma} \mathbf{m} > 0$ and $|\boldsymbol{u}_i^T \mathbf{m}| \leq L_2$ for all $i = 1, \dots, r$ uniformly on k. Let $P(\mathbf{z}^T \boldsymbol{\Sigma} \mathbf{z} = 0) = 0$. Then, under (A1) and (A2) it holds that the asymptotic distribution of $\mathbf{m}^T \mathbf{A} \mathbf{z}$ is given by

$$\sqrt{n}\sigma^{-1} \left(\frac{1}{n} \mathbf{m}^T \mathbf{A} \mathbf{z} - \mathbf{m}^T \boldsymbol{\Sigma} \boldsymbol{\mu} \right) \xrightarrow{d} \mathcal{N}(0, 1),$$

where

$$\sigma^2 = \left(\mathbf{m}^T \boldsymbol{\Sigma} \boldsymbol{\mu} \right)^2 + \mathbf{m}^T \boldsymbol{\Sigma} \mathbf{m} \left[\kappa tr(\boldsymbol{\Sigma}^2) + \boldsymbol{\mu}^T \boldsymbol{\Sigma} \boldsymbol{\mu} \right] + \frac{\kappa}{c} \mathbf{m}^T \boldsymbol{\Sigma}^3 \mathbf{m}.$$

Asymptotic distribution(cont...)

Theorem 4[Bodnar et al., 2016]

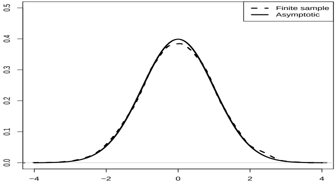
Assume $\frac{r}{n} = c + o(n^{-1/2}), c \in [0, 1)$ and $\kappa r = O(1)$ as $n \to \infty$. Let $\mathbf{M} = (\mathbf{m}_1, \ldots, \mathbf{m}_p)^T : p \times k$ be a matrix of constants of rank p such that $rank((\mathbf{M}^T, \mathbf{z})^T \boldsymbol{\Sigma}) = p + 1 \leq r$ with probability one and let $|\boldsymbol{u}_i^T \mathbf{m}_j| \leq L_2$ for all $i = 1, \ldots, r$ and $j = 1, \ldots, p$ uniformly on k. Then under (A1) and (A2) the asymptotic distribution of $\mathbf{MAz}$ under the double asymptotic regime is given by

$$\sqrt{n}\boldsymbol{\Omega}^{-1/2}\left(\frac{1}{n}\mathbf{MAz} - \mathbf{M}\boldsymbol{\Sigma}\mathbf{z}\right) \xrightarrow{d} \mathcal{N}(0, \mathbf{I})$$

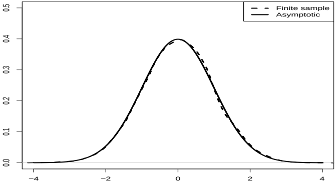
where

$$\boldsymbol{\Omega} = \mathbf{M}\boldsymbol{\Sigma}\boldsymbol{\mu}\boldsymbol{\mu}^T\boldsymbol{\Sigma}\mathbf{M}^T + \mathbf{M}\boldsymbol{\Sigma}\mathbf{M}^T\left[\kappa tr(\boldsymbol{\Sigma}^2) + \boldsymbol{\mu}^T\boldsymbol{\Sigma}\boldsymbol{\mu}\right] + \frac{\kappa}{c}\mathbf{M}\boldsymbol{\Sigma}^3\mathbf{M}^T.$$

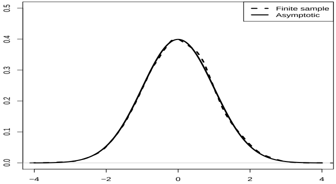
Finite sample performance



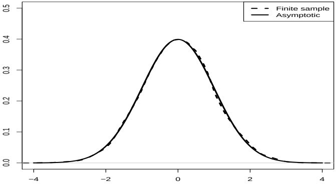
(a) $k = 750, c = 0.1$



(b) $k = 750, c = 0.5$



(c) $k = 750, c = 0.8$



(d) $k = 750, c = 0.95$

Possible extension

- Mean variance efficiency of TP in high dimension

Reference

Bodnar, T., Mazur, S., Muhinyuza, S., and Parolya, N. (2016).
On the product of a singular wishart matrix and a singular gaussian vector in high dimension.
arXiv preprint arXiv:1611.03042.

Bodnar, T., Mazur, S., and Okhrin, Y. (2013).
On the exact and approximate distributions of the product of a wishart matrix with a normal vector.
Journal of Multivariate Analysis, 122:70–81.

Bodnar, T., Mazur, S., and Okhrin, Y. (2015).
Distribution of the product of a singular wishart matrix and a normal vector.
Theory of Probability and Mathematical Statistics, 91:1–15.

Bodnar, T., Mazur, S., and Okhrin, Y. (2017).
Bayesian estimation of the global minimum variance portfolio.
European Journal of Operational Research, 256:292–307.

Bodnar, T. and Okhrin, Y. (2011).
On the product of inverse wishart and normal distributions with applications to discriminant analysis and portfolio theory.
Scandinavian Journal of Statistics, 38(2):311–331.

Bodnar, T. and Schmid, W. (2008).
Estimation of optimal portfolio compositions for gaussian returns.
Statistics & Decisions, 26(3):179–201.

Bodnar, T. and Schmid, W. (2009).
Econometrical analysis of the sample efficient frontier.
The European journal of finance, 15(3):317–335.

Bodnar, T. and Zabolotskyy, T. (2017).
How risky is the optimal portfolio which maximizes the sharpe ratio?

Tack så mycket!

Thank you!