Optimal Shrinkage Estimators for Large Covariance Matrices

Nestor Parolya
Leibniz Universität Hannover, Institute of Empirical Economics

Taras Bodnar
Humboldt Universität zu Berlin, Department of Mathematics

Arjun K. Gupta
Bowling Green State University, Department of Mathematics and Statistics

Random Matrix Theory, Kraków 03.07.2014
Outline

Introduction

Large Dimensional Data Analysis (Motivation)
Assumptions

Theoretical Findings and Applications
Large Covariance Matrices


Asymptotics

Two main types of asymptotics in multivariate statistics:

- **standard asymptotics**
  - fixed dimension $p$ and large sample size $n \to \infty$;
  - classical limit theorems hold

- **large dimensional asymptotics**
  - both the dimension $p$ and the sample size $n$ tend to infinity;
  - the ratio $p/n$ tends to a positive constant $c > 0$;
  - classical limit theorems do not hold anymore (the curse of dimensionality).
Assumptions

Define the $p \times n$ matrix $X_n$ which contains independent and identically distributed (i.i.d.) real random variables with zero mean and unit variance such that

$$Y_n \overset{d}{=} \Sigma_n^{1/2} X_n + \mu_n 1_n', \quad (1)$$

where $\Sigma_n$ is the covariance matrix and $Y_n$ is called the observation matrix.

The corresponding sample covariance matrix is then given by

$$S_n = \frac{1}{n} (Y_n - \bar{y}_n 1_n') (Y_n - \bar{y}_n 1_n')' = \frac{1}{n} Y_n Y_n' - \bar{y}_n \bar{y}_n' \quad (2)$$

with the sample mean vector given by

$$\bar{y}_n = \frac{1}{n} Y_n 1_n. \quad (3)$$
Assumptions contd.

- **(A1)** The covariance matrix of asset returns $\Sigma_n$ is nonrandom $p$-dimensional positive definite matrix.

- **(A2)** Only the matrix $Y_n = \frac{1}{\sqrt{n}} X_n$ is an observable one.

- **(A3)** The elements of the matrix $X_n$ have uniformly bounded $4 + \varepsilon$ moments.

- **(A4)** The largest eigenvalue of the covariance matrix $\Sigma_n$ is at most of the order $O(\sqrt{p})$. Moreover, we assume that the order of only finite number of eigenvalues could depend on $p$. 

Random Matrix Theory, Kraków 03.07.2014
Theoretical Findings and Applications

Strong Convergence of the Frobenius norm of $S_n$

Theorem

[Bodnar, Gupta and Parolya (2013a)] Under assumptions (A1)-(A4) for $\frac{p}{n} \to c \in (0, +\infty)$ it holds that

$$\frac{1}{p} \left\| S_n \right\|^2_F - \left[ \left\| \Sigma_n \right\|^2_F + \frac{c}{p} \left\| \Sigma_n \right\|_{tr} \right] \longrightarrow 0 \quad a.s. \ for \ n \to \infty$$ \hspace{1cm} (4)

where $\left\| \Sigma \right\|^2_F = tr(\Sigma^2_n)$ and $\left\| \Sigma \right\|^2_{tr} = \left( tr(\Sigma_n) \right)^2$ are the squared Frobenius and trace norms.

Additionally, let the matrix $\Theta$ be a symmetric positive definite matrix with bounded trace norm, it holds that

$$| tr(S_n \Theta) - tr(\Sigma_n \Theta) | \longrightarrow 0 \quad a. s. \ for \ \frac{p}{n} \to c \in (0, +\infty).$$ \hspace{1cm} (5)
Strong Convergence of the Frobenius norm of $S_n^{-1}$

Theorem

[Bodnar, Gupta and Parolya (2013b)] Let the assumptions (A1)-(A4) hold and $\frac{p}{n} \to c \in (0, 1)$. Then as $n \to \infty$,

$$\frac{1}{p} \left\| S_n^{-1} \right\|_F^2 - \left[ \frac{1}{(1-c)^2} \left\| \Sigma_n^{-1} \right\|_F^2 + \frac{c}{p(1-c)^3} \left\| \Sigma_n^{-1} \right\|_{tr}^2 \right] \xrightarrow{a.s.} 0. \quad (6)$$

Additionally, let $\Theta$ be a symmetric positive definite matrix with uniformly bounded trace norm as $n \to \infty$ the norm

$$\left| tr(S_n^{-1} \Theta) - \frac{1}{1-c} tr(\Sigma_n^{-1} \Theta) \right| \xrightarrow{a.s.} 0 \quad \text{for} \quad \frac{p}{n} \to c \in (0, 1). \quad (7)$$
General Shrinkage Estimator for Covariance Matrix

The general linear shrinkage estimator (GLSE) for the covariance matrix is given by

$$\hat{\Sigma}_{GLSE} = \alpha_n S_n + \beta_n \Sigma_0 \quad \text{with} \quad \|\Sigma_0\|_{tr} \leq M. \quad (8)$$

where the symmetric positive definite matrix $\Sigma_0$ has bounded trace norm at infinity, i.e., there exists $M > 0$ such that

$$\sup_n \|\Sigma_0\|_{tr} = \sup_n \text{tr}(\Sigma_0) \leq M.$$
Optimization problem

Aim: find the optimal shrinkage intensities $\alpha_n$ and $\beta_n$ which minimize the Frobenius norm for a given nonrandom target matrix $\Sigma_0$

$$L^2_F = ||\hat{\Sigma}_{GLSE} - \Sigma_n||_F^2 = ||\Sigma_n||_F^2 + ||\hat{\Sigma}_{GLSE}||_F^2 - 2\text{tr}(\hat{\Sigma}_{GLSE}\Sigma_n).$$

The optimal shrinkage intensities $\alpha_n^*$ and $\beta_n^*$ are given by

$$\alpha_n^* = \frac{\text{tr}(S_n\Sigma_n)||\Sigma_0||_F^2 - \text{tr}(\Sigma_n\Sigma_0)\text{tr}(S_n\Sigma_0)}{||S_n||_F^2||\Sigma_0||_F^2 - (\text{tr}(S_n\Sigma_0))^2},$$

$$\beta_n^* = \frac{\text{tr}(\Sigma_n\Sigma_0)||S_n||_F^2 - \text{tr}(S_n\Sigma_n)\text{tr}(S_n\Sigma_0)}{||S_n||_F^2||\Sigma_0||_F^2 - (\text{tr}(S_n\Sigma_0))^2}.$$
Asymptotics of Optimal Shrinkage Intensities

Proposition (Bodnar, Gupta and Parolya(2013a))

Assume that (A1)-(A4) are fulfilled. Then for $\frac{p}{n} \to c \in (0, +\infty)$ as $n \to \infty$ the optimal shrinkage intensities $\alpha_n^*$ and $\beta_n^*$ satisfy

\[ |\alpha_n^* - \alpha^*| \to 0 \quad \text{a.s. for } n \to \infty \tag{12} \]

with

\[ \alpha^* = 1 - \frac{c}{p} \frac{||\Sigma||^2_{tr} ||\Sigma_0||^2_F}{(||\Sigma_n||^2_F + c \frac{||\Sigma_n||^2_{tr}}{p} ||\Sigma_0||^2_F - (tr(\Sigma_n \Sigma_0))^2) ||\Sigma_0||^2_F} \] \tag{13} 

and

\[ |\beta_n^* - \beta^*| \to 0 \quad \text{a.s. for } n \to \infty \tag{14} \]

with

\[ \beta^* = \frac{tr(\Sigma_n \Sigma_0)}{||\Sigma_0||^2_F} (1 - \alpha^*) \] \tag{15}
The theoretical findings and applications of large covariance matrices are discussed. The optimal linear shrinkage estimator (OLSE) for the covariance matrix $\Sigma_n$ is given by (ct. Bodnar, Gupta and Parolya(2013a))

$$\hat{\Sigma}_{OLSE} = \hat{\alpha}^* S_n + \hat{\beta}^* \Sigma_0 \quad \text{with} \quad \|\Sigma_0\|_{tr} \leq M,$$

where

$$\hat{\alpha}^* = 1 - \frac{1}{n} \frac{\|S_n\|_{tr}^2 \|\Sigma_0\|_F^2}{\|S_n\|_F^2 \|\Sigma_0\|_F^2 - (\text{tr}(S_n \Sigma_0))^2}, \quad \text{(17)}$$

and

$$\hat{\beta}^* = \frac{\text{tr}(S_n \Sigma_0)}{\|\Sigma_0\|_F^2} \left(1 - \hat{\alpha}^*\right). \quad \text{(18)}$$

This is the bona fide estimator for the covariance matrix $\Sigma_n$. 

Random Matrix Theory, Kraków 03.07.2014
The optimal linear shrinkage estimator (OLSE) for the precision matrix is given by (ct. Bodnar, Gupta and Parolya (2013b))

$$\hat{\Pi}_{OLSE} = \hat{\alpha}_n S_n^{-1} + \hat{\beta}_n \Pi_0 \quad \text{with} \quad \| \Pi_0 \|_{tr} \leq M,$$

(19)

where

$$\hat{\alpha}_n = 1 - p/n - \frac{1}{n} \frac{\| S_n^{-1} \|_F^2 \| \Pi_0 \|_F^2}{\| S_n^{-1} \|_F^2 \| \Pi_0 \|_F^2 - (\text{tr}(S_n^{-1} \Pi_0))^2}.$$

(20)

and

$$\hat{\beta}_n = \frac{\text{tr}(S_n^{-1} \Pi_0)}{\| \Pi_0 \|_F^2} (1 - p/n - \hat{\alpha}_n).$$

(21)
Percentage Relative Improvement in Average Loss (PRIAL)

For an arbitrary estimator of the covariance matrix, \( \hat{M} \), the PRIAL (Percentage Relative Improvement in Average Loss) is defined as

\[
PRIAL(\hat{M}) = \left( 1 - \frac{E\| \hat{M} - \Sigma_n \|_F^2}{E\| S_n - \Sigma_n \|_F^2} \right) \cdot 100\%.
\] (22)

For an arbitrary estimator of the precision matrix, \( \hat{N} \), the PRIAL is defined by

\[
PRIAL(\hat{N}) = \left( 1 - \frac{E\| \hat{N} - \Sigma_n^{-1} \|_F^2}{E\| S_n^{-1} - \Sigma_n^{-1} \|_F^2} \right) \cdot 100\%.
\] (23)
OLSE for Covariance Matrix. Normal Distribution

Random Matrix Theory, Kraków 03.07.2014
OLSE for Precision Matrix. Normal Distribution

Random Matrix Theory, Kraków 03.07.2014
References


THANK YOU FOR YOUR ATTENTION!