AL KUT JOURNAL OF ECONOMIC AND ADMINISTRATIVE SCIENCES
Publisher: College of Economics and Management - Wasit University

Nonparametric Shrinkage Estimator for Covariance Matrix Under Heterogeneity and High Dimensions Conditions

Ahmed Mahdi Salih
Lecturer
Statistics Department
College of Administration and Economics
Wasit University
amahdi@uowasit.edu.iq

Abstract

In this paper, we discuss different kinds of covariance matrix estimators and their behavior under the conditions of heterogeneity and high dimensions. Covariance matrix estimation that is well-conditioned matrix is very important procedure for many statistical applications which require that. Sometimes, the common estimator of covariance matrix - the sample covariance matrix- suffers from ill conditions and in many cases be invertible and without good qualities of estimator as dimensions of matrix go larger. Here, we view a shrinkage estimator for covariance matrix which is a combination of unbiased estimator and minimum variance estimator with different types of shrinkage factors parametric and non-parametric ones. Simulation study have been made by using Heterogeneous Autoregressive Process ***ARH(1)*** *as a structure covariance matrix for population, moreover, a comparison has been made among different types of covariance estimators by using minimum mean square errors* ***MMSE.***

Key Words: Nonparametric, Shrinkage, Covariance Matrix

Introduction

Estimating covariance matrix is a main task for many statistical applications and problems. Lately, estimating covariance matrix under high dimensions condition attracts many researchers to find new estimators or improve the old ones. In the cases; such as high dimensions the common estimator of covariance – The Sample Covariance Matrix- [7]

$$S = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})\,(x_i - \bar{x})' \qquad \dots (1)$$

Where x_i represents observation from p-dimensional distribution vector, and $\bar{x}$ is the p-dimensional mean vector. This estimator under high dimensions suffers from ill conditions such as the differences between the large and small eigenvalues grows to infinity and thus makes it a low quality estimator.
High dimensions challenges give the researchers the intensive interest to develop new techniques to estimate the covariance matrix such as robust and shrinkage and nonparametric estimators. Shrinkage estimation for covariance matrix was first preseted by [9]Stein (1956) and then it was improved by many other authors such as [1]Bai and Yin (1993) , [2]Bickel and Levina (2008), [8]Ledoit and Wolf (2004) and [10]Touloumis (2014).

Shrinkage Estimator

A common improve for Stein type estimator under high dimensions condition presented by [4]Efron (1975) and [5] Efron and Morris (1975) which describe the shrinkage estimator as a linear combination between unbiased estimator with another estimator of minimum variance.

$$\hat{\Sigma} = (1-\lambda)S + \lambda T \qquad .. (2)$$

Where $\hat{\Sigma}$ is the shrinkage estimator for covariance matrix, and λ is the shrinkage intensity, and T is the shrinkage target $T = \mu I_p = \frac{tr(\Sigma)}{p}$. The theoretical idea of shrinkage technique is to shrink the eigenvalues of $\boldsymbol{S}$ to the eigenvalues of $\boldsymbol{T}$. The problem of estimating is to choose shrinkage intensity $\lambda \in (0,1)$which minimizes risk function $E\left\{\left\|\hat{\Sigma} - \Sigma\right\|_F^2\right\} = tr(\hat{\Sigma} - \Sigma)^2$

which is a quadratic Frobenius distance. [8]Lediot and Wolf present shrinkage intensity by rewriting the risk function as following:-

$$
\begin{aligned}
E\left\{\left\|\hat{\Sigma}-\Sigma\right\|_F^2\right\} &= E\{\|(1-\lambda)S+\lambda T-\Sigma\|_F^2\} \\
&= \lambda^2 E\{\|\Sigma-\mathrm{T}\|_F^2\}+(1-\lambda)^2 E\{\|S-\Sigma\|_F^2\}
\end{aligned}
$$

And by increasing the matrix dimensions $E(S)=\Sigma$ and by take the derivative to λ and make it equal to zero, as following:-

$$2\lambda E\{\|\Sigma-T\|_F^2\}-2(1-\lambda)E\{\|S-\Sigma\|_F^2\}$$

That will lead to

$$\lambda=\frac{E\{\|S-\Sigma\|_F^2\}}{E\{\|S-\Sigma\|_F^2\}+E\{\|\Sigma-T\|_F^2\}}$$

[8]Lediot and wolf presented some explanation for the value of the denominator as following:-

$$
\begin{aligned}
E\{\|S-T\|_F^2\} &= E\{\|S-\Sigma+\Sigma-T\|_F^2\} \\
&= E\{\|S-\Sigma\|_F^2\}+E\{\|\Sigma-T\|_F^2\}+2\langle E[S-\Sigma],\Sigma-T\rangle
\end{aligned}
$$

Recalling that $E(S)=\Sigma$ then the last term in the above equation will be equal to zero thus

$$E\{\|S-T\|_F^2\}= E\{\|S-\Sigma\|_F^2\}+E\{\|\Sigma-T\|_F^2\}$$

Hence,the shrinkage intensity will be

$$\lambda=\frac{E\{\|S-\Sigma\|_F^2\}}{E\{\|S-T\|_F^2\}} \qquad \ldots (3)$$

Hence, the shrinkage intensity selection is an important part of covariance matrix estimation therefore many authors change the shrinkage target matrix to improve shrinkage estimators in case of large dimensions.

Fisher and Sun Estimator

This parametric estimator was presented by [6]Fisher and Sun which depended on [7]Ledoit and Wolf definition of shrinkage intensity in equation (3) but first let us recall some basic definitions.

$$E\{Tr(S)\} = Tr(\Sigma)$$
$$E\{Tr(S^2)\} = \frac{n+1}{n} Tr(\Sigma^2) + \frac{1}{n} Tr^2(\Sigma) \qquad \dots (4)$$
$$E\{Tr^2(S)\} = Tr^2(\Sigma) + \frac{2}{n} Tr(S^2)$$

By expanding the numerator in (3) and using results in equations (4) we have

$$E\{\|S-\Sigma\|_F^2\} = E\{\|S\|_F^2\} - 2E\{\langle S,\Sigma\rangle\} + \frac{Tr(\Sigma^2)}{p}$$
$$= \frac{n+1}{np} Tr(\Sigma^2) + \frac{1}{np} Tr^2(\Sigma) - \frac{2}{p} Tr(\Sigma^2) + \frac{1}{p} Tr(\Sigma^2)$$

For making things simpler, we use $b_j = \frac{1}{p} Tr(\Sigma^j)$

$$= \frac{n+1}{n} b_2 + \frac{p}{n} b_1^2 - 2b_2 + b_2$$
$$= \frac{1}{n} b_2 + \frac{p}{n} b_1^2$$

And by expanding the denominator in (3) we get:-

$$E\{\|S-T\|_F^2\} = E\{\|S\|_F^2\} - 2\mu E\{\langle S,\mathrm{I}\rangle\} + \mu^2 \|I\|^2$$
$$= \frac{n+1}{np} Tr(\Sigma^2) + \frac{1}{np} Tr^2(\Sigma) - \frac{2\mu}{p} Tr(\Sigma) + b_1^2$$
$$= \frac{n+1}{n} b_2 + \frac{p}{n} b_1^2 - 2b_1^2 + b_1^2$$
$$= \frac{n+1}{n} b_2 + \frac{p-n}{n} b_1^2$$

Then, the Fisher and Sun shrinkage intensity and covariance shrinkage estimator will be:-

$$\lambda_{FS} = \frac{\frac{1}{n}b_2+\frac{p}{n}b_1^2}{\frac{n+1}{n}b_2+\frac{p-n}{n}b_1^2} \qquad \dots (5)$$

$$\hat{\Sigma}_{FS} = (1-\lambda_{FS})S + \lambda_{FS}T \qquad \dots (6)$$

And by changing the shrinkage target $T = I$ the denominator of the shrinkage intensity in (3) will expand like

$$\begin{aligned} E\{\|S-I\|_F^2\} &= E\{\|S\|_F^2\} - 2E\{\langle S,I\rangle\} + \|I\|_F^2 \\ &= \frac{n+1}{np}Tr(\Sigma^2) + \frac{1}{np}Tr^2(\Sigma) - \frac{2}{p}Tr(\Sigma) + 1 \\ &= \frac{n+1}{n}b_2 + \frac{p}{n}b_1^2 - 2b_1 + 1 \end{aligned}$$

Then in this case the Fisher and Sun shrinkage intensity and covariance shrinkage estimator will be

$$\lambda_{FS2} = \frac{\frac{1}{n}b_2+\frac{p}{n}b_1^2}{\frac{n+1}{n}b_2+\frac{p}{n}b_1^2-2b_1+1} \qquad \dots (7)$$

$$\hat{\Sigma}_{FS2} = (1-\lambda_{FS2})S + \lambda_{FS2}I \qquad \dots (8)$$

Nonparametric Shrinkage Estimator

Under the large dimensions of the covariance matrix [10]Touloumis presented a nonparametric shrinkage estimator improving Stien type of shrinkage estimators. In the parametric estimation the expanding of the numerator and the denominator of the shrinkage intensity in equation (3) depending on the multivariate normal distribution assumption, so here a new estimator has been presented according to the following nonparametric model.

Let $X, X_2, \dots\dots X_n$ be a p-dimensional random vectors and

$$X_i = \Sigma^{\frac{1}{2}}Z_i + \mu \qquad \dots (9)$$

Where $\Sigma = cov[X_i] = \Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}}$ is a $p \times p$ dimensional matrix and Z_i be p-dimensional random vectors and instead of putting a distributional assumption here we use restrictions on Z_i concerning with moments. Let Z_{ib} be the ib_{th} random variable in Z_i with $E[Z_{ib}] = 0$ and $E\left[{Z_{ib}}^2\right] = 1$ and let $E\left[{Z_{ib}}^4\right] = 3 + G$ where $-2 < G < \infty$. Here the nonparametric model in equation (9) include p-dimensional multivariate normal distribution in Z_i so G can be used as a measure of departure from the fourth moment of Z_i to the multivariate normal distribution.

The expanding of the numerator and denominator in equation (3) will be according to the nonparametric model in equation (9) the author has put some basic definitions according to the nonparametric model as following:-

$$E\{Tr(S)\} = Tr(\Sigma)$$

$$E\{Tr(S^2)\} = \frac{n}{n-1}Tr(\Sigma^2) + \frac{1}{n-1}Tr^2(\Sigma) + \frac{G}{n-1}Tr(D_\Sigma^2) \quad \dots (10)$$

$$E\{Tr^2(S)\} = Tr^2(\Sigma) + \frac{2}{n-1}Tr(\Sigma^2) + \frac{G}{n-1}Tr(D_\Sigma^2)$$

Where D_Σ is the diagonal matrix of eigenvalues of Σ then by using results in equations (10), the shrinkage intensity in equation (3) could expand to be

$$\lambda_{NB} = \frac{Tr(\Sigma^2)+Tr^2(\Sigma)+G\ Tr(D_\Sigma^2)}{n\ Tr(\Sigma^2)+\frac{p-n+1}{p}Tr^2(\Sigma)} \quad \dots (11)$$

The nonparametric shrinkage estimator for covariance matrix will be

$$\hat{\Sigma}_{NB} = (1-\lambda_{NB})S + \lambda_{NB}T \quad \dots (12)$$

Simulation Study

We present here a simulation by using $p = 50$ and different sample sizes and we use Heterogeneous Autoregressive Process from the first order ARH(1) [] to generate covariance matrix for population $\sigma_i \sigma_j \rho^{|i-j|} = \Sigma$ where $\sigma_i \sigma_j$ will be generated from standard normal distribution and $\rho \in (0,1)$ is the autocorrelation factor We suggest two values for the autocorrelation factor 0.35 , 0.75 and make comparison among different estimators of the covariance matrix depending on the risk function or the quadratic frobenius distance to be the Minimum mean Square Error in the form [6] $MMSE = Tr\left(\hat{\Sigma} - \Sigma\right)^2$ by using MATLAB we replicate the experiment 1000 times.

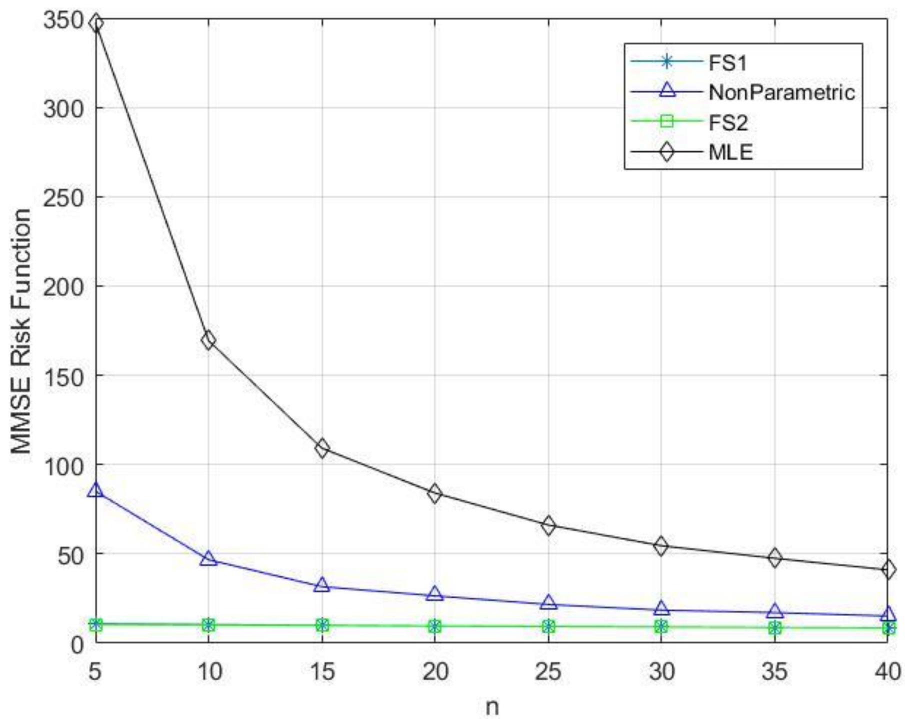

Figure (1) MMSE for the Estimators $\rho = 0.35$

n	FS1	Nonparametric	FS2	MLE
5	10.8499	85.1117	10.5029	347.3563
10	10.8499	46.7562	10.1697	169.3461
15	10.8499	31.6614	9.8723	109.2322
20	10.8499	26.5375	9.5301	84.0571
25	10.8499	21.7175	9.2910	66.1971
30	10.6134	18.5234	9.0581	54.5597
35	9.8862	17.1506	8.7924	47.5355
40	9.3673	15.2480	8.5870	41.0693

Table (1) the values of MMSE for estimators $\rho = 0.35$

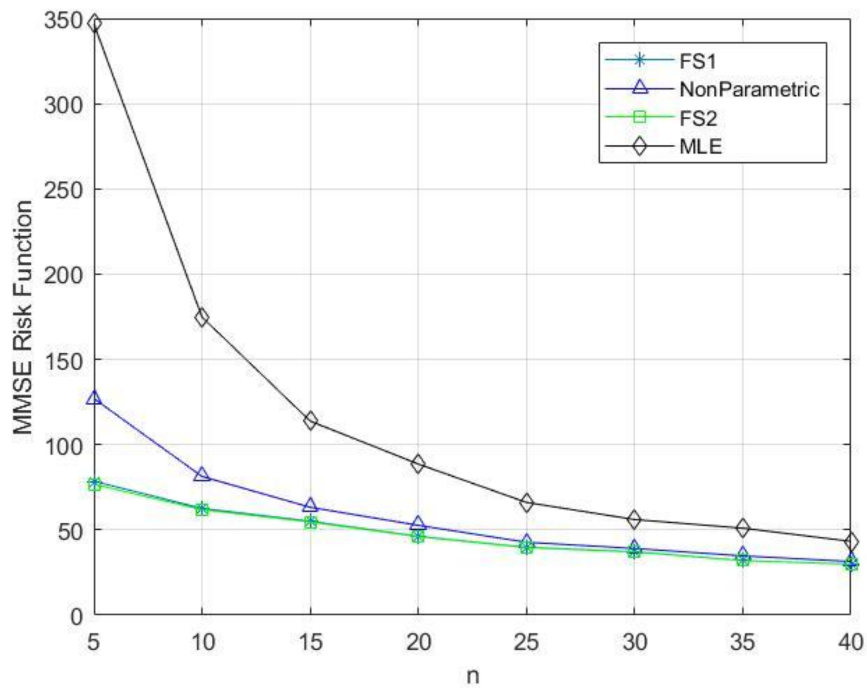

Figure (2) MMSE for the Estimators $\rho = 0.75$

n	FS1	Nonparametric	FS2	MLE
5	82.2657	127.0119	76.7242	346.7863
10	62.7013	81.3294	61.8817	174.7821
15	54.8423	63.2621	54.6135	113.7959
20	46.0324	52.6706	45.9980	88.5948
25	39.6482	42.7299	39.5659	66.0964
30	37.0793	39.0999	37.0300	56.0535
35	31.9125	34.7397	31.9325	51.0080
40	29.8318	31.4532	29.8279	43.2726

Table (2) the values of MMSE for estimators $\rho = 0.75$

Conclusions

From simulation study in Table 1,2 We can make conclusions that when sample size is very small comparing to matrix dimensions, Fisher and Sun estimators were the best to estimate covariance matrix. And when sample size be close to the matrix dimension, We can notice that the nonparametric estimator work fine and be as good as Fisher and Sun Estimators, while MLE estimator has no good performance at all.

References

[1]. Bai, P . Yin, Q. 1993. Limit of the smallest eigenvalues of a large-dimensional Sample covariance matrix . Ann. Prob. 21(3), 1275-1294.

[2]. Bickel, P. Levina, E. 2008. Regularized estimation of large covariance matrices. Ann. Statist. 36(1), 199-227.

[3]. Chen, Y. Wiesel, A. Hero, A. 2009. Shrinkage estimation of high dimensional covariance matrices. IEEE International Conference on 0. pp. 2937-2940.

[4]. Efron, B. 1975. Biased versus unbiased estimation. Advances in Math. 16, 259-277.

[5]. Efron, B. Morris, C. 1975. Data analysis using Stein's estimators and its generalizations. J. Amer. Statst. Assoc. 70(350), 311-319.

[6]. Fisher, T.Sun, X. 2010. Improved Stein-type shrinkage estimator for the high-dimensional multivariate normal covariance matrix. Comp. Statist. 55(2011), 1909-1918.

[7]. Ledoit, O. Wolf, M. 2004. Honey, I shrunk the sample covariance matrix. J. Portfolio Manage. 31(1), 110-119.

[8]. Ledoit, O. Wolf, M. 2004. A well-conditioned estimator for large-dimensional covariance matrices. J. Multivariate Anal. 88(2), 365-411.

[9]. Stein, C. 1956. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Proceeding of the Third Berkeley Symposium on Mathematical Statistics and Probability. 1954-1955, vol. I. University of California Press, Berkeley and Los Angeles, pp. 197-206.

[10]. Touloumis, A. 2014. Nonparametric Stein-type Shrinkage Covariance Matrix Estimators in High-Dimensional Settings. Computational Statistics & Data Analysis 83:251-261