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

Comparison Among Different Shrinkage Covariance Estimators Under Multicollinearity and High Dimensions Conditions

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

Keywords: Shrinkage, Covariance Estimation, Large p, Small n

Abstract

*Covariance matrix estimation is a very important process for many multivariate applications like canonical analysis and multivariate hypotheses testing. Many data conditions require unusual estimation for covariance matrix that be different from the sample covariance matrix because the last (latter) is very weak under conditions like multicollinearity and high dimensions. Here, we introduce a comparison among three kinds of covariance matrix estimators under multicollinearity and high dimension conditions. Three estimators were submitted for covariance matrix: the **Oracle** estimator(**OE**), **Chen** estimator **CE** and sample covariance estimator **MLE** under **Fractional Brownian motion FBM** structure covariance matrix to simulate the multicollinearity and the high dimensions conditions. A comparison was made by using Frobenius distance as a measure of goodness for estimators.*

Introduction

The estimator of covariance matrix plays a main role for many statistical issues. But (However), estimating covariance matrix under conditions like multicollinearity and high dimensions get (attract) the attention for many researchers to find good estimators or develop the old ones. The sample Covariance Matrix- [7]

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

Where x_i represents a p-dimensional distribution vector, and $\bar{x}$ is the p-dimensional mean vector. This estimator will be too weak and far away from the properties of good estimator such as the unbiasness and consistency and thus makes it not really good estimator. High dimension problems make researchers trying to develop new ways to estimate the covariance matrix such as robust, shrinkage and nonparametric estimators. The early Shrinkage estimator for covariance matrix was presented by [10]Stein (1956) and then developed by many authors such as [5]Efron (1975) , [1]Bai and Yin (1993) , [2]Bickel and Levina (2008), [8]Ledoit and Wolf (2004). A significant improvement to Stein estimator with high dimensions condition was presented by [5]Efron (1975) and [6] Efron and Morris (1975) as seen below/ as in the following equation:-

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

Where $\hat{\Sigma}$ represents shrinkage estimator for the covariance matrix, and U stands for **shrinkage intensity**, and F represents **shrinkage target** $F = \frac{tr(\Sigma)}{p}$ where p is the matrix dimension.

The idea of shrinkage estimation is to make the eigenvalues of $\boldsymbol{S}$ close to the eigenvalues of $\boldsymbol{F}$. To estimate the covariance matrix by shrinkage method we must choose shrinkage intensity $U \in (0,1)$ which minimizes risk function $E\left\{\left\|\hat{\Sigma} - \Sigma\right\|_F^2\right\} = tr(\hat{\Sigma} - \Sigma)^2$ which is a Frobenius distance. [8]Lediot and Wolf present shrinkage intensity as in the following equation:-

$$E\left\{\left\|\hat{\Sigma} - \Sigma\right\|_F^2\right\} = E\{\|(1-U)S + UF - \Sigma\|_F^2\}$$
$$= U^2 E\{\|\Sigma - F\|_F^2\} + (1-U)^2 E\{\|S - \Sigma\|_F^2\}$$

By increasing the matrix dimensions, the authors assume $E(S) = \Sigma$, and by taking the derivative to U and equalize it to zero[8], as in the following:-

$$2UE\{\|\Sigma - F\|_F^2\} - 2(1-U)E\{\|S - \Sigma\|_F^2\}$$

That will lead to

$$U = \frac{E\{\|S - \Sigma\|_F^2\}}{E\{\|S - \Sigma\|_F^2\} + E\{\|\Sigma - F\|_F^2\}}$$

[8]Lediot and Wolf presented some definitions for the value of the denominator as in the following:-

$$E\{\|S - F\|_F^2\} = E\{\|S - \Sigma + \Sigma - F\|_F^2\}$$
$$= E\{\|S - \Sigma\|_F^2\} + E\{\|\Sigma - F\|_F^2\} + 2\langle E(S - \Sigma), (\Sigma - F)\rangle$$

Recalling the assumption $E(S) = \Sigma$ then the last term will be equal to zero thus (as shown below):-

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

Hence, the shrinkage intensity will be

$$U = \frac{E\{\|S - \Sigma\|_F^2\}}{E\{\|S - F\|_F^2\}} \qquad \dots (3)$$

Thus, choosing the shrinkage intensity is a main part to estimate the covariance matrix, this is the reason why many researchers introduce many estimates for shrinkage intensity with huge collections of shrinkage target matrices.

Oracle Estimator

This nonlinear estimator was presented by [8]Lediot& Wolf. It is an extension for the shrinkage intensity in (2) which restricts the risk function so if we substitute equation (2) inside the risk function as in the following equation:-

$$E\left\{\left\|\hat{\Sigma}-\Sigma\right\|_F^2\right\} = E\{\|(1-U)S+UF-\Sigma\|_F^2\}$$
$$= E\{\|(S-\Sigma)-U(S-F)\|_F^2\}$$
$$= E\{\|S-\Sigma\|_F^2\} - 2UE\{\|\langle(S-\Sigma),(S-F)\rangle\|_F^2\}$$
$$+U^2E\left\{\left\|\hat{S}-F\right\|_F^2\right\}$$

And by taking derivative with respect to U and equalize to zero we get; -

$$2UE\left\{\left\|\hat{S}-F\right\|_F^2\right\} - 2E\{\|\langle(S-\Sigma),(S-F)\rangle\|_F^2\} = 0$$
$$U = \frac{E\{\|\langle(S-\Sigma),(S-F)\rangle\|_F^2\}}{E\left\{\left\|\hat{S}-F\right\|_F^2\right\}}$$

Then by defining the risk function, we get: -

$$U = \frac{E\{tr((S-\Sigma)(S-F))\}}{E\{tr(S-F)^2\}}$$

This estimation of the Shrinkage Intensity is an expectation, we can simplify it by using the following expectation results [7].

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

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

Then we can expand the expectations in the shrinkage intensity estimation so we can get a simple formula for it
As for the denominator, we get the following formula: -

$$E\{tr((S-\Sigma)(S-F))\} = E\{tr(S^2)\} - \frac{E\{tr^2(S)\}}{p} - E\{tr(\Sigma S)\} + \frac{tr(\Sigma)}{p} E\{tr(S)\}$$

and as for the numerator, we get the following formula: -

$$\begin{aligned} E\{tr(S-F)^2\} &= E\{tr(S^2)\} - 2E\{tr(SF)\} + E\{tr(F^2)\} \\ &= E\{tr(S^2)\} - \frac{E\{tr^2(S)\}}{p} \end{aligned}$$

Therefore, by using the results of expectations in (4), we get the following equation: -

$$U_{OE} = \frac{(1-{}^{2}/_{p})tr(\Sigma^2)+tr^2(\Sigma)}{(n+1-{}^{2}/_{p})tr(\Sigma^2)+(1-{}^{2}/_{p})tr^2(\Sigma)} \quad \ldots (5)$$

Hence, the Oracle estimator for the covariance matrix will be as the following

$$\hat{\Sigma}_{OE} = (1-U_{OE})S - U_{OE}F \quad \ldots (6)$$

Chen Estimator

This robust estimator for covariance matrix was presented by Chen [4] in case of high dimensions and under the assumptions of normal distribution, the author uses the Normalized Samples instead of real data directly $z_i = \frac{x_i}{\|x_i\|_F^2}$ then in that case the MLE estimator (Sample Covariance) of the covariance matrix will be like $C = \frac{p}{n}\sum_{i=1}^{n} \frac{z_i z_i'}{z_i' S z_i}$ so here we have Chen estimator as follows.

$$\Sigma_C = (1 - U_C)C + U_C I \quad \ldots (7)$$

The U_C represents the Chen shrinkage intensity which minimizes the risk function as we derive it we substitute the result in (7) inside the risk function we get:

$$E\{\|\Sigma_C - \Sigma\|_F^2\} = E\{\|(1 - U_C)C + U_C I - \Sigma\|_F^2\}$$

$$= E\{\|(C - \Sigma) - U_C(C - I)\|_F^2\}$$
$$= E\{\|C - \Sigma\|_F^2\} - 2U_C E\{\langle (C - \Sigma), (C - I)\rangle\} + U_C^2 E\{\|C - I\|_F^2\}$$

And by taking the first derivative and equalizing it to zero, we get the shrinkage intensity as in the following [4]

$$U_C = \frac{E\{tr(C^2)\} - E\{tr(C\Sigma)\} - E\{tr(\mathrm{C})\} + tr(\Sigma)}{E\{tr(C^2)\} - 2E\{tr(C)\} + p}$$

Here, we have three expectations that determine the shrinkage intensity such as: $E\{tr(C^2)\}$, ,$E\{tr(C\Sigma)\}$ and $E\{tr(C)\}$ [4]. By using decomposition theories, Chen puts the values of those three expectations as in the following

$$E\{tr(C^2)\} = \left(1 - \frac{1}{n} + \frac{2}{n\left(1 + {}^2/_p\right)}\right) tr(\Sigma^2) + \frac{tr^2(\Sigma)}{n\left(1 + {}^2/_p\right)}$$
$$E\{tr(C\Sigma)\} = tr(\Sigma^2)$$
$$E\{tr(\mathrm{C})\} = tr(\Sigma)$$

consequently, the final form of the shrinkage intensity will be like:

$$U_C = \frac{p^2 + \left(1 - {}^2/_p\right) tr(\Sigma^2)}{(p^2 - np - 2n) + \left(n + 1 + 2^{(n-1)}/_p\right) tr(\Sigma^2)} \quad \ldots (8)$$

which minimizes the shrinkage estimator of covariance matrix in (7)

Simulation Study

In the following section, we make a comparison among three estimators of covariance matrix: the sample covariance estimator in (1) , the Oracle estimator in (6) and Chen estimator in (7) under quadratic risk function

$tr(\hat{\Sigma} - \Sigma)^2$. Here, we select Σ to be the result of the increment fractional Brownian motion FBM as in the following [9].

$$\Sigma_{ij} = \frac{1}{2}[(|i-j|+1)^{2h} - 2|i-j|^{2h} + (|i-j|-1)^{2h}]$$

As we can see in the above equation, h is Hurst parameter $0.5 < h < 1$. In this paper, we choose three values of h which is 0.5 for the normal case , 0.7 for little value of autocorrelation and 0.9 for the higher case of autocorrelation condition and multi values of small sample sizes n and selected values as high dimensions p we select (10 , 20 ,30) as sample sizes and (50, 100 , 150) as covariance matrix dimensions and (0.5 , 0.7 , 0.9) as ***h*** value and calculate different matrix estimators as in (10 , (6) and (7) and replicate this experiment 1000 times by using MATLAB program.

			Risk Function		
h	*p*	*n*	*MLE*	*Oracle*	*Chen*
0.5	50	10	3.0661	0.1877	0.4412
		20	0.5731	0.1000	0.1921
		30	0.2744	0.0558	0.1208
	100	10	10.4785	0.2029	0.3662
		20	1.5288	0.0830	0.1964
		30	0.5620	0.0757	0.1327
	150	10	24.1726	0.1496	0.4201
		20	3.3883	0.1179	0.2467
		30	0.9712	0.0747	0.1316
0.7	50	10	25.3597	22.2345	22.9809
		20	21.2610	20.3385	20.9322
		30	19.4774	19.2392	19.4953
	100	10	61.9029	51.6173	52.5846
		20	50.6852	48.5998	49.3020
		30	47.3779	46.7650	47.1682
	150	10	106.1724	82.4768	83.3433
		20	82.3982	79.5426	80.0835
		30	78.0428	76.6464	76.1610
0.9	50	10	237.3167	206.6654	207.0329
		20	131.5094	119.4785	119.5319
		30	95.4083	87.2448	88.3404

	100	10	729.0392	641.4572	640.8629
		20	425.8982	394.1195	391.4799
		30	364.9119	328.8244	322.2903
	150	10	1608.987	1423.265	1415.898
		20	1058.587	965.0074	964.045
		30	743.4043	679.6709	644.1907

Values of risk function under different p and n

Conclusions

From the simulation study, we notice that in the normal circumstances, the Oracle estimator is the best estimator with least risk function value with *h=0.5*.

And in the case of a weak autocorrelation condition when *h=0.7*, we see that Oracle estimator is still better but it makes strong competition with Chen estimator which gets close to Oracle estimator as p goes larger.

In the case of high autocorrelation condition when *h=0.9*, we see that Chen estimator is still the best estimator of covariance matrix as p goes larger

This gives a) good evidence that robust estimators are better in the cases of higher autocorrelation and high dimensions conditions.

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]. Chen, Y. Wiesel, A. Hero, A.2011. Robust shrinkage estimation of high dimensions matrix. IEEE Transections on signal processing 59:4097-

4107.

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

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

[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]. Nualart, D. 2000. Fractional Brownian motion; stochastic calculus and application. Mathematica subject classification p60h30 , s60g18

[10]. 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.