Asymptotic Variance of Sample Vector Variance of Standardized Variables

Erna T. Herdiani 1,3, Maman A. Djauhari 2, Saleh A.F. 1, Annisa 1, Nasrah Sirajang 1
1 Hasanuddin University, Makassar, Indonesia
2 University Teknologi Malaysia, Johor Bahru, Malaysia
3 Corresponding author: Erna T. Herdiani, email : herdiani.erna@unhas.ac.id

Abstract

The stability of the correlation matrices is noteworthy. Usually to testing stability of correlation matrices used to statistics M Box, Jennrich and G. Its statistics However, M Box and G statistic involves matrix inversion. The former needs the condition that all sample correlation matrices are positive definite which is not always satisfied in practice. This condition is not apt for high dimension data sets because its computational efficiency becomes low. To handle this obstacles, we proposed a new statistical test based on what we call vector variance of standardized variables (VVSV). The proposed test is constucted based on vector variance (VV).This is evidenced by several papers describing the correlation matrix, Vector variance of standardized variables sample used a statistical formula variance vector. In practice there are difficulties in the calculation to determine variance of Vector Variance of Sample Variance of Standardized Variables. In this paper, by utilizing the vec operator and properties of the matrix will be investigated alternative formulation of asymptotic variance of Vector Variance of Sample Variance of Standardized Variables.

Keywords: vec operator, vector variance, covariance matrix, correlation matrix

1. Introduction

The stability of correlation matrix is one of the most important problems in economic development and financial industry. We can see in the literature that, since the last decade, that problem can be found in a wide spectrum of applications. Its applications in property, real estate and asset businesses can be seen, for example, in eichholtz (1995), Lee(1998), Cooka el al. (2002) and Fischer (2007). Other applications such as in equity market, global market and risk management are presented by Meric and meric (1997), Tang (1998), Gande and Parsley (2002), Annaert et al. (2003), Ragea (2003), Da costa et al. (2005) and Goetzmann et al. (2005), Michelle et al. (2010). We can also find its application in parallel computation of high dimensional robust correlation matrices in Chilson et al. (2006).

Djauhari and Herdiani (2010) proposed a new statistical test based on what we call vector variance of standardized variables (VVSV). The proposed test is constucted based on vector variance (VV). The role of VV could be increasing the computational efficiency of fast minimum covariance determinant (FMCD) algorithm, Herwindiati et al. (2008). Let P be the correlation matrix of the population under studi. Vector Variance of Standardized Variable (VVSV) is the trace of the squared correlation matrix population, i.e. \( WSV = Tr(P^2) = \|\text{vec}(P)\|^2 \).

In practice there are difficulties in calculating variance of vector variance of standardized variables. Therefore, in this paper we investigate about on asymptotic
variance of vector variance of standardized variables. The rest of the paper is organized as follows. In section 2, we discuss on the Asymptotic Variance of Sample Vector Variance of Standardized Variables. Our approach will be based on the notions of vec operator and commutation matrix the problem formulation will be presented. Later on, in section 3, Additional remarks will close this paper.

2. Asymptotic Variance of Sample Vector Variance of Standardized Variables

Let X is a random vector p dimension with definite positive covariance matrix \( \Sigma \). By using vec operator, see Muirhead (1982), El Maache and Lepage (1998), Schott (1997, 2007) and Djauhari (2007), vector variance (VV) of X is simply \( \| \text{vec}(\Sigma) \|^2 \). It is a multivariate variability measure like Generalized Variance (GV). See Djauhari (2007) for an in depth discussion on VV and its asymptotic behavior. In what follows our attention will be focused on the case where all variables are standardized.

Let \( Z \) be the random vector where its k-th component is the standardized version of the k-th component of X; \( k=1, 2, ..., p \). The covariance matrix \( P \) of \( Z \) is the correlation matrix of \( X \). We call the parameter \( \| \text{vec}(Z) \| \) vector variance of the standardized variables (VVSV). Now, let \( Z_1, Z_2, ..., Z_n \) be a random sample of size n from \( Z \) with covariance matrix \( P \). If \( R \) is the sample correlation matrix, then \( \| \text{vec}(R) \|^2 \) is a representation of \( R \) in vector form. Neudecker (1996) and Schott (2007) wrote about asymptotic distribution of \( \| \text{vec}(R) \|^2 \) as follows.

**Theorem 1**

Let \( \bar{X}_1, \bar{X}_2, ..., \bar{X}_n \) be a random vector of size n from the distribution normal p-variat, \( \bar{X} \sim N_p(\mu, \Sigma) \) then

\[
\sqrt{n} - 1 \{ \text{vec}(R) - \text{vec}(P) \} \xrightarrow{d} N_p^2(0, I)
\]

Where :

a. \( \Gamma = 2M_p \Phi M_p, \ M_p = \frac{1}{2} (I_p^2 + K_{pp}) \)

b. \( K_{pp} \) is commutation matrix of size \( p^2 \times p^2 \)

c. \( \Phi = \{ I_p^2 - (I_p \otimes P) \Lambda_p \} (P \otimes P) \{ I_p^2 - \Lambda_p (I_p \otimes P) \} \)

d. \( \Lambda_p = \sum_{i=1}^p (\bar{e}_i \bar{e}_i^T \otimes \bar{e}_i \bar{e}_i^T) \), \( \bar{e}_i \) is i-th column of identity matrix \( I_p \) of size \( p \times p \).

Value approximation of \( \text{vec}(R) \), as follows

\( \text{vec}(R) \approx \text{vec}(P + A - \frac{1}{2} (PD_A + D_A P)) \)

It’s approximation will be used to find mean and variance of \( \text{vec}(R) \). Mean of \( \text{vec}(R) \) is \( \text{vec}(P) \) and variance of \( \text{var}(\text{vec}(R)) \)

\[
\left[ M_p \left[ I_p - \frac{1}{2} (I_p \otimes P + P \otimes I_p) \right] \Lambda_p \right]
\]

Cause \( \left[ I_p^2 - \frac{1}{2} (I_p \otimes P) + (P \otimes I_p) \right] \Lambda_p \) \( M_p = M_p \left[ I_p^2 - (I_p \otimes P) \Lambda_p \right] \), and \( M_p \)
symmetry, see Schott (1997, p. 415 and 282), then
\[ \text{var}(\text{vec}(R)) \approx \frac{2}{n-1} M_p \left\{ I_{p^2} - (I_p \otimes P) \Lambda_p \right\} (P \otimes P) \left\{ I_{p^2} - \Lambda_p (I_p \otimes P) \right\} M_p. \]

Therefore,

\[ \{ I_{p^2} - (I_{p^2} \otimes P) \Lambda_p \} (P \otimes P) \{ I_{p^2} - \Lambda_p (I_{p^2} \otimes P) \} = \Phi. \]

We can get,

\[ \text{var}(\text{vec}(R)) \approx \frac{2}{n-1} M_p \Phi M_p. \]

Further, let \( \bar{Z}_1, \bar{Z}_2, \ldots, \bar{Z}_n \) be a random sample of size \( n \) from the distribution having covariance matrix \( P \). If \( R \) be a sample correlation matrix then vec(\( R \)) is a representation of \( R \) in vector form with dimension \( p^2 \). Thus, VVSV sample be a \( \| \text{vec}(R) \|^2 \) or Tr(\( R^2 \)). It is number of elements diagonal matrix of \( R^2 \), namely

\[ \text{Tr}(R^2) = \sum_{i=1}^{p} \sum_{j=1}^{p} r_{ij}^2. \]

Asymptotic distribution of \( \| \text{vec}(R) \|^2 \) can be written as follows.

**Theorem 2**

If \( \sqrt{n-1} (\text{vec}(R) - \text{vec}(P)) \overset{d}{\rightarrow} N_p(0, \Gamma) \) and \( (\text{vec}(R))^T (\text{vec}(R)) = \| \text{vec}(R) \|^2 \) then \( \| \text{vec}(R) \|^2 \overset{d}{\rightarrow} N(\eta, \theta^2) \), with mean \( \eta = \| \text{vec}(P) \|^2 \) and variance \( \theta^2 = \frac{8}{n-1} (\text{vec}(P))^T M_p \Phi M_p (\text{vec}(P)) \).

Where:
- \( P \) = Population correlation matrix
- \( R \) = Sample correlation matrix
- \( p \) = Number of variable
- \( M_p = \frac{1}{2} (I_{p^2} + K_{pp}) \), \( K_{pp} \) be a commutation matrix with size \( p^2 \times p^2 \)
- \( \Phi = \left\{ I_{p^2} - (I_p \otimes P) \Lambda_p \right\} (P \otimes P) \left\{ I_{p^2} - \Lambda_p (I_p \otimes P) \right\} \)

(Djaouhari & Erna 2008)

Formulation variance of VVSV was been caused difficulty to compute, therefore its need an alternative formulation by vec operator and commutation matrix to become simply. The result is written in theorem form, as follows.

**Proposition**

If \( \text{var}([\| \text{vec}(R) \|^2]) = \frac{8}{n-1} (\text{vec}(P))^T M_p \Phi M_p (\text{vec}(P)), \) then

\( (\text{vec}(P))^T M_p \Phi M_p (\text{vec}(P)) = \text{Tr}(P^4) - 2 \text{Tr}(D_{p^2} P^3) + \text{Tr}(D_{p^2} P^3)^2 \)

**Proof:**

\( (\text{vec}(P))^T M_p \Phi M_p (\text{vec}(P)) = \)

\[ = (\text{vec}(P)) \left[ \frac{1}{2} (I_{p^2} + K_{pp}) \right] \left\{ I_{p^2} - (I_p \otimes P) \Lambda_p \right\} (P \otimes P) \left\{ I_{p^2} - \Lambda_p (I_p \otimes P) \right\} \frac{1}{2} (I_{p^2} + K_{pp}) (\text{vec}(P)) \]

\[ = \frac{1}{4} (\text{vec}(P)) \left[ I_{p^2} + K_{pp} \right] \left\{ I_{p^2} - (I_p \otimes P) \Lambda_p \right\} (P \otimes P) \left\{ I_{p^2} - \Lambda_p (I_p \otimes P) \right\} (I_{p^2} + K_{pp}) (\text{vec}(P)) \]

\[ = \frac{1}{4} \left( (\text{vec}(P))^T I_{p^2} + (\text{vec}(P))^T K_{pp} \right) \left\{ I_{p^2} - (I_p \otimes P) \Lambda_p \right\} (P \otimes P) \left\{ I_{p^2} - \Lambda_p (I_p \otimes P) \right\} (\text{vec}(P)) \]

\[ + \left[ (\text{vec}(P))^T I_{p^2} + K_{pp} (\text{vec}(P)) \right] \text{ where } K_{pp} \text{ be a symmetry} \]
\[ = \frac{1}{4} \left( (\text{vec}(P))^\prime \right) \left[ (I_p) - (I_p \otimes P) \Lambda_p \right] (P \otimes P) \left[ (I_p) - \Lambda_p (I_p \otimes P) \right] + (\text{vec}(P)) \right) + \text{vec}(P^\prime) \right) \right] = \text{vec}(P) \]

\[ = \frac{1}{4} \left( (\text{vec}(P))^\prime \right) \left[ (I_p) - (I_p \otimes P) \Lambda_p \right] (P \otimes P) \left[ (I_p) - \Lambda_p (I_p \otimes P) \right] (2 \text{vec}(P)) \]

\[ = \left\{ (\text{vec}(P))^\prime \right\} \left[ (I_p) - (I_p \otimes P) \Lambda_p \right] (P \otimes P) \left[ (I_p) - \Lambda_p (I_p \otimes P) \right] \text{vec}(P) \right\} \]

\[ = \text{Tr} \left( P^4 \right) - (\text{vec}(P))^\prime \right\} \left[ (I_p) - (I_p \otimes P) \Lambda_p \right] (P \otimes P) \text{vec}(P) - \]

\[ = \text{Tr} \left( P^4 \right) - (\text{vec}(P))^\prime \right\} \left[ (I_p) - (I_p \otimes P) \Lambda_p \right] (P \otimes P) \text{vec}(P) \]

\[ = \text{Tr} \left( P^4 \right) - (\text{vec}(P))^\prime \right\} \left[ (I_p) - (I_p \otimes P) \Lambda_p \right] (P \otimes P) \text{vec}(P) \]

\[ = \text{Tr} \left( P^4 \right) - (\text{vec}(P))^\prime \right\} \left[ (I_p) - (I_p \otimes P) \Lambda_p \right] (P \otimes P) \text{vec}(P) \]

\[ = \text{Tr} \left( P^4 \right) - (\text{vec}(P))^\prime \right\} \left[ (I_p) - (I_p \otimes P) \Lambda_p \right] (P \otimes P) \text{vec}(P) \]

Selanjutnya,

karena \( \text{vec} \left( D_{p^3} \right)^\prime \left( P \otimes P \right) \text{vec} \left( D_{p^3} \right) = \text{Tr} \left( D_{p^3} \otimes D_{p^3} \right) \) lihat persamaan (2.6)

sehingga,

\[ = \text{Tr} \left( P^4 \right) - \text{Tr} \left( D_{p^3} \otimes P \right) - \text{Tr} \left( P^3 \otimes D_{p^3} \right) + \text{Tr} \left( D_{p^3} \otimes D_{p^3} \right) \]

karena \( D_{p^3} \otimes D_{p^3} = \left( D_{p^3} \right)^2 \)

\[ = \text{Tr} \left( P^4 \right) - \text{Tr} \left( D_{p^3} \right)^2 - \text{Tr} \left( P^3 \otimes D_{p^3} \right) + \text{Tr} \left( D_{p^3} \otimes D_{p^3} \right)^2 \]

\( P^3 \) dan \( D_{p^3} \) simetri maka \( D_{p^3} \otimes P^3 \) juga simetri sehingga

\[ = \text{Tr} \left( P^4 \right) - 2 \text{Tr} \left( D_{p^3} \right)^2 + \text{Tr} \left( D_{p^3} \right)^2 \]

Akhirnya terbukti bahwa

\[ \left( \text{vec}(P) \right)^\prime \right\} M_p \Phi M_p \left( \text{vec}(P) \right) = \text{Tr} \left( P^4 \right) - 2 \text{Tr} \left( D_{p^3} \right)^2 + \text{Tr} \left( D_{p^3} \right)^2 \]

3. Additional Remarks

Alternative formulation of variance of VVVS is

\[ \frac{1}{n} \left\{ \text{Tr} \left( P^4 \right) - 2 \text{Tr} \left( D_{p^3} \right)^2 + \text{Tr} \left( D_{p^3} \right)^2 \right\} \]

References


