Abstract—In recent years, the use of vector variance as a measure of multivariate variability has received much attention in wide range of statistics. This paper deals with a more economic measure of multivariate variability, defined as vector variance minus all duplication elements. For high dimensional data, this will increase the computational efficiency almost 50% compared to the original vector variance. Its sampling distribution will be investigated to make its applications possible.

Keywords—Asymptotic distribution, covariance matrix, likelihood ratio test, vector variance.

I. INTRODUCTION

HYPOTHESIS testing about the stability of covariance structure is one of the fundamental issues in multivariate analysis. It is usually realized based on likelihood ratio test (LRT). See, for example, [2], [12], [18], [20], and [21] for the details of simultaneous test and [1], [7], [19], and the references therein for repeated test. Its wide range of applications can be easily found in literature. To mention some, see [3] for an early development; or [29] and [18] for its application in MANOVA; or [1], [35], [31], [36], [19], [15], [7], and [32], and the references there in, for historical background and its development in manufacturing industry; or [26] and [2] in biological research.

Under Normality, LRT means that one has to use covariance determinant (CD) as the measure of multivariate variability. This implies that LRT can only be used when the number of variables p is limited. In practice, it is rare that the number of variables p is large. See, for example, [34], [26], and [4], for the discussion when the sample size n > p and [17] and [15]-[16] for the case n < p. This is a serious problem because, when p is large, the computation of CD is quite cumbersome and tedious. Its computational complexity is of order \(O(p^3)\). Due to that limitation of CD, very recently, in [13] we propose to use vector variance (VV) as an alternative measure of multivariate variability. It is derived from the notion of vector covariance presented and used in [5], and its development in manufacturing industry; or [26] and [2] in biological research.

Under Normality, LRT means that one has to use covariance determinant (CD) as the measure of multivariate variability. This implies that LRT can only be used when the number of variables p is limited. In practice, it is rare that the number of variables p is large. See, for example, [34], [26], and [4], for the discussion when the sample size n > p and [17] and [15]-[16] for the case n < p. This is a serious problem because, when p is large, the computation of CD is quite cumbersome and tedious. Its computational complexity is of order \(O(p^3)\). Due to that limitation of CD, very recently, in [13] we propose to use vector variance (VV) as an alternative measure of multivariate variability. It is derived from the notion of vector covariance presented and used in [5], and its development in manufacturing industry; or [26] and [2] in biological research.

Under Normality, LRT means that one has to use covariance determinant (CD) as the measure of multivariate variability. This implies that LRT can only be used when the number of variables p is limited. In practice, it is rare that the number of variables p is large. See, for example, [34], [26], and [4], for the discussion when the sample size n > p and [17] and [15]-[16] for the case n < p. This is a serious problem because, when p is large, the computation of CD is quite cumbersome and tedious. Its computational complexity is of order \(O(p^3)\). Due to that limitation of CD, very recently, in [13] we propose to use vector variance (VV) as an alternative measure of multivariate variability. It is derived from the notion of vector covariance presented and used in [5], and its development in manufacturing industry; or [26] and [2] in biological research.

II. PROBLEM FORMULATION

Let X is a random vector with mean vector \(\mu\) and definite positive covariance matrix \(\Sigma\). Consider X as the superposition of two random vectors \(X^{(1)}\) and \(X^{(2)}\) of dimensions p and q, respectively,

\[ X = (X^{(1)})^T X^{(2)} \]  

If

\[ \mu^{(i)} = E(X^{(i)}) \]  

\[ i = 1, 2 \]  

then

\[ \Sigma_{ij} = E \left( (X^{(i)} - \mu^{(i)}) (X^{(j)} - \mu^{(j)})^T \right) \]  

\[ i, j = 1, 2 \]. Then \(\Sigma\) can be written in form of partitioned matrix

\[ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \]  

[5] uses \(Tr(\Sigma_{12} \Sigma_{21})\) to measure the linear relationship between the two random vectors \(X^{(1)}\) and \(X^{(2)}\). He calls this parameter vector covariance. It is the sum of square of all diagonal elements of \(\Sigma_{12} \Sigma_{21}\). Thus \(Tr(\Sigma_{12}^2)\) and \(Tr(\Sigma_{21}^2)\) are called vector variance (VV) of \(X^{(1)}\) and \(X^{(2)}\) respectively. In a special case, where \(p = q = 1\), vector covariance is the square of the classical covariance.

According to the above point of view, thus, VV of X is simply \(Tr(\Sigma^2)\), i.e., the sum of square of all elements of \(\Sigma\).
But, by using the vec operator, see [20] and [27], it can also be represented as \( ||\text{vec}(\Sigma)||^2 \). The vec operator transforms \( \Sigma \) into the vector \( \text{vec}(\Sigma) \) of \( p^2 \) dimension by stacking its column one after another. We see that if \( VV \), \( ||\text{vec}(\Sigma)||^2 \), is a quadratic form, covariance determinant (CD), \( |\Sigma| \), is a multilinear form. Thus, the computational complexity of \( VV \) is of order \( O(p^2) \) whereas that of \( CD \), as mentioned previously, is of order \( O(p^3) \). This advantage of \( VV \) is very promising especially when we work with multivariate data of high dimension. However, as \( \Sigma \) is symmetric, there are \( \frac{(p+1)p}{2} \) elements of \( \Sigma \) which are doubly counted in \( ||\text{vec}(\Sigma)||^2 \). This is the first problem that we want to discuss in this present paper. More specifically, instead of using the vec operator, we propose to use further operator which will transform the lower triangular part \( \Sigma_L \) of \( \Sigma \) into the vector \( \text{vec}(\Sigma_L) \) of dimension \( p(p+1)/2 \) by stacking its column one after another. From now on we call the paramete

\[
\text{vec}(\Sigma) = \begin{bmatrix}
\text{vec}(\Sigma_{11})
\vdots
\text{vec}(\Sigma_{pp})
\end{bmatrix}
\]

This transformation is also valid for all symmetric matrices.

\[\sqrt{n-1}(\text{vec}(S) - \text{vec}(\Sigma)) \xrightarrow{d} N(p\sigma^2, I)\]  \hspace{1cm} (11)

where

\[\Gamma = \left( I_{p^2} + K \right) (\Sigma \otimes \Sigma), \hspace{1cm} (12)\]

\[K = \sum_{i=1}^{p} \sum_{j=1}^{p} (H_{ij} \otimes H_{ij}) \hspace{1cm} (13)\]

and \( H_{ij} \) is defined in the previous section, i.e., a matrix of size \((p x p)\) where its \((i,j)\)-th element is equal to 1 and 0 otherwise. From this result, if the transformation (1) is used on \( S \), by using the result in [20] we have

\[\sqrt{n-1}(\text{vec}(S_L) - \text{vec}(\Sigma_L)) \xrightarrow{d} N_k(0, A)\]  \hspace{1cm} (14)

where

\[k = \frac{p(p+1)}{2} \text{ and} \]

\[A = \text{var}(\text{vec}(S_L)) = D_p^\dagger \text{var}(\text{vec}(S)) (D_p^\dagger)^T = D_p^\dagger \Gamma (D_p^\dagger)^T \]

Further, based on corollary 3.2. and Proposition 3.3. in [28], if we define \( u(\text{vec}(S_L)) = ||\text{vec}(S_L)||^2 \) arrive at the following proposition about the asymptotic distribution of sample MVV.

\[\sqrt{(n-1)||\text{vec}(S_L)||^2 - ||\text{vec}(\Sigma_L)||^2} \xrightarrow{d} N(0, \sigma^2) \hspace{1cm} (15)\]

where \( \sigma^2 = 4(\text{vec}(\Sigma_L))^T D_p^\dagger \Gamma (D_p^\dagger)^T (\text{vec}(\Sigma_L)) \)

This proposition is seemingly complicated to be used in application because the variance of sample MVV, \( ||\text{vec}(S_L)||^2 \), involves multiplication of large size matrix \( \Gamma (p^2 \times p^2) \), size even for moderate value of \( p \). However, the following proposition helps us to simplify the computation of that variance. The proof is only a matter of algebraic
manipulation using the properties of vec operator and commutation matrix.

**Proposition 2**

Let $\Omega$ be a matrix of size $(p \times p)$ such that

$$vec(\Omega) = (D_p^*)^T D_p^* vec(\Sigma)$$

(16)

then

$$\sigma^2 = 8\|vec(\Omega \Sigma)\|^2$$

(17)

**Proof**

Since $\Gamma = (I_{p^2} + K)(\Sigma \otimes \Sigma)$, then $\sigma^2$ can be written in the form

$$\sigma^2 = 4\left(\sum \Omega \right) (I_{p^2} + K)(\Sigma \otimes \Sigma)(D_p^*)^T (\sum \Omega)$$

(18)

$$\sigma^2 = 8\left(\sum \Omega \right) (D_p^*)^T N_p (\Sigma \otimes \Sigma) N_p^{-1} (\sum \Omega)$$

(19)

where

$$N_p = \frac{1}{2} \left( I_{p^2} + K \right)$$

(20)

But,

$$N_p (\Sigma \otimes \Sigma) = N_p (\Sigma \otimes \Sigma) N_p^{-1}$$

(21)

Hence,

$$\sigma^2 = 8\left(\sum \Omega \right) (D_p^*)^T N_p (\Sigma \otimes \Sigma) N_p^{-1} (D_p^*)^T (\sum \Omega)$$

(22)

Finally, since $D_p^* N_p = D_p^*$ and $N_p$ is symmetric, we get

$$\sigma^2 = 8\left(\sum \Omega \right) (D_p^*)^T (\Sigma \otimes \Sigma) (D_p^*) (\sum \Omega)$$

(23)

$$\sigma^2 = 8\left(vec(\Omega)\right)^T (\Sigma \otimes \Sigma) vec(\Omega)$$

(24)

Because

$$(D_p^*)^T (\sum \Omega) = (D_p^*)^T D_p^* vec(\Sigma)$$

(25)

$$\sigma^2 = 8\|vec(\Omega \Sigma)\|^2$$

(26)

**IV. Conclusion**

If vector variance $vec(\Sigma)$ is of dimension $p^2$, $\Sigma(\ell)$ is of dimension $k = (p + 1)/2$. This gain is too good to be neglected. Furthermore, Proposition 1 and 2 have made possible the application of modified vector variance, $\nu(\Sigma)$, where $\sigma^2$ simply eight is times the sum of square of all elements of $\Omega \Sigma$.

**ACKNOWLEDGMENT**

This research is partially supported by the Hasanuddin University 2012 research basis program study the authors thank the hasanuddin university, for that support.

**REFERENCES**


[34] M. Werner, Identification of multivariate outliers in large data sets, PhD dissertation, University of Colorado at Denver, 2003

