

A Test for Equality Form of Covariance Matrices of Multivariate Normal Populations

Hea-Jung Kim*

ABSTRACT

Given a set of data $p \times N_i$ matrices X_i observed from p -variate normal populations $\Pi_i \sim N(\mu_i, \Sigma_i)$ for $i=1, \dots, K$, the test for equality form of the covariance matrices is to choose a hypothetical model which best explains the homogeneity/heterogeneity structure across the covariance matrices among the hypothesized class of models. This paper describes a test procedure for selecting the best model. The procedure is based on a synthesis of Bayesian and a cross-validation or sample reuse methodology that makes use of a one-at-a-time schema of observational omissions. Advantages of the test are argued on two grounds, and illustrative examples and simulation results are given.

1. Introduction

Let $\{x_{ij}\}$, $j=1, \dots, N_i$; $i=1, \dots, K$, be p -variate observations from i -th multivariate normal population $N(\mu_i, \Sigma_i)$, where parameters are unknown. Then the test of the hypothesis $H: \Sigma_1 = \dots = \Sigma_K$ is called as testing for equality for covariance matrices. The test strategy is not uniquely defined, leading to a confusion of riches. Since Bartlett(1937) suggested a likelihood ratio test statistic modified for unbiased estimates, Roy(1957), Nagao(1973), Perlman(1980), and Gupta and Tang(1984) among others, proposed several testing procedures.

As in the most inferences, the rejection or acceptance of the hypothesis H is only the first step in the inference of the covariance matrices. When it is rejected, it is usually of interest to know the reason the null hypothesis was rejected and to establish true homogeneity/heterogeneity relationships among the subgroups of the covariance matrices(multiple homogeneities). However, a test which resolves this second step has not been seen yet. Our object of this study is to construct a test which simultaneously achieves first and second steps of the inference.

This paper introduces and studies yet another test procedure, and place it as an alternative approach. The methodology involved in the test is similar to that of Geisser and Eddy(1979),

* Department of Statistics, Dongguk University, Seoul 100, Korea.

Stone(1974) and Allen(1974) and is based on the reuse of the sample to estimate the likelihood function of interest. It uses a generalized version of predictive sample reuse(PSR) quasi-Bayes criterion. Application of the criterion to our problem can be well justified by the works of Murray (1977), Ng(1980), and Larimore(1983). Advantages of the suggested test are argued on two grounds. First, it is a method that simultaneously takes account of a set of all possible alternative hypotheses. Secondly, it is a criterion which involves direct comparison of quasi-Bayes likelihoods under each hypothesis. It is shown that this criterion is asymptotically equivalent to Akaike's criterion.

In the next section we formulate a multivariate version of the predictive sample reuse quasi-Bayes criterion first suggested by Geisser and Eddy(1979). In Section 3 the generalized version is applied to construct the test procedure for the equality form of the covariance matrices under the multivariate normal distribution assumption. Some test strategies and asymptotic properties of the test are also discussed. Section 4 gives illustrative examples and simulation results to depict the advantages of the suggested test. Finally, in Section 5, concluding remarks including some further research topics of interest, related with this test are given.

2. A Quasi-Bayes Likelihood Criterion

If the process generating observations is random and the models M_δ , $\delta=1, \dots, r$, are specified up to a known distribution function $F_\delta(\cdot | M_\delta)$, then the likelihoods under the alternative models can be ranked and the most likely model for the process, given the data, can be chosen. But usual statistical model identification problems are not this simple. At least they involve the model M_δ only specifies the sampling distribution up to a set of unknown parameters θ_δ . Geisser and Eddy(1979) offered a synthesis of Bayesian and sample reuse approach to those problems. The criterion is to choose a model which maximizes the joint conditional predictive densities of the observations, and is named as the predictive sample reuse(PSR) quasi-Bayes likelihood criterion. This criterion can be extended to the problem of a multivariate model identification, and is constructed as follows.

Let $Y(j)=[y_1, \dots, y_{j-1}, y_{j+1}, \dots, y_N]$ represents a $p \times (N-1)$ data matrix Y with the j -th observation vector y_j omitted, and assume that Y are generated from one of models M_δ with distribution $F_\delta(\cdot | \theta_\delta, M_\delta)$, $\delta=1, \dots, r$ with θ_δ unknown. Suppose that $f(y | Y, z, M_\delta)$ is the Bayesian predictive density for a future observation y when M_δ is true:

$$f(y | Y, z, M_\delta) \propto \int f(y | \theta_\delta, z, M_\delta) \prod_{j=1}^N f(y_j | \theta_\delta, z_j, M_\delta) P(\theta_\delta) d\theta_\delta, \quad (1)$$

$$\delta=1, \dots, r,$$

where z, z_1, \dots, z_N denote associate covariates of the corresponding observations and $P(\theta_\delta)$ is an invariant vague prior of the model M_δ . When the given density (1) is modified so that $f_j(y_j | Y(j), z_j, M_\delta)$ is taken as the Bayesian predictive density for y_j , we can construct the product of the predictive densities of y_j , $j=1, \dots, N$;

$$L_\delta = \prod_{j=1}^N f_j(y_j | Y(j), z_j, M_\delta), \quad \alpha=1, \dots, r. \quad (2)$$

It is obtained by a sample reuse methodology that makes use of a one-at-time schema of observational omissions (the cross-validation method). Geisser and Eddy (1979) showed that (2) can be used as an estimate of the likelihood for independent y_j , $j=1, \dots, N$:

$$g(Y | Z, M_\delta) = \prod_{j=1}^N g_j(y_j | z_j, M_\delta), \quad \delta = 1, \dots, r, \quad (3)$$

where $Z = [z_1, \dots, z_N]$.

We call (2) as the PSR quasi-Bayes likelihood of M_δ , and the PSR quasi-Bayes criterion is to choose a model M_{δ^*} which associates with the value δ^* , satisfying

$$L_{\delta^*} = \text{Max} (L_1, L_2, \dots, L_r) \quad (4)$$

as the most appropriate of the models being considered.

When we have M_δ 's that describe multivariate normal populations, the PSR quasi-Bayes criterion can be readily applicable to the problem of selecting the best model among them. For estimating a multivariate normal densities, the use of the Bayesian predictive density with vague prior as a surrogate for the sampling density was suggested by Aitchison (1975), and Murray (1977) and Ng (1980) demonstrated its superiority for the estimation of the multivariate normal distribution, based on the Kullback-Liebler information measure. They noted that for the case of devising an estimation method for the multivariate normal density, the choice of minimizing the negentropy in the class of invariance distributions coincides with the Bayesian predictive density using the vague prior. Hence, under the multivariate normal distribution assumption, the PSR criterion can be viewed as a model identification criterion based on a quasi-likelihood principle in a sense that it makes explicit use of the likelihood principle to the PSR quasi-Bayes likelihoods of M_δ for the identification.

3. A Multiple Homogeneity Test

When we have K multivariate normal populations and we are interested in determining the homogeneity/heterogeneity structure among the K population covariance matrices, usual likelihood ratio test cannot resolve this problem. Here we suggest a test procedure to solve this problem (multiple homogeneity test) and its properties are studied.

3.1 Test Criterion

Let $X_i = [x_{i1}, \dots, x_{iN_i}]$ be $p \times N_i$ data matrix from i -th population, $i=1, \dots, K$, with density

$$\eta(x | \mu_i, \Sigma_i) = (2\pi)^{-1/2} |\Sigma_i|^{-1/2} \exp\left\{-\frac{1}{2}(x - \mu_i)' \Sigma_i^{-1} (x - \mu_i)\right\}. \quad (5)$$

Suppose, however, that there are r possible underlying populations models:

$$M_1 : \Sigma_1 = \Sigma_2 = \dots = \Sigma_K \quad (6)$$

$$\begin{aligned}
 &M_2 : \Sigma_1 \neq \Sigma_2, \Sigma_1 = \Sigma_3 = \dots = \Sigma_K \\
 &\vdots \\
 &M_r : \Sigma_i \neq \Sigma_j, \text{ for all } i \neq j,
 \end{aligned}$$

the problem of selecting a model which best explains a given set of data is the same as the selecting the homogeneity/heterogeneity structure which exists across the K population covariance matrices. The number of ways to partition K population covariance matrices into m homogeneous groups are a Stirling number of the second kind,

$$S_K(m) = \sum_{j=0}^m (-1)^{m-j} j^K / \{j! (m-j)!\}. \tag{7}$$

Thus the total numbers of possible models with K populations can be formidable :

$$r = \sum_{m=1}^K S_K(m), \tag{8}$$

for example, for K=3, there are r=5 ; for K=5, there are r=52. Nonetheless, the PSR criterion for any intermediate model can be computed.

Suppose that the K populations are, in sequence, grouped into m homogeneous covariance matrix groups. Then the corresponding model, say M_* , can be expressed as

$$M_* : \Sigma_1 = \dots = \Sigma_{k_1}, \Sigma_{k_1+1} = \dots = \Sigma_{k_1+k_2}, \dots, \Sigma_{k^*} = \dots = \Sigma_K, \tag{9}$$

where $k^* = \sum_{i=1}^{m-1} k_{i+1}$, $K = \sum_{i=1}^m k_i$.

If we let \bar{X}_i and S_i be unbiased sample mean and covariance matrices of i-th population, under the usual vague prior

$$P(\mu_i, \Sigma_i) \propto |\Sigma_i|^{-(p+1)/2}, \tag{10}$$

the following predictive density for M_* is found(Cf. Geisser, 1964) :

$$\begin{aligned}
 f(x | \mathbf{X}, z_i, M_*) &= St_p(N(t) - k_i, \bar{X}_i, (N_i + 1)S(k_i)/N_i), \\
 i &= a(t), \dots, b(t), t = 1, \dots, m,
 \end{aligned} \tag{11}$$

where $\mathbf{X} = [X_1 | \dots | X_K]$, is p x N augmented total sample matrix, and z_i denotes an indicator covariate specifying one of K possible sampling distributions, and where

$$a(t) = \sum_{j=1}^{t-1} k_j + 1, b(t) = \sum_{j=1}^t k_j, N(t) = \sum_{i=a(t)}^{b(t)} N_i, N = \sum_{i=1}^K N_i,$$

and $S(k_i) = \sum_{i=a(t)}^{b(t)} (N_i - 1)S_i / (N(t) - k_i)$.

Here $St_p(a, b, C)$ denotes a p-dimensional variate t density function defined on $Y = R^p$ by the density at y (Cf. Press, 1982)

$$\frac{\Gamma\{(a+1)/2\}}{\Gamma\{(a-p+1)/2\}} \pi^{-p/2} |aC|^{-1/2} \{1+(y-b)'(aC)^{-1}(y-b)\}^{-(a+1)/2} \tag{12}$$

If we write $\mathbf{X}(\mathbf{ij})$ for the data matrix \mathbf{X} with j -th observation vector x_{ij} from i -th population omitted, according to (2), the PSR quasi-Bayes likelihood of M_* can be expressed as a product of the multivariate t densities :

$$L_* = \prod_{i=1}^K \prod_{j=1}^{N_i} f(x_{ij} | \mathbf{X}(\mathbf{ij}), z_i, M_*) \tag{13}$$

$$= \prod_{t=1}^m \prod_{i=a(t)}^{b(t)} \prod_{j=1}^{N_i} St_p(N(t)-k_t-1, \bar{X}_i(j), N_i S^*(k_i)/(N_i-1)), \tag{14}$$

where

$$\bar{X}_i(j) = \frac{1}{N_i-1} \sum_s^{(j)} x_{is},$$

$$S^*(k_i) = \frac{1}{N(t)-k_t-1} \sum_{u=a(t), u \neq i}^{b(t)} \{(N_u-1)S_u + \sum_s^{(j)} (x_{is}-\bar{X}_i(j))(x_{is}-\bar{X}_i(j))'\},$$

and $\sum_s^{(j)}$ represents the sum over all values s except for j .

L_* can be used to construct the PSR quasi-Bayes likelihood of any possible model considered in (6). This can be done by changing the subscripts of the statistics (sample covariance matrix, sample mean, etc.) involved in a model of interest in conformance with those of M_* . For example, the model M_2 in (6) may be the same as M_* , if we interchange the subscripts of those from population 1 and population 2, and set $k_1=1$ and $k_2=K-1$ in the expression of M_* . Under these changes, (13) gives the PSR quasi-Bayes likelihood of the subscript changed version of the model M_2 . Upon restoring the subscripts in the quasi-Bayes likelihood in conformance with original M_2 , we get

$$L_2 = \prod_{j=1}^{N_2} St_p\{N_2-2, \bar{X}_2(j), N_2 S_2/(N_2-2)\} \times \prod_{i=1, i \neq 2}^K \prod_{j=1}^{N_i} St_p\{N(t)-k_t-1, \bar{X}_i(j), N_i S^*(k_i)/(N_i-1)\}, \tag{15}$$

where $N(t) = \sum_{u=1, u \neq 2}^K N_u$, $k_t = K-1$, and

$$S^*(k_i) = \frac{1}{N(t)-k_t-1} \left\{ \sum_{u=1, u \neq 2}^K (N_u-1)S_u + \sum_s^{(j)} (x_{is}-\bar{X}_i(j))(x_{is}-\bar{X}_i(j))' \right\}.$$

Similarly, L_1 and L_r of the models M_1 and M_r in (6) can be obtained from (13) by letting $k_1=K$, $k_2=\dots=k_m=0$ and $k_1=\dots=k_m=1$, respectively.

Once we calculate the PSR quasi-Bayes likelihoods of all possible models in (6) from (13), the PSR quasi-Bayes likelihood criterion chooses the best model which have the largest of L_1 ,

..., L_r . Such that our test criterion is to choose a model which have

$$\text{Max} \{L_1, \dots, L_r\} \quad (16)$$

as the best underlying model of all the possible multiple homogeneity models. When the total number of possible models is too large for enumeration, an interesting subest should be chosen, and an appropriate model should be selected from the limited subest. This method can be done successively, using a stepwise enumeration methodology, to choose a sub-optimal model.

3.2 Asymptotic properties

Stone(1977) derived a result which is of interest here. He showed that, under a key assumption, the sample reuse criterion based on the cross-validation is asymptotically equivalent to the Akaike(1973) information criterion assuming maximum likelihood is used within each model. Using the similar procedure as in Stone's derivation, we can derive an asymptotic property of PSR quasi-Bayes likelihood criterion. Specifically, the PSR quasi-Bayes likelihood under the general model M_* is a product of multivariate t densities :

$$L_* = \prod_{t=1}^m \prod_{i=a(t)}^{\delta(t)} \prod_{j=1}^{N_i} St_p(N(t) - k_t - 1, \bar{X}_i(j), \frac{N_i}{N_i - 1} S^*(k_t)), \quad (17)$$

as in (13). With the same notations in (9), if M_* is true, the log-density assessment(log of the PSR quasi-Bayes likelihood) $\log L_*$ converges to

$$A = \sum_{i,j} \log \eta_p(\bar{X}_i(j), S_i(k_t)). \quad (18)$$

which is the sum of log of the multivariate normal densities $\eta_p(\cdot)$, where $S_i(k_t) = S^*(k_t) \{(N(t) - k_t - 1)/(N_i - 1)\}$ (Cf. Press, 1982, p.142). Without loss of generality, we omit subscript t from the notation, and we write $l(\cdot)$ for $\log \eta_p(\cdot)$ and $\hat{\phi}(-ij)$ to be a vector which denotes the reconstructed vector form of $\{\bar{X}_i(j), S_i(k_t)\}$ involved in A. Then A can be expressed as

$$A = \sum_{i,j} l(x_{ij} | \hat{\phi}(-ij), M_*). \quad (19)$$

Let us write the true log-likelihood of M_* as

$$\ell(\Phi, M_*) = \sum_{i,j} l(x_{ij} | \phi(i), M_*), \quad (20)$$

where $\Phi = \{\cup_{i=1}^K \phi(i)\} = \{\phi_1, \dots, \phi_{d_*}\}$ Φ an open region of \mathbf{R}^{d_*} , with $\phi(i)$ denotes the vector form of i-th multivariate normal population parameter, $\{\mu_i, \Sigma_i\}$. Here d_* denotes total number of parameters in the K multivariate normal populations.

With $\ell(\Phi, M_*)$, let define $\hat{\Phi}$ and $\hat{\Phi}_{-ij}$ as the unique solutions(MLEs) of

$$\ell'(\Phi, M_*) = 0$$

and

$$\ell'(\Phi, M_*) - \ell'(x_{ij} | \phi(i), M_*) = 0$$

respectively, so that $\hat{\Phi}_{-ij} = \bigcup_{i=1}^k \hat{\phi}(-ij)$ denotes MLE of the M_* under the data information $\mathbf{X}(ij)$ defined in (13).

If we write

$$l' = \left(\frac{\partial l}{\partial \phi_1}, \dots, \frac{\partial l}{\partial \phi_{d^*}} \right)^T, \quad l'' = \left(\frac{\partial^2 l}{\partial \phi_\beta \partial \phi_r} \right), \quad (21)$$

by Taylor's theorem, we have

$$A = \ell(\hat{\Phi}, M_*) + \sum_{i,j} (\hat{\phi}(-ij) - \hat{\phi}(j))^T l' \{x_{ij} | \hat{\phi}(i) + a_{ij}(\hat{\phi}(-ij) - \hat{\phi}(i))\} \quad (22)$$

and

$$\ell'(\hat{\Phi}_{-ij}, M_*) = \sum_{i,j} l'' \{x_{ij} | \hat{\phi}(i) + b_{ij}(\hat{\phi}(-ij) - \hat{\phi}(i))\} (\hat{\phi}(-ij) - \hat{\phi}(i)) \quad (23)$$

with $|a_{ij}| \leq 1$, $|b_{ij}| \leq 1$.

Also we get

$$\ell'(\hat{\Phi}_{-ij}, M_*) = \ell'(x_{ij} | \hat{\phi}(-ij), M_*). \quad (24)$$

Above relations, (22), (23), and (24), give

$$A = \ell(\hat{\Phi}, M_*) + \sum_{i,j} l' \{x_{ij} | \hat{\phi}(-ij), M_*\}^T \quad (25)$$

$$\left[\sum_{i,j} l'' \{x_{ij} | \hat{\phi}(i) + b_{ij}(\hat{\phi}(-ij) - \hat{\phi}(i))\} \right]^{-1} l' \{x_{ij} | \hat{\phi}(i) + a_{ij}(\hat{\phi}(-ij) - \hat{\phi}(i))\}.$$

Since we can expect the following approximations :

- (1) $\hat{\phi}(i) \rightarrow \phi(i)$ as $N_i \rightarrow \infty$;
- (2) $\hat{\phi}(-ij) \rightarrow \phi(i)$ as $N_i \rightarrow \infty$ for all i and j ;
- (3) $\sum_{i=1}^K \frac{1}{N_i} \sum_{j=1}^{N_i} l'' \{x_{ij} | \hat{\phi}(i) + b_{ij}(\hat{\phi}(-ij) - \hat{\phi}(i))\} \rightarrow \mathbf{E}\{l''(\mathbf{x}_1, \dots, \mathbf{x}_K | \Phi, M_*)\} = \mathbf{L}_2$, say ;
- (4) $\sum_{i=1}^K \frac{1}{N_i} \sum_{j=1}^{N_i} l'' \{x_{ij} | \hat{\phi}(i) + a_{ij}(\hat{\phi}(-ij) - \hat{\phi}(i))\} l' \{x_{ij} | \hat{\phi}(-ij), M_*\}^T$
 $\rightarrow \mathbf{E}\{l'(\mathbf{x}_1, \dots, \mathbf{x}_K | \Phi, M_*) l'(\mathbf{x}_1, \dots, \mathbf{x}_K | \Phi, M_*)^T\} = \mathbf{L}_1$, say,

where $l(\mathbf{x}_1, \dots, \mathbf{x}_K | \Phi, M_*)$ denotes the joint log-density of K populations under M_* , we have, using the well-known identity $\mathbf{L}_1 = \mathbf{L}_2^{-1}$, established that log of L_* in (17) is asymptotically

$$\ell(\hat{\Phi}, M_*) + trace(L_2^{-1}L_1) = \ell(\hat{\Phi}, M_*) + d_* \tag{26}$$

This is identical to Akaike's (1973) criterion when the missing t 's are restored.

If we consider two models M_s, M_{s^*} , of type (6) are nested so that

$$M_s \subseteq M_{s^*} \tag{27}$$

and suppose that both are true, then it is well known that, under regularity conditions,

$$2\{\ell(\hat{\Phi}_{s^*}, M_{s^*}) - \ell(\hat{\Phi}_s, M_s)\} \tag{28}$$

is asymptotically Chi-square with $d = d_{s^*} - d_s$ degrees of freedom, where d_{s^*} and d_s denote respective dimensions of parameter space of M_{s^*} and M_s . This gives, by (26), that the difference between two PSR log-quasi-Bayes likelihoods $\log L_{s^*} - \log L_s$ is asymptotically $\frac{1}{2} \chi_{(d)}^2 - d$.

Hence considering the null hypothesis that M_s is true versus the alternative that M_{s^*} is true, the quasi-Bayes likelihood criterion has an asymptotic significance level

$$\alpha = Pr(\chi_d^2 > 2d). \tag{29}$$

4. Numerical Examples

In this section we give limited data examples and simulation results to show the advantages of the suggested test procedure over the likelihood ratio test. As shown before, since the PSR quasi-Bayes likelihood criterion is asymptotically equivalent to Akaike criterion, our prime interest in the following examples are the behavior of the suggested test under small and moderated sample sizes.

4.1 Data Example

To illustrate the use of the PSR criterion, two examples are presented. For each example, under the multivariate normality assumption, the criterion is compared with the Box's (1949) approximation result for the usual likelihood ratio test. The data which follow are "Remote Sensing Data" from SAS/STAT guide. They are grouped into three crops - Corn, Soybean, and Cotton. Four measures of them are denoted y_1, \dots, y_4 which make up the descriptive variables :

Table 1. Remote Sensing Data

Variables	Sample 1(Corn)	Sample 2(Soybean)	Sample 3(Cotton)
y_1	16 15 16 18 15 15 12	20 24 21 27 12 22	31 29 34 26 53 34
y_2	27 23 27 20 15 32 15	23 24 25 45 13 32	32 24 32 25 48 35
y_3	31 30 27 25 31 32 16	23 25 23 24 15 31	33 26 28 23 75 25
y_4	33 30 26 23 32 15 73	25 32 24 12 42 43	34 28 45 24 26 78

Example 1. For sample 1 and sample 2, the log values of the PSR quasi-Bayes likelihoods, $\log L$, were computed under the two alternative models :

$$M_1 : \Sigma_1 = \Sigma_2, \quad M_2 : \Sigma_1 \neq \Sigma_2.$$

They are computed as $\log L_1 = -187.20128$ and $\log L_2 = -224.8175$. The PSR criterion clearly opts for model M_1 , where $\Sigma_1 = \Sigma_2$. In this two populations test case, the likelihood ratio test does permit comparison of the two models. The comparison of M_1 and M_2 yields the likelihood ratio test statistic value of 15.258468 for Chi-square statistic with 10 degrees of freedom. Since p-value for the test is 0.1229, the test does not reject M_1 . Thus we can see that suggested test and the likelihood ratio test lead to the same choice: M_1 .

Example 2. For testing the equality form of the three group covariance matrices a total of five possible models can be assumed. The possible models are as follows :

$$\begin{aligned} M_3 : \Sigma_1 &= \Sigma_2 = \Sigma_3 \\ M_4 : \Sigma_1 &= \Sigma_2 \neq \Sigma_3 \\ M_5 : \Sigma_1 &= \Sigma_3 \neq \Sigma_2 \\ M_6 : \Sigma_1 &\neq \Sigma_2 = \Sigma_3 \\ M_7 : \Sigma_1 &\neq \Sigma_2 \neq \Sigma_3. \end{aligned}$$

Log values of the corresponding PSR quasi-Bayes likelihood under the five models, M_3 to M_7 , are $\log L_3 = -289.243$, $\log L_4 = -284.389$, $\log L_5 = -313.749$, $\log L_6 = -303.044$, $\log L_7 = -321.988$. Thus the PSR criterion choose the model M_4 as the most appropriate of the five models being considered.

However, the likelihood ratio test does not permit direct comparison of the five models. Instead following conditional test may be possible. As given in example 1, the comparison of M_1 and M_2 does not reject the null hypothesis M_1 . Conditional on this result, the likelihood ratio test for M_3 yields a value of 39.334515 for the Chi-square statistic with 20 degrees of freedom. Now the null hypothesis is easily rejected (p-value = 0.0061). This means that the PSR test procedure and the conditional hypothesis test lead to the same choice, M_4 . Hence, example 2 shows that, unlike the likelihood ratio test, the PSR test procedure makes the selection directly.

4.2 A simulation Study

The aim of this study is to show the suggested PSR test criterion gives good test result. For the study we consider 7 simulation situations with three populations $\Pi_i \sim N_p(\mu_i, \Sigma_i)$, $i = 1, 2, 3$, with $\Sigma_1 \neq \Sigma_2 = \Sigma_3 = \Sigma$. Using a non-singular linear transformation H such that $H\Sigma_i H' = I_p$, and $H\Sigma H' = D_p$, diagonal matrix, 30 samples for each of simulation characterized by the set

$$\{\mu_1, \mu_2, \mu_3, H, D, p, N_1 = N_2 = N_3 = J\}$$

were generated. Table 2 provides details of the simulation situations.

Table 2. Simulation Situations

p	J	H	D	μ_1	μ_2	μ_3
3	5, 10, 20	Q_3	D_3	$\mu_{1(3)}$	$\mu_{2(3)}$	$\mu_{3(3)}$
5	10, 20	Q_5	D_5	$\mu_{1(5)}$	$\mu_{2(5)}$	$\mu_{3(5)}$
7	10, 20	Q	D	μ_1	μ_2	μ_3

Note : $D = \text{diag}(2, 3, 4, 5, 6, 7, 8, \dots)$, Q is the same matrix used in Press(1982, p.275). D_i and Q_i , $i=3, 5$, denote principal submatrices with first i diagonal elements of D and H respectively. $\mu_{1(i)}$, $\mu_{2(i)}$, and $\mu_{3(i)}$ are $i \times 1$ subvectors composed of first i components of $\mu_1 = (0, -3, 0, -3, 0, -3, 0, \dots)'$, $\mu_2 = (-3, 5, -3, 5, -3, 5, -3, \dots)'$, and $\mu_3 = (1, 1, 1, 1, 1, 1, 1, \dots)'$, respectively.

Under this simulation scheme, we assume that our interesting subset of models are M_a (true model) and M_b ;

$$M_a : \Sigma_1 \neq \Sigma_2 = \Sigma_3,$$

$$M_b : \Sigma_1 \neq \Sigma_2 \neq \Sigma_3.$$

Hence, it will be sufficient to compare M_a and M_b to show the performance of the PSR quasi-Bayes criterion. The PSR quasi-Bayes likelihoods for the models are calculated for each sampling experiment. The resulting averaged values of log-quasi-Bayes likelihoods are given below.

Table 3. Averaged Values of Log-Quasi-Bayes Likelihoods(AQL)

p	3			5			7	
J	5	10	20	10	20	10	20	
AQL M_a	-121.145	-227.351	-423.480	-369.070	-659.595	-503.970	-897.490	
AQL M_b	-154.151	-232.642	-427.566	-392.100	-688.643	-585.010	-931.170	

As expected, in this table, uniformly larger values of the PSR quasi-Bayes likelihoods for M_a indicate that the criterion works well for detecting a best model among considered.

5. Concluding Remarks

In this paper a Bayesian predictive sample reuse approach to the problem of testing the multiple homogeneities of the multivariate normal covariance matrices has been suggested. We show that the suggested testing method is different from the usual likelihood ratio test in that it enables us to test all possible homogeneity/heterogeneity forms of the multivariate normal covariance matrices. Limited examples of Section 4 depict this fact and advocate the suggested test.

The PSR quasi-Bayes likelihood criterion is shown to be asymptotically equivalent to Akaike's criterion, and hence asymptotic significance level of the suggested test between two nested models comes out to be $\alpha = Pr(\chi_{(d)}^2 > 2d)$, where d denotes the difference in numbers of unknown parameters between the two models. However, the issues of more developments pertaining

to the small sample behavior of the significance level are not done here and left for continuing study.

As pointed out before, when the number of populations to be considered takes large value, the number of all possible homogeneity/heterogeneity multivariate normal covariance matrices models becomes enormous. In this case, it is inevitable to devise a systematic method which makes the application of the suggested test simple. For this aspect, a stepwise procedure for the suggested test is being in study.

Applications of the suggested test to the multivariate models are appreciable. A model to which the test may be immediately applicable is the discriminant analysis model.

Acknowledgemnt

The detailed and constructive comments of referees were very useful in making the paper more readable and in correcting some errors. Their efforts are greatly appreciated.

References

1. Aitchison, J.(1975). Goodness of prediction fit. *Biometrika*, 62, 547-554.
2. Akaike, H.(1973). Information theory and an extension of the maximum likelihood principle. 2nd International Symposium on Information Theory (B.N. petrov and F. Craki, eds), pp.267-281. Budapest : Akademiai Kiado.
3. Allen, David M.(1974). The relationships between variable selection and data augmentation and a method of prediction. *Technometrics*, 16, 125-127.
4. Bartlett, M.S.(1937). Properties of sufficiency and statistical tests. *Proceedings of Royal Society of London, A*, 160, 268-282.
5. Box, G.E.P.(1949). A general distribution theory for a class of likelihood ratio criteria. *Biometrika*, 36, 317-346.
6. Geisser, S. and Eddy, W.F.(1979). A predictive approach to model selection. *Journal of American Statistical Association*, 74, 152-160.
7. Gupta, A.K. and Tang, J.(1984). Distribution of likelihood ratio statistic for testing equality of covariance matrices of multivariate Gaussian models. *Biometrika*, 71, 555-559.
8. Larimore, W.E.(1983). Predictive inference, sufficiency, entropy and an asymptotic likelihood principle. *Biometrika*, 70, 175-181.
9. Murray, G.D.(1977). A note on the estimation of probability density functions. *Biometrika*, 64, 150-152.
10. Nagao, H.(1973). On some test criteria for covariance matrix. *Annlas of Statistics*, 1, 700-709.
11. Ng, V.M.(1980). On the estimation of parametric density functions. *Biometrika*, 67, 505-506.
12. Perlman, M.D.(1980). Unbiasedness of the likelihood ratio tests for equality of several covariance matrices and equality of several multivariate populations. *Annals of Statistics* 8, 247-263.
13. Press, S.J.(1982). *Applied multivariate analysis*. 2nd edition. Krieger Publishing Company, Florida.
14. Roy, S.N.(1957). *Some aspects of multivariate analysis*. John Wiley & Sons, New York.
15. Stone, M.(1974). Cross-validatory choice and assessment of statistical predictions (with discussion). *Journal of Royal Statistical Society, B*, 36, 111-147.