

## 5 Conclusions

Most people, in practice, select few principal components or factors, which have larger corresponding eigenvalues to discriminate data among populations. However, from our simulation study in Section 4.1, we can see that the principal component or factor having the largest variance is not necessary having good discriminant power. Thus, depending on the purpose of further analysis, we may have to select principal components or factors based on other criterion instead of eigenvalues.

In this research, we apply two variable selection methods given by Mardia et al. and Chang to select principal components and factors for discrimination, respectively. The classification error rates based on those principal components or factors, selected by either one of the two variable selection methods, are better (i.e smaller) than that based on the principal components or factors, selected traditionally by the corresponding larger eigenvalues.

In Section 4.2, we generate the random data which are from one of three multivariate-normal populations with the same covariance matrix. We apply the variable selection method of Mardia et al. on these random data and find that the selected principal components are not the same as those principal components having the larger eigenvalues. Furthermore, the classification error rate based on the principal components, selected by Mardia's statistic is better (i.e smaller) than that based on the principal components, selected by the order of eigenvalues.

The main contribution of this research is to use the two variable selection methods of Mardia et al. and Chang to select principal components and factors, which have better discriminant power than the traditional ones. Moreover, we prove that the test statistic of Mardia et al. method obtained by principal components and that by corresponding factors (obtained by the principal component method) are the same, and the Mahalanobis distance of Chang's method obtained by principal components and that by corresponding factors are also the same. Hence we can

conclude that if we use either one of these two methods on principal components or factors selection, the orders of principal components and factors to be selected are the same. Moreover, we discuss how to apply the selection method of Mardia et al. on discriminating data from three groups. By simulation experiment, the discriminant result (error rate) obtained by the principal components selected by this method is better (i.e smaller) than that obtained by the first and second principal components selected by corresponding larger eigenvalues.