Advanced Science and Technology Letters
Vol.108 (Mechanical Engineering 2015), pp.99-102
http://dx.doi.org/10.14257/astl.2015.108.21

Application of Adjusted Discrimination Rule in Balanced Discriminant Analysis

Kyubark Shim[1], Jaegeol Yim[2]

[1]Department of Applied Statistics, Dongguk University at Gyeongju, Korea
E-mail: shim@dongguk.ac.kr
[2]Department of Computer Engineering, Dongguk University at Gyeongju, Korea
E-mail: yim@dongguk.ac.kr

Abstract. If the number of samples is smaller or equal to the order of the samples, the analysis results will have large error values. In discriminant analysis, the problem is called ill-posed problem when the number of samples is smaller than the order of the samples. In this paper, we applied adjusted method to solve the ill-posed problem, which can occur at discriminant analysis of balanced quadratic classification rule on which a growth curve model is applied.

Keywords: Balanced discriminant analysis, Growth curve model, Ill-posed problem, Regularization parameter

1 Introduction

Potthoff et al(1964) proved that the reliability of discriminant analysis is closely related to efficient estimation of covariance. Haff(1980) and Dey et al(1985) used several loss functions and proposed the covariance estimation value in relation with eigenvalue estimation. However, their covariance estimation value is based on the assumption that sample covariance is a nonsingular matrix. If the dimension of estimation parameter is similar to the number of observed value, parameter estimation values will vary in a large range, which can cause bias, and singular problem can occur because of decreased efficiency. To solve this problem, Cornfield(1967) proposed estimation of covariance using regularization parameter which is regulating bias for population parameter using James-Stein shrinkage. For adjusted parameter, a proper value is given according to the magnitude of bias. Friedman(1989) proved that misclassification error is decreasing when this method is applied in covariance estimation.

2 Adjusted Discrimination Rule in Balanced Discriminant Analysis

The simplest method Friedman applied in discriminant analysis is the division of pooled sample covariance matrix and the estimation value is defined as below.

ISSN: 2287-1233 ASTL

$$\widehat{R}_k = {}^{S}/_{n} \tag{1}$$

Here, $S = \sum_{k=1}^{K} S_k$, n : Total number of data.

Estimator proposed in the equation (1) sometimes brings a good result by decreasing estimator of covariance matrix, but since this does not always happen, (1) is not widely used. Covariance matrix estimated by using regularization parameter $\lambda\ (0 \le \lambda \le 1)$ is in below form.

$$\widehat{R}_k(\lambda) = \frac{S_k(\lambda)}{n_k(\lambda)} \tag{2}$$

Here, $S_k(\lambda) = (1-\lambda)S_k + \lambda S$, $n_k(\lambda) = (1-\lambda)n_k + \lambda n$ (3)

and, n_k and S_k are number of sample data and Covariance matrix.

In the equation (2),(3) the regularization parameter λ regularizes singularity among covariance matrix estimators of each group which can occur because of using pooled sample covariance matrix as estimator of covariance matrix. The meaning of regularize is that adjust a degree of singularity of pooled sample covariance matrix by using λ, and λ varies with respect to degree of singularity, this is called the degree of shrinkage. Especially, when λ=1, it can be proved from equation (2),(3) that the result is equal to that of Linear Discriminant Analysis, and for this case, the analysis can be done with $\widehat{R}_k(1) = \widehat{R}(1)$. Titterington(1985) insisted that if a proper regularization is used for estimator defined in equation (2), singularity of high degree is eliminated, which brings a good effect for accuracy of estimator

$$\widehat{R}(\lambda,\nu) = (1-\nu)\widehat{R}(\lambda) + ({}^{\nu}/_{p})\mathrm{tr}\big(\widehat{R}(\lambda)\big)I \tag{4}$$

Above estimator is that a solution of singular problem by using a new regularization parameter $\nu\,(0 \le \nu \le 1)$ rather than using only ν . ν is a regularization parameter that regularize the degree of shrinkage of product of identity matrix.

When adjusted linear discriminant analysis is done, by using this estimator, with appropriate value of ν and under multivariate normal distribution assumption, discriminant score of kth group, the standerd of discriminant is shown below.

$$d_k(X) = (X-\overline{X}_k)^T\widehat{R}^{-1}(\lambda,\nu)(X-\overline{X}_k) + \ln\big|\widehat{R}(\lambda,\nu)\big| - 2\ln\pi_k \tag{5}$$

From simulations under multivariate normal distribution, Friedman showed that the value of regularized parameter ν used in adjusted linear discriminant analysis should be set with respect to misclassification error.

When K=2, and q_k, k=1,2 are prior distribution of each growth curve model, classification range of a new classification rule of pxq sample matrix V is,

$$R_1 : \frac{g_1(V)}{g_2(V)} \ge \frac{q_2}{q_1}\ ,\quad R_2 : \frac{g_1(V)}{g_2(V)} < \frac{q_2}{q_1} \tag{6}$$

Here, $g_k(V) = (2\pi)^{-\frac{pq}{2}} |R_k(\lambda,\nu)|^{\frac{q}{2}} \times \exp\left\{-\frac{1}{2} \mathrm{tr} R_k^{-1}(\lambda,\nu)(V - X\alpha_k F_k)(V - X\alpha_k F_k)^T\right\}$ (7)

Therefore, the classification rule sorts V to the first growth curve model when it comes under R_1, and to the second model when R_2.

By natural logarithm on both sides in equation (1), it becomes this form.

$$q\{|R_2(\lambda,\nu)| - |R_1(\lambda,\nu)|\} + \mathrm{tr}(V - X\alpha_2 F_2)^T R_1(\lambda,\nu)(V - X\alpha_2 F_2) \\ -\mathrm{tr}(V - X\alpha_1 F_1)^T {R_1}^{-1}(\lambda,\nu)(V - X\alpha_1 F_1) \ \geq 2\ln\left(\frac{q_2}{q_1}\right) \quad (8)$$

The optimal classification rule of both growth curve models is that when the inequality (8) satisfies, V is sorted to group 1.

If $R_1(\lambda,\nu) = R_2(\lambda,\nu) = R(\lambda,\nu)$ and $q = 1$, then the equation (8) is abbreviated to below form.

$$V^T R(\lambda,\nu) X(\alpha_1 F_1 - \alpha_2 F_2) - \frac{1}{2}(\alpha_1 F_1 - \alpha_2 F_2)^T X^T {R_1}^{-1}(\lambda,\nu) X(\alpha_1 F_1 - \alpha_2 F_2) \\ \geq 2\ln\left(\frac{q_2}{q_1}\right) \quad (9)$$

Assume that two growth curve model follows below distribution.

$$G(V|\alpha_i, R_i(\lambda,\nu), \pi_i) \backsim N(V|X\alpha_i F_i, R_i(\lambda,\nu) \otimes I_{N_i}), \ i = 1,2$$

Adjusted balanced quadratic classification rule is that when below condition is satisfied, V is sorted to π_i.

$$(V - X\alpha_2 F_2)^T R_2(\lambda,\nu)^{-1}(V - X\alpha_2 F_2) - (V - X\alpha_1 F_1)^T R_1(\lambda,\nu)^{-1}(V - X\alpha_1 F_1) \\ -\log(|R_1(\lambda,\nu)|/|R_2(\lambda,\nu)|) \ \geq 2\ln\left(\frac{q_2}{q_1}\right) \quad (10)$$

Here, $q_1 = (M_1/M_2)$, $q_2 = 1 - q_1$. (11)

$$\text{And, } M_1 = \log(|R_2(\lambda,\nu)|/|R_1(\lambda,\nu)|) + p - (\alpha_1 F_1 - \alpha_2 F_2)^T X^T R_1(\lambda,\nu)^{-1} \\ (\alpha_1 F_1 - \alpha_2 F_2) - \mathrm{tr}(R_2(\lambda,\nu) R_1(\lambda,\nu)^{-1}) \quad (12)$$

$$M_2 = 2p - (\alpha_1 F_1 - \alpha_2 F_2)^T X^T (R_1(\lambda,\nu)^{-1} + R_2(\lambda,\nu)^{-1}) X(\alpha_1 F_1 - \alpha_2 F_2) \\ -\mathrm{tr}(R_1(\lambda,\nu) R_2(\lambda,\nu)^{-1}) - \mathrm{tr}(R_2(\lambda,\nu) R_1(\lambda,\nu)^{-1}) . \quad (13)$$

According to adjusted balanced quadratic classification rule estimated from sample by equation (10) and (11), if below condition is satisfied, V is sorted to π_i.

$$(V - X\hat{\alpha}_2 F_2)^T \hat{R}_2(\lambda,\nu)^{-1}(V - X\hat{\alpha}_2 F_2) - (V - X\hat{\alpha}_1 F_1)^T \hat{R}_1(\lambda,\nu)^{-1}(V - X\hat{\alpha}_1 F_1) \\ -\log(|\hat{R}_1(\lambda,\nu)|/|\hat{R}_2(\lambda,\nu)|) \ \geq 2\ln\left(\frac{\hat{q}_2}{\hat{q}_1}\right) \quad (14)$$

Here, $\hat{q}_1 = (\widehat{M}_1/\widehat{M}_2)$, $\hat{q}_2 = 1 - \hat{q}_1$

And, $$M_1 = \log(|\hat{R}_2(\lambda,\nu)|/|\hat{R}_1(\lambda,\nu)|) + p - (\hat{\alpha}_1 F_1 - \hat{\alpha}_2 F_2)^T X^T R_1(\lambda,\nu)^{-1} (\hat{\alpha}_1 F_1 - \hat{\alpha}_2 F_2) - \mathrm{tr}(\hat{R}_2(\lambda,\nu)\hat{R}_1(\lambda,\nu)^{-1}) \quad (15)$$

$$M_2 = 2p - (\hat{\alpha}_1 F_1 - \hat{\alpha}_2 F_2)^T X^T (\hat{R}_1(\lambda,\nu)^{-1} + \hat{R}_2(\lambda,\nu)^{-1}) X (\hat{\alpha}_1 F_1 - \hat{\alpha}_2 F_2) - \mathrm{tr}(\hat{R}_1(\lambda,\nu)\hat{R}_2(\lambda,\nu)^{-1}) - \mathrm{tr}(\hat{R}_2(\lambda,\nu)\hat{R}_1(\lambda,\nu)^{-1}) \,. \quad (16)$$

The cutoff point for classification rule in the adjusted linear discriminant analysis of growth curve model using balanced rule is $C = 2\log(\hat{q}_2/\hat{q}_1)$.

3 Conclusion

A balanced discriminant analysis modified with adjusted method to solve the ill-posed problem is proposed in this paper. Performing adjusted discrimination rule in balanced discriminant analysis in the early stage of computer system design is very important. Therefore, the proposed method will contribute to save the cost of system design. For the further research, we are applying the method on practical applications.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (NRF-2011-0006942) and by 'Development of Global Culture and Tourism IPTV Broadcasting Station' Project through the Industrial Infrastructure Program for Fundamental Technologies funded by the Ministry of Knowledge Economy (10037393).

Reference

1. Cornfield,J, Discriminant Function. Review of the Intenational Statistical Institute,35,142-153 (1967)
2. Dey,D.K. and Srmivasan,C, Estimation of a Covariance Matrix Under Stein's Loss. The Annals of Statistics,13,pp.1581-1591 (1985)
3. Friedman, J.H, Regularized Discriminant Analysis. Journal of American Statistical Association, 84, pp.165 - 175 (1989)
4. Haff,L.R, Empirical Bayes Estimation of Multivariate Normal Covariance Matrix. The Annals of Statistics ,8, pp.586-597 (1980)
5. Potthoff,R.R and Roy,S.N, A generalized multivariate analysis of Covariance model useful especially for growth curve problems ,Biometrika, 51, 313-326 (1964)
6. Titterington,D.M, Common Structure of Smoothing Techniques in Statistics. International Statistical Review ,53,pp.141-170 (1985)

