ANNALES UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN–POLONIA
VOL. XXXV, 11 SECTIO A 1981

Statistics Department
University of Adelaide, South Australia

Kerwin W. MORRIS

The Asymptotic Distribution of Certain Eigenvalues Occurring in Discriminant Analysis – non-normal Theory

Rozkład asymptotyczny pewnych wartości własnych występujących w analizie dyskryminacji – niegaussowska teoria

Асимптотические распределения некоторых собственных значений в анализе дискриминации – негауссовская теория

1. Introduction. In a recent paper [5], some non-normal asymptotic results in MANOVA were derived using a theory of convergence in distribution of multiply-indexed arrays. With the notation of that paper, $M(k \times p)$ denotes the matrix of means of the k p-variate populations, Σ their common convariance matrix, and H the hypotesis that M has the form X_1B_1, where X_1 $(k \times r)$ is given of and rank r. When H is not true, M can be written uniquely the form

$$M = X_1B_1 + M_2 \tag{1}$$

where $X_1B_1 = P_1M$, $M_2 = (I - P_1)M \neq 0$, and $P_1 = X_1(X_1'X_1)^{-1}X_1'$ is the orthogonal projector ($o.p.$) matrix onto the r-dimensional subspace $\Omega_1 = \mathcal{R}(X_1) \subset R^k \cdot (2)$.

In the associated discriminant analysis, which is discussed in the case when $X_1 = 1$ by Kshirsagar [4] and generally by Bartlett [2], the number of useful discriminant functions is equal to the rank of M_2. Bartlett's test of the hypothesis

$$H_q : r(M_2) = q\ ,$$

where q is a given integer, $0 < q < p = \min(p, k-r)$, is based on the $p - q$ smallest e.values of S_1S^{-1}, where S_1 is the $S.S.P.$ matrix used for testing H and S the withinclass estimate of Σ. Hsu [3] has obtained the asymptotic distribution of these e.values when H_q is true, in the case when the populations are normal and the sample sizes $n_1, .., n_k$ maintain the proportions as $n = \Sigma n_i$ increases. We show below that in the non-normal case the e.values converge in distribution (in the generalized sense of [5]) to Hsu's limiting distribution, viz. the distribution of the smallest $p-q$ e.values of $W'W$, where

$$\underset{(p-q)\times(k-r-q)}{W} \sim N(0, I_{(p-q)(k-r-q)}),$$

and discuss some approximate testes of H_q when all the sample sizes are large.

Hsu [3] and Anderson [1] also discuss the asymptotic distribution of the q largest e.values of $S_1 S^{-1}$ and the associated e.vectors. In the present context, however, these quantities do not seem to have much practical importance, since the corresponding population quantities depend on $n_1, .., n_k$. The definition of discriminant functions that depend only on M and Σ, their estimation, and the associated asymptotic theory in the non-normal case are discussed in [6] and [7].

2. Initial transformations. We begin by making a series of linear transformations of the data.

First, we transform to $Z_1 = YA$, where A is a symmetric matrix such that $A^2 = \Sigma^{-1}$. Then $\mathrm{Var}(Z_1) = I_{np}$, and the matrix of means of the new variates is

$$MA = X_1 B_1 A + M_2 A.$$

We now assume that H_q is true, i.e. that the unknown matrix M_2 has rank q. Then $r(M_2 A) = q$, and there exists an orthogonal elementary operation matrix E such that the first q columns of $M_2 AE$ are linearly independent, i.e. such that $M_2 AE$ has the form

$$M_2 AE = (M_3, M_3 C), \tag{2}$$

where M_3 is $k \times q$ of rank q, C is $q \times q_1$ and $q_1 = p - q$.

Transforming now to $Z_2 = Z_1 E$, when $\mathrm{Var}(Z_2) = I_{np}$, and the corresponding matrix of means is

$$MAE = X_1 B_1 AE + (M_3, M_3 C)$$

Next, we construct a $p \times p$ orthogonal matrix as follows. Since $I_{q_1} + C'C > 0$, there exists a $q_1 \times q_1$ symmetric matrix B_2 such that $B_2 (I + C'C) B_2 = I$.

Writing now

$$C_1 = \begin{pmatrix} C \\ -I_{q_1} \end{pmatrix} \tag{4}$$

and $H_2 = C_1 B_2$, then $H_2' H_2 = I_{q_1}$, and there exists $H_1 (p \times q)$ such that $H = (H_1, H_2)$ is orthogonal.

Finally, we transform to $Z = Z_2 H$. Then $\mathrm{Var}(Z) = I_{np}$, and the corresponding matrix of means is $MAEH = X_1 B_1 AEH + (M_2 AEH_1, M_2 AEH_2)$.

From (3) and (4),

$$M_2AEH_2 = M_3(I, C)C_1B_2 = 0\,,$$

whence, writing $M_0 = MAEH$, $B_0 = B_1AEH$, and

$$M_1 = M_2AEH_1 \tag{5}$$

then

$$M_0 = X_1B_0 + (M_1, 0)\,. \tag{6}$$

Summing up,

the rows of $Z = YAEH$ are independent, $E(Z) = XM_0$ and $\mathrm{Var}(Z) = I_{np}$. (7)

Writing now

$$T_1 = Z'(P - P_0)Z = A_1'S_1A_1$$

and

$$T = \frac{1}{n-k}\,Z'(I-P)Z = A_1'SA_1\,,$$

where

$$A_1 = AEH\,,$$

then

$$T_1T^{-1} = A_1'(S_1S^{-1})(A_1')^{-1}\,,$$

whence T_1T^{-1} and S_1S^{-1} have the same e.values.

Furthemore, since EH is orthogonal, then T_1 and $S_1\Sigma^{-1}$ have the same e.values. (8)

3. The e.values of $S_1\Sigma^{-1}$. We now recall from [5], § 4.3, that

$$P - P_0 = XN^{-1/2}(I - P_N)N^{-1/2}X'$$

where $P_N = N^{1/2}X_1(X_1'NX_1)^{-1}X_1'N^{1/2}$ is the o.p. matrix onto the r-dimensional subspace $\Omega_N = \mathcal{R}(N^{1/2}X_1) \subset R^k$, and also that $(I - P_N)$ was written in the form

$$(I - P_N) = H_NH_N'\,,$$

where H_N is $k \times (k-r)$ and $H_N'H_N = I_{k-r}$.

Then

$$T_1 = Z'XN^{-1/2}(I - P_N)N^{-1/2}X'Z = \bar{Z}_N'N^{1/2}(I - P_N)N^{1/2}\bar{Z}_N\,,$$

where $N\bar{Z}_N = X'Z$.

Write now

$$\begin{aligned} W_N &= N^{1/2}(\bar{Z}_N - M_0) \\ U_N &= H_N'W_N \end{aligned} \tag{9}$$

Then

$$T_1 = (W_N + N^{1/2}M_0)' H_N H_N' (W_N + N^{1/2}M_0) =$$

$$= U_N' U_N + (U_N' H_N' N^{1/2} M_0 + M_0' N^{1/2} H_N U_N) + M_0' N^{1/2} H_N H_N' N^{1/2} M_0 \; .$$

Since by definition of H_N, $H_N'(N^{1/2}X_1) = 0$, it follows from (6) that

$$H_N' N^{1/2} M_0 = (B, 0), \text{ where } B = H_N' N^{1/2} M_1 \; . \tag{10}$$

Thus, if we now write

$$U_N = (U_1, U_2) \tag{11}$$

where U_1 is $(k-r) \times q$ and U_2 is $(k-r) \times q_1$, then

$$T_1 = \begin{pmatrix} U_1'U_1 + (U_1'B + B'U_1) + B'B & U_1'U_2 + B'U_2 \\ U_2'U_1 + U_2'B & U_2'U_2 \end{pmatrix}$$

We now show that $r(B) = q$ for every N,

Note first from (1) and (5) thet each column of M_1 is contained in $\Omega_1^{\perp}$. Further, since the columns of H_N are a basis of $\Omega_N^{\perp}$, then the columns of $N^{1/2}H_N$ are a basis of $\Omega_1^{\perp}$. It follows that M_1 can be written in the form $M_1 = N^{1/2}H_N C_N$, and since $N^{1/2}H_N$ has full column rank, $r(C_N) = r(M_1) = q$. Thus $B = H_N'N\ H_N C_N$ also has rank q, since $H_N' N H_N$ is non-singular.

It follows that for each N there exists a $q \times q$ symmetric matrix F such that

$$F B' B F = I_q \; . \tag{12}$$

Now consider the e.values of $S_1 \Sigma^{-1}$. From (8), these are the solutions of $|T_1 - \lambda I| = 0$, and hence also the solutions of

$$\left| \begin{pmatrix} F & 0 \\ 0 & I_{q_1} \end{pmatrix} (T_1 - \lambda I) \begin{pmatrix} F & 0 \\ 0 & I_{q_1} \end{pmatrix} \right| = 0,$$

which, after simplification, has the form

$$\begin{vmatrix} V_{11} + I_q - \lambda F^2 & V_{12} \\ V_{12}' & U_2'U_2 - \lambda I_{q_1} \end{vmatrix} = 0$$

where $V_{11} = FU_1'U_1F + F(U_1'B + B'U_1)F$ and $V_{12} = FU_1'U_2 + FB'U_2$.

Finally, premultiplying by

$$\begin{vmatrix} I_q & 0 \\ -V_{12}' & I_{q_1} \end{vmatrix}.$$

The e.values of T_1 are the roots of $g_N(\lambda) = 0$, where

$$g_N(\lambda) = \begin{vmatrix} V_{11} + I_q - \lambda F^2 & V_{12} \\ V_{12}'(\lambda F^2 - V_{11}) & U_2'(I_{k-r} - BF^2B')U_2 - V_{22} - \lambda I_{q_1} \end{vmatrix} \tag{13}$$

and $V_{22} = U_2'(U_1F^2U_1' + U_1F^2B' + BF^2U_1')U_2$.

4. **The asymptotic distribution of the e.values of $S_1\Sigma^{-1}$.** Let $\lambda_1 \geqslant \lambda_2 \ldots \geqslant \lambda_p \geqslant 0$ denote the e.values of $S_1\Sigma^{-1}$. Since $r(S_1) \leqslant \rho = \min(p, k-r)$ for every N, then $\lambda_{\rho+1} = \ldots = \lambda_p = 0$ for every N. We determine here the asymptotic distribution of $\lambda_{q+1}, \ldots, \lambda_\rho$ when H_q is true.

(i) We show first that $\lim_{N\to\infty} F = 0$. From §3, $(I - P_N)N^{1/2}M_1 = N^{1/2}M_1 - N^{1/2}X_1(X_1'NX_1)^{-1}X_1'NM_1 = N^{1/2}(M_1 - \bar{M}_1)$ where $\bar{M}_1 = X_1(X_1'NX_1)^{-1}X_1'NM_1$. Thus, from (10), $B'B = M_1'N^{1/2}(I-P_N)N^{1/2}M_1 = (M_1 - \bar{M}_1)'N(M_1 - \bar{M}_1)$. For given N, write now $n_0 = \min(n_1, \ldots, n_k)$, $F = (\underset{\sim}{f}_1, \ldots, \underset{\sim}{f}_q)$, and let π_N denote the smallest e.value of $B'B$. Then

$$\pi_N = \inf_{\underset{\sim}{x}} \frac{\underset{\sim}{x}'B'B\underset{\sim}{x}}{\underset{\sim}{x}'\underset{\sim}{x}} \geqslant n_0 \inf_{\underset{\sim}{x}} \frac{\underset{\sim}{x}'(M_1 - \bar{M}_1)'(M_1 - \bar{M}_1)\underset{\sim}{x}}{\underset{\sim}{x}'\underset{\sim}{x}}.$$

But since the columns of $\bar{M}_1$ and M_1 are respectively in Ω_1 and $\Omega_1^{\perp}$, then $(M_1 - \bar{M}_1)'(M_1 - \bar{M}_1) = M_1'M_1 + \bar{M}_1'\bar{M}_1$, and hence

$$\pi_N \geqslant n_0 \inf_{\underset{\sim}{x}} \frac{\underset{\sim}{x}'M_1'M_1\underset{\sim}{x}}{\underset{\sim}{x}'\underset{\sim}{x}} = n_0\nu,$$

where ν is the smallest e.value of $M_1'M_1$. Since $r(M_1) = q$, then $\nu > 0$ and $\lim_{N\to\infty} \pi_N = \infty$. Finally, from (12), for $i = 1, \ldots, q$, $1 = \underset{\sim}{f}_i'B'B\underset{\sim}{f}_i \geqslant \pi_N \underset{\sim}{f}_i'\underset{\sim}{f}_i$, whence $\lim_{N\to\infty} F = 0$.

(ii) Next, we determine the asymptotic distribution of $V_N = FB'U_2$.

By inspection of the proof of theorem 4 in [5], it follows from (7) and (9) that $U_N \xrightarrow{D} N(0, I_{p(k-r)})$, and hence, from (11)

$$U_1 \xrightarrow{D} N(0, I_{q(k-r)}) \text{ and } U_2 \xrightarrow{D} N(0, I_{q_1(k-r)}) \tag{14}$$

Thus from theorem 1 in [5], the c.f. $\zeta_N(T_1)$ of U_2 is given by

$$\zeta_N(T_1) = E\,[\exp(i\,Tr\,(T_1' U_2))] = \exp(-1/2\,Tr\,(T_1' T_1)) + f_N\,(T_1),$$

where $\lim_{N \to \infty} f_N\,(T_1) = 0$ uniformly in any bounded region

$$C \subset R^{q_1 (k-r)}. \tag{15}$$

Consider now the c.f. $\phi_N(T_2)$ of V_N, viz.

$$\phi_N(T_2) = E\,[\exp(i\,Tr\,(T_2' V_N))] = \zeta_N\,(BFT_2) =$$

$$= \exp(-1/2\,Tr\,(T_2' FB'BFT_2)) + f_N\,(BFT_2) =$$

$$= \exp(-1/2\,Tr\,(T_2' T_2)) + f_N\,(BFT_2)\,, \qquad \text{using (12).}$$

For fixed T_2, choose in (15) $C = \left\{ T_1;\ Tr\,(T_1'\,T_1) \leq Tr\,(T_2'\,T_2) \right\}$. Since $Tr\,((BFT_2')(BFT_2)) = Tr\,(T_2' T_2)$ for every N, it follows from (15) that, for fixed T_2, $\lim_{N \to \infty} \phi_N\,(T_2) = \exp(-1/2\,Tr\,(T_2'\,T_2))$, and hence, from theorem 1 in [5], that $FB'U_2 \xrightarrow{D} V \sim N\,(0, I_{q_1\,q})$. A similar argument shows that $FB'U_1 \xrightarrow{D} V_0 \sim N\,(0, I_{q_1^2})$.

(iii) Consider now

$$U_2'\,(I_{k-r} - BF^2B')\,U_2 = U_2'\,(I - Q_N)\,U_2\,,$$

where, from (12), $Q_N = BF^2B'$ is a $(k-r) \times (k-r)$ o.p. matrix of rank q. A repetition of the proof of theorem 4 in [5] then shows that $U_2'\,(I - BF^2B')\,U_2 \xrightarrow{D} W'W$, where $\underset{(k-r-q) \times (p-q)}{W} \sim N\,(0, I_{(p-q)\,(k-r-q)})$.

(iv) Finally, consider the polynomial $g_N(\lambda)$, of degree p, in (13). Using (14), the results of (i) and (ii), and theorem 2 of [5] it follows that

$$V_{12} = F\,(U_1' U_2) + FB'U_2 \xrightarrow{D} 0 + V = V\,.$$

Similarly $V_{11} \xrightarrow{D} 0$, $V_{22} \xrightarrow{D} 0$, whence $V_{12}'\,F^2 \xrightarrow{D} 0$ and $V_{12}'\,V_{11} \xrightarrow{D} 0$, and, for fixed λ, $g_N(\lambda) \xrightarrow{D} g(\lambda)$, where

$$g(\lambda) = \begin{vmatrix} I_q & V \\ 0 & W'W - \lambda I_{q_1} \end{vmatrix} = \left| W'W - \lambda I_{q_1} \right| \tag{16}$$

Since $g(\lambda)$ has degree $p - q$, it follows that the q largest zeros $\lambda_1, .., \lambda_q$ of $g_N(\lambda)$

converge in probability to $+\infty$, and $\lambda_{q+1}, ..., \lambda_\rho$ converge in distribution to the $\rho - q$ largest e.values $L_1 \geq L_2 ... \geq L_{\rho-q}$ of $W'W$.

5. The asymptotic distribution of the e.values of $S_1 S^{-1}$. As in § 4.4 of [5], denote by $\ell_1 \geq \ell_2 ... \geq \ell_\rho$ the ρ largest e.values of $S_1 S^{-1}$, and write

$$\underset{\sim}{\ell}_q = (\ell_{q+1}, .., \ell_\rho)'$$

Theorem. *When H_q is true,*

$$\underset{\sim}{\ell}_q \xrightarrow{D} \underset{\sim}{L}_q$$

where $\underset{\sim}{L}_q = (L_1, ..., L_{\rho-q})'$, $L_1 \geq L_2 ... \geq L_{\rho-q}$ are the largest e.values of $W'W$ and

$$\underset{(k-r-q)\times(p-q)}{W} \sim N(0, I_{(p-q)(k-r-q)}).$$

Proof. From (8), $T_1 T^{-1}$ has the same e.values as $S_1 S^{-1}$. Let A_N be a precisely defined $p \times p$ symmetric matrix such that $A_N^2 = T^{-1}$ (i.e. when $|T| \neq 0, A_N = \phi(T)$ for some well-defined ϕ).

Then $A_N T_1 A_N$ also has the same e.values as $S_1 S^{-1}$. Write now

$$A_N = \begin{pmatrix} A_{11} & A_{12} \\ A'_{12} & A_{22} \end{pmatrix}, \text{ where } A_{11} \text{ is } q \times q,$$

$$B_N = \begin{pmatrix} I & -A_{11}^{-1} A_{12} \\ 0 & I \end{pmatrix}$$

and $F = A_{11}^{-1} F$.

The e.values of $S_1 S^{-1}$ are then the roots of the equation.

$$\left| \begin{pmatrix} F' & 0 \\ 0 & I \end{pmatrix} B'_N (A_N T_1 A_N - \lambda I) B_N \begin{pmatrix} F & 0 \\ 0 & I \end{pmatrix} \right| = 0,$$

which, after simplification has the form

$$\begin{vmatrix} V_{11} + I_q - \lambda F'F & V_{12} + \lambda W_{12} \\ V_{12} + \lambda W'_{12} & U'_2 U_2 - \lambda (I + A'_{12} A_{11}^{-2} A_{12}) \end{vmatrix} = 0,$$

where

$$V_{11} = V_{11} + F'A_{12}V'_{12} + V_{12}A'_{12}F + F'A_{12}U'_2U_2A'_{12}F,$$

$$V_{12} = (V_{12} + F'A_{12}U'_2U_2)\,C,$$

$$C = A_{22} - A'_{12}A_{11}^{-1}A_{12},$$

and $$W_{12} = -F'A_{11}^{-1}A_{12}.$$

Finally, premultiplying by

$$\begin{matrix} I & 0 \\ -V'_{12} & I \end{matrix}$$

and simplifying, the e.values of S_1S^{-1} are the solutions of

$$h_N(\lambda) = \begin{vmatrix} V_{11} + I_q - \lambda F'F & V_{12} + \lambda W_{12} \\ \lambda(V'_{12}FF' + W'_{12}) - V'_{12}V_{11} & [CU'_2(I - BF^2B')\,U_2C - V_{22} - \\ & -\lambda(I + A'_{12}A_{11}^{-2}A_{12} + V'_{12}W_{12}] \end{vmatrix} = 0$$

where $V_{22} = C\,[(FB'U_2)'V_0 + V'_0\,(FB'U_2)]\,C$ and $V_0 = FU'_1 + F'A_{12}U'_2$.

Now consider T. From (7)–(8) and theorem 5 of [5], $T \xrightarrow{D} I_p$. It follows then from (17) and theorem 2 of [5] that $A_N \xrightarrow{D} I_p$. Using now the results of § 4, and repeated use of theorem 2 of [5], it follows that

$$A_{11} \xrightarrow{D} I_q,\ A_{12} \xrightarrow{D} 0,\ A_{22} \xrightarrow{D} I_{q_1},\ F \xrightarrow{D} 0,\ V_{11} \xrightarrow{D} 0,\ V_{12} \xrightarrow{D} V,\ W_{12} \xrightarrow{D} 0,$$

$$C \xrightarrow{D} I_{q_1},\ V_0 \xrightarrow{D} 0,\ V_{22} \xrightarrow{D} 0, \text{ and, for fixed } \lambda,\ h_N(\lambda) \xrightarrow{D} g(\lambda) \text{ of (16)}$$

and the theorem follows as in § 4 (iv).

6. **Significance tests of H_q.** Since the limiting distribution obtained above is the same as that of theorem 6 in [5] with p and k replaced by $p - q$ and $k - q$, it follows from theorem 7 of [5] that one can write down various statistics for testing H_q, each of which converges in distribution to $\chi^2_{(p-q)(k-r-q)}$. Writing $g_i = 1 + (\ell_i/n - k)$ and $\ell'_i = \ell_i/g_i$, so that $\ell'_1, \ldots, \ell'_p$ are the ordered e.values of $S_1S_0^{-1}$, these statistics are

$$\sum_{q+1}^{p} \ell_i,\ \sum_{q+1}^{p} \ell'_i,\ (n-k)\Big(\prod_{q+1}^{p} g_i - 1\Big) \text{ and } n\sum_{q+1}^{p} \ln g_i.$$

The last of these is essentially Bartlett's statistic, which replaces n by a correction factor (on the same order) that is appropriate when normality is assumed.

REFERENCES

[1] Anderson, T. W., *The Asymptotic Distribution of Certain Characteristic Roots and Vectors*, Proceedings of the Second Berkely Symposium on Mathematical Statistics and Probability, University of California Press, Berkely and Los Angeles, (1951), 103–130.

[2] Bartlett, M. S., *Multivariate Analysis*, J. Roy. Statist. Soc., Supplc., 9, (1947), 176–197.

[3] Hsu, P. L., *On the Limiting Distribution of Roots of a Determinantal Equation*, J. London Math. Soc., 16, (1941), 183–194.

[4] Kshirsagar, A. M., *Multivariate Analysis*, Marcel Dekker Inc., New York, (1972).

[5] Morris, K. W., Szynal, D., *Convergence in Distribution of Multiply-indexed Arrays with Application in MANOVA*, Annales Univ. Mariae Curie-Skłodowska Sect. A, vol. 34 (1980), 83–95.

[6] Morris, K. W., *Some Asymptotic Results in Multiple Discriminant Analisis-non-normal Theory*, Technical Paper No. 10, Department of Statistic, University of Adelaide South Australia, (1980).

[7] Morris, K. W., *Multiple Discriminant Analysis with Multinomial Sampling-non-normal Asymptotic Results*, Technical Paper No. 12, Department of Statistics, University of Adelaide (1981).

STRESZCZENIE

W przypadku rozkładu różnego od normalnego bada się zbieżność według rozkładu (w sensie wieloskładnikowym, określonym w [5] wartości własnych w teście Bartletta [2] dla liczby funkcji dyskryminacyjnych w multidyskryminacyjnej analizie. Ponadto rozważa się asymptotyczny rozkład statystyk testowych gdy wszystkie próby są duże.

РЕЗЮМЕ

В случае распределения разного от нормального исследуется сходимость по распределению (в мультииндексном смысле определенном в [5]) собственных значений из теста Бартлетта [2] для дискриминантных функций в мульти дискриминантном анализе. Кроме того рассматривается асимптотические распределения тестовых статистик, когда все объемы быборки являются большими.