International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 11(Special Issue), 2011, pp 145-157

Asymptotic Distribution of The Jackknife Statistics for Canonical Correlations Under Nonnormal Populations

Arjun K. Gupta
Department of Mathematics & Statistics
Bowling Green State University
Bowling Green, Ohio-43403, USA

Hisao Nagao
Dept. of Mathematical Sciences
College of Engineering
Osaka Prefecture University
Osaka, Japan

[Received December 27, 2010; Revised May 25, 2011; Accepted June 23, 2011]

Abstract

In this paper the limiting distribution of the jackknife statistics of canonical correlation coefficient has been studied when the parent population is nonnormal with finite fourth moments. The limiting distribution of the jackknife statistic of a function of the sample canonical correlations is also derived.

Keywords and Phrases: Canonical correlation coecient; Jackknife statistic; Limiting distribution; Nonnormal population;Implicit function theorem;Perturbation method.

AMS Classification: 62H10; 62H05.

1 Introduction

This paper is concerned with limiting distributions of jackknife statistics of canonical correlation coefficients for a sample from a nonnormal distribution having finite fourth moments.

Canonical correlation analysis is used to study the relationship between two random vectors $U_1 = (x_1, \cdots, x_p)'$ and $U_2 = (x_{p+1}, \cdots, x_{p+q})', (p \leq q)$. Let $1 > \rho_1^2 \geq \cdots \geq \rho_p^2 \geq 0$ be the characteristic roots of the matrix $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, where

$$\mathrm{Var}(U_1) = \Sigma_{11},\ \mathrm{Var}(U_2) = \Sigma_{22},\ \mathrm{Cov}(U_1, U_2) = \Sigma_{12} = \Sigma_{21}'. \tag{1}$$

Their positive square roots $1 > \rho_1 \geq \cdots \geq \rho_p \geq 0$ are the population canonical correlation coefficients. It is well-known that ρ_1 is the maximum correlation between two linear functions $y_{11} = a_1' U_1$ and $y_{12} = b_1' U_2$ subject to the condition that $\mathrm{Var}(y_{11}) = \mathrm{Var}(y_{12}) = 1$. Also ρ_2 is the maximum correlation between $y_{21} = a_2' U_1$ and $y_{22} = b_2' U_2$ subject to the conditions that y_{21} and y_{22} are uncorrelated with both y_{11} and y_{12}, and have unit variances and so on. The variables y_{i1}, y_{i2} are the ith canonical variables $(i = 1, \cdots, p)$. Let the $(p+q) \times 1$ vectors $X_1, \cdots, X_N$ denote a random sample from a $(p+q)$-variate nonnormal distribution with mean $\mu = (\mu_1, \cdots, \mu_{p+q})'$, covariance matrix Σ and finite fourth moments, where

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}. \tag{2}$$

The sample canonical correlation coefficients $1 > r_1(S/n) \geq \cdots \geq r_p(S/n) > 0$ are the positive square roots of the characteristic roots of $S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}$, where $n = N - 1$. The $S_{11}(p \times p), S_{22}(q \times q)$ and $S_{12} = S_{21}'(p \times q)$ are the submatrices of S partitioned in the same manner as (2) of Σ. That is,

$$S = \sum_{a=1}^{N} (X_a - \bar{X})(X_a - \bar{X})' = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \tag{3}$$

where $\bar{X} = (\bar{x}_{\cdot 1}, \cdots, \bar{x}_{\cdot\, p+q})' = N^{-1} \sum_{a=1}^{N} X_a$, and $\bar{x}_{\cdot j} = N^{-1} \sum_{k=1}^{N} x_{kj}, (j = 1, 2, \cdots, p+q)$. The main purpose of this paper is to derive the limiting distribution of the jackknife statistic of $r_j^2(S/n)$. The jackknife statistic was originally defined by Quenouille [13] to reduce the bias of an estimate. Then Tukey [15] proposed the general use of this technique to obtain approximate confidence intervals for problems where standard statistical procedures may not exist or are difficult to apply. This paper is related to Tukey's confidence interval or testing hypothesis rather than Quenouille's bias reduction.

A nice review has been written by Miller [7]. Also Parr and Schucany [12] and Frangos [5] have given a list of references on jackknife statistics. It seems that the jackknife statistics which deal with eigenvalue problems of covariance matrices have not been studied by many authors. Dempster [4] derived the bias correction of the canonical correlation under the multivariate normal distribution. More than 20 years ago, Nagao [9], [10], and [11] gave the limiting distribution of eigenvalue problems of a covariance matrix and also of a correlation matrix under the nonnormal situation. Also Beran and Srivastava [1] have treated some problems of eigenvalues and eigenvectors of a covariance matrix without normality by using bootstrap method. Recently Das and Sen [2], [3] wrote interesting papers about the nervous system, where they applied the resampling method to the canonical correlation analysis.

2 The limiting distribution of the canonical correlation

The pseudo-values and the jackknife statistic of the jth root $r_j^2(S/n)$ are given by

$$r_{j,a}^2 = r_j^2(S/n) + (N-1)\{r_j^2(S/n) - r_j^2(S_{-a}/(n-1))\}, (a = 1, \cdots, N) \tag{4}$$

and

$$\bar{r}_j^2 = \frac{1}{N}\sum_{a=1}^{N} r_{j,\,a}^2,$$

respectively, where S_{-a} corresponding to S is a $(p+q)\times(p+q)$ matrix which is obtained by deleting $X_a = (x_{a1}, \cdots, x_{a\,p+q})'$ from the random sample $X_1, \cdots, X_N$. Then S_{-a} is given by

$$S_{-a} = (s_{k\ell}^{-a}) = S - \frac{N}{N-1}(X_a - \bar{X})(X_a - \bar{X})'. \tag{5}$$

Also let $S_{ij,a}(i,j=1,2)$ denote the submatrices of S_{-a} partitioned as in (3). Here we give the limiting distribution of $\bar{r}_j^2$. Since the problem is concerned with eigenvalues of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, we can assume the covariance matrix Σ such that $\Sigma_{11} = I_p$, $\Sigma_{22} = I_q$, where I_p denotes a $p\times p$ identity matrix, and $\Sigma_{12} = (P, 0)$, where $P = \mathrm{diag}(\rho_1, ..., \rho_p)$. Then we have

$$\sqrt{n}(\bar{r}_j^2 - \rho_j^2) = \sqrt{n}(r_j^2(S/n) - \rho_j^2) + \frac{\sqrt{n}(N-1)}{N}\sum_{a=1}^{N}\{r_j^2(S/n) - r_j^2(S_{-a}/(n-1))\}. \tag{6}$$

When ρ_j is a nonzero simple root, Muirhead and Waternaux [8] have shown that the first term of (6) converges in law to a normal distribution. Thus we will show that the second term of the R.H.S. of (6) converges in probability to zero. In order to expand $r_j^2(S_{-a}/(n-1))$ about S/n, we apply the implicit function theorem. We consider the equation $F(S_{-a}/(n-1), r^2) = |S_{11,\,a}^{-1}S_{12,\,a}S_{22,\,a}^{-1}S_{21,\,a} - r^2 I_p| = 0$. At first we will show that the equation can be solved for r^2 around $(S/n, r_j^2(S/n))$. Then $F(S/n, r_j^2(S/n)) = 0$. The partial derivative of $F(\cdot)$ with respect to r^2 under $(S/n, r_j^2(S/n))$ is given by

$$F_{r^2}(S/n, r_j^2(S/n)) = \sum_{i=1}^{p}|C_i|, \tag{7}$$

where C_i is the matrix obtained from $C = S_{11}^{-1}S_{12}S_{22}^{-1}S_{21} - r_j^2(S/n)I_p$ by replacing ith column with (0,...,-1,0,...,0)' having -1 for the ith element. Since $r_i^2(S/n) \to \rho_i^2$ $(i =$

$1, ..., p$) in probability, the expression (7) converges to

$$\begin{vmatrix} -1 & 0 & & \mathbf{0} \\ 0 & \rho_2^2-\rho_j^2 & & \\ \vdots & & \ddots & \\ 0 & \mathbf{0} & & \rho_p^2-\rho_j^2 \end{vmatrix} + \cdots + \begin{vmatrix} \rho_1^2-\rho_j^2 & & \mathbf{0} \\ & \ddots & \\ & \rho_{p-1}^2-\rho_j^2 & \\ \mathbf{0} & & -1 \end{vmatrix} \qquad (8)$$

$$= -\prod_{i=1, i\neq j}^{p} (\rho_i^2 - \rho_j^2).$$

Thus if ρ_j^2 is a simple root, for large N we have $F_{r^2}(S/n, r_j^2(S/n)) \neq 0$ in the region $||S/n - \Sigma|| \leq \epsilon$ for a suitable norm $||\cdot||$ and some $\epsilon > 0$. Let $(t_{k\ell}^a) = S_{-a}/(n-1) - S/n$. Then by Chebyshev's inequality, since the fourth order moments are finite, we can easily show that

$$\mathrm{P}(\sup_{1\leq a\leq N} |t_{k\ell}^a| \geq \epsilon) \to 0 \quad \text{as} \quad N \to \infty. \qquad (9)$$

Thus, for all a ($1 \leq a \leq N$), we can expand the function $r_j^2(S_{-a}/(n-1))$ around S/n, and get

$$r_j^2(S_{-a}/(n-1)) = r_j^2(S/n) + \sum_{k\leq\ell} g_{k\ell}^j(S/n) t_{k\ell}^a + \frac{1}{2} < t_{k\ell}^a > C_a^{(j)} < t_{k\ell}^a >', \qquad (10)$$

where $g_{k\ell}^j(S/n) = -F_{s_{k\ell/n}}(S/n, r_j^2(S/n))$, each element of a $w \times w$ matrix $C_a^{(j)}$ with $w = (p+q)(p+q+1)/2$ is the derivative of $g_{k\ell}^j(S/n)$ evaluated at the elements of some matrix between S/n and $S_{-a}/(n-1)$ and $< a_{k\ell} >= (a_{11}, \cdots, a_{pp}, a_{12}, \cdots, a_{p-1,p})$. Since

$$\begin{aligned} \textstyle\sum_{a=1}^N t_{k\ell}^a &= \textstyle\sum_{a=1}^N \{ \frac{s_{k\ell}}{(N-1)(N-2)} - \frac{N}{(N-1)(N-2)} (x_{ak} - \bar{x}_{\cdot k})(x_{a\ell} - \bar{x}_{\cdot \ell}) \} \\ &= \frac{N s_{k\ell}}{(N-1)(N-2)} - \frac{N s_{k\ell}}{(N-1)(N-2)} \\ &= 0, \end{aligned} \qquad (11)$$

we only have to show that the expression

$$\begin{aligned} &\frac{\sqrt{n}(N-1)}{N} \sum_{a=1}^N \{ r_j^2(S/n) - r_j^2(S_{-a}/(n-1)) \} \\ &= -\frac{\sqrt{n}(N-1)}{2N} \sum_{a=1}^N < t_{k\ell}^a > C_a^{(j)} < t_{k\ell}^a >' \end{aligned} \qquad (12)$$

converges in probability to zero. Since for large N, each element of $C_a^{(j)}$ is bounded, by the continuity of the derivatives of $g_{k\ell}^j(\cdot)$ and the convergence of $S_{-a}/(n-1)$ uniformly in a, by using Schwarz's inequality, we have

$$\begin{aligned}&\text{the absolute value of R.H.S. of (12)} \le C(P, p+q)\\ &\times \frac{\sqrt{n}(N-1)}{N}\sum_{a=1}^{N}\mathrm{tr}(\frac{S_{-a}}{n-1}-\frac{S}{n})^2,\end{aligned} \tag{13}$$

where $C(P, p+q)$ is a function of P and $p+q$. Then from (5) we have

$$\begin{aligned}&\text{R.H.S. of (13)} = C(P, p+q)(N-1)^{-1/2}(\frac{N}{N-2})^2\\ &\times\{\frac{1}{N}\sum_{a=1}^{N}\mathrm{tr}[(X_a-\bar{X})(X_a-\bar{X})']^2-\mathrm{tr}(S/n)^2\}.\end{aligned} \tag{14}$$

Since the fourth moments of X_a are finite, the above expression (14) converges in probability to zero. Thus we have

THEOREM 2.1. Let the $(p+q)\times 1, (p\le q)$ vectors $X_1,\cdots,X_N$ denote a random sample from a $(p+q)$-variate distribution with mean μ, covariance matrix Σ and finite fourth moments. If the jth canonical correlation ρ_j is a non-zero simple root, then we have

$$\sqrt{n}(\bar{r}_j^2-\rho_j^2)\to N(0,\tau_{jj}), \tag{15}$$

where

$$\tau_{jj}=\rho_j^4(\kappa_{jj}^{jj}+\kappa_{p+j,p+j}^{p+j,p+j}+2\kappa_{jj}^{p+j,p+j})-4\rho_j^3(\kappa_{jj}^{j,p+j}+\kappa_{p+j,p+j}^{j,p+j})+4\rho_j^2\kappa_{j,p+j}^{j,p+j} \tag{16}$$

with $\kappa_{ab}^{cd}=\mathrm{Cov}((x_a-\mu_a)(x_b-\mu_b),(x_c-\mu_c)(x_d-\mu_d))$.

In case of normal distribution, since $\kappa_{jj}^{jj}=\kappa_{p+j,p+j}^{p+j,p+j}=2$, $\kappa_{j,p+j}^{j,p+j}=1+\rho_j^2$, $\kappa_{jj}^{p+j,p+j}=2\rho_j^2$ and $\kappa_{jj}^{j,j+p}=\kappa_{p+j,p+j}^{j,p+j}=2\rho_j$, we have $\tau_{jj}=4\rho_j^2(1-\rho_j^2)^2$.

3. The convergence of $\sum_{a=1}^{N}(r_{j,a}^2-\bar{r}_j^2)^2/(N-1)$

We will show that

$$\sum_{a=1}^{N}(r_{j,a}^2-\bar{r}_j^2)^2/(N-1)\to\tau_{jj}\ \text{ in probability.} \tag{17}$$

Now we have

$$\textstyle\sum_{a=1}^{N}(r_{j,a}^2-\bar{r}_j^2)^2/(N-1)$$

$$= (N-1)\textstyle\sum_{a=1}^{N}(r_j^2(S_{-a}/(n-1)) - \dfrac{1}{N}\displaystyle\sum_{a=1}^{N} r_j^2(S_{-a}/(n-1)))^2. \tag{18}$$

If ρ_j is a simple root, by the implicit function theorem, we have

$$r_j^2(S_{-a}/(n-1)) = r_j^2(S/n) + \sum_{k\le\ell} g_{k\ell}^j(\xi_a^j)t_{k\ell}^a = r_j^2(S/n) + U_a + V_a, \tag{19}$$

where ξ_a^j is some matrix between $S_{-a}/(n-1)$ and S/n. U_a and V_a in (19) are given by

$$U_a = \sum_{k\le\ell} g_{k\ell}^j(S/n)t_{k\ell}^a$$

and (20)

$$V_a = \sum_{k\le\ell}(g_{k\ell}^j(\xi_a^j) - g_{k\ell}^j(S/n))t_{k\ell}^a.$$

Thus we have

$$\text{R. H. S. of (18)} = (N-1)\sum_{a=1}^{N}(U_a+V_a-\bar{V})^2, \tag{21}$$

where $\bar{V} = N^{-1}\sum_{a=1}^{N} V_a$. First we consider the sum $(N-1)\sum_{a=1}^{N} U_a^2$. After some simplification we have

$$(N-1)\textstyle\sum_{a=1}^{N} U_a^2 = \displaystyle\sum_{k\le\ell}\sum_{t\le u} g_{k\ell}^j(S/n)g_{tu}^j(S/n)$$

$$\times[(\frac{N}{N-2})^2\sum_{a=1}^{N}(x_{ak}-\bar{x}_{\cdot k})(x_{a\ell}-\bar{x}_{\cdot\ell}) \tag{22}$$

$$\times(x_{at}-\bar{x}_{\cdot t})(x_{au}-\bar{x}_{\cdot u})/(N-1) - \frac{N}{N-1}(s_{tu}/(N-2))(s_{k\ell}/(N-2))].$$

Next we will give the values to which $g_{k\ell}^j(S/n)$ converges in probability. In order to get them, we need a well-known lemma (see for example, Srivastava and Khatri [14, p. 28]).

LEMMA 3.1. Let $S = (s_{ij})$ be a $p \times p$ nonsingular symmetric matrix. Then

$$\frac{\partial S^{-1}}{\partial s_{ij}} = \begin{cases} -(s_i s_j' + s_j s_i') & i \neq j \\ -s_i s_i' & i = j \end{cases} \tag{23}$$

where $S^{-1} = (s_1, \cdots, s_p)$.

After some tedious calculations we see that $g^j_{k\ell}(S/n)$ converges in probability to the following values:

$$g^j_{k\ell}(S/n) \to \begin{cases} -\rho_j^2 & k = \ell = j, \quad k = \ell = p + j \\ 2\rho_j & k = \ell - p = j \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

Thus we have

$$(N-1)\sum_{a=1}^{N} U_a^2 \to \tau_{jj} \ \text{ in probability.}$$

Next we consider $(N-1)\sum_{a=1}^{N} V_a^2$. Since ξ_a^j converge in probability to Σ for all $1 \leq a \leq N$ when $N \to \infty$, for large N we can get $|g^j_{k\ell}(\xi^j_a) - g^j_{k\ell}(S/n)| \leq \epsilon$ for any $\epsilon > 0$. Thus from Schwarz's inequality, we have

$$\begin{aligned} (N-1)\sum_{a=1}^{N} V_a^2 &\leq \epsilon^2 (N-1)\frac{1}{2}(p+q)(p+q+1) \\ &\times \mathrm{tr}\{N^2 \sum_{a=1}^{N}[(X_a - \bar{X})(X_a - \bar{X})']^2 - NS^2\}[\frac{1}{(N-1)(N-2)}]^2. \end{aligned} \tag{25}$$

Since the fourth moments are finite, (25) converges in probability to zero. Also from $\sum_{a=1}^{N}(V_a - \bar{V})^2 \leq \sum_{a=1}^{N} V_a^2$ and Schwarz's inequality, the other terms also converge to zero. Thus we get the following:

THEOREM 3.2. Under the assumptions of Theorem 2.1,

$$\sum_{a=1}^{N}(r^2_{j,a} - \bar{r}^2_j)^2/(N-1) \to \tau_{jj} \quad \text{in probability.} \tag{26}$$

From Theorems 2.1 and 3.2, follows

THEOREM 3.3. Let the $(p+q)\times 1$ vectors $X_1,\cdots,X_N$ denote a random sample from a $(p+q)$-variate distribution with mean μ, covariance matrix Σ and finite fourth moments. If the jth canonical correlation ρ_j is a non-zero simple root, then we have

$$\frac{n(\bar{r}_j^2-\rho_j^2)}{\sqrt{\sum_{a=1}^N (r_{j,a}^2-\bar{r}_j^2)^2}} \to \mathrm{N}(0,1), \quad \text{as} \quad N\to\infty. \tag{27}$$

3 The jackknife statistic of a function of the sample canonical correlation coefficients

In this section we generalize the above results for a function of eigenvalues of $S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$. Let $f(\cdot)$ be a real-valued function with the second continuous derivatives in some neighborhood of $(\rho_1^2,\cdots,\rho_p^2)$. By using the notations of the previous sections, the pseudo-values and the jackknife statistic of $f(r_1^2(S/n),\cdots,r_p^2(S/n))$ are given by

$$\begin{aligned} f^a &= f(r_1^2(S/n),\cdots,r_p^2(S/n)) + (N-1)\{f(r_1^2(S/n),\cdots,r_p^2(S/n)) \\ &\quad - f(r_1^2(S_{-a}/(n-1)),\cdots,r_p^2(S_{-a}/(n-1)))\}, \quad (a=1,\cdots,N) \end{aligned}$$

and (28)

$$\bar{f} = \frac{1}{N}\sum_{a=1}^N f^a,$$

respectively. First we will derive the limiting distribution of $\bar{f}$. Since the method used is similar to the above, we only sketch the proof. Expanding $f(r_1^2(S_{-a}/(n-1)),\cdots,r_p^2(S_{-a}/(n-1)))$ around $(r_1^2(S/n),\cdots,r_p^2(S/n))$, we have

$$\begin{aligned} & f(r_1^2(S_{-a}/(n-1)),\cdots,r_p^2(S_{-a}/(n-1))) = f(r_1^2(S/n),\cdots,r_p^2(S/n)) \\ & + \sum_{j=1}^p A_{aj} f_j(r_1^2(S/n),\cdots,r_p^2(S/n)) + \frac{1}{2}\sum_{j,k=1}^p A_{aj}A_{ak}f_{jk}(\xi_a), \end{aligned} \tag{29}$$

where $A_{aj}=r_j^2(S_{-a}/(n-1))-r_j^2(S/n)$, $f_j(\rho_1,\cdots,\rho_p)=\dfrac{\partial}{\partial\rho_j}f(\rho_1,\cdots,\rho_p)$, $f_{jk}(\rho_1,\cdots,\rho_p)=\dfrac{\partial}{\partial\rho_j}f_k(\rho_1,\cdots,\rho_p)$ and ξ_a is a point on the line segment between the vectors $(r_1^2(S/n),\cdots,r_p^2(S/n))$ and $(r_1^2(S_{-a}/(n-1)),\cdots,r_p^2(S_{-a}/(n-1)))$. Then the term eliminating $\sqrt{n}(f(r_1^2(S/n),\cdots,r_p^2(S/n))-f)$ from $\sqrt{n}(\bar{f}-f)$ with $f=f(\rho_1^2,\cdots,\rho_p^2)$ converges in probability to zero if all ρ_j^2 are non-zero simple roots. Since

$r_j^2(\cdot)$ is a continuous function, ξ_a converge in probability to $(\rho_1^2, \cdots, \rho_p^2)$ uniformly $1 \leq a \leq N$. Hence by the assumption, $|f_{jk}(\xi_a)|$ are bounded $(a = 1, ..., N)$ in some neighborhood of $(\rho_1^2, \cdots, \rho_p^2)$. Thus we only show that for $j = 1, \cdots, p$,

$$\frac{\sqrt{n}(N-1)}{N} \sum_{a=1}^{N} \{r_j^2(S_{-a}/(n-1)) - r_j^2(S/n)\}^2 \to 0 \quad \text{in probability.} \tag{30}$$

Let $\sum_{a=1}^{N} \{r_j^2(S_{-a}/(n-1)) - r_j^2(S/n)\}^2 = Q_j + R_j$, where

$$Q_j = \sum_{a=1}^{N} \{r_j^2(S_{-a}/(n-1)) - N^{-1} \sum_{a=1}^{N} r_j^2(S_{-a}/(n-1))\}^2$$

and (31)

$$R_j = N\{r_j^2(S/n) - N^{-1} \sum_{a=1}^{N} r_j^2(S_{-a}/(n-1))\}^2.$$

Then from Theorem 3.2, $\sqrt{n}(N-1)N^{-1}Q_j \to 0$ in probability. Since $\sqrt{n}(N-1)N^{-1}R_j = (N-1)^{-1/2}(\bar{r}_j^2 - r_j^2(S/n))^2$, we have the desired conclusion (30). Therefore we have

THEOREM 4.1. Let the $(p+q) \times 1$ vectors $X_1, \cdots, X_N$ denote a random sample from a $(p+q)$-variate distribution with mean μ, covariance matrix Σ and finite fourth moments. For any function $f(\cdot)$ with continuous second derivatives in an open set about $(\rho_1^2, \cdots, \rho_p^2)$, if the population canonical correlations satisfy $1 > \rho_1 > \cdots > \rho_p > 0$, then we have

$$\sqrt{n}(\bar{f} - f) \to \mathrm{N}(0, \tau^2), \tag{32}$$

where $\tau^2 = \sum_{i,j=1}^{p} f_i f_j \tau_{ij}$, $f_j = \dfrac{\partial}{\partial \rho_j^2} f(\rho_1^2, \cdots, \rho_p^2)$, $f = f(\rho_1^2, \cdots, \rho_p^2)$ and τ_{ij} $(i, j = 1, \cdots, p)$ are given by

$$\begin{aligned} \tau_{ij} = \; & \rho_i^2 \rho_j^2 (\kappa_{ii}^{jj} + \kappa_{ii}^{p+j,p+j} + \kappa_{p+i,p+i}^{jj} + \kappa_{p+i,p+i}^{p+j,p+j}) \\ & -2\rho_i^2 \rho_j (\kappa_{ii}^{j,p+j} + \kappa_{p+i,p+i}^{j,p+j}) - 2\rho_i \rho_j^2 (\kappa_{i,p+i}^{jj} + \kappa_{i,p+i}^{p+j,p+j}) \\ & +4\rho_i \rho_j \kappa_{i,p+i}^{j,p+j}. \end{aligned} \tag{33}$$

Next we shall investigate the convergence of $\sum_{a=1}^{N} (f^a - \bar{f})^2/(N-1)$. We have

$$f(r_1^2(S_{-a}/(n-1)), \cdots, r_p^2(S_{-a}/(n-1))) = f(r_1^2(S/n), \cdots, r_p^2(S/n)) + \sum_{j=1}^{p} A_{aj} f_j(\xi_a), \tag{34}$$

where ξ_a is a point on the line segment between the vectors $(r_1^2(S/n), \cdots, r_p^2(S/n))$ and $(r_1^2(S_{-a}/(n-1)), \cdots, r_p^2(S_{-a}/(n-1))$. Then we have

$$\frac{1}{N-1}\sum_{a=1}^{N}(f^a - \bar{f})^2 = (N-1)\{\sum_{a=1}^{N}(\sum_{j=1}^{p}(U_{aj} - \bar{U}_{\cdot j} + V_{aj} - \bar{V}_{\cdot j}))^2\}, \qquad (35)$$

where $U_{aj} = A_{aj} f_j(r_1^2(S/n), \cdots, r_p^2(S/n))$, $V_{aj} = A_{aj}(f_j(\xi_a) - f_j(r_1^2(S/n), \cdots, r_p^2(S/n)))$, $\bar{U}_{\cdot j} = N^{-1}\sum_{a=1}^{N} U_{aj}$ and $\bar{V}_{\cdot j} = N^{-1}\sum_{a=1}^{N} V_{aj}$. Using (24), after some tedious calculations, we have

$$(N-1)\sum_{a=1}^{N}\{\sum_{j=1}^{p}(U_{aj} - \bar{U}_{\cdot j})\}^2 \to \tau^2 \quad \text{in probability.} \qquad (36)$$

Also $(N-1)\sum_{a=1}^{N}(\sum_{j=1}^{p}(V_{aj} - \bar{V}_{\cdot j}))^2$ converges in probability to zero since $(f_j(\xi_a) - f_j(r_1^2(S/n), \cdots, r_p^2(S/n)))$ converges in probability to zero for all a when $N \to \infty$. The convergence in probability of the term to zero follows from Schwarz's inequality. Therefore by using the notation of Theorem 4.1, we have

THEOREM 4.2. The statistic $\sum_{a=1}^{N}(f^a - \bar{f})^2/(N-1)$ converges in probability to τ^2, if all nonzero ρ_j are distinct.

Accordingly from Theorems 4.1 and 4.2 we have

THEOREM 4.3. Let the $(p+q)\times 1$ vectors $X_1, \cdots, X_N$ denote a random sample from a $(p+q)$-variate continuous distribution with mean μ, covariance matrix Σ and finite fourth moments. If non-zero canonical correlations ρ_j are all distinct, then for any function $f(\cdot)$ with continuous second derivatives about $(\rho_1^2, \cdots, \rho_p^2)$, we have

$$\frac{n(\bar{f} - f)}{\sqrt{\sum_{a=1}^{N}(f^a - \bar{f})^2}} \to \mathrm{N}(0,1), \qquad (37)$$

where $f = f(\rho_1^2, \cdots, \rho_p^2)$ and $n = N-1$.

Similarly we have

THEOREM 4.4. If non-zero canonical correlations ρ_j are all distinct, then we have

$$n(\bar{r}_1^2 - \rho_1^2, \cdots, \bar{r}_p^2 - \rho_p^2)\hat{\Omega}^{-1}(\bar{r}_1^2 - \rho_1^2, \cdots, \bar{r}_p^2 - \rho_p^2)' \to \chi_p^2,$$

where χ_p^2 is chi-square distribution with p degrees of freedom and the elements $\hat{\tau}_{ij}$ of the $p \times p$ matrix $\hat{\Omega} = (\hat{\tau}_{ij})$ are given by

$$\hat{\tau}_{ij} = \frac{1}{N-1}\sum_{a=1}^{N}(r_{i,a}^2 - \bar{r}_i^2)(r_{j,a}^2 - \bar{r}_j^2). \qquad (38)$$

4 The jackknife statistic when sample size $N = gh$ with fixed g and $h \to \infty$

Finally we consider the case of the size $N = gh$ where the group g is fixed and h is a sample size in each group with $h \to \infty$. The previous problem is the case $g = N \to \infty$ and $h = 1$. Let the $(p+q) \times 1$ vectors $X_1, \cdots, X_h, \cdots, X_{(g-1)h+1}, \cdots, X_{gh}$ denote a random sample from a $(p+q)$-variate distribution with mean μ, covariance matrix Σ and finite fourth moments. S^{-i}/n_g denotes the unbiased estimator of Σ based on a random sample obtained by deleting the ith group sample $X_{(i-1)h+1}, \cdots, X_{ih}$, where $n_g = h(g-1) - 1$. Then the pseudo-values and the jackknife statistic of $r_j^2(S/n)$ are given by

$$r_{i,j}^2 = g r_j^2(S/n) - (g-1) r_j^2(S^{-i}/n_g), \quad (i = 1, \cdots, g)$$

and

$$\bar{r}_j^2 = \frac{1}{g} \sum_{i=1}^{g} r_{i,j}^2, \tag{39}$$

respectively. Then S^{-i} is given by

$$S^{-i} = \sum_{a \notin \mathcal{A}_i, a=1}^{N} (X_a - \bar{X})(X_a - \bar{X})' - \frac{h^2}{N-h} (\bar{X}^i - \bar{X})(\bar{X}^i - \bar{X})', \tag{40}$$

where $\bar{X}^i = h^{-1} \sum_{a \in \mathcal{A}_i} X_a$ and $\mathcal{A}_i$ denotes the set $\{(i-1)h+1, ..., ih\}$. Now, if non-zero ρ_j is a simple root, using perturbation method (for example, see Dempster [4], Konishi and Gupta [6]), then for any sample covariance matrix S/n, we have

$$\begin{aligned} r_j^2(S/n) = \ & \rho_j^2 + 2\rho_j\left(\frac{s_{j,p+j}}{n} - \rho_j\right) - \rho_j^2\left(\frac{s_{j+p,j+p}}{n} - 1\right) \\ & -\rho_j^2\left(\frac{s_{j,j}}{n} - 1\right) + O_p(n^{-1}). \end{aligned} \tag{41}$$

On the other hand, as $h \to \infty$,

$$\frac{\sqrt{h} h^2}{(N-h) n_g} (\bar{X}^i - \bar{X})(\bar{X}^i - \bar{X})' \to 0, \quad \text{in probability.} \tag{42}$$

Thus from (40), after some calculations, we have

$$\begin{aligned} r_{i,j}^2 = \ & \rho_j^2 + 2\rho_j\left(\frac{v_{j,j+p}^i}{h} - \rho_j\right) - \rho_j^2\left(\frac{v_{p+j,p+j}^i}{h} - 1\right) \\ & -\rho_j^2\left(\frac{v_{j,j}^i}{h} - 1\right) + O_p(n_g^{-1}), \end{aligned} \tag{43}$$

where $v^i_{k\ell} = \sum_{a\in\mathcal{A}_i}(x_{ak} - \bar{x}_{\cdot k})(x_{a\ell} - \bar{x}_{\cdot\ell})$. Now putting $\bar{x}^i_{\cdot a} = h^{-1}\sum_{j=1}^{h} x_{(i-1)g+j,a}$, we have

$$v^i_{k\ell} = \sum_{a\in\mathcal{A}_i}(x_{ak} - \bar{x}^i_{\cdot k})(x_{a\ell} - \bar{x}^i_{\cdot\ell}) + h(\bar{x}^i_{\cdot k} - \bar{x}_{\cdot k})(\bar{x}^i_{\cdot\ell} - \bar{x}_{\cdot\ell}). \tag{44}$$

From (42), the second term of (44) times $1/\sqrt{h}$ converges in probability to zero. Thus for each j we have

$$\sqrt{h}(r^2_{i,j} - \rho^2_j) \to \mathrm{N}(0, \tau_{jj}), \tag{45}$$

where τ_{jj} is given by (16). Noting that $\mathcal{A}_i \cap \mathcal{A}_j = \phi$ $(i \neq j)$, we can show that $r^2_{i,j}$ $(i = 1, ..., g)$ are asymptotically independent. Therefore we have

THEOREM 5.1. Let the $(p+q)\times 1$ vectors $X_1, \cdots, X_h, \cdots, X_{(g-1)h+1}, \cdots, X_{gh}$ denote a random sample from a $(p+q)$-variate distribution with mean μ, covariance matrix Σ and finite fourth moments. If ρ^2_j is a non-zero simple root, then we have

$$\frac{\sqrt{g}(\bar{r}^2_j - \rho^2_j)}{\sqrt{\sum_{i=1}^{g}(r^2_{i,j} - \bar{r}^2_j)^2/(g-1)}} \to t_{g-1}, \tag{46}$$

where t_{g-1} denotes a t-distribution with $(g-1)$ degrees of freedom.

References

[1] Beran, R. and Srivastava, M. S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. Ann. Statist., 13, 95-115.

[2] Das, S. and Sen, P. K. (1995). Simultaneous spike-trains and stochastic dependence. Sankhyā, Ser. B. 57, 32-47.

[3] Das, S. and Sen, P. K. (1996). Asymptotic distribution of restricted canonical correlations and relevant resampling methods. J. Mult. Analy. 56, 1-19.

[4] Dempster, A. (1966). Estimation in multivariate analysis. In P. R. Krishnaiah (ed.), Multivariate Analysis, 1, 315-334, Academic Press, New York.

[5] Frangos, C. C. (1987). An updated bibliography on the jackknife method. Commun. Statist., 16, 1543-1584.

[6] Konishi, S. and Gupta, A. K. (1989). Testing the equality of several intraclass correlation coefficients. J. Statist. Plan. Inference, 19, 93-105.

[7] Miller, R. G. (1974). The jackknife: A review. Biometrika, 61, 1-15.

[8] Muirhead, R.J. and Waternaux, C. M. (1980). Asymptotic distributions in canonical correlation analysis and other multivariate procedures for nonnormal populations. Biometrika 67, 31-43.

[9] Nagao, H. (1985). On the limiting distributions of the jackknife statistics for eigenvalues of a sample covariance matrix. Commun. Statist., 14, 1547-1569.

[10] Nagao, H. (1988). On the jackknife statistics for eigenvalues and eigenvectors of a covariance matrix. Ann. Inst. Statist. Math., 40, 477-489.

[11] Nagao, H. (1990). On jackknife statistics of eigenvectors of a covariance matrix. Bull. Univ. Osaka Prefect., Ser. A, 39, 89-94.

[12] Parr, W. C. and Schucany, W. R. (1980). The jackknife: A bibliography. Inter. Statist. Review, 48, 73-78.

[13] Quenouille, M. (1956). Notes on bias in estimation. Biometrika, 43, 353-360.

[14] Srivastava, M. S. and Khatri, C. G. (1979). An Introduction to Multivariate Statistics, North-Holland, New York.

[15] Tukey, J. W. (1958). Bias and confidence in not quite large samples (abstract). Ann. Math. Statist., 29, 614.