On Jackknife Statistics of Eigenvectors of a Covariance Matrix

On Jackknife Statistics of Eigenvectors of a Covariance Matrix

Hisao NAGAO*

(Received June 16, 1990)

This paper is concerned with jackknife statistics of eigenvector of a covariance matrix. We shall give the limiting distributions of jackknife statistics under a general population, which results are useful in principal component analysis.

1. Introduction

Jackknife statistic has been defined by Quenouille[1] to reduce a bias of estimator. Also, Tukey[2] proposed some method to get confidence interval and testing hypothesis by applying pseudo-values in jackknife statistic. Miller[3] has written a nice review and given some guides which we should proceed in future. Frangos[4] has written recent references on jackknife method. One of many problems treated so far was to ask for the limiting distribution of jackknife statistic. The statistics treated so far were analytic function in neighbourhood of unknown parameter. But the eigenvalues and eigenvectors of a covariance matrix are not analytic function. So Nagao[5] has dealt with the problem of eigenvalues. The main purpose of the note is to give the limiting distribution of jackknife statistic of eigenvector by using implicit function theorem for several variables.

2. Limiting Distribution

Let $p\times 1$ vectors $X_1,\cdots,X_N$ be a random sample from a p-variate distribution with mean μ and covariance matrix Σ. Let $\lambda_1\geqq\cdots\geqq\lambda_p>0$ be eigenvalues of a covariance matrix Σ and let $p\times 1$ vector t_j stand for eigenvector corresponding to λ_j. In a previous paper[5] we have derived the limiting distributions of jackknife statistics of eigenvalues of a covariance matrix.

Let $l_1(S/n)\geqq\cdots\geqq l_p(S/n)$ be eigenvalues of a sample covariance matrix S/n, where

* Department of Mathematical Sciences, College of Engineering

$S=\sum_{\alpha=1}^{N}(X_\alpha-\bar{X})(X_\alpha-\bar{X})'$ with $n=N-1$ and let $p\times 1$ vector $h_j=(h_{1j},\cdots,h_{pj})'$ be an eigenvector corresponding to the j-th eigenvalue $l_j(S/n)$, where the norm of h_j is 1. At first we shall give the limiting distribution of the jackknife statistics of the component of h_j. A vector $h_j^\alpha=(h_{1j}^\alpha,\cdots,h_{pj}^\alpha)'$ stands for an eigenvector corresponding to $l_j(S_{-\alpha}/(n-1))$, where

$$S_{-\alpha}=S-\frac{N}{N-1}(X_\alpha-\bar{X})(X_\alpha-\bar{X})'. \tag{2.1}$$

Then the pseudo-values and the jackknife statistic of the β-th component $h_{\beta j}$ are given by

$$\tilde{h}_{\beta j}^\alpha=h_{\beta j}+(N-1)(h_{\beta j}-h_{\beta j}^{\alpha}),\quad (\alpha=1,\cdots,N) \tag{2.2}$$

and

$$\tilde{h}_{\beta j}=\frac{1}{N}\sum_{\alpha=1}^{N}\tilde{h}_{\beta j}^\alpha. \tag{2.3}$$

Let $y_\alpha=(y_{1\alpha},\cdots,y_{p,\alpha})'=T'x_\alpha$ $(\alpha=1,\cdots,N)$, then $h_j=Th_j^*$, where h_j^* is an eigenvector corresponding to $l_j(S^*/n)(=l_j(S/n))$ with $S^*=\sum_{\alpha=1}^{N}(y_\alpha-\bar{y})(y_\alpha-\bar{y})'$ and $\bar{y}=N^{-1}\sum_{\alpha=1}^{N}y_\alpha$, where T is an orthogonal matrix such that $T'\Sigma T=\mathrm{diag}(\lambda_1,\cdots,\lambda_p)$ with $\lambda_1\geqq\cdots\geqq\lambda_p>0$. Also $h_j^{*\alpha}=(h_{1j}^{*\alpha},\cdots,h_{pj}^{*\alpha})'$ denotes an eigenvector of $l_j(S_{-\alpha}^*/(n-1))$, where $S_{-\alpha}^*$ is a $p\times p$ matrix corresponding to $S_{-\alpha}$.
Thus we have

$$\tilde{h}_{\beta j}^\alpha=t_\beta' h_j^*+(N-1)(t_\beta' h_j^*-t_\beta' h_j^{*\alpha}), \tag{2.4}$$

where t_β' is the β-th row vector of the orthogonal matrix T. Accordingly we shall show that since t_β is a fixed vector,

$$\frac{n^{3/2}}{N}\sum_{\alpha=1}^{N}(h_j^*-h_j^{*\alpha})\longrightarrow 0 \text{ in probability.} \tag{2.5}$$

In order to prove the above, first of all, we need the following lemma, which plays a key role in this note.

Lemma, 2.1. Let a $p\times p$ matrix A be a real symmetric and put

$$A=\begin{bmatrix}A_{11} & a_p\\ a_p' & a_{pp}\end{bmatrix}, \tag{2.6}$$

where A_{11} is a square matrix of order $(p-1)$. We suppose that the rank of $(A_{11}-\lambda I)$ is $(p-1)$, where λ is an eigenvalue of A. Then a necessary and sufficient condition for $=(x_1',x_p)'$ with $(p-1)\times 1$ vector x_1 to be an eigenvector corresponding to the eigenvalue λ is $(A_{11}-\lambda I)x_1+a_p x_p=0$ and $x_p\neq 0$.

Proof. The necessary condition is obvious. For the sufficient condition, we have only to show that $a_p' x_1+(a_{pp}-\lambda)x_p=0$. Then we obtain $a_p' x_1+(a_{pp}-\lambda)x_p=\{-a_p'(A_{11}-\lambda I)^{-1}a_p+(a_{pp}-\lambda)\}\,x_p$. But $0=|A-\lambda I|=|A_{11}-\lambda I|\,\{(a_{pp}-\lambda)-a_p'(A_{11}-\lambda I)^{-1}a_p\}$. Therefore we obtain the desired conclusion.

For the sake of simplicity, we shall only prove the case of $j=p$, that is,

$$\frac{n^{3/2}}{N}\sum_{\alpha=1}^{N}(h_p^*-h_p^{*\alpha})\longrightarrow 0 \text{ in probability.} \tag{2.7}$$

To prove it, we shall use the implicit function theorem for several variables. Let us consider the following equation by Lemma 2.1.

$$F(S^*_{-\alpha}/(n-1),x_1)=(S^*_{11,\alpha}/(n-1)-l_p(S^*_{-\alpha}/(n-1))I)x_1+\frac{s^*_{p,\alpha}}{n-1}x_p=0, \tag{2.8}$$

where a matrix $S^*_{11,\alpha}$ and a vector $s^*_{p,\alpha}$ are submatrix and subvector of $S^*_{-\alpha}$ partitioned by the same manner as (2.6). By Lemma 2.1 we can take $h_p^*=(h_{1p}^{*\prime},x_p)'$ as the eigenvector corresponding to $l_p(S^*/n)$. That is, we can choose the same value for the p-th component of two eigenvectors h_p^* and $h_p^{*\alpha}$. Then we shall show that the solution of equation (2.8) on x_1 are analytic around $(S^*/n,\ h_{1p}^*)$. At first we have $F(S^*/n,\ h_{1p}^*)=0$. The partial derivative of $F(S^*/n,\ h_{1p}^*)$ with respect to h_{1p}^* is given by

$$\frac{\partial}{\partial h_{1p}^*}F(S^*/n,\ h_{1p}^*)=(S_{11}^*/n-l_p(S^*/n)I), \tag{2.9}$$

where S_{11}^* is a submatrix of $S^*=(s_{ij}^*)$ corresponding to $S^*_{11,\alpha}$. Since $S_{11}^*/n\longrightarrow\mathrm{diag}(\lambda_1,\cdots,\lambda_{p-1})$ and $l_p(S^*/n)\longrightarrow\lambda_p$ in probability, the matrix (2.9) converges to $\mathrm{diag}(\lambda_1-\lambda_p,\cdots,\lambda_{p-1}-\lambda_p)$ in probability. Hence if λ_p is a simple root, this diagonal matrix is a nonsingular.

By the implicit function theorem, we have

$$h_{1p}^{*\alpha}=h_{1p}^*+A_p^\alpha\langle t_{kl}^\alpha\rangle+\frac{1}{2}\begin{bmatrix}\langle t_{kl}^\alpha\rangle\ C_1^\alpha\ \langle t_{kl}^\alpha\rangle'\\ \cdots\cdots\cdots\cdots\\ \langle t_{kl}^\alpha\rangle\ C_{p-1}^\alpha\ \langle t_{kl}^\alpha\rangle'\end{bmatrix} \tag{2.10}$$

where

$$A_p^\alpha=-\Big(\frac{\partial}{\partial h_{1p}^*}F(S^*/n,\ h_{1p}^*)\Big)^{-1}\Big(\frac{\partial}{\partial(s_{kl}^*/n)}F(S^*/n,\ h_{1p}^*)\Big) \tag{2.11}$$

and s_{kl}^* is a (k,l)-element of the matrix S^* and each matrix C_i^α with order $p(p+1)/2$ is derivatives of the i-th row of A_p^α at some point. By the similar argument as in the previous paper[5], we can get the formula (2.7). Thus we obtain that $\sqrt{n}(\bar{h}_{\beta j}-t_\beta' h_j^*)$ converges to zero in probability.

By the perturbation method, if λ_j is a simple root, we have

$$h_j^* = -(s_{1j}^*/n(\lambda_1-\lambda_j),\cdots,1,\cdots,s_{pj}^*/n(\lambda_p-\lambda_j))' + o_p(n^{-1}), \tag{2.12}$$

where h_j^* is $p\times 1$ vector and the position of one is at the j-th component. Thus the limiting distribution of $\sqrt{n}(t_\beta' h_j^* - t_{\beta j})$ is a normal distribution with mean 0 and variance $\tau_{\beta j} = \sum_{k,l\neq j} \sigma_{kl;j} t_{k\beta} t_{l\beta}$ and

$$\sigma_{kl;j} = (\lambda_k-\lambda_j)^{-1}(\lambda_l-\lambda_j)^{-1}\mathrm{cov}((y_{\alpha k}-Ey_{\alpha k})(y_{\alpha j}-Ey_{\alpha j}),\ (y_{\alpha l}-Ey_{\alpha l})(y_{\alpha j}-Ey_{\alpha j})). \tag{2.13}$$

Theorem 2.1. Let $h_j = (h_{1j},\cdots,h_{pj})'$ $(h_{jj}>0)$ be an eigenvector of the length one corresponding to the j-th eigenvalue $l_j(S/n)$ of S/n. If the j-th eigenvalue λ_j of a covariance matrix Σ is a simple root, then we have

$$\sqrt{n}(\bar{h}_{\beta j} - t_{\beta j}) \longrightarrow N(0,\tau_{\beta j}), \tag{2.14}$$

where $\tau_{\beta j} = \sum_{k,l\neq j} \sigma_{kl;j} t_{k\beta} t_{l\beta}$ and $\sigma_{kl;j}$ is given by (2.13). Also $T=(t_{ij})$ with $t_{ii}>0$ $(i=1,\cdots,p)$ is an orthogonal matrix such that $T'\Sigma T = \mathrm{diag}(\lambda_1,\cdots,\lambda_p)$ for $\lambda_1 \geqq \cdots \geqq \lambda_p > 0$.

We note that though $t_{\beta j}$ is unique value because of the simplicity of λ_j, other terms $t_{\alpha\beta}$ contained in $\tau_{\beta j}$ may depend on the determination of an orthogonal matrix T. As we can see later, it turns out that the limiting distribution of our problem does not depend on T.

Anderson[6] has considered the testing hypothesis on an eigenvector in a multivariate normal case under the assumption that λ_j is a simple root.

Next we shall show that

$$\sum_{\alpha=1}^{N} (\bar{h}_{\beta j}^{\alpha} - \bar{h}_{\beta j})^2/(N-1) \longrightarrow \tau_{\beta j} \quad \text{in probability.} \tag{2.15}$$

As in (2.7), we show the case of $j=p$. Then we have only to show that

$$(N-1)\sum_{\alpha=1}^{N} (h_{ip}^{*\alpha} - \bar{h}_{ip}^{*})(h_{lp}^{*\alpha} - \bar{h}_{lp}^{*}) \longrightarrow \sigma_{il;p} \text{ in probability,} \tag{2.16}$$

where $\bar{h}_{ip}^{*} = N^{-1}\sum_{\alpha=1}^{N} h_{ip}^{*\alpha}$. Since $A_p^\alpha \longrightarrow DJ$ in probability, where

$$D = \mathrm{diag}((\lambda_1-\lambda_p)^{-1},\cdots,(\lambda_{p-1}-\lambda_p)^{-1}) \tag{2.17}$$

and

$$J = \begin{bmatrix} 0 & \vdots & \overset{(1,p)}{1} & \overset{(2,p)}{} & \cdots & \overset{(p-1,p)}{} \\ & \vdots & & 1 & & \\ & \vdots & & & & 1 \end{bmatrix}, \tag{2.18}$$

that is, the elements of $(p-1)\times p(p+1)/2$ matrix J are 1 at the i-th component of the (i,p) column, $(i=1,\cdots,p-1)$ and zero otherwise. By the similar calculation as section 3 in the previous paper[5] and tedious argument, we can get (2.15). Hence we can obtain

Theorem 2.2. Under the same assumption as Theorem 2.1, for $\tilde{h}^{\alpha}_{\beta j}$ *and* $\bar{h}_{\beta j}$ defined by (2.2) and (2.3), we have

$$\sum_{\alpha=1}^{N}(\tilde{h}^{\alpha}_{\beta j}-\bar{h}_{\beta j})^2/(N-1)\longrightarrow\tau_{\beta j} \text{ in probability.} \tag{2.19}$$

Hence we have

Theorem 2.3. Let $h_j=(h_{1j},\cdots,h_{pj})'$ $(h_{jj}>0)$ be the eigenvector with the length one corresponding to an eigenvalue $l_j(S/n)$ of S/n. If the j-th eigenvalue λ_j of Σ is a simple root, then we have

$$\frac{n(\bar{h}_{\beta j}-t_{\beta j})}{\sqrt{\sum_{\alpha=1}^{N}(\tilde{h}^{\alpha}_{\beta j}-\bar{h}^{\alpha}_{\beta j})^2}}\longrightarrow N(0,1), \tag{2.20}$$

where $t_{\beta j}$ is the β-th element of eigenvector corresponding to λ_j $(t_{jj}>0)$.

Finally we shall give the limiting distribution of the jackknife statistic with respect to an eigenvector h_j corresponding to an eigenvalue $l_j(S/n)$. Let us define the pseudo-values $\tilde{h}^{\alpha}_j$ $(\alpha=1,\cdots,N)$ and jackknife statistic $\bar{h}_j$ as follows:

$$\tilde{h}^{\alpha}_j=h_j+(N-1)(h_j-h^{\alpha}_j),\ (\alpha=1,\cdots,N) \tag{2.21}$$

and

$$\bar{h}_j=\frac{1}{N}\sum_{\alpha=1}^{N}\tilde{h}^{\alpha}_j.$$

After some tedious calculation, we have

Theorem 2.4. Let $\bar{h}_j$ and t_j be the jackknife statistics defined by (2.21) and an eigenvector corresponding to an eigenvalue λ_j of Σ. If λ_j is a simple root, then we have

$$n(\bar{h}_j-\tilde{t}_j)'U_n^{-1}(\bar{h}_j-\tilde{t}_j)\xrightarrow{\text{in law}}x^2_{(p-1)}, \tag{2.22}$$

where $(p-1)\times 1$ vectors $\bar{h}_j$, $\tilde{t}_j$ and $\tilde{h}^{\alpha}_j(j)$ obtaining by omitting the j-th components $\bar{h}_j$, t_j and $\tilde{h}^{\alpha}_j$ and $U_n=\sum_{\alpha=1}^{N}(\tilde{h}^{\alpha}_j(j)-\bar{h}_j)(\tilde{h}^{\alpha}_j(j)-\bar{h}_j)'/(N-1)$. $x^2_{(p-1)}$ stands for a chi-square distribution with $(p-1)$ degrees of freedom.

References

1) M. Quenouille, Biometrika, **43**, 353 (1956).
2) J.W. Tukey, Ann. Math. Statist., **29**, 614 (1958).
3) R.G. Miller Jr., Biometrika, **61**, 1 (1974).
4) C.C. Frangos, Commun. Statist.-Theory Math., **16**, 1543 (1987).
5) H. Nagao, Commun. Statist.-Theory Meth., **14**, 1547 (1985).
6) T.W. Anderson, Ann. Math. Statist., **34**, 122 (1963).