JOURNAL OF MULTIVARIATE ANALYSIS 47, 283–300 (1993)

# Asymptotics of Eigenvalues and Unit-Length Eigenvectors of Sample Variance and Correlation Matrices

TÕNU KOLLO

*University of Tartu, Estonia*

AND

HEINZ NEUDECKER

*University of Amsterdam, The Netherlands*

Multivariate asymptotic (normal) distributions for eigenvalues and unit-length eigenvectors of sample variance and correlation matrices are derived. Beside the general case, when existence of the (finite) fourth-order moments of the population distribution is assumed, formulae for the asymptotic variance matrices in the cases of normal and elliptical populations are also derived. It is assumed throughout that population variance and correlation matrices are nonsingular and without multiple eigenvalues. 

## INTRODUCTION

The first noteworthy results on the topic of this paper were obtained by Girshick [5]. In the case of the $p$-dimensional normal population, $x \sim N(\mu, \Sigma)$, he derived asymptotic variances and covariances of eigenvalues of the sample variance matrix $S$ and the sample correlation matrix $R$ assuming that the population variance matrix $\Sigma$ is nonsingular and without multiple roots: $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$. Under the same assumptions he got asymptotic variances and covariances of coordinates of eigenvectors (with

Received January 29, 1992; revised October 5, 1992.

AMS 1991 subject classification numbers: 15A09, 15A18, 15A52, 15A69, 62E20, 62H10.

Keywords and phrases: random matrices, multivariate asymptotic (normal) distributions, matrix differential calculus, Kronecker product, Hadamard product, vec operator, commutation matrix.



0047-259X/93 $5.00

length $\sqrt{\lambda_i}$) of the matrix $S$. Anderson [1] generalized Girshick's results to the case of multiple roots, $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p \geqslant 0$, and unit-length eigenvectors of $S$. Waternaux [14] found the asymptotic (normal) distribution of eigenvalues of $S$ for a nonnormal population. She assumed the existence of the (finite) fourth-order moments of the population distribution (let us denote this assumption by $M_4(x) < \infty$) and supposed that $\lambda_1 > \cdots > \lambda_p > 0$. Fujikoshi [4] derived the asymptotic expansion for the distribution function of an eigenvalue of $S$ under the assumptions of Waternaux [14]. Fang and Krishnaiah [3] generalized previous results and found asymptotic expansions for functions of $\lambda_i$ ($\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p \geqslant 0$) assuming the existence of the moments of the distribution, occurring in the expansions.

In matrix form the asymptotic distribution of eigenvalues of $S$ was presented by Kollo [7] for a class of distributions, including $N(\mu, \Sigma)$. In this paper we give in matrix form Waternaux's [14] result (in Theorem 5).

The asymptotic distribution of coordinates of unit-length eigenvectors of $S$ under the assumption $M_4(x) < \infty$ was obtained by Davis [2]. He presented the distribution assuming that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p > 0$. In Kollo [7] the asymptotic distribution of eigenvectors of $S$ was presented in matrix form for a class, including $N(\mu, \Sigma)$.

The asymptotic behaviour of eigenvalues and eigenvectors of the sample correlation matrix $R$ is more complicated. For a normal population $x \sim N(\mu, \Sigma)$ the distribution of eigenvalues of $R$ was obtained by Girshick [5] under the assumptions indicated above. Kollo [6] presented the asymptotic variance matrix of eigenvalues of $R$ for a normal population $N(\mu, \Sigma)$. Konishi [9] got the asymptotic distribution and first terms of the Edgeworth expansion of the distribution function of eigenvalues of matrix $R$ for a normal population $x \sim N(\mu, \Sigma)$ under the assumptions $\lambda_1 \geqslant \cdots \geqslant \lambda_p \geqslant 0$. Kollo [8] presented the multivariate asymptotic (normal) distribution of eigenvalues of $R$ assuming $\lambda_1 > \cdots > \lambda_p > 0$ and $M_4(x) < \infty$. Fang and Krishnaiah [3] found asymptotic expansions for functions of eigenvalues of $R$, from which as a special case we get the asymptotic (normal) distribution of eigenvalues of $R$ under the assumptions $M_4(x) < \infty$ and $\lambda_1 \geqslant \cdots \geqslant \lambda_p \geqslant 0$.

For the unit-length eigenvectors of $R$ the asymptotic (normal) distribution was obtained by Kollo [7] under the assumption $\lambda_1 > \cdots > \lambda_p > 0$ for a class of distributions including $N(\mu, \Sigma)$. Konishi [9] derived asymptotic (normal) distributions and asymptotic expansions of coordinates of eigenvectors of $R$ and found expansions of asymptotic covariances between the coordinates assuming that $x \sim N(\mu, \Sigma)$ and $\lambda_1 \geqslant \cdots \geqslant \lambda_p \geqslant 0$.

## 2. Basic Convergence Results

Let $X' = (x_1 \cdots x_n)$ be a sample of $n$ independent and identically distributed observations from a $p$-dimensional population that is characterized by the $p \times 1$ random vector $x$ with mean $E(x) = \mu$, (nonsingular) variance matrix $D(x) = \Sigma$, and fourth-order central moment $M_4(x) = E[(x-\mu)(x-\mu)' \otimes (x-\mu)(x-\mu)'] < \infty$. Let $\bar{x}$ and $S$ be the usual unbiased estimators of $\mu$ and $\Sigma$, respectively. Let us denote the correlation matrix for $x$ by $P$ and its estimator by $R$; then

$$P = \Sigma_{\mathrm{d}}^{-1/2} \Sigma \Sigma_{\mathrm{d}}^{-1/2} \qquad \text{and} \qquad R = S_{\mathrm{d}}^{-1/2} S S_{\mathrm{d}}^{-1/2}.$$

(The notation $A_{\mathrm{d}}$ is used to indicate the diagonal $p \times p$ matrix with diagonal elements $a_{11} \cdots a_{pp}$ of the $p \times p$ matrix $A$.)

In this paper we derive in matrix form the asymptotic (normal) distributions for eigenvalues and unit-length eigenvectors of $S$ and $R$. For the definition of asymptotic distribution see Theorem 1.

The main basic results used in the paper are the two following theorems.

Theorem 1. *Let $\{y_{(n)}\}$ be a sequence of random vectors $y_{(n)}$ and $b$ a compatible fixed vector. Assume that $\sqrt{n}\,[y_{(n)} - b] \xrightarrow{\mathrm{D}} N(0, T)$ (convergence in distribution), or equivalently $\sqrt{n}\, y_{(n)}$ is asymptotically normally distributed with mean $\sqrt{n}\, b$ and variance $T$.*

*Let $f(z)$ be a vector function of a vector $z$ with first and second derivatives existing in a neighbourhood of $z = b$. Then*

$$\sqrt{n}\,[f(y_{(n)}) - f(b)] \xrightarrow{\mathrm{D}} N(0, \Phi T \Phi'),$$

*where*

$$\Phi = \left.\frac{\delta f(z)}{\delta z'}\right|_{z=b}$$

*is a matrix derivative.*

Theorem 2 (Parring [13], Neudecker and Wesselman [12]. *Let $X' = (x_1 \cdots x_n)$ be a sample of $n$ independent and identically distributed observations from a p-dimensional population with $E(x) = \mu$, $D(x) = \Sigma$, $M_4(x) < \infty$.*

*Let $S_{(n)} = (n-1)^{-1} X'NX$, $N = I - n^{-1}U$, $U = \iota\iota'$, $\iota = (1 \cdots 1)'$. Then, when $n \to \infty$,*

$$\sqrt{n}\, \mathrm{vec}(S_{(n)} - \Sigma) \xrightarrow{\mathrm{D}} N(0, V), \tag{1}$$

*where*

$$V = M_4(x) - (\operatorname{vec} \Sigma)(\operatorname{vec} \Sigma)'. \tag{2}$$

The statement of Theorem 2 can be phrased alternatively

$$\sqrt{n}\, v(S_{(n)} - \Sigma) \xrightarrow{\mathrm{D}} N(0, D^+ V D^{+\prime}), \tag{3}$$

where $D$ is the appropriate duplication matrix and $D^+$ its Moore–Penrose inverse.

As usual $v(\cdot)$ is the shortened version of $\operatorname{vec}(\cdot)$ with

$$v(\cdot) = D^+ \operatorname{vec}(\cdot)$$

and

$$\operatorname{vec}(\cdot) = Dv(\cdot)$$

for symmetric matrices.

On matrix theory and especially duplication matrix, $\operatorname{vec}(\cdot)$, and $v(\cdot)$ operators see Magnus and Neudecker [10]. Reference is also made to Appendix II of the present paper.

It is well known that the finite-sample variance matrix of $\sqrt{n} \operatorname{vec} S_{(n)}$ is

$$M_4(x) + (n-1)^{-1} (I+K)(\Sigma \otimes \Sigma) - (\operatorname{vec} \Sigma)(\operatorname{vec} \Sigma)'. \tag{4}$$

(A proof can be found in Appendix I.)

For the *normally* distributed random vector $x \sim N(\mu, \Sigma)$, the (asymptotic variance) matrix $V$ has the form

$$V = (I+K)(\Sigma \otimes \Sigma), \tag{5}$$

where $K$ is the appropriate commutation matrix.

If it is assumed that the random vector is *elliptically* distributed (see Muirhead [11, Sect. 1.5], for example), then the asymptotic variance matrix is

$$V = (1+\kappa)(I+K)(\Sigma \otimes \Sigma) + \kappa(\operatorname{vec} \Sigma)(\operatorname{vec} \Sigma)', \tag{6}$$

where $\kappa$ is the common kurtosis coefficient.

For the correlation matrix $R_{(n)}$ a similar convergence takes place:

THEOREM 3 (Neudecker and Wesselman [12]). *Under the assumptions and definitions of the foregoing theorem we have, when $n \to \infty$,*

$$\sqrt{n} \operatorname{vec}(R_{(n)} - P) \xrightarrow{\mathrm{D}} N(0, \Psi), \tag{6}$$

*where*

$$\begin{aligned}\Psi &= \{I-\tfrac{1}{2}(I+K)(I\otimes P)\,K_{\mathrm{d}}\}(\Sigma_{\mathrm{d}}^{-1/2}\otimes\Sigma_{\mathrm{d}}^{-1/2})\\ &\quad\cdot V(\Sigma_{\mathrm{d}}^{-1/2}\otimes\Sigma_{\mathrm{d}}^{-1/2})\{I-\tfrac{1}{2}K_{\mathrm{d}}(I\otimes P)(I+K)\}\\ &= \{\Sigma_{\mathrm{d}}^{-1/2}\otimes\Sigma_{\mathrm{d}}^{-1/2}-\tfrac{1}{2}(I\otimes P\Sigma_{\mathrm{d}}^{-1}+P\Sigma_{\mathrm{d}}^{-1}\otimes I)\,K_{\mathrm{d}}\}\\ &\quad\cdot V\{\Sigma_{\mathrm{d}}^{-1/2}\otimes\Sigma_{\mathrm{d}}^{-1/2}-\tfrac{1}{2}K_{\mathrm{d}}(I\otimes\Sigma_{\mathrm{d}}^{-1}P+\Sigma_{\mathrm{d}}^{-1}P\otimes I)\}.\end{aligned}$$

*The equality follows readily from the properties* $KK_{\mathrm{d}}=K_{\mathrm{d}}$, $KV=V$, *and* $(M\otimes N)\,K_{\mathrm{d}}=K_{\mathrm{d}}(M\otimes N)=(I\otimes MN)\,K_{\mathrm{d}}$ *for* diagonal *matrices* $M$ *and* $N$.

For the *normal* case we have

$$\Psi = A_1 - A_2 - A_2' + A_3, \tag{7}$$

where

$$\begin{aligned}A_1 &= (I+K)(P\otimes P)\\ A_2 &= (P\otimes P)\,K_{\mathrm{d}}(I\otimes P+P\otimes I)\\ A_3 &= \tfrac{1}{2}(I\otimes P+P\otimes I)\,K_{\mathrm{d}}(P\otimes P)\,K_{\mathrm{d}}(I\otimes P+P\otimes I).\end{aligned} \tag{8}$$

For the *elliptical* case we have

$$\Psi = B_1 - B_2 - B_2' + B_3, \tag{9}$$

where

$$\begin{aligned}B_1 &= (1+\kappa)(I+K)(P\otimes P)+\kappa(\operatorname{vec}P)(\operatorname{vec}P)'\\ B_2 &= [(1+\kappa)(P\otimes P)-\tfrac{1}{2}\kappa(\operatorname{vec}P)(\operatorname{vec}P)']\,K_{\mathrm{d}}(I\otimes P+P\otimes I)\\ B_3 &= \tfrac{1}{2}(I\otimes P+P\otimes I)\,K_{\mathrm{d}}[(1+\kappa)(P\otimes P)\\ &\quad+\tfrac{1}{2}\kappa(\operatorname{vec}P)(\operatorname{vec}P)']\,K_{\mathrm{d}}(I\otimes P+P\otimes I).\end{aligned} \tag{10}$$

Having recapitulated these basic results we now look into the general problem of asymptotic distributions of eigenvalues and unit-length eigenvectors of a random matrix.

## 3. Asymptotics of Eigenvalues and Unit-Length Eigenvectors of a Random Matrix

Let $M$ be a real symmetric $p\times p$ matrix with eigenvalues $\lambda_1>\lambda_2>\cdots>\lambda_p>0$ and associated orthonormal (i.e., orthogonal with unit length) eigenvectors $w_i$ $(i=1\cdots p)$. Thus

$$\begin{aligned}MW &= W\Lambda\\ W'W &= I,\end{aligned} \tag{11}$$

where $W = (w_1 \cdots w_p)$ and $\Lambda$ is an appropriate diagonal matrix of eigenvalues of $M$.

Let us then consider a real symmetric random matrix $\hat{M}_{(n)}$ ($n$ again being sample size) with eigenvalues $\hat{\lambda}_{(n)i}$ and orthonormal eigenvectors $\hat{w}_{(n)i}$, these being estimators of the aformentioned parameters. (We omit the sample-size indicator in the further development, whenever that is sensible.)

We have

$$\begin{aligned} \hat{M}\hat{W} &= \hat{W}\hat{\Lambda} \\ \hat{W}'\hat{W} &= I. \end{aligned} \tag{12}$$

The following convergence is *assumed* to hold

$$\sqrt{n}\,\mathrm{vec}(\hat{M}_{(n)} - M) \xrightarrow{\mathrm{D}} N(0, G)$$

or alternatively phrased

$$\sqrt{n}\,v(\hat{M}_{(n)} - M) \xrightarrow{\mathrm{D}} N(0, D^{+}GD^{+\prime}). \tag{13}$$

Consider now the vector functions $f_i\{v(Z)\}$ and the scalar functions $\psi_i\{v(Z)\}$ of a vector $v(Z)$ with first and second derivatives existing in a neighbourhood of $v(Z) = v(M)$, such that

$$\begin{aligned} \hat{w}_i &= f_i\{v(\hat{M})\}, \qquad \hat{\lambda}_i = \psi_i\{v(\hat{M})\}, \qquad \text{and} \\ w_i &= f_i\{v(M)\}, \qquad \lambda_i = \psi_i\{v(M)\}. \end{aligned}$$

By using Theorem 1 we can then derive the asymptotic distributions of $\hat{\lambda}_i$ and $\hat{w}_i$. For this purpose we need the first derivatives. These are presented in the following lemma.

LEMMA 1.

$$\left.\frac{\delta f_i\{v(Z)\}}{\delta v'(Z)}\right|_{Z=M} = [w_i' \otimes W(\lambda_i I - \Lambda)^{+} W']D.$$

*If further* $\psi = (\psi_1 \cdots \psi_p)'$, *then*

$$\left.\frac{\delta \psi\{v(Z)\}}{\delta v'(Z)}\right|_{Z=M} = J'(W' \otimes W')D,$$

*where* $J = (e_1 \otimes e_1 \cdots e_p \otimes e_p)$, $e_i$ *being the compatible* $i$*th unit vector.*

*Proof.* We differentiate $Zf_i = \psi_i f_i$ in the point $Z = M$, the perturbations being symmetric. This leads to the following sequence of results:

$$(dZ)\, w_i + M\, df_i = (d\psi_i)\, w_i + \lambda_i\, df_i$$
$$w_j'(dZ)\, w_i + w_j' M\, df_i = (d\psi_i)\, w_j' w_i + \lambda_i w_j'\, df_i \qquad (j \neq i)$$
$$w_j'(dZ)\, w_i = (\lambda_i - \lambda_j)\, w_j'\, df_i \qquad \text{as} \quad w_j' w_i = 0.$$

Further

$$\begin{aligned} df_i &= I\, df_i = \sum_j w_j w_j'\, df_i = \sum_{j \neq i} w_j w_j'\, df_i \qquad (\text{as } w_i'\, df_i = 0) \\ &= \sum_{j \neq i} (\lambda_i - \lambda_j)^{-1} (\lambda_i - \lambda_j)\, w_j w_j'\, df_i = \sum_{j \neq i} (\lambda_i - \lambda_j)^{-1}\, w_j w_j' (dZ)\, w_i \\ &= \left\{ w_i' \otimes \sum_{j \neq i} (\lambda_i - \lambda_j)^{-1}\, w_j w_j' \right\} D\, dv(Z) \\ &= \{ w_i' \otimes W(\lambda_i I - \Lambda)^+\, W' \}\, D\, dv(Z). \end{aligned}$$

From this we derive

$$\left. \frac{\delta f_i\{v(Z)\}}{\delta v'(Z)} \right|_{Z = M} = [w_i' \otimes W(\lambda_i I - \Lambda)^+\, W']\, D.$$

Using $d\psi_i = w_i'(dZ)\, w_i$ we get

$$\begin{aligned} d\psi &= (w_1 \otimes w_1 \cdots w_p \otimes w_p)'\, d \operatorname{vec} Z \\ &= \{(W \otimes W)(e_1 \otimes e_1 \cdots e_p \otimes e_p)\}'\, D\, dv(Z) \\ &= J'(W' \otimes W')\, D\, dv(Z). \end{aligned}$$

From this follows

$$\left. \frac{\delta \psi_i\{v(Z)\}}{\delta v'(Z)} \right|_{Z = M} = J'(W' \otimes W')D.$$

As immediate corollaries of Theorem 1 and Lemma 1 we have the convergence results listed in the following theorem:

THEOREM 4. *Under the assumptions and definitions of Lemma* 1 *and convergence assumption* (13) *we have, when* $n \to \infty$,

1. $\sqrt{n}\,(\hat{w}_{(n)i} - w_i) \xrightarrow{\mathrm{D}} N(0, T_i)$,

*where*

$$T_i = [w_i' \otimes W(\lambda_i I - \Lambda)^+\, W']\, G[w_i \otimes W(\lambda_i I - \Lambda)^+\, W'];$$

2. $\sqrt{n}\,(\hat{\lambda}_{(n)}-\lambda)\xrightarrow{\mathrm{D}} N[0,\ J'(W'\otimes W')\ G(W\otimes W)J]$,

*where*

$$\lambda=(\lambda_1\cdots\lambda_p)' \text{ and } \hat{\lambda}_{(n)}=(\hat{\lambda}_{(n)1}\cdots\hat{\lambda}_{(n)p})'.$$

From Theorem 1 we obtain by the convergence (13) and Lemma 1 the equality

$$T_i=(w_i'\otimes W(\lambda_i I-A)^+\ W']\ DD^+GD^{+\prime}D'[w_i\otimes W(\lambda_i I-A)^+\ W'].$$

As $DD^+A=A$ for any matrix $A$ that is the variance matrix of a symmetric stochastic matrix, we get the equality

$$DD^+GD^{+\prime}D'=G,$$

which proves the first statement.

A consequence of this fact is that there is no need to replace differentials such as $d\operatorname{vec} Z$ by $D\,dv(Z)$ in the differentiation process set out in Lemma 1.

In the next section we consider eigenvalues and unit-length eigenvectors of the sample variance matrix $S$.

## 4. Asymptotic Distributions of Eigenvalues and Unit-Length Eigenvectors of the Sample Variance Matrix $S$

The results of Section 3 can be immediately applied to the sample variance matrix $S$. This yields the following result for the eigenvalues, obtained by Waternaux [14] in the elementwise form:

THEOREM 5. *Let the population variance matrix $\Sigma$ have eigenvalues $\lambda_1>\lambda_2>\cdots>\lambda_p$ and associated orthonormal eigenvectors $w_i$ $(i=1\cdots p)$. The latter are assembled in the orthogonal matrix $W$. Let the sample variance matrix $S_{(n)}$—where $n$ is the sample size—have eigenvalues $\hat{\lambda}_{(n)i}$ $(i=1,\ldots,p)$, which are estimators of the $\lambda_i$.*

*Let Theorem 2 apply. Then, when $n\to\infty$,*

$$\sqrt{n}\,(\hat{\lambda}_{(n)}-\lambda)\xrightarrow{\mathrm{D}} N[0,\ J'(W'\otimes W')\ V(W\otimes W)J],$$

*where $V$ is the asymptotic variance matrix of $\sqrt{n}\operatorname{vec} S_{(n)}$, $J=(e_1\otimes e_1\cdots e_p\otimes e_p)$, $W=(w_1\cdots w_p)$, $\lambda=(\lambda_1\cdots\lambda_p)'$, and $\hat{\lambda}_{(n)}=(\hat{\lambda}_{(n)1}\cdots\hat{\lambda}_{(n)p})'$.* Hence *$J'(W'\otimes W')\ V(W\otimes W)J$ is the asymptotic variance matrix of $\sqrt{n}\,\hat{\lambda}_{(n)}$.*

Using the equality $V=M_4(x)-(\operatorname{vec}\Sigma)(\operatorname{vec}\Sigma)'$ we can write the asymptotic variance matrix as $J'(W'\otimes W')\ M_4(x)(W\otimes W)\,J+\Lambda U\Lambda$. (See Theorem 2 for the definition of $U$.)

For the eigenvectors of $S_{(n)}$ we have the result:

THEOREM 6. *Under the assumptions and definitions of Theorem* 5, *when further $\hat{w}_{(n)i}$ are orthonormal eigenvectors of $S_{(n)}$ associated with $\hat{\lambda}_i$, we have for $n \to \infty$*

$$\sqrt{n}\,(\hat{w}_{(n)i} - w_i) \xrightarrow{\mathrm{D}} N(0, T_i),$$

*where*

$$\begin{aligned} T_i &= \{w_i' \otimes W(\lambda_i I - \Lambda)^+ W'\}\, V\{w_i \otimes W(\lambda_i I - \Lambda)^+ W'\} \\ &= \{w_i' \otimes W(\lambda_i I - \Lambda)^+ W'\}\, M_4(x)\{w_i \otimes W(\lambda_i I - \Lambda)^+ W'\}. \end{aligned}$$

*Comment.* The second equality is obtained from

$$\begin{aligned} \{w_i' \otimes W(\lambda_i I - \Lambda)^+ W'\} \operatorname{vec} \Sigma &= W(\lambda_i I - \Lambda)^+ W'\Sigma w_i \\ &= \lambda_i W(\lambda_i I - \Lambda)^+ W' w_i = 0. \end{aligned}$$

It is in order to specialize Theorems 5 and 6 to the normal and elliptical cases. We get

COROLLARY 1 (Kollo [7]). *For the normal case, when $n \to \infty$,*

$$\begin{aligned} &\sqrt{n}\,(\hat{\lambda}_{(n)} - \lambda) \xrightarrow{\mathrm{D}} N(0, 2\Lambda^2), \\ &\sqrt{n}\,(\hat{w}_{(n)i} - w_i) \xrightarrow{\mathrm{D}} N[0, \lambda_i W\Lambda(\lambda_i I - \Lambda)^{+^2} W']. \end{aligned}$$

*Proof.* Now

$$\begin{aligned} J'(W' \otimes W')\, V(W \otimes W) J &= J'(W' \otimes W')(I + K)(\Sigma \otimes \Sigma)(W \otimes W) J \\ &= 2J'(W' \otimes W')(\Sigma \otimes \Sigma)(W \otimes W) J \\ &= 2J'(W'\Sigma W \otimes W'\Sigma W) J \\ &= 2J'(\Lambda \otimes \Lambda)\, J = 2\Lambda \times \Lambda = 2\Lambda^2, \end{aligned}$$

as

$$(I + K)\, J = (I + K)(e_1 \otimes e_1 \cdots e_p \otimes e_p) = 2(e_1 \otimes e_1 \cdots e_p \otimes e_p) = 2J.$$

We further use

$$\begin{aligned} &\{w_i' \otimes W(\lambda_i I - \Lambda)^+ W'\}(I + K)(\Sigma \otimes \Sigma) \\ &\qquad = w_i'\Sigma \otimes W(\lambda_i I - \Lambda)^+ W'\Sigma + W(\lambda_i I - \Lambda)^+ W'\Sigma \otimes w_i'\Sigma \\ &\qquad = \lambda_i w_i' \otimes W(\lambda_i I - \Lambda)^+ \Lambda W' + W\Lambda(\lambda_i I - \Lambda)^+ W' \otimes \lambda_i w_i', \end{aligned}$$

which upon postmultiplication by $w_i \otimes W(\lambda_i I - \Lambda)^+ W'$ yields $\lambda_i W\Lambda(\lambda_i I - \Lambda)^{+2} W'$, as $(\lambda_i I - \Lambda)^+ e_i = 0$.

COROLLARY 2. *For the elliptical case, when $n \to \infty$,*

$$\sqrt{n}\,(\hat{\lambda}_{(n)} - \lambda) \xrightarrow{\mathrm{D}} N[0, 2(1+\kappa)\,\Lambda^2 + \kappa\Lambda U\Lambda],$$

$$\sqrt{n}\,(\hat{w}_{(n)i} - w_i) \xrightarrow{\mathrm{D}} N[0, \lambda_i(1+\kappa)\,W\Lambda(\lambda_i I - \Lambda)^{+2}\,W'].$$

*Proof.* Substitute $V = (1+\kappa)(I+K)(\Sigma \otimes \Sigma) + \kappa(\operatorname{vec}\Sigma)(\operatorname{vec}\Sigma)'$ in the expression for the asymptotic variance matrix $J'(W' \otimes W')\,V(W \otimes W)J$ of Theorem 5.

Further substitute $V = (1+\kappa)(I+K)(\Sigma \otimes \Sigma) + \kappa(\operatorname{vec}\Sigma)(\operatorname{vec}\Sigma)'$ in the expression for the variance matrix $T_i$ of Theorem 6.

This finishes the section on asymptotics of eigenvalues and unit-length eigenvectors of the sample variance matrix $S$.

## 5. ASYMPTOTICS OF EIGENVALUES AND UNIT-LENGTH EIGENVECTORS OF THE SAMPLE CORRELATION MATRIX $R$

We defined the sample correlation matrix $R$ and the population correlation matrix $P$,

$$R = S_{\mathrm{d}}^{-1/2} S S_{\mathrm{d}}^{-1/2}$$

$$P = \Sigma_{\mathrm{d}}^{-1/2} \Sigma \Sigma_{\mathrm{d}}^{-1/2},$$

where $S(\Sigma)$ is the sample (population) variance matrix and $S_{\mathrm{d}}(\Sigma_{\mathrm{d}})$ is the diagonal matrix displaying the diagonal elements of $S(\Sigma)$. We do the same analysis as in Section 4, but now applied to the eigenvalues and unit-length eigenvectors of $R$.

Although this may be confusing, the same notation is used for eigenvalues and eigenvectors as before. This leads in the first instance to

THEOREM 7 (Kollo [8]). *Let the population correlation matrix $P$ have eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$ and associated orthonormal eigenvectors $w_i$ $(i = 1 \cdots p)$. Let the sample correlation matrix $R_{(n)}$—where $n$ is the sample size—have eigenvalues $\hat{\lambda}_{(n)i}$ in decreasing order. It is assumed that Theorem 3 applies. Then, when $n \to \infty$,*

$$\sqrt{n}\,(\hat{\lambda}_{(n)} - \lambda) \xrightarrow{\mathrm{D}} N[0, J'(W \otimes W)'\,\Psi(W \otimes W)J],$$

*where $\Psi$ is the asymptotic variance matrix of* $\sqrt{n}\,\operatorname{vec} R_{(n)}$ as derived in Theorem 3, $J=(e_1\otimes e_1 \cdots e_p\otimes e_p)$, $\lambda=(\lambda_1\cdots\lambda_p)'$; $\hat{\lambda}_{(n)}=(\hat{\lambda}_{(n)1}\cdots\hat{\lambda}_{(n)p})'$, and $W=(w_1\cdots w_p)$.

The following result is

THEOREM 8. *Under the assumptions and definitions of Theorem 7 we have, when $n\to\infty$,*

$$\sqrt{n}\,(\hat{w}_{(n)i}-w_i)\xrightarrow{\text{D}} N(0,T_i),$$

*where*

$$T_i=\{w_i'\otimes W(\lambda_i I-\Lambda)^+\,W'\}\,\Psi\{w_i\otimes W(\lambda_i I-\Lambda)^+\,W'\}.$$

Again we give the specializations for the *normal* and *elliptical* cases.

COROLLARY 3. *For the normal case, when $n\to\infty$,*

$$\sqrt{n}\,(\hat{\lambda}_{(n)}-\lambda)\xrightarrow{\text{D}} N(0,\Gamma),$$

*where*

$$\begin{aligned}\Gamma=2\Lambda\{&I-\Lambda(W\times W)'\,(W\times W)-(W\times W)'\,(W\times W)\,\Lambda\\&+(W\times W)'\,(P\times P)(W\times W)\}\,\Lambda.\end{aligned}$$

*The expression $W\times W$ is the elementwise (Hadamard) matrix product.*

*Proof.* We use $\Psi=A_1-A_2-A_2'+A_3$ (see (7) for the definitions). This yields for the asymptotic variance of $\sqrt{n}\,(\hat{\lambda}_{(n)})$, in rough form,

$$J'(W'\otimes W')(A_1-A_2-A_2'+A_3)(W\otimes W)J.$$

The following simplifications hold:

$$\begin{aligned}&J'(W'\otimes W')\,A_1(W\otimes W)J\\&\quad=J'(W'\otimes W')(I+K)(P\otimes P)(W\otimes W)J\\&\quad=J'(I+K)(W'PW\otimes W'PW)J\\&\quad=2J'(\Lambda\otimes\Lambda)\,J=2\Lambda\times\Lambda=2\Lambda^2;\\&J'(W'\otimes W')\,A_2(W\otimes W)J\\&\quad=J'(W'\otimes W')(P\otimes P)\,K_{\text{d}}(I\otimes P+P\otimes I)(W\otimes W)J\\&\quad=J'(\Lambda W'\otimes\Lambda W')\,K_{\text{d}}(W\otimes W\Lambda+W\Lambda\otimes W)J\end{aligned}$$

$$= J'(\Lambda \otimes \Lambda)(W' \otimes W')\, K_{\mathrm{d}}(W \otimes W)(I \otimes \Lambda + \Lambda \otimes I)J$$

$$= J'(\Lambda \otimes \Lambda)\, JJ'(W' \otimes W')\, JJ'(W \otimes W)\, JJ'(I \otimes \Lambda + \Lambda \otimes I)J$$

$$= \Lambda^2(W' \times W')(W \times W)(I \times \Lambda + \Lambda \times I)$$

$$= 2\Lambda^2(W \times W)'\,(W \times W)\Lambda;$$

$$J'(W' \otimes W')A_3(W \otimes W)J$$

$$= \tfrac{1}{2}J'(W' \otimes W')(I \otimes P + P \otimes I)\, K_{\mathrm{d}}(P \otimes P)$$

$$\cdot K_{\mathrm{d}}(I \otimes P + P \otimes I)(W \otimes W)J$$

$$= \tfrac{1}{2}J'(I \otimes \Lambda + \Lambda \otimes I)(W \otimes W)'\, JJ'(P \otimes P)$$

$$\cdot JJ'(W \otimes W)(I \otimes \Lambda + \Lambda \otimes I)J$$

$$= \tfrac{1}{2}J'(I \otimes \Lambda + \Lambda \otimes I)\, JJ'(W \otimes W)'\, JJ'(P \otimes P)$$

$$\cdot JJ'(W \otimes W)\, JJ'(I \otimes \Lambda + \Lambda \otimes I)J$$

$$= 2\Lambda(W \times W)'\,(P \times P)(W \times W)\Lambda.$$

For definitions and properties used in this proof see Appendix II.

Collecting the four terms we get the asymptotic variance matrix of $\sqrt{n}\,\hat{\lambda}_{(n)}$:

$$2\Lambda\{I - \Lambda(W \times W)'\,(W \times W) - (W \times W)'\,(W \times W)\Lambda$$
$$+ (W \times W)'\,(P \times P)(W \times W)\}\Lambda.$$

COROLLARY 4. *For the normal case, when $n \to \infty$,*

$$\sqrt{n}\,(\hat{w}_{(n)i} - w_i) \xrightarrow{\mathrm{D}} N(0, \Xi_i),$$

where

$$\Xi_i = \lambda_i W\Lambda_{(i)}^2 \Lambda W' - \lambda_i W\Lambda_{(i)}\Lambda W'\Delta^2(w_i)\, W(\Lambda + \lambda_i I)\, \Lambda_{(i)} W'$$
$$- \lambda_i W\Lambda_{(i)}(\Lambda + \lambda_i I)\, W'\Delta^2(w_i)\, W\Lambda\Lambda_{(i)} W'$$
$$+ \tfrac{1}{2} W\Lambda_{(i)}(\Lambda + \lambda_i I)\, W'\Delta(w_i)(P \times P)\,\Delta(w_i)\, W(\Lambda + \lambda_i I)\, \Lambda_{(i)} W',$$

*where $\Lambda_{(i)} = (\lambda_i I - \Lambda)^+$ and $\Delta(w_i)$ is the diagonal-matrix representation of $w_i$.*

*Proof.* Again we use $\Psi = A_1 - A_2 - A_2' + A_3$.

We evaluate the four terms. The first is

$$
\begin{aligned}
&(w_i' \otimes WA_{(i)}W')\, A_1 (w_i \otimes WA_{(i)}\, W') \\
&\quad = (w_i' \otimes WA_{(i)} W')(I+K)(P \otimes P)(w_i \otimes WA_{(i)} W') \\
&\quad = (w_i' \otimes WA_{(i)} W')(I+K)(\lambda_i w_i \otimes WAA_{(i)} W') \\
&\quad = \lambda_i (w_i' \otimes WA_{(i)} W')(w_i \otimes WAA_{(i)} W' + WAA_{(i)} W' \otimes w_i) \\
&\quad = \lambda_i WA_{(i)} W'WAA_{(i)} W' + \lambda_i w_i' WAA_{(i)} W' \otimes WA_{(i)} W' w_i \\
&\quad = \lambda_i WA_{(i)} AA_{(i)} W' + \lambda_i e_i' AA_{(i)} W' \otimes WA_{(i)} e_i \\
&\quad = \lambda_i WA_{(i)}^2 AW', \qquad \text{as} \quad A_{(i)} e_i = 0 \qquad \text{and} \qquad A_{(i)} AA_{(i)} = A_{(i)}^2 A.
\end{aligned}
$$

For the evaluation of the other three terms it is essential to look into the expression

$$
\begin{aligned}
&K_{\mathrm{d}}(I \otimes P + P \otimes I)(w_i \otimes WA_{(i)} W') \\
&\quad = JJ'(w_i \otimes WAA_{(i)} W' + \lambda_i w_i \otimes WA_{(i)} W') \\
&\quad = J\Delta(w_i)\, W(A + \lambda_i I)\, A_{(i)}\, W'.
\end{aligned}
$$

This result yields for the negative of the second and transposed third terms

$$
\begin{aligned}
&(w_i' \otimes WA_{(i)}\, W')\, A_2 (w_i \otimes WA_{(i)} W') \\
&\quad = (w_i' \otimes WA_{(i)} W')(P \otimes P)\, K_{\mathrm{d}}(I \otimes P + P \otimes I)(w_i \otimes WA_{(i)}\, W') \\
&\quad = \lambda_i (w_i' \otimes WA_{(i)} AW')\, J\Delta(w_i)\, W(A + \lambda_i I)\, A_{(i)}\, W' \\
&\quad = \lambda_i WA_{(i)} AW' \Delta^2(w_i)\, W(A + \lambda_i I)\, A_{(i)} W'.
\end{aligned}
$$

Finally

$$
\begin{aligned}
&(w_i' \otimes WA_{(i)} W')\, A_3 (w_i \otimes WA_{(i)} W') \\
&\quad = \tfrac{1}{2}(w_i' \otimes WA_{(i)} W')(I \otimes P + P \otimes I)\, K_{\mathrm{d}}(P \otimes P) \\
&\quad\quad \cdot K_{\mathrm{d}}(I \otimes P + P \otimes I)(w_i \otimes WA_{(i)} W') \\
&\quad = \tfrac{1}{2} WA_{(i)}(A + \lambda_i I)\, W'\Delta(w_i)\, J'(P \otimes P)\, J\Delta(w_i)\, W(A + \lambda_i I)\, A_{(i)} W' \\
&\quad = \tfrac{1}{2} WA_{(i)}(A + \lambda_i I)\, W'\Delta(w_i)(P \times P)\, \Delta(w_i)\, W(A + \lambda_i I)\, A_{(i)} W'
\end{aligned}
$$

yields the fourth term. This finishes the proof.

COROLLARY 5. *For the elliptical case, when $n \to \infty$,*

$$
\sqrt{n}\,(\hat{\lambda}_{(n)} - \lambda) \xrightarrow{\mathrm{D}} N(0, \Gamma),
$$

*where*

$$\begin{aligned}\Gamma = {} & 2(1+\kappa)\,\Lambda^2 + 4\kappa\Lambda U\Lambda \\ & - 2(1+\kappa)\,\Lambda[\Lambda(W'\times W')(W\times W) + (W'\times W')(W\times W)\Lambda]\,\Lambda \\ & + 2(1+\kappa)\,\Lambda(W'\times W')(P\times P)(W\times W)\Lambda.\end{aligned}$$

*Proof.* We replace $\Psi$ by $B_1 - B_2 - B_2' + B_3$ in Theorem 7. (See (9) for the definitions.) This yields for the asymptotic variance of $\sqrt{n}\,\hat{\lambda}_{(n)}$

$$J'(W'\otimes W')(B_1 - B_2 - B_2' + B_3)(W\otimes W)J.$$

It is easy to see that

$$\begin{aligned}& J'(W'\otimes W')\,B_1(W\otimes W)J \\ & \quad = (1+\kappa)\,J'(W'\otimes W')(I+K)(P\otimes P)(W\otimes W)J \\ & \qquad + \kappa J'(W'\otimes W')(\operatorname{vec} P)(\operatorname{vec} P)'\,(W\otimes W)J \\ & \quad = 2(1+\kappa)\,J'(\Lambda\otimes\Lambda)\,J + \kappa J'(\operatorname{vec}\Lambda)(\operatorname{vec}\Lambda)'J \\ & \quad = 2(1+\kappa)\,\Lambda^2 + \kappa\Lambda U\Lambda.\end{aligned}$$

Further

$$\begin{aligned}& J'(W'\otimes W')\,B_2(W\otimes W)J \\ & \quad = (1+\kappa)\,J'(W'\otimes W')(P\otimes P)\,K_d(I\otimes P + P\otimes I)(W\otimes W)J \\ & \qquad - \tfrac{1}{2}\kappa J'(W'\otimes W')(\operatorname{vec} P)(\operatorname{vec} P)'\,K_d(I\otimes P + P\otimes I)(W\otimes W)J \\ & \quad = 2(1+\kappa)\,J'(\Lambda\otimes\Lambda)(W'\otimes W')\,JJ'(W\otimes W)(I\otimes\Lambda)J \\ & \qquad - \kappa J'(\operatorname{vec}\Lambda)(\operatorname{vec} P)'\,JJ'(W\otimes W)(I\otimes\Lambda)J \\ & \quad = 2(1+\kappa)\,J'(\Lambda\otimes\Lambda)\,JJ'(W'\otimes W')\,JJ'(W\otimes W)\,JJ'(I\otimes\Lambda)J \\ & \qquad - \kappa J'(\operatorname{vec}\Lambda)(\operatorname{vec} P)'\,JJ'(W\otimes W)\,JJ'(I\otimes\Lambda)J \\ & \quad = 2(1+\kappa)\,\Lambda^2(W'\times W')(W\times W)\,\Lambda - \kappa\Lambda\imath(\operatorname{vec} I)'\,(W\otimes W)\,J\Lambda \\ & \quad = 2(1+\kappa)\,\Lambda^2(W'\times W')(W\times W)\,\Lambda - \kappa\Lambda U\Lambda.\end{aligned}$$

Finally

$$\begin{aligned}& J'(W'\otimes W')\,B_3(W\otimes W)J \\ & \quad = \tfrac{1}{2}(1+\kappa)\,J'(W'\otimes W')(I\otimes P + P\otimes I)\,K_d(P\otimes P) \\ & \qquad \cdot K_d(I\otimes P + P\otimes I)(W\otimes W)J \\ & \qquad + \tfrac{1}{4}\kappa J'(W'\otimes W')(I\otimes P + P\otimes I)\,K_d(\operatorname{vec} P)(\operatorname{vec} P)' \\ & \qquad \cdot K_d(I\otimes P + P\otimes I)(W\otimes W)J\end{aligned}$$

$$= 2(1+\kappa)\, \Lambda(W' \times W')(P \times P)(W \times W)\Lambda$$
$$+ \kappa \Lambda J'(\operatorname{vec} I)(\operatorname{vec} I)' J\Lambda$$
$$= 2(1+\kappa)\, \Lambda(W' \times W')(P \times P)(W \times W)\Lambda$$
$$+ \kappa \Lambda U \Lambda, \qquad \text{in a similar way.}$$

COROLLARY 6. *For the elliptical case, when $n \to \infty$,*

$$\sqrt{n}\,(\hat{w}_{(n)i} - w_i) \xrightarrow{\mathrm{D}} N(0, \Xi_i),$$

*where*

$$\begin{aligned}\Xi_i = {} & (1+\kappa)\, \lambda_i W \Lambda_{(i)}^2 \Lambda W' - (1+\kappa)\, \lambda_i W \Lambda_{(i)} \Lambda W' \Delta^2(w_i)\, W(\Lambda + \lambda_i I)\, \Lambda_{(i)} W' \\ & - (1+\kappa)\, \lambda_i W \Lambda_{(i)} (\Lambda + \lambda_i I)\, W' \Delta^2(w_i)\, W \Lambda \Lambda_{(i)} W' \\ & + \tfrac{1}{2}(1+\kappa)\, W \Lambda_{(i)} (\Lambda + \lambda_i I)\, W' \Delta(w_i)(P \times P)\, \Delta(w_i)\, W(\Lambda + \lambda_i I)\, \Lambda_{(i)} W',\end{aligned}$$

*where $\Lambda_{(i)} = (\lambda_i I - \Lambda)^+$.*

*Proof.* We insert $\Psi = B_1 - B_2 - B_2' + B_3$ in Theorem 8. We then get the expression

$$(w_i' \otimes W \Lambda_{(i)} W')(B_1 - B_2 - B_2' + B_3)(w_i \otimes W \Lambda_{(i)} W').$$

Consider

$$\begin{aligned}&(w_i' \otimes W \Lambda_{(i)} W')\, B_1 (w_i \otimes W \Lambda_{(i)} W') \\ &\quad = (1+\kappa)(w_i' \otimes W \Lambda_{(i)} W')(I + K)(P \otimes P)(w_i \otimes W \Lambda_{(i)} W') \\ &\qquad + \kappa (w_i' \otimes W \Lambda_{(i)} W')(\operatorname{vec} P)(\operatorname{vec} P)'\, (w_i \otimes W \Lambda_{(i)} W') \\ &\quad = (1+\kappa)\, \lambda_i W \Lambda_{(i)}^2 \Lambda W'.\end{aligned}$$

Further

$$\begin{aligned}&(w_i' \otimes W \Lambda_{(i)} W')\, B_2 (w_i \otimes W \Lambda_{(i)} W') \\ &\quad = (1+\kappa)(w_i' \otimes W \Lambda_{(i)} W')(P \otimes P) \\ &\qquad \cdot K_{\mathrm{d}} (I \otimes P + P \otimes I)(w_i \otimes W \Lambda_{(i)} W') \\ &\qquad - \tfrac{1}{2}\kappa (w_i' \otimes W \Lambda_{(i)} W')(\operatorname{vec} P)(\operatorname{vec} P)' \\ &\qquad \cdot K_{\mathrm{d}} (I \otimes P + P \otimes I)(w_i \otimes W \Lambda_{(i)} W') \\ &\quad = (1+\kappa)\, \lambda_i W \Lambda_{(i)} \Lambda W' \Delta^2(w_i)\, W \Lambda \Lambda_{(i)} W' \\ &\qquad + (1+\kappa)\, \lambda_i^2\, W \Lambda_{(i)} \Lambda W' \Delta^2(w_i)\, W \Lambda_{(i)} W' \\ &\quad = (1+\kappa)\, \lambda_i W \Lambda_{(i)} \Lambda W' \Delta^2(w_i)\, W(\Lambda + \lambda_i I)\, \Lambda_{(i)} W'.\end{aligned}$$

Finally

$$
\begin{aligned}
&(w_i' \otimes WA_{(i)}W')\, B_3(w_i \otimes WA_{(i)}W') \\
&\quad = \tfrac{1}{2}(w_i' \otimes WA_{(i)}W')(I \otimes P + P \otimes I)\, K_d[(1+\kappa)(P \otimes P) \\
&\qquad + \tfrac{1}{2}\kappa(\operatorname{vec} P)(\operatorname{vec} P)']\, K_d(I \otimes P + P \otimes I)(w_i \otimes WA_{(i)}W') \\
&\quad = \tfrac{1}{2}(WA_{(i)}\Lambda W'\Delta(w_i) + \lambda_i WA_{(i)}W'\Delta(w_i))[(1+\kappa)\, P \times P + \tfrac{1}{2}\kappa U] \\
&\qquad \cdot (\Delta(w_i)\, W\Lambda A_{(i)}W' + \lambda_i \Delta(w_i)\, WA_{(i)}W') \\
&\quad = \tfrac{1}{2}[WA_{(i)}(\Lambda + \lambda_i I)\, W'\Delta(w_i)][(1+\kappa)\, P \times P + \tfrac{1}{2}\kappa U] \\
&\qquad \cdot [\Delta(w_i)\, W(\Lambda + \lambda_i I)\, A_{(i)}W'] \\
&\quad = \tfrac{1}{2}(1+\kappa)\, WA_{(i)}(\Lambda + \lambda_i I)\, W'\Delta(w_i)(P \times P)\, \Delta(w_i)\, W(\Lambda + \lambda_i I)\, A_{(i)}W'.
\end{aligned}
$$

## APPENDIX I: The Finite-Sample Variance Matrix of $\sqrt{n}\operatorname{vec} S_{(n)}$: A Matrix Derivation

We define

$$
y := x - \mu, \qquad y_i := x_i - \mu,
$$

$$
Y' := (y_1, ..., y_n) = X' - \mu\iota'.
$$

Clearly $X'NX = Y'NY$, as $NY = N(X - \iota\mu') = NX$. Hence $\sqrt{n}\, S_{(n)} = \sqrt{n}\,(n-1)^{-1}\, Y'NY$. Obviously $D(\operatorname{vec} Y'NY) = E[(\operatorname{vec} Y'NY)(\operatorname{vec} Y'NY)'] - [E(\operatorname{vec} Y'NY)][E(\operatorname{vec} Y'NY)]'$. We have immediately

$$
E(\operatorname{vec} Y'NY) = (n-1)\operatorname{vec} \Sigma.
$$

The expectation of the product term is more difficult to obtain. We write

$$
Y'NY = \sum_{ij} n_{ij}\, y_i y_j', \qquad \text{where} \quad \begin{cases} n_{ii} = 1 - n^{-1} \\ n_{ij} = -n^{-1} \quad (i \neq j). \end{cases}
$$

Hence $\operatorname{vec} Y'NY = \sum_{ij} n_{ij}(y_j \otimes y_i)$ and

$$
\begin{aligned}
&E(\operatorname{vec} Y'NY)(\operatorname{vec} Y'NY') \\
&\quad = \sum_{ijkl} n_{ij} n_{kl} E(y_j y_l' \otimes y_i y_k') \\
&\quad = \sum_{i \neq j} n_{ii} n_{ij} E(y_i y_j' \otimes y_i y_j') + \sum_{i \neq j} n_{ij}^2 E(y_i y_j' \otimes y_j y_i') \\
&\qquad + \sum_{i} n_{ii}^2 E(y_i y_i' \otimes y_i y_i') + \sum_{i \neq j} n_{ij}^2 E(y_i y_i' \otimes y_j y_j')
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{i \neq j} n_{ii} n_{jj} E[(\operatorname{vec} y_i y_i')(\operatorname{vec} y_j y_j')'] + \sum_{i \neq j} n_{ij}^2 E(y_i y_i' \otimes y_j y_j') K \\
&\quad + \sum_{i} n_{ii}^2 E(y_i y_i' \otimes y_i y_i') + \sum_{i \neq j} n_{ij}^2 E(y_i y_i' \otimes y_j y_j') \\
&= (n-1)^3 n^{-1} (\operatorname{vec} \Sigma)(\operatorname{vec} \Sigma)' + (n-1) n^{-1} K(\Sigma \otimes \Sigma) \\
&\quad + (n-1)^2 n^{-1} M_4(y) + (n-1) n^{-1} (\Sigma \otimes \Sigma) \\
&= (n-1)^3 n^{-1} (\operatorname{vec} \Sigma)(\operatorname{vec} \Sigma)' + (n-1)^2 n^{-1} M_4(x) \\
&\quad + (n-1) n^{-1} (I + K)(\Sigma \otimes \Sigma),
\end{aligned}
$$

as the $y_i$ are independent with zero mean. This yields

$$
\begin{aligned}
D(\operatorname{vec} Y'NY) &= (n-1)^3 n^{-1} (\operatorname{vec} \Sigma)(\operatorname{vec} \Sigma)' + (n-1)^2 n^{-1} M_4(x) \\
&\quad + (n-1) n^{-1} (I + K)(\Sigma \otimes \Sigma) - (n-1)^2 (\operatorname{vec} \Sigma)(\operatorname{vec} \Sigma)' \\
&= -(n-1)^2 n^{-1} (\operatorname{vec} \Sigma)(\operatorname{vec} \Sigma)' + (n-1)^2 n^{-1} M_4(x) \\
&\quad + (n-1) n^{-1} (I + K)(\Sigma \otimes \Sigma),
\end{aligned}
$$

and finally

$$
\begin{aligned}
D(\sqrt{n} \operatorname{vec} S_{(n)}) &= n(n-1)^{-2} D(\operatorname{vec} Y'NY) \\
&= M_4(x) + (n-1)^{-1} (I + K)(\Sigma \otimes \Sigma) - (\operatorname{vec} \Sigma)(\operatorname{vec} \Sigma).
\end{aligned}
$$

## APPENDIX II: Some Useful Matrix Definitions and Properties

In the main text we used most of the following definitions and properties:

1. The $i$th *unit vector* $e_i$ $(i = 1, \dots, p)$, i.e. the $i$th column of the identity matrix $I_p$;

2. The $p \times 1$ *summation vector* $\imath = \sum_{i=1}^{p} e_i = (1 \cdots 1)'$;

3. The $p \times q$ unit matrix $E_{ij} = e_i e_j'$ $(i = 1, \dots, p;\ j = 1, \dots, q)$;

4. The *Kronecker product* $A \otimes B = [a_{ij} B]$, where $A = [a_{ij}]$; the *Hadamard* (or *Schur*) product $A \times B = [a_{ij} b_{ij}]$. In the first definition $B$ has arbitrary order, in the second definition $B$ and $A$ are compatible.

5. The vec *operator*, with property $\operatorname{vec} ABC = (C' \otimes A) \operatorname{vec} B$.

6. The *commutation matrix* $K = \sum_{ij} (E_{ij} \otimes E_{ij}')$, with properties $K \operatorname{vec} C = \operatorname{vec} C'$ and $K(A \otimes B) K = B \otimes A$ for square matrices $A$, $B$, and $C$ of equal order.

7. The $p^2 \times p$ matrix $J = (\operatorname{vec} E_{11}, \dots, \operatorname{vec} E_{pp})$, where the unit matrices are square of order $p$, with property $J'J = I_p$. Further $J\imath = \operatorname{vec} I_p$.

8. The $p \times p$ diagonal matrix $A_d$ with diagonal elements $a_{11}, ..., a_{pp}$ of the $p \times p$ matrix $A$. Clearly $J'$ vec $A = A_d \iota$.

9. The $p \times p$ matrix $K_d = \sum_i (E_{ii} \otimes E_{ii})$. Clearly $K_d = JJ'$.

Remember that $K_d$ is the diagonal matrix obtained from the commutation matrix $K$.

10. $J'(A \otimes B) J = A \times B$ for compatible *square* matrices $A$ and $B$.

11. $K_d(A \otimes B) J = (A \otimes B) J$ for compatible *diagonal* matrices $A$ and $B$.

12. The *diagonal* matrix $\Delta(w)$, implicitly defined by $\Delta(w)\, \iota = w$. It has the useful property $(w' \otimes A) J = A\Delta(w)$ for compatible $w$, $A$, and $J$ (with $p$ elements).

## References


[1] Anderson, T. W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Statist.* **34** 122–148.

[2] Davis, A. W. (1977). Asymptotic theory for principal component analysis: Non-normal case. *Austral. J. Statist.* **19** (3) 206–212.

[3] Fang, C. and Krishnaiah, P. R. (1982). Asymptotic distributions of functions of the eigenvalues of some random matrices for nonnormal populations. *J. Multivar. Anal.* **12** 39–63.

[4] Fujikoshi, Y. (1980). Asymptotic expansions for the distributions of the sample roots under nonnormality. *Biometrika* **67** 45–51.

[5] Girshick, M. A. (1939). On the sampling theory of roots of determinantal equations. *Ann. Math. Statist.* **10** 203–224.

[6] Kollo, T. (1977). Some notions of matrix calculus with applications in mathematical statistics. *Proc. Comput. Centre Tartu State Univ.* **40** 30–51. [In Russian]

[7] Kollo, T. (1979). *Asymptotic Distributions of Sample Covariance and Correlation Matrices, their Eigenvalues and Eigenvectors.* Dissertation, Tartu. [In Russian]

[8] Kollo, T. (1981). On using asymptotic distributions in the principal component model. In *The Second Conference on Applications of Multivariate Analysis in Economics and Quality Control Theory, Tartu*, pp. 282–283. [In Russian]

[9] Konishi, S. (1979). Asymptotic expansions for the distributions of statistics based on the sample correlation matrix in principal component analysis. *Hiroshima Math. J.* **9** 647–700.

[10] Magnus, J. R., and Neudecker H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Wiley, New York.

[11] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* Wiley, New York.

[12] Neudecker, H., and Wesselman, A. M. (1990). The asymptotic variance matrix of the sample correlation matrix. *Linear Algebra Appl.* **127** 589–599.

[13] Parring, A.-M. (1979). Calculation of asymptotic characteristics of sample functions. *Acta comment. Univ. Tartu.* **492** 86–90. [In Russian]

[14] Waternaux, C. (1976). Asymptotic distributions of the sample roots for a non-normal population. *Biometrika* **63** 639–664.