ASYMPTOTIC DISTRIBUTION OF THE LARGEST EIGENVALUE

Martin Bilodeau

Centre de recherches mathématiques and
Département de mathématiques et de statistique,
Université de Montréal,
C.P. 6128, Succursale Centre-ville,
Montréal, Canada H3C 3J7

E-mail: bilodeau@dms.umontreal.ca

This paper studies the asymptotic distribution of the largest eigenvalue of the sample covariance matrix. The multivariate distribution for the population is assumed to be elliptical with finite kurtosis 3κ. An expression as an expectation is obtained for the distribution function of the largest eigenvalue regardless of the multiplicity, m, of the population's largest eigenvalue. The asymptotic distribution function and density function are evaluated numerically for $m = 2, 3, 4, 5$. The bootstrap of the average of the m largest eigenvalues is shown to be consistent for any underlying distribution with finite fourth-order cumulants.

Key Words: asymptotic distributions, elliptical distributions, kurtosis, largest eigenvalue, multiplicity

1. INTRODUCTION

The largest eigenvalue of random symmetric matrices is often encountered in statistical analyses. Assume $\boldsymbol{V} : p \times p$ is distributed as a Wishart matrix with scaling matrix $\boldsymbol{\Sigma}$ and n degrees of freedom, i.e. $\boldsymbol{V} \sim W_p(n, \boldsymbol{\Sigma})$. The union-intersection test procedure for the hypotheses $H : \boldsymbol{\Sigma} = \boldsymbol{I}$ versus $A : \boldsymbol{\Sigma} \neq \boldsymbol{I}$ rejects H whenever $\mathrm{ch}_1(\boldsymbol{V}) \geq c_1$ or $\mathrm{ch}_p(\boldsymbol{V}) \leq c_2$, where $\mathrm{ch}_1(\boldsymbol{V}) \geq \cdots \geq \mathrm{ch}_p(\boldsymbol{V})$ are the ordered eigenvalues of $\boldsymbol{V}$. In principal component analysis, the largest eigenvalue is the variance of the first principal component. The exact distribution of the largest eigenvalue of $\boldsymbol{S}_n = \boldsymbol{V}/n$ follows from a result of Constantine (1963)

$$P_{\boldsymbol{\Sigma}}(\mathrm{ch}_1(\boldsymbol{S}_n) < x) = \frac{\Gamma_p[(p+1)/2]}{\Gamma_p[(n+p+1)/2]} \det(nx\boldsymbol{\Sigma}^{-1}/2) \cdot {}_1F_1(n/2; (n+p+1)/2; -nx\boldsymbol{\Sigma}^{-1}/2). \qquad (1)$$

The hypergeometric ${}_1F_1$ function is a zonal polynomial series which converges very slowly, even for small values of n and p. Percentage points using (1) were calculated by Sugiyama (1972) for $p = 2$ and 3.

Seminal papers on the asymptotic distributions of eigenvalues and eigenvectors are those of Girshick (1939) and Anderson (1963). The perturbation method of Bellman (1960) also provides an asymptotic normal distribution given by $v_1 \equiv n^{1/2}\left(\mathrm{ch}_1(\boldsymbol{S}_n) - d_1\right)/(\sqrt{2}d_1) \xrightarrow{L} N(0,1)$ when n is large and d_1, the largest eigenvalue of $\boldsymbol{\Sigma}$, is distinct. More accurate approximations were obtained by Sugiyama (1972), Muirhead and Chikuse (1975), and Muirhead (1974). However, if the largest population eigenvalue is not distinct, then the asymptotic distribution of v_1 is more complicated. When

$\boldsymbol{\Sigma} = d_1 \boldsymbol{I}_2$, the asymptotic distribution of v_1 can be expanded as

$$P(v_1 \leq t) = \Phi(t\sqrt{2}) - \sqrt{\pi}\phi(t)\Phi(t) + O(n^{-1/2}), \tag{2}$$

whereas if $\boldsymbol{\Sigma} = d_1 \boldsymbol{I}_3$, then the expression becomes

$$P(v_1 \leq t) = \Phi(t\sqrt{2})\Phi(t) - 2t\phi(t)\Phi(t\sqrt{2}) - \pi^{-1/2}\phi(t\sqrt{3}) + O(n^{-1/2}), \tag{3}$$

where ϕ and Φ are, respectively, the density function and distribution function of a standard normal variable. Further terms in these non-normal asymptotic distributions can be found in Muirhead (1974). Sugiura (1976) also derived an asymptotic expansion for the joint density of the roots of a Wishart matrix with multiple population roots; see also Fujikoshi (1977) and Khatri and Srivastava (1978) for related results.

All these previous results, however, are valid when sampling from a multivariate normal distribution. Waternaux (1976) obtained the asymptotic joint distribution of the roots of the sample covariance matrix when sampling from a non-normal population with finite fourth cumulants and all population roots simple. In the same context and assuming higher moments, Fujikoshi (1980, 1986) calculated the Edgeworth expansions. Asymptotic joint distributions of eigenvalues and normalized eigenvectors of the sample covariance and correlation matrices were derived by Kollo and Neudecker (1993). For non-normal populations, only the case of distinct eigenvalues had been studied before Tyler (1983) who derived the asymptotic distribution of eigenvalues under local alternatives to multiple roots. Eaton and Tyler (1991) used Wielandt's inequality to obtain the asymptotic distribution of eigenvalues of a random symmetric matrix under multiple roots.

The present paper proposes to investigate the effect of the kurtosis of the population distribution and of the multiplicity of the largest population eigenvalue on the asymptotic distribution of v_1. Specifically, assuming an underlying elliptical distribution with finite kurtosis 3κ, an expression, as an expectation, is obtained for the asymptotic distribution function of v_1 regardless of the multiplicity, say m, of d_1. The asymptotic distribution function and density function are then evaluated numerically for various values of m and κ. The asymptotic distribution of the average of the m largest eigenvalue is also derived for any underlying distribution with finite fourth-order moments. The bootstrap of this average is shown to be consistent.

2. MAIN RESULTS

Consider a sample $\boldsymbol{x_1}, \ldots, \boldsymbol{x_n}$ form a p-dimensional elliptical distribution with finite kurtosis 3κ and covariance matrix $\boldsymbol{\Sigma} = \text{diag}(d_1 \boldsymbol{I}_{p_1}, \ldots, d_k \boldsymbol{I}_{p_k})$, $p = p_1 + \cdots + p_k$, where $d_1 > \cdots > d_k$ are the distinct eigenvalues of multiplicity p_i, $i = 1, \ldots, k$. The sample covariance matrix is $\boldsymbol{S}_n = \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})'$, where $\bar{\boldsymbol{x}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{x}_i$. For simplicity of notation, denote by m the multiplicity p_1 of the largest eigenvalue d_1. The following theorem establishes an expression, as an expectation, for the asymptotic distribution of $v_1 = n^{1/2} \left(\text{ch}_1(\boldsymbol{S}_n) - d_1\right) / (\sqrt{2} d_1)$. The proofs are given in the Appendix.

THEOREM 2.1. *The asymptotic distribution function, as $n \to \infty$, of v_1 is given by*

$$P(v_1 \le t) \;\to\; D_m(\kappa)E\Big\{\big[\prod_{i<j}(z_i - z_j)\big] \cdot \exp\left[\frac{(1+\kappa)^{-1}\kappa(\sum_i z_i)^2}{4[1+\kappa(1+\kappa)^{-1}m/2]}\right] \cdot I\left(z_1 \le t/(1+\kappa)^{1/2}\right)\Big\}, \qquad (4)$$

where $I(\cdot)$ is an indicator function,

$$D_m(\kappa) = \frac{\pi^{m(m+1)/4}}{m!\Gamma_m(m/2)}[1+\kappa(1+\kappa)^{-1}m/2]^{-1/2},$$

and $z_1 > \cdots > z_m$ is the order statistics of m independently and identically distributed observations from a standard normal distribution.

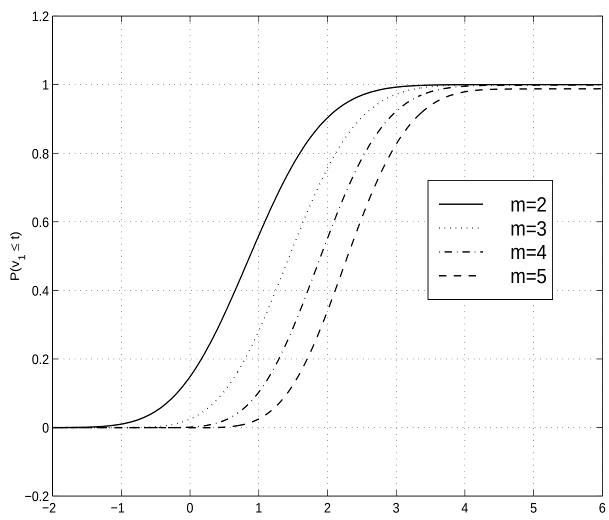

FIG. 1. Asymptotic distribution function of the largest eigenvalue for multiplicity $m = 2, 3, 4, 5$ and underlying multivariate normality ($\kappa = 0$).

Obvious modifications to the proof of Theorem 2.1 yield the following corollary on the asymptotic distribution of the smallest eigenvalue. Here, m denotes the multiplicity p_k of the smallest eigenvalue d_k.

COROLLARY 2.1. *The asymptotic distribution function, as $n \to \infty$, of $v_p \equiv n^{1/2}\left(ch_p(\boldsymbol{S}_n) - d_k\right)/(\sqrt{2}d_k)$ is given by*

$$P(v_p \le t) \;\to\; D_m(\kappa)E\Big\{\big[\prod_{i<j}(z_i - z_j)\big] \cdot \exp\left[\frac{(1+\kappa)^{-1}\kappa(\sum_i z_i)^2}{4[1+\kappa(1+\kappa)^{-1}m/2]}\right] \cdot I\left(z_m \le t/(1+\kappa)^{1/2}\right)\Big\},$$

where $z_1 > \cdots > z_m$ is the order statistics of m independently and identically distributed observations from a standard normal distribution.

The next result establishes an ordering between the asymptotic distribution functions of v_1 as the multiplicity increases.

THEOREM 2.2. *As $m = 1, \ldots, p$ increases, the asymptotic distribution function of v_1, evaluated at t, decreases for all values of t.*

A consequence of theorem 2.2 is that as the multiplicity increases the asymptotic probability distribution of the largest eigenvalue of the sample covariance matrix shifts to the right. This can be seen from Figures 1 and 2 for the special case of null kurtosis.

The asymptotic distribution of v_1, when $m \ge 2$, is not normal for any elliptical distribution with finite kurtosis. However, for underlying multivariate normal distribution, Anderson (1963) established that the maximum likelihood estimate of d_1, which is $\bar{d} = \left(\mathrm{ch}_1(\boldsymbol{S}_n) + \cdots + \mathrm{ch}_m(\boldsymbol{S}_n)\right)/m$, is

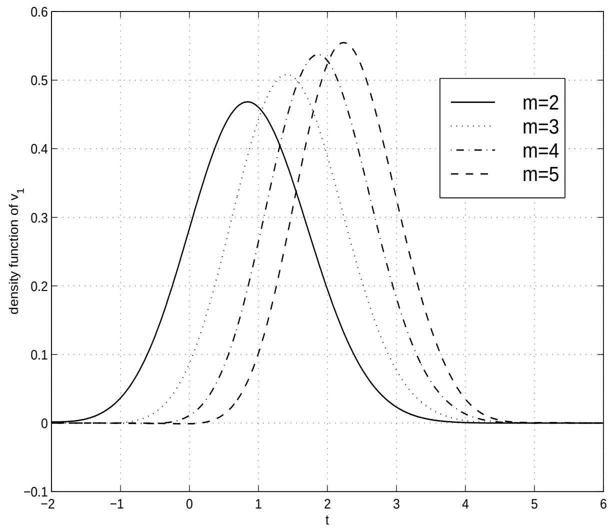

FIG. 2. Asymptotic density function of the largest eigenvalue for multiplicity $m = 2, 3, 4, 5$ and underlying multivariate normality ($\kappa = 0$).

asymptotically normal,

$$n^{1/2}(\bar{d} - d_1) \xrightarrow{L} N(0, 2d_1^2/m).$$

The next result is a generalisation to any underlying distribution with finite fourth-order cumulants.

THEOREM 2.3. *If the underlying distribution has finite fourth-order cumulants, then*

$$n^{1/2}(\bar{d} - d_1) \xrightarrow{L} N(0, \gamma^2 2d_1^2/m),$$

where

$$\gamma^2 = 1 + \left[\sum_{i=1}^{m} k_4^i + 2 \sum_{1 \le i < j \le m} k_{22}^{ij} \right] / (2d_1^2 m)$$

and k_4^i, k_{22}^{ij} are fourth-order cumulants of the joint distribution of variables i and j.

These cumulants can be evaluated for elliptical distributions to obtain $\gamma^2 = (1 + 3\kappa/2) + (m-1)\kappa/2$. Anderson's (1963) result is the special case where $\kappa = 0$.

Beran and Srivastava (1985) established the consistency of the bootstrap in the case of simple roots. The inconsistency of the bootstrap in presence of multiple roots was shown by Beran and Srivastava (1987) for $p = 2$. They also established the consistency when bootstraping a sample of size $o(n)$. Eaton and Tyler (1991) generalized these bootstrap results to any $p \geq 2$.

It is now established that the bootstrap is consistent for $\bar{d}$. Note that this bootstrap still requires knowledge of the multiplicity m. Let F_n be the sample distribution function of $\{\boldsymbol{x}_i; 1 \leq i \leq n\}$. A bootstrap sample is a random sample $\{\boldsymbol{x}_i^*; 1 \leq i \leq n\}$ from F_n. Let $\boldsymbol{S}_n^*$ be the sample covariance from the bootstrap sample and $\bar{d}^* = (\mathrm{ch}_1(\boldsymbol{S}_n^*) + \cdots + \mathrm{ch}_m(\boldsymbol{S}_n^*))/m$. The distribution of $n^{1/2}(\bar{d} - d_1)$ under F can be approximated by the distribution of $n^{1/2}(\bar{d}^* - \bar{d})$ under F_n comes as a consequence of equation (5.5) in Eaton and Tyler (1991).

COROLLARY 2.2. *If the distibution F has finite fourth-order cumulants, then $n^{1/2}(\bar{d}^* - \bar{d}) \xrightarrow{L^*} N(0, \gamma^2 2d_1^2/m)$ a.s.*

The notation $\xrightarrow{L^*}$ refers to the weak convergence of the distribution of $n^{1/2}(\bar{d}^* - \bar{d})$ under F_n.

3. NUMERICAL ASPECTS

Evaluation of the expectation in Theorem 2.1 for $\kappa = 0$ and $m = 2, 3$, using integration by parts, leads, respectively, to the expressions in (2) and (3). These calculations become intractable for $m \geq 4$ even for symbolic computer software such as Maple or Mathematica. The approach adopted here is to evaluate, for given m and κ, the expectation in (4) by simulation. The algorithm starts with a discretization of the support into a large number of subintervals, $[t_s, t_{s+1}]$, $s = 1, \ldots, S$. Then, a large table of dimension $B \times m$ is generated where, independently, each line consists of an ordered sample of size m, $z_1 > \cdots > z_m$, from a standard normal distribution. The figures included used $S = 200$ subintervals and $B = 5$ millions. Then, at each point t_s the variable

$$u_b(t_s) = [\prod_{i<j}(z_i - z_j)] \exp\left[\frac{(1+\kappa)^{-1}\kappa(\sum_i z_i)^2}{4[1+\kappa(1+\kappa)^{-1}m/2]}\right] I\left(z_1 \leq t_s/(1+\kappa)^{1/2}\right)$$

is evaluated along each line to obtain $u_1(t_s), \ldots, u_B(t_s)$. The estimate of the asymptotic distribution function (4), evaluated at t_s, is $D_m(\kappa)\bar{u}(t_s)$, where $\bar{u}(t_s) = \sum_{b=1}^{B} u_b(t_s)/B$, with a standard deviation of

$$B^{-1/2} D_m(\kappa) \left\{ \sum_{b=1}^{B} [u_b(t_s) - \bar{u}(t_s)]^2/(B-1) \right\}^{1/2}. \tag{5}$$

The points $(t_s, D_m(\kappa)\bar{u}(t_s))$ can then be joined to produce curves as those in Figures 1 and 5.

As the density function is also of interest, rather than joining the points by line segments, a cubic smoothing spline was used which yielded a differentiable curve with two continuous derivatives. The simulation was done

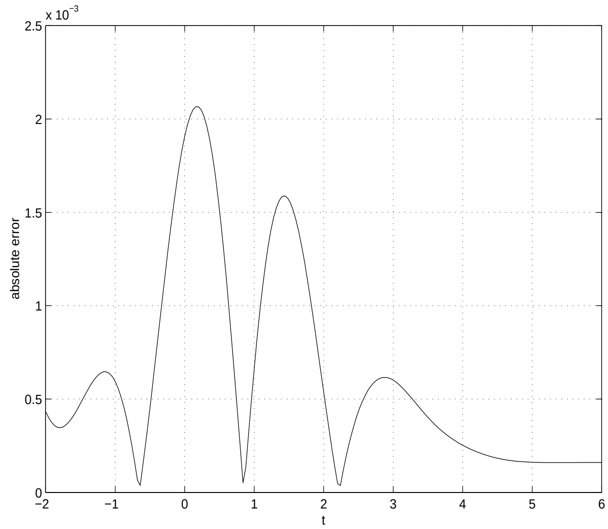

FIG. 3. Absolute error of the spline estimate of the asymptotic distribution function for $m = 2$ and $\kappa = 0$.

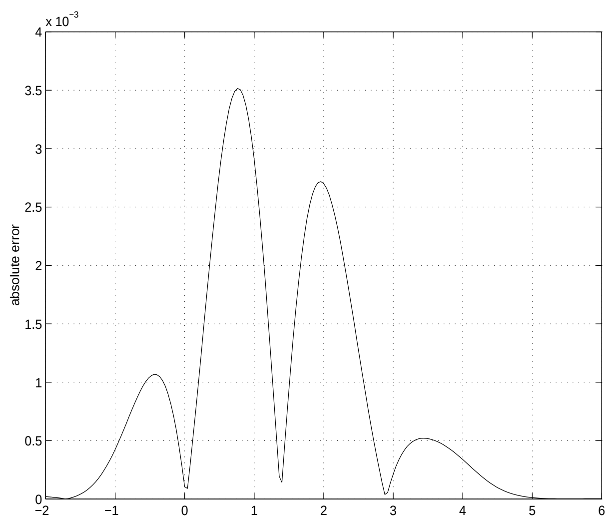

FIG. 4. Absolute error of the spline estimate of the asymptotic distribution function for $m = 3$ and $\kappa = 0$.

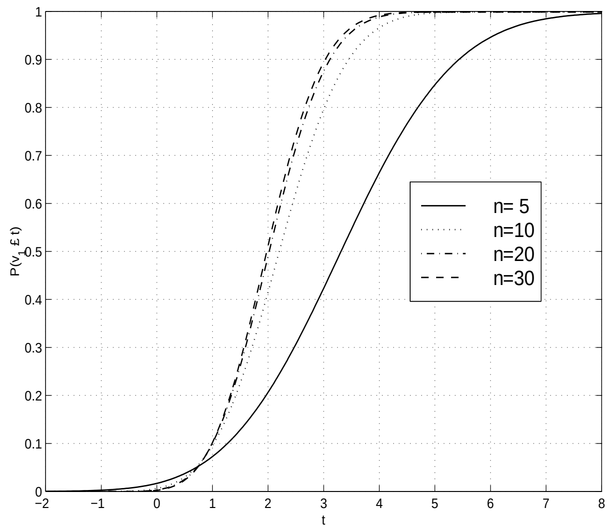

FIG. 5. Asymptotic distribution function of the largest eigenvalue for multiplicity $m = 4$ and underlying multivariate Student distribution with degrees of freedom $\nu = 5, 10, 20, 30$.

with the Fortran NAG library and the smoothing was performed with the cubic smoothing spline function "csaps" of Matlab. It includes a smoothing parameter p in the interval $[0, 1]$. The value $p = 0$ gives the least-squares line (maximum amount of smoothing), whereas $p = 1$ produces an interpolation (minimum amount of smoothing). The value $p = 0.95$ was used here to get a smooth curve very close to an interpolation. The density function estimate is just the derivative of the cubic smoothing spline. Figures 5 and 6 show the effect of kurtosis on the asymptotic distribution of the largest eigenvalue in a non trivial situation where the multiplicity is 4. Non zero kurtosis was obtained from the relation $3\kappa = 6/(\nu - 4)$ which describes the kurtosis of a multivariate Student with ν degrees of freedom. It can be observed that as ν decreases, i.e. the underlying distribution has more probability in the tails, then the probability that the largest eigenvalue takes on large values increases. Also, the variance increases with the kurtosis; an observation which is consistent with the variance expression (3) of Waternaux (1976) for the case $m = 1$. Calculations leading to Figures 5 and 6 were produced in approximately 33 minutes on a Sun Ultra 2 with 4 CPUs of 400 MHz.

The accuracy of the estimated asymptotic distribution function could be determined exactly in the case $m = 2$, 3, and $\kappa = 0$ using the exact asymptotic expressions (2) and (3). The absolute deviations between the exact asymptotic expressions and the cubic smoothing splines are reported

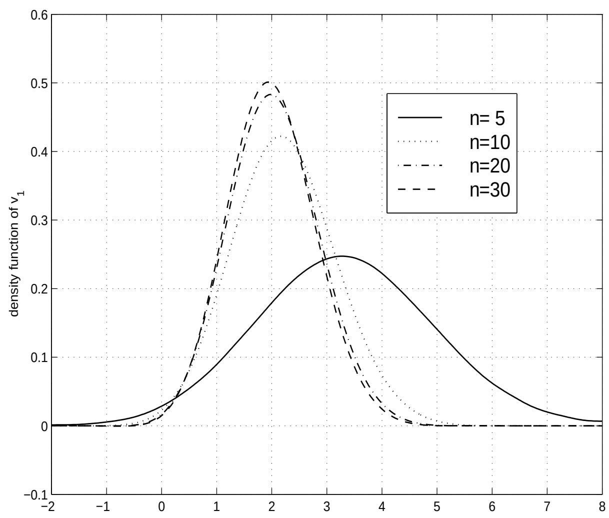

FIG. 6. Asymptotic density function of the largest eigenvalue for multiplicity $m = 4$ and underlying multivariate Student distribution with degrees of freedom $\nu = 5, 10, 20, 30$.

TABLE 1.

Maximum standard deviations when $\kappa = 0$ of the estimates in Figure 1.

m	Maximum standard deviation
2	3.377E-04
3	7.982E-04
4	1.913E-03
5	5.358E-03

in Figures 3 and 4. These values indicate that the estimated probabilities for the two curves corresponding to $m = 2$ and 3 are accurate to 2 decimals.

For $m > 3$ or $\kappa \neq 0$, a possible measure of accuracy could be the standard deviation in (5). For each curve, Tables 1 and 2 report the maximum standard deviations over all points t_s, $s = 1, \ldots, S$.

TABLE 2.

Maximum standard deviations when $m = 4$ of the estimates in Figure 5.

ν	Maximum standard deviation
5	2.882E-03
10	2.038E-03
20	1.937E-03
30	1.923E-03

It appears that the asymptotic distribution of the largest eigenvalue is very much affected by the kurtosis of the population distribution and by the multiplicity of the largest population eigenvalue.

APPENDIX: PROOFS

Proof (*Theorem* 2.1). By invariance with respect to orthogonal transformations one may assume at the onset that $\boldsymbol{\Sigma} = \mathrm{diag}(d_1\boldsymbol{I}_{p_1}, \ldots, d_k\boldsymbol{I}_{p_k}) \equiv \boldsymbol{D}$, $p = p_1 + \cdots + p_k$, where $d_1 > \cdots > d_k$ are the distinct eigenvalues of multiplicity p_i, $i = 1, \ldots, k$. From Muirhead (1982) or Tyler (1982), the asymptotic distribution of the sample covariance matrix $\boldsymbol{S}_n = \frac{1}{n}\sum_{i=1}^n(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})'$, where $\bar{\boldsymbol{x}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{x}_i$, is given by $n^{1/2}(\boldsymbol{S}_n - \boldsymbol{D}) \xrightarrow{L} \boldsymbol{W}$, with

$$\boldsymbol{W} \sim N_p^p\left(\boldsymbol{0}, (1+\kappa)(\boldsymbol{I} + \boldsymbol{K}_p)(\boldsymbol{D} \otimes \boldsymbol{D}) + \kappa \ \mathrm{vec}(\boldsymbol{D})\mathrm{vec}(\boldsymbol{D})'\right).$$

The matrix $\boldsymbol{K}_p = \sum_{i=1}^p \sum_{i=1}^p \boldsymbol{e}_i\boldsymbol{e}_j' \otimes \boldsymbol{e}_j\boldsymbol{e}_i' : p^2 \times p^2$ is the commutation matrix of Magnus and Neudecker (1979), where $\boldsymbol{e}_i$ is a p-vector with a one in position i and zeros elsewhere. Consider the partition $\boldsymbol{W} = (\boldsymbol{W}_{ij})$, where $\boldsymbol{W}_{ij} : p_i \times p_j$. Given $\boldsymbol{S} : p \times p$ symmetric, let $\boldsymbol{\phi}(\boldsymbol{S}) = (\mathrm{ch}_1(\boldsymbol{S}), \ldots, \mathrm{ch}_p(\boldsymbol{S}))$ be the ordered eigenvalues of $\boldsymbol{S}$, $\mathrm{ch}_1(\boldsymbol{S}) \geq \cdots \geq \mathrm{ch}_p(\boldsymbol{S})$. From Eaton and Tyler (1991), $n^{1/2}(\boldsymbol{\phi}(\boldsymbol{S}_n) - \boldsymbol{\phi}(\boldsymbol{D})) \xrightarrow{L} \boldsymbol{H}(\boldsymbol{W})$, where the function $\boldsymbol{H}(\cdot)$ is defined as

$$\boldsymbol{H}(\boldsymbol{W}) = \begin{pmatrix} \boldsymbol{\phi}(\boldsymbol{W}_{11}) \\ \vdots \\ \boldsymbol{\phi}(\boldsymbol{W}_{kk}) \end{pmatrix}.$$

For simplicity, denote by m the multiplicity p_1 of the largest eigenvalue. However, since

$$\boldsymbol{Y} \equiv \frac{\boldsymbol{W}_{11}}{\sqrt{2}d_1} \sim N_m^m\left(\mathbf{0}, (1+\kappa)(\boldsymbol{I}+\boldsymbol{K}_m)/2 + \kappa\ \mathrm{vec}(\boldsymbol{I}_m)\mathrm{vec}(\boldsymbol{I}_m)'/2\right),$$

it follows that $v_1 \xrightarrow{L} \mathrm{ch}_1(\boldsymbol{Y})$. The asymptotic distribution of v_1 can thus be obtained by writing the density function of $\boldsymbol{Y}$, applying Theorem 3.2.17 of Muirhead (1982) to get the density function of the eigenvalues of $\boldsymbol{Y}$, and, finally, integrating to reach the distribution function of $\mathrm{ch}_1(\boldsymbol{Y})$. Following Kent and Tyler (1991), define

$$\begin{aligned} \boldsymbol{A}_j &= (\mathbf{0}, \boldsymbol{I}_j)' : m \times j,\ j = 1, \ldots, m, \\ \boldsymbol{M}_m &= \mathrm{diag}(\boldsymbol{A}_m, \ldots, \boldsymbol{A}_1) : m^2 \times m(m+1)/2. \end{aligned}$$

Then, for any symmetric matrix $\boldsymbol{A} : m \times m$, $\boldsymbol{M}_m'\mathrm{vec}(\boldsymbol{A})$ is just the $m(m+1)/2$-dimensional vector formed by stacking the columns of $\boldsymbol{A}$ after deleting the upper triangular part. It is easy to check that

$$\boldsymbol{M}_m'\boldsymbol{M}_m = \boldsymbol{I},\ \boldsymbol{M}_m'\boldsymbol{K}_m\boldsymbol{M}_m = \mathrm{diag}(\boldsymbol{a}_m),\ \boldsymbol{M}_m'\mathrm{vec}(\boldsymbol{I}_m) = \boldsymbol{a}_m,$$

where

$$\begin{aligned} \boldsymbol{\alpha}_j &= (1, 0, \ldots, 0)' : j \times 1,\ j = 1, \ldots, m, \\ \boldsymbol{a}_m &= (\boldsymbol{\alpha}_m', \ldots, \boldsymbol{\alpha}_1')' : m(m+1)/2 \times 1. \end{aligned}$$

Thus,

$$\boldsymbol{y} \equiv \boldsymbol{M}_m'\mathrm{vec}(\boldsymbol{Y}) \sim N_{m(m+1)/2}\left(\mathbf{0}, (1+\kappa)\mathrm{diag}(\boldsymbol{a}_m + \mathbf{1})/2 + \kappa\ \boldsymbol{a}_m\boldsymbol{a}_m'/2\right),$$

where $\mathbf{1} = (1, \ldots, 1)'$ is a $m(m+1)/2$-vector of ones. Define $\mathbf{\Omega} = (1 + \kappa)\text{diag}(\boldsymbol{a}_m + \mathbf{1})/2 + \kappa\ \boldsymbol{a}_m \boldsymbol{a}_m'/2$ and use the representation of a perturbed matrix (by a rank one matrix) to find

$$\mathbf{\Omega}^{-1} = 2(1+\kappa)^{-1}\text{diag}^{-1}(\boldsymbol{a}_m + \mathbf{1}) - \frac{(1+\kappa)^{-2}\kappa\ \boldsymbol{a}_m \boldsymbol{a}_m'}{2[1+\kappa(1+\kappa)^{-1}m/2]}$$

and

$$\det(\mathbf{\Omega}) = (1+\kappa)^{m(m+1)/2}(1/2)^{m(m-1)/2}[1+\kappa(1+\kappa)^{-1}m/2].$$

Straightforward algebra yields

$$\boldsymbol{y}'\mathbf{\Omega}^{-1}\boldsymbol{y} = (1+\kappa)^{-1}\text{tr}(\boldsymbol{Y}^2) - \frac{(1+\kappa)^{-2}\kappa\ \text{tr}^2(\boldsymbol{Y})}{2[1+\kappa(1+\kappa)^{-1}m/2]}.$$

The density function of $\boldsymbol{Y}$ thus takes the final form

$$f(\boldsymbol{Y}) = C_m(\kappa)\exp\left\{-\frac{1}{2}\left[(1+\kappa)^{-1}\text{tr}(\boldsymbol{Y}^2) - \frac{(1+\kappa)^{-2}\kappa\ \text{tr}^2(\boldsymbol{Y})}{2[1+\kappa(1+\kappa)^{-1}m/2]}\right]\right\},$$

where

$$C_m(\kappa) = [2\pi(1+\kappa)]^{-m(m+1)/4}(1/2)^{-m(m-1)/4}[1+\kappa(1+\kappa)^{-1}m/2]^{-1/2}$$

Theorem 3.2.17 of Muirhead (1982), which is valid for nonsingular w.p.1 symmetric matrix, gives the density function of the ordered eigenvalues, $y_1 > \cdots > y_m$, of $\boldsymbol{Y}$

$$g(y_1, \ldots, y_m) = \frac{\pi^{m^2/2}}{\Gamma_m(m/2)}\prod_{i<j}(y_i - y_j)\int_{\boldsymbol{O}(m)} f(\boldsymbol{GLG}')(d\boldsymbol{G}),$$

where $\boldsymbol{L} = \mathrm{diag}(y_1, \ldots, y_m)$ and $\boldsymbol{O}(m)$ is the orthogonal group of order m. Since $f(\boldsymbol{Y})$ is invariant with respect to $\boldsymbol{O}(m)$,

$$\begin{aligned} g(y_1, \ldots, y_m) &= \frac{\pi^{m^2/2}}{\Gamma_m(m/2)} C_m(\kappa) \prod_{i<j} (y_i - y_j) \\ & \exp\left\{ -\frac{1}{2} \left[(1+\kappa)^{-1} \sum_i y_i^2 - \frac{(1+\kappa)^{-2}\kappa \ \left(\sum_i y_i\right)^2}{2[1+\kappa(1+\kappa)^{-1} m/2]} \right] \right\}. \end{aligned}$$

Recognizing $(2\pi)^{-m/2}(1+\kappa)^{-m/2} \exp[-\frac{1}{2}(1+\kappa)^{-1} \sum_i y_i^2] m!$ as the density of the order statistics of m independently and identically distributed observations from a $N(0, (1+\kappa))$ distribution, the distribution function of y_1 can thus be expressed as

$$\begin{aligned} P(y_1 \le t) &= D_m(\kappa) E \left\{ \prod_{i<j} (z_i - z_j) \exp\left[\frac{(1+\kappa)^{-1}\kappa (\sum_i z_i)^2}{4[1+\kappa(1+\kappa)^{-1} m/2]} \right] \cdot \right. \\ & \left. I\left(z_1 \le t/(1+\kappa)^{1/2} \right) \right\}, \end{aligned}$$

where $I(\cdot)$ is an indicator function,

$$D_m(\kappa) = \frac{\pi^{m(m+1)/4}}{m! \Gamma_m(m/2)} [1 + \kappa(1+\kappa)^{-1} m/2]^{-1/2}$$

and $z_1 > \cdots > z_m$ is the order statistics of m independently and identically distributed observations from a $N(0,1)$ distribution. ▮

Proof (*Theorem* 2.2). Let

$$\boldsymbol{Y}_{m+1} \sim N_{m+1}^{m+1}\left(\boldsymbol{0}, (1+\kappa)(\boldsymbol{I} + \boldsymbol{K}_{m+1})/2 + \kappa \ \ \mathrm{vec}(\boldsymbol{I}_{m+1}) \mathrm{vec}(\boldsymbol{I}_{m+1})'/2\right)$$

and consider the partition

$$\boldsymbol{Y}_{m+1} = \begin{pmatrix} \boldsymbol{Z} & \boldsymbol{z} \\ \boldsymbol{z}' & z \end{pmatrix} : (m+1) \times (m+1),$$

where $\boldsymbol{z}$ is a m-vector. Note that since $\boldsymbol{Z} \overset{d}{=} \boldsymbol{Y}_m$ (equality in distribution) and

$$\mathrm{ch}_1(\boldsymbol{Y}_{m+1}) = \sup_{|\boldsymbol{h}|=1} \boldsymbol{h}'\boldsymbol{Y}_{m+1}\boldsymbol{h} \geq \mathrm{ch}_1(\boldsymbol{Z}),$$

then the result follows. ▮

Proof (*Theorem* 2.3). Along the lines in the proof of Theorem 2.1 the asymptotic distribution $n^{1/2}(\boldsymbol{S}_n - \boldsymbol{D}) \xrightarrow{L} \boldsymbol{W}$ holds, where $\boldsymbol{W}$ is multivariate normal with mean $\mathbf{0}$ and covariances, expressed with cumulants,

$$\mathrm{cov}\,(\boldsymbol{W}(i,k), \boldsymbol{W}(j,l)) = k_{1111}^{ijkl} + k_{11}^{kl}k_{11}^{ij} + k_{11}^{il}k_{11}^{jk}.$$

However, since $n^{1/2}[(\mathrm{ch}_1(\boldsymbol{S}_n), \ldots, \mathrm{ch}_m(\boldsymbol{S}_n)) - d_1\mathbf{1}_m] \xrightarrow{L} \boldsymbol{\phi}(\boldsymbol{W}_{11})$, then

$$n^{1/2}(\bar{d} - d_1) \xrightarrow{L} \mathrm{tr}(\boldsymbol{W}_{11})/m. \tag{A.1}$$

The matrix $\boldsymbol{W}_{11}$ being normally distributed, so is the linear combination $\mathrm{tr}(\boldsymbol{W}_{11})$. Straightforward calculations of $\mathrm{var}\,(\boldsymbol{W}_{11}(1,1) + \cdots + \boldsymbol{W}_{11}(m,m))$ leads to γ^2 in Theorem 2.3. ▮

Proof (*Corollary* 2.2). The proof uses (5.5) of Eaton and Tyler (1991) which asserts

$$\boldsymbol{X}_n^* - [\boldsymbol{H}(\tilde{\boldsymbol{W}}_n^{*0} + \tilde{\boldsymbol{A}}_n) - \boldsymbol{H}(\tilde{\boldsymbol{A}}_n)] \xrightarrow{L^*} \mathbf{0} \text{ a.s.}, \tag{A.2}$$

where $\boldsymbol{X}_n^* = n^{1/2}[\boldsymbol{\phi}(\boldsymbol{S}_n^*) - \boldsymbol{\phi}(\boldsymbol{S}_n)]$, and $\tilde{\boldsymbol{W}}_n^{*0}$, $\tilde{\boldsymbol{A}}_n$ are defined in Eaton and Tyler (1991). Hence, taking the average of the first m components in (A.2) yields

$$n^{1/2}(\bar{d}^* - \bar{d}) - [\frac{1}{m}\mathrm{tr}(\boldsymbol{W}_{n,11}^{*0} + \boldsymbol{A}_{n,11}) - \frac{1}{m}\mathrm{tr}(\boldsymbol{A}_{n,11})] \xrightarrow{L^*} 0 \text{ a.s.},$$

and subsequently

$$n^{1/2}(\bar{d}^* - \bar{d}) - \frac{1}{m}\mathrm{tr}(\boldsymbol{W}^{*0}_{n,11}) \xrightarrow{L^*} 0 \text{ a.s.}$$

Since $\boldsymbol{W}^{*0}_{n,11} \xrightarrow{L^*} \boldsymbol{W}_{11}$ a.s. which also appears in Eaton and Tyler (1991), the conclusion follows from (A.1). ▮

ACKNOWLEDGMENT

I would like to thank Claude Beauchamps for the efficient computer programs used in the simulations.

REFERENCES

1. Anderson, T.W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Statist.* **34**, 122–148.
2. Bellman, R. (1960). *Introduction to matrix analysis*, McGraw-Hill, New York, 1960.
3. Beran, R. and Srivastava, M.S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix, *Ann. Statist.* **13**, 95–115.
4. Beran, R. and Srivastava, M.S. (1987). Correction: Bootstrap tests and confidence regions for functions of a covariance matrix, *Ann. Statist.* **15**, 470–471.
5. Constantine, A.G. (1963). Some noncentral distribution problems in multivariate analysis, *Ann. Math. Statist.* **34**, 1270–1285.
6. Eaton, M.L. and Tyler, D.E. (1991). On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix, *Ann.*

Statist. **19**, 260–270.

7. Fujikoshi, Y. (1986). Amendments and corrections: Asymptotic expansions for the distributions of the sample roots under nonnormality, *Biometrika* **73**, 245.

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

9. Fujikoshi, Y. (1977). An asymptotic expansion for the distributions of the latent roots of the Wishart matrix with multiple population roots, *Ann. Inst. Statist. Math.* **29**, 379–387.

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

11. Kent, J.T. and Tyler, D.E. (1991). Redescending M-estimates of multivariate location and scatter, *Ann. Statist.* **19**, 2102–2019.

12. Khatri, C.G. and Srivastava, M.S. (1978). Asymptotic expansions for distributions of characteristic roots of covariance matrices, *South African Statist. J.* **12**, 161–186.

13. Kollo, T. and Neudecker, H. (1993). Asymptotics of eigenvalues and unit-length eigenvectors of sample variance and correlation matrices, *J. Multivariate Anal.* **47**, 283–300.

14. Magnus, J.R. and Neudecker, H. (1979). The commutation matrix: Some properties and applications, *Ann. Statist.* **7**, 381–394.

15. Muirhead, R.J. (1974). Powers of the largest root test of $\Sigma = I$, *Comm. Statist.* **3**, 513–524.

16. Muirhead, R.J. (1982). *Aspects of multivariate statistical theory*, John Wiley & Sons, New York.

17. Muirhead, R.J. and Chikuse, Y. (1975). Asymptotic expansions for the joint and marginal distributions of the latent roots of the covariance matrix, *Ann. Statist.* **3**, 1011–1017.

18. Sugiura, N. (1973). Derivatives of the characteristic root of a symmetric or a Hermitian matrix with two applications in multivariate analysis, *Comm. Statist.* **1**, 393–417.

19. Sugiura, N. (1976). Asymptotic expansions of the distributions of the latent roots and the latent vector of the Wishart and multivariate F matrices, *J. Multivariate Anal.* **6**, 500–525.

20. Sugiyama, T. (1972). Percentile points of the largest latent root of a matrix and power calculations for testing hypothesis $\boldsymbol{\Sigma} = \boldsymbol{I}$, *J. Japan Statist. Soc.* **3**, 1–8.

21. Tyler, D.E. (1982). Radial estimates and the test for sphericity, *Biometrika* **69**, 429–436.

22. Tyler, D.E. (1983). The asymptotic distribution of principal components roots under local alternatives to multiple roots. *Ann. Statist.* **11**, 1232–1242.

23. Waternaux, C.M. (1976). Asymptotic distribution of the sample roots for a nonnormal population, *Biometrika* **63**, 639–645.