数理解析研究所講究録
231巻 1975年 47-57

Asymptotic Expansions for the Joint and Marginal Distributions of the Latent Roots of the Covariance Matrix

Yale大　筑瀬 靖子

1. Introduction.

Let nS be an $m \times m$ matrix having the Wishart distribution $W_m(n, \Sigma)$. Let $\ell_1 > \ell_2 > \cdots > \ell_m > 0$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m > 0$ denote the latent roots of S and Σ respectively. For large n and simple latent roots of Σ, it is known that the latent roots of S are asymptotically independently normal. We assume throughout this paper that all the roots of Σ are simple. In this paper an expansion, up to and including the term of order n^{-1}, is given for the joint density function of $\ell_1, \cdots, \ell_m$ in terms of normal density functions. Expansions for the marginal distributions of the roots are also given, valid when the corresponding roots of Σ are simple.

2. Expansions for the extreme root distributions.

We consider first the largest root ℓ_1. From the exact expression for the distribution function of l_1 by Sugiyama [15], [16], the distribution function of $x_1=(n/2)^{1/2}(l_1/\lambda_1-1)$ can be written as

(2.1) $P(x_1<x)=[\Gamma_m(p)/\Gamma_m(\frac{1}{2}n+p)](\det R)^{n/2}\,{}_1F_1(\frac{1}{2}n;\frac{1}{2}n+p;-R)$,

where $p=\frac{1}{2}(m+1)$, $\Gamma_m(a)=\pi^{m(m-1)/4}\prod_{i=1}^{m}\Gamma(a-(i-1)/2)$, $R=\mathrm{diag}(r_1,r_2,\cdots,r_m)$ with $r_i=[n/2+(n/2)^{1/2}x]z_i$, $z_i=\lambda_1/\lambda_i$ $(i=1,\cdots,m)$ and ${}_1F_1$ is a confluent hypergeometric function of matrix argument (see Herz [7], Constantine [5]). A system of partial differential equations (pde's) satisfied by the ${}_1F_1$ function has been given by Muirhead [12]. Starting with this system it can be readily verified that $P\equiv P(x_1<x)$ satisfies each of the m pde's

(2.2)
$$\frac{\partial^2 P}{\partial x^2}+x\frac{\partial P}{\partial x}+\left(\frac{2}{n}\right)^{\frac{1}{2}}\Big[2x\frac{\partial^2 P}{\partial x^2}+(1+x^2-\tfrac{1}{2}A_1)\frac{\partial P}{\partial x}-x\sum_{k=2}^{m}z_k\frac{\partial P}{\partial z_k}$$
$$-2\sum_{k=2}^{m}z_k\frac{\partial^2 P}{\partial x\,\partial z_k}\Big]+\frac{2}{n}\Big[x^2\frac{\partial^2 P}{\partial x^2}+x(1-\tfrac{1}{2}A_1)\frac{\partial P}{\partial x}$$
$$+\sum_{k=2}^{m}z_k\Big(1+\tfrac{1}{2}A_1-\frac{1}{2(1-z_k)}\Big)\frac{\partial P}{\partial z_k}-2x\sum_{k=2}^{m}z_k\frac{\partial^2 P}{\partial x\,\partial z_k}$$
$$+\sum_{k=2}^{m}\sum_{j=2}^{m}z_j z_k\frac{\partial^2 P}{\partial z_k\,\partial z_j}\Big]=0$$

and

(2.3) $$(z_i-1)\frac{\partial P}{\partial z_i}+\left(\frac{2}{n}\right)^{\frac{1}{2}}\Big[\frac{1}{2(1-z_i)}\frac{\partial P}{\partial x}+xz_i\frac{\partial P}{\partial z_i}\Big]+\frac{2}{n}\Big[z_i\frac{\partial^2 P}{\partial z_i^2}$$

$$+\frac{x}{2(1-z_i)}\frac{\partial P}{\partial x}+(1-\tfrac{1}{2}A_i)\frac{\partial P}{\partial z_i}-\frac{1}{2(1-z_i)}\sum_{k=2}^{m} z_k\frac{\partial P}{\partial z_k}$$

$$-\frac{1}{2}\sum_{\substack{j=2\\ j\neq i}}^{m}\frac{z_j}{z_i-z_j}\frac{\partial P}{\partial z_j}\Big]=0 \qquad (i=2,3,\cdots,m),$$

where

$$(2.4)\qquad A_i=\sum_{\substack{j=1\\ j\neq i}}^{m}\frac{z_i}{z_j-z_i}\qquad (i=1,2,\cdots,m).$$

We now look for a solution of these m pde's (2.2) and (2.3) of the form

$$(2.5)\qquad P=\Phi(x)+\sum_{k=1}^{\infty}(2/n)^{k/2}Q_k,$$

where the Q_k are functions of $x, z_2, \cdots, z_m$. (That P possesses such an expansion follows from results in the next section.) We substitute the series (2.5) into (2.2) and (2.3) and equate coefficients of powers of $(2/n)^{1/2}$ on the L.H.S.'s to zero. Equating the coefficient of $(2/n)^{1/2}$ in (2.2) and (2.3) to zero and using the boundary conditions $P(x,<\infty)=1$ and $P(x,<-\infty)=0$, we have

$$(2.6)\qquad Q_1=-(1/6)\varphi(x)[2H_2(x)+3A_1H_0(x)],$$

where $H_j(x)$ denotes the Hermite polynomial of degree j (see Kendall and Stuart [9], p. 155). Similarly, equating the coefficient of $2/n$ in (2.2) and (2.3) to zero and solving the resulting equations gives

(2.7) $Q_2 = -(1/72)\varphi(x)[4H_5(x)+18H_3(x)+12A_1H_3(x)-18B_1H_1(x)+9A_1^2H_1(x)]$,

where $A_1 = \sum_{j=2}^{m}(z_j-1)^{-1}$, $\quad B_1 = \sum_{j=2}^{m}(z_j-1)^{-2}$.

Coefficients of higher powers of $(2/n)^{1/2}$ in (2.5) may be obtained in a similar manner if required. The expansion is summarized in the following

Theorem 2.1. The distribution function of $x_1 = (n/2)^{1/2}(\ell_1/\lambda_1 - 1)$, when the latent roots of Σ are simple, can be expanded for large n as

(2.8) $P(x_1 < x) = \Phi(x) + (2/n)^{1/2}Q_1 + (2/n)Q_2 + O(n^{-3/2})$,

where Q_1 and Q_2 are given by (2.6) and (2.7) respectively.

Consider now the distribution of the smallest root ℓ_m. Since nS is $W_m(n,\Sigma)$ we have

(2.9) $P(\ell_m > y) = [(\tfrac{1}{2}n)^{\frac{1}{2}mn}(\det)^{-\frac{1}{2}n}/\Gamma_m(\tfrac{1}{2}n)]\int_{S>yI}\exp(-\tfrac{1}{2}n\,\mathrm{tr}(\Sigma^{-1}S))\det S^{\frac{1}{2}n-p}\,dS.$

Making the transformation $T = y^{-1}S - I$ it is easily seen that (2.9) becomes

(2.10) $P(\ell_m > y) = [\Gamma_m(p)/\Gamma_m(\tfrac{1}{2}n)]\det(\tfrac{1}{2}ny\Sigma^{-1})^{\frac{1}{2}n}\exp(-\tfrac{1}{2}ny\,\mathrm{tr}\,\Sigma^{-1})$

$\cdot\Psi(p, \tfrac{1}{2}n+p; \tfrac{1}{2}ny\Sigma^{-1})$, where $\Psi(a, c; R) \overset{def.}{\equiv} [1/\Gamma_m(a)] \int_{S>0} \exp(-tr(RS)) (\det)^{a-p} \det(I+S)^{c-a-p} dS.$

The function Ψ is another confluent hypergeometric function of matrix argument (see Muirhead [13]).

Putting $x_m = (n/2)^{1/2}(l_m/\lambda_m - 1)$ and using the system of pde's satisfied by the Ψ function given by Muirhead [13] it can readily be shown that the distribution function of x_m, $P \equiv P(x_m < x)$, satisfies each of the m pde's (2.2) and (2.3). The only difference here is that now $z_i = \lambda_m/\lambda_{m-i+1}$ instead of λ_1/λ_i as it was in the largest root distribution. Hence

Theorem 2.2. The distribution function of $x_m = (n/2)^{1/2}(l_m/\lambda_m - 1)$, when the latent roots of Σ are simple, can be expanded for large n as

$$P(x_m < x) = \Phi(x) + (2/n)^{1/2} Q_1 + (2/n) Q_2 + O(n^{-3/2}),$$

where $z_i = \lambda_m/\lambda_{m-i+1}$ in Q_1 and Q_2 given by (2.6) and (2.7) respectively.

3. Expansion for the joint distribution.

The joint density function of $l_1, \cdots, l_m$ can be expressed in the form (see James [8])

(3.1) $\pi^{\frac{1}{2}m^2}(\frac{1}{2}n)^{\frac{1}{2}mn}[\Gamma_m(\frac{1}{2}n)\Gamma_m(\frac{1}{2}m)]^{-1}\prod_{i=1}^{m} l_i^{\frac{1}{2}n-p}\lambda_i^{-\frac{1}{2}n}\prod_{i<j}^{m}(l_i-l_j)\,{}_0F_0(-\frac{1}{2}nL,\Sigma^{-1})$,

where $p=\frac{1}{2}(m+1)$, $L=\mathrm{diag}(l_1,\cdots,l_m)$, $\Sigma=\mathrm{diag}(\lambda_1,\cdots,\lambda_m)$ and ${}_0F_0$ is a hypergeometric function with two argument matrices. The ${}_0F_0$ function in (3.1) has been expanded for large n by G. Anderson [1] by expressing it as an integral over the orthogonal group. In [1] it is shown that the joint density function can be expressed as

(3.2) $k_1\prod_{i=1}^{m}[\lambda_i^{(m-n-1)/2}l_i^{n/2-p}\exp(-nl_i/2\lambda_i)]\prod_{i<j}^{m}[(l_i-l_j)/(\lambda_i-\lambda_j)]^{1/2}\cdot G$,

where $k_1=(n/2)^{mn/2-m(m-1)/4}/\prod_{i=1}^{m}\Gamma((n-i+1)/2)$

and

(3.3) $G=1+(2n)^{-1}\sum_{i<j}^{m}\lambda_i\lambda_j(\lambda_i-\lambda_j)^{-1}(l_i-l_j)^{-1}+O(n^{-2})$.

(Anderson did not show in general that the remainder term in (3.3) is of order n^{-2}; this has been shown by the author in her Ph.D. thesis at Yale University in the more general case.) Now putting $x_i=(n/2)^{1/2}(l_i/\lambda_i-1)$ $(i=1,\cdots,m)$, from (3.2) the joint density function of $x_1,\cdots,x_m$ can be expressed as

$$(3.4)\qquad k_2 F_1 F_2 \{1 + (2n)^{-1} \sum_{i<j}^{m} \lambda_i \lambda_j / (\lambda_i - \lambda_j)^2 + O(n^{-3/2})\},$$

where
$$k_2 = (n/2)^{mn/2 - m(m+1)/4} \exp(-mn/2) / \prod_{i=1}^{m} \Gamma((n-i+1)/2),$$
$$F_1 = \prod_{i=1}^{m} \{(1 + (\tfrac{2}{n})^{1/2} x_i)^{n/2-p} \exp(-(n/2)^{1/2} x_i)\}$$

and
$$F_2 = \prod_{i<j}^{m} \{1 + (\tfrac{2}{n})^{1/2} (x_i \lambda_i - x_j \lambda_j) / (\lambda_i - \lambda_j)\}^{1/2}.$$

It remains to expand k_2, F_1 and F_2 in (3.4) for large n. For example, by expanding the gamma functions for large n it follows that

$$k_2 = (2\pi)^{-m/2} \{1 - (24n)^{-1} m (2m^2 + 3m - 1) + O(n^{-2})\}.$$

The functions F_1 and F_2 can be easily expanded in terms of powers of $n^{-1/2}$; however these expansions, up to and including the terms of order n^{-1}, are quite lengthy and are omitted here. Substituting these expansions in (3.4) gives an expansion of the joint density function of $x_1, \cdots, x_m$. The final result is summarized in the following

<u>Theorem 3.1.</u> The joint density function of $x_i = (n/2)^{1/2} (l_i / \lambda_i - 1)$ $(i = 1, \cdots, m)$, where $\lambda_1, \cdots, \lambda_m$ are simple roots of Σ, may be expanded for large n as

$$(3.5)\quad \prod_{i=1}^{m} \varphi(x_i)\cdot\Big\{1+(2/n)^{1/2}\sum_{i=1}^{m} P_{1i}(x_i)+(2/n)\Big(\sum_{i=1}^{m} P_{2i}(x_i) + \sum_{i<j}^{m} P_{1i}(x_i)P_{1j}(x_j)+\frac{1}{2}\sum_{i<j}^{m}\frac{x_i x_j \lambda_i \lambda_j}{(\lambda_i-\lambda_j)^2}\Big)+O(n^{-3/2})\Big\},$$

where

$$(3.6)\quad P_{1i}(x)=(1/6)\{2H_3(x)+3A_i H_1(x)\},$$

$$(3.7)\quad P_{2i}(x)=(1/72)\{4H_6(x)+18H_4(x)+12A_i H_4(x)-18B_i H_2(x)+9A_i^2 H_2(x)\},$$

$H_j(x)$ is the Hermite polynomial of degree j, and

$$(3.8)\quad A_i=\sum_{\substack{j=1\\ j\neq i}}^{m} \lambda_j/(\lambda_i-\lambda_j)\ ,\qquad B_i=\sum_{\substack{j=1\\ j\neq i}}^{m} \lambda_j^2/(\lambda_i-\lambda_j)^2.$$

Note that A_i is the same as in (2.4).

By integrating out the other variables in (3.5) an expansion of the marginal density function of x_i can be obtained.

<u>Corollary</u> The marginal density function of $x_i=(n/2)^{1/2}(l_i/\lambda_i-1)$, where λ_i is a simple root of Σ, may be expanded for large n as

$$(3.9)\quad \varphi(x_i)\{1+(2/n)^{1/2}P_{1i}(x_i)+(2/n)P_{2i}(x_i)+O(n^{-3/2})\},$$

where $P_{1i}(x_i)$ and $P_{2i}(x_i)$ are given by (3.6) and (3.7) respectively.

The expansion (3.9), in the cases $i=1$ and m, agrees with the expansions for the extreme root distributions given in the previous section. Sugiura [14] has also obtained (3.9) using another method.

Asymptotic moments of l_i can be obtained from (3.9); we obtain

$$E(l_i) = \lambda_i + A_i\lambda_i/n + O(n^{-2}),$$

(3.10) $$\mathrm{Var}(l_i) = 2\lambda_i^2/n - 2\lambda_i^2 B_i/n^2 + O(n^{-3}),$$

$$\kappa_3(l_i) = 8\lambda_i^3/n^2 + O(n^{-3}), \quad \kappa_4(l_i) = 48\lambda_i^4/n^3 + O(n^{-4}),$$

where $\kappa_3(l_i)$ and $\kappa_4(l_i)$ denote the third and fourth cumulants of l_i and A_i, B_i are given by (3.8). From (3.5) we obtain

(3.11) $$\mathrm{Cov}(l_i, l_j) = 2[\lambda_i\lambda_j/(\lambda_i-\lambda_j)]^2/n^2 + O(n^{-3}).$$

These expansions agree with results obtained by Lawley [11] without using the asymptotic normality.

Acknowledgment: This is a portion of the author's Ph. D. thesis written at Yale University and supervised by Professor Robb J. Muirhead.

REFERENCES

[1] ANDERSON, G. A. (1965). An asymptotic expansion for the distribution of the latent roots of the estimated covariance matrix. Ann. Math. Statist. 36 1153-1173.

[2] ANDERSON, T. W. (1958). An Introduction to Multivariate Statistical Analysis. Wiley, New York.

[3] ANDERSON, T. W. (1963). Asymptotic theory for principal component analysis. Ann. Math. Statist. 34 122-148.

[4] CHAMBERS, J. M. (1967). On methods of asymptotic approximation for multivariate distributions. Biometrika 54 367-383.

[5] CONSTANTINE, A. G. (1963). Some noncentral distribution problems in multivariate analysis. Ann. Math. Statist. 34 1270-1285.

[6] GIRSHICK, M. A. (1939). On the sampling theory of roots of determinantal equations. Ann. Math. Statist. 10 203-224.

[7] HERZ, C. S. (1955). Bessel functions of matrix argument. Ann. of Math. 61 474-523.

[8] JAMES, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. Ann. Math. Statist. 35 475-501.

[9] KENDALL, M. G. and STUART, A. (1969). The Advanced Theory of Statistics. Vol. 1. Hafner, New York.

[10] KHATRI, C. G. (1972). On the exact finite series distribution of the smallest or the largest root of matrices in three situations. J. Multivariate Anal. 2 201-207.

[11] LAWLEY, D. N. (1956). Tests of significance for the latent roots of covariance and correlation matrices. Biometrika 43 128-136.

[12] MUIRHEAD, R. J. (1970). Systems of partial differential equations for hypergeometric functions of matrix argument. Ann. Math. Statist. 41 991-1001.

[13] MUIRHEAD, R. J. (1970). Asymptotic distributions of some multivariate tests. Ann. Math. Statist. 41 1002-1010.

[14] SUGIURA, N. (1973). Derivatives of the characteristic root of a symmetric or a Hermitian matrix with two applications in multivariate analysis. Communications in Statistics 1 393-417.

[15] SUGIYAMA, T. (1967). On the distribution of the largest latent root of the covariance matrix. Ann. Math. Statist. 38 1148-1151.

[16] SUGIYAMA, T. (1972). Approximation for the distribution of the largest latent root of a Wishart matrix. Austral. J. Statist. 14 17-24.