SIAM J. MATH. ANAL.
Vol. 16, No. 3, May 1985

018

THE LIMITING EIGENVALUE DISTRIBUTION OF A MULTIVARIATE F MATRIX*

JACK W. SILVERSTEIN†

Abstract. Let X_{ij}, Y_{ij} $i,j=1,2,\cdots$ be i.i.d. $N(0,1)$ random variables and for positive integers p,m,n, let $\bar{X}_p=(X_{ij})$ $i=1,2,\cdots,p$; $j=1,2,\cdots,m$, and $\bar{Y}_p=(Y_{ij})$ $i=1,2,\cdots,p$; $j=1,2,\cdots,n$. Suppose further that $p/m\to y>0$ and $p/n\to y'\in(0,\frac{1}{2})$ as $p\to\infty$. In [5], [6] it is shown that the empirical distribution function of the eigenvalues of $(1/m\bar{X}_p\bar{X}_p^T)(1/n\bar{Y}_p\bar{Y}_p^T)^{-1}$ converges i.p. as $p\to\infty$ to a nonrandom d.f.

In the present paper the limiting d.f. is derived.

1. Introduction. Let X_{ij}, $i,j=1,2,\cdots$ be i.i.d. $N(0,1)$ random variables, and for any positive integers p,m, let $W_p=\bar{X}_p\bar{X}_p^T$, $\bar{X}_p=(X_{ij})$ $i=1,2,\cdots,p$; $j=1,2,\cdots,m$, be the $p\times p$ Wishart matrix $W(I,m)$. It is well known [1], [2], [4] that if $p/m\to y>0$ as $p\to\infty$, then the empirical distribution function F_p of the eigenvalues of $(1/m)W_p$ (i.e. $F_p(x)=(1/p)$ (# of eigenvalues of $(1/m)W_p\leqq x$)) converges a.s. for every $x\geqq 0$ to a nonrandom d.f. F_y, where for $0<y\leqq 1$, F_y has density

$$(1.1)\qquad f_y(x)=\begin{cases}\dfrac{1}{2\pi yx}\sqrt{\left(x-(1-\sqrt{y})^2\right)\left((1+\sqrt{y})^2-x\right)} & \text{for } (1-\sqrt{y})^2<x<(1+\sqrt{y})^2,\\ 0 & \text{otherwise,}\end{cases}$$

and for $1<y<\infty$ F_y has mass $1-1/y$ at zero and density f_y on $((1-\sqrt{y})^2,(1+\sqrt{y})^2)$.

In [6] it is shown that the empirical d.f. of $(1/m)W_pT_p$, under certain conditions on the $p\times p$ matrix T_p, converges in probability to a nonrandom d.f. $\bar{F}$. The specific conditions on T_p are the following:

1) T_p is symmetric positive definite a.s.
2) W_p and T_p are independent.
3) If G_p is the empirical d.f. of the eigenvalues of T_p, then for every positive integer k, $\int x^k dG_p(x)$ converges in L^2 to a nonrandom value H_k, where $\sum_{k=1}^{\infty}H_{2k}^{-1/2k}=\infty$.

The moments $\{E_k\}_{k=1}^{\infty}$ of $\bar{F}$ are also derived. They are given by

$$(1.2)\qquad E_k=\sum_{w=1}^{k}y^{k-w}\sum_{\substack{n_1+\cdots+n_w=k-w+1,\\ n_1+2n_2+\cdots+wn_w=k}}\frac{k!}{n_1!\cdots n_w!w!}H_1^{n_1}\cdots H_w^{n_w}.$$

No further information of $\bar{F}$ is given.

* Received by the editors March 4, 1983, and in revised form August 8, 1983. This research was supported by the National Science Foundation under grant MCS-8101703.

† Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27650.

In [5] it is shown the conditions are satisfied for $T_p = ((1/n)\ \underline{W}_p)^{-1}$ where $\underline{W}_p$ is $W(I,n)$, W_p and $\underline{W}_p$ are independent, and $p/n \to y' \in (0, 1/2)$ as $p \to \infty$. In particular, 3) is verified by showing

$$\int x^k dG_p(x) \overset{L^2}{\to} \int_{(1-\sqrt{y'})^2}^{(1+\sqrt{y'})^2} \frac{1}{x^k}\, dF_{y'}(x).$$

The matrix $((1/m)W_p)((1/n)\ \underline{W}_p)^{-1}$ is seen to be a multivariate F matrix, fundamental to statistical work in multivariate analysis.

In this paper we will derive the limiting empirical d.f. of $((1/m)W_p)((1/n)\ \underline{W}_p)^{-1}$. We will show for any $y' \in (0,1)$, if

$$H_k = \int_{(1-\sqrt{y'})^2}^{(1+\sqrt{y'})^2} \frac{1}{x^k}\, dF_{y'}(x), \qquad k = 1, 2, \cdots,$$

then $\{E_k\}_{k=1}^{\infty}$ are the moments of the d.f. $F_{y,y'}$, where for $0 < y \leqq 1$ $F_{y,y'}$ has density

$$f_{y,y'}(x)$$

$$= \begin{cases} \dfrac{(1-y')\sqrt{\left(x - \left(\dfrac{1-\sqrt{1-(1-y)(1-y')}}{1-y'}\right)^2\right)\left(\left(\dfrac{1+\sqrt{1-(1-y)(1-y')}}{1-y'}\right)^2 - x\right)}}{2\pi x(xy'+y)} \\ \qquad \text{for } \left(\dfrac{1-\sqrt{1-(1-y)(1-y')}}{1-y'}\right)^2 < x < \left(\dfrac{1+\sqrt{1-(1-y)(1-y')}}{1-y'}\right)^2, \\ 0 \qquad \text{otherwise.} \end{cases}$$

and for $1 < y < \infty$ $F_{y,y'}$ has mass $1 - 1/y$ at zero and density $f_{y,y'}$ on

$$\left(\left(\frac{1-\sqrt{1-(1-y)(1-y')}}{1-y'}\right)^2, \left(\frac{1+\sqrt{1-(1-y)(1-y')}}{1-y'}\right)^2\right).$$

The derivation of $F_{y,y'}$ will be handled in the next section by first evaluating a general expression for $E(e^{sX})$ $s \in \mathbb{C}$, where X is a random variable having moments $\{E_k\}$, and $\{H_k\}$ are the moments of a random variable Y having support on a closed interval on $\mathbb{R}^+$ bounded away from zero. This expression will be seen to involve an integral of a function in the complex plane depending on the generating function of the moments of Y^{-1}. Then $F_{y,y'}$ will be determined by evaluating the integral when Y^{-1} has d.f. $F_{y'}$.

2. Derivation of $F_{y,y'}$. Assume that $\{H_k\}$ are the moments of the random variable Y having support on $[a,b]$ with $0 < a < b < \infty$. Let $G(z) = E((1-zY)^{-1})$, $z \in \mathbb{C}$. Then G is analytic on $\mathbb{C} - [1/b, 1/a]$ and for $|z| < 1/b$, $G(z) = \sum_{k=0}^{\infty} H_k z^k$ $(H_0 \equiv 1)$. Let $G_I(z) = E((1-zY^{-1})^{-1})$. Then G_I is analytic on $\mathbb{C} - [a,b]$. Moreover, we have $G_I(z) = 1 - G(1/z)$, $z \in \mathbb{C} - [a,b]$.

Let X be a random variable having moments $\{E_k\}$ given by (1.2). We may ignore the question of whether $\{E_k\}$ are the moments of a random variable since the following steps will be reversible and we will wind up with $F_{y,y'}$, a proper probability d.f. Expanding $E(e^{sX})$, $s\in\mathbb{C}$, in a formal power series around $s=0$ we have

(1.3)

$$E(e^{sX})=\sum_{k=0}\frac{E_k s^k}{k!}=1+\sum_{k=1}^{\infty}s^k\sum_{w=1}^{k}\frac{y^{k-w}}{w!}\sum_{\substack{n_1+\cdots+n_w=k-w+1,\\ n_1+\cdots+wn_w=k}}\frac{H_1^{n_1}\cdots H_w^{n_w}}{n_1!\cdots n_w!}$$

$$=1+\sum_{k=1}^{\infty}s^k\sum_{w=1}^{k}\frac{y^{k-w}}{w!}\sum_{\substack{n_2+2n_3+\cdots+(w-1)n_w=w-1,\\ k-(2n_2+\cdots+wn_w)\geqq 0}}\frac{H_2^{n_2}\cdots H_w^{n_w}}{n_2!\cdots n_w!}$$

$$\cdot\frac{H_1^{(k-(2n_2+\cdots+wn_w))}}{(k-(2n_2+\cdots+wn_w))!}$$

$$=1-\frac{1}{y}+\frac{1}{y}e^{ysH_1}+\sum_{w=2}^{\infty}\frac{y^{-w}}{w!}\sum_{n_2+\cdots+(w-1)n_w=w-1}\frac{H_2^{n_2}\cdots H_w^{n_w}}{n_2!\cdots n_w!}$$

$$\sum_{k\geqq\max(w,2n_2+\cdots+wn_w)}\frac{(sy)^k H_1^{(k-(2n_2+\cdots+wn_w))}}{(k-(2n_2+\cdots+wn_w))!}.$$

Notice when $w\geqq 2$ and $n_2+\cdots+(w-1)n_w=w-1$, $2n_2+\cdots+wn_w\geqq w$. Therefore

(1.4)

$$E(e^{sX})=1-\frac{1}{y}+\frac{1}{y}e^{ysH_1}$$

$$+e^{ysH_1}\sum_{w=2}^{\infty}\frac{y^{-w}}{w!}\sum_{n_2+\cdots+(w-1)n_w=w-1}\frac{H_2^{n_2}\cdots H_w^{n_w}}{n_2!\cdots n_w!}(sy)^{2n_2+\cdots+wn_w}$$

$$=1-\frac{1}{y}+\frac{1}{y}e^{ysH_1}\sum_{n=0}^{\infty}\frac{s^n}{(n+1)!}\sum_{m_1+2m_2+\cdots+nm_n=n}\frac{(ysH_2)^{m_1}\cdots(ysH_{n+1})^{m_n}}{m_1!\cdots m_n!}.$$

Notice that

$$\sum_{m_1+\cdots nm_n=n}\frac{(ysH_2)^{m_1}\cdots(ysH_{n+1})^{m_n}}{m_1!\cdots m_n!},$$

defined to be 1 when $n=0$, is the coefficient of z^n in the series expansion about $z=0$ of $\exp(ys\Sigma_{k=1}^{\infty}H_{k+1}z^k)=\exp(ys((G(z)-1)/z-H_1))$. Note also that $1/(n+1)!$ is the coefficient of z^n in the expansion about $z=0$ of $(e^z-1)/z$. Both functions are analytic in a neighborhood of the origin, independent of y and s. Therefore we can write ([3, p. 158])

$$(1.5)\quad E(e^{sX})=1-\frac{1}{y}+\frac{1}{y2\pi i}e^{ysH_1}\oint_{|z|=r<1/b}\frac{(e^{s/z}-1)}{s/z}e^{ys((G(z)-1)/z-H_1)}\left(\frac{1}{z}\right)dz$$

$$=1-\frac{1}{y}+\frac{1}{sy2\pi i}\oint_{|z|=r<1/b}e^{s/z}e^{ys((G(z)-1)/z)}dz.$$

Making the substitution $z \to 1/z$ we have

$$E(e^{sX}) = 1 - \frac{1}{y} + \frac{1}{sy2\pi i}\oint_{|z|=r>b} e^{sz} e^{ysz(G(1/z)-1)} z^{-2}\,dz \tag{1.6}$$

$$= 1 - \frac{1}{y} + \frac{1}{sy2\pi i}\oint_{|z|=r>b} e^{sz - yszG_I(z)} z^{-2}\,dz.$$

Using integration by parts we have

$$E(e^{sX}) = 1 - \frac{1}{y} + \frac{1}{y2\pi i}\oint_{|z|=r>b} \frac{d}{dz}\left(z(1-yG_I(z))\right)e^{sz(1-yG_I(z))}\left(\frac{1}{z}\right)dz. \tag{1.7}$$

Provided

$$v = z(1 - yG_I(z)) \tag{1.8}$$

is invertible along $|z| = r$, we make the substitution (1.8) and arrive at

$$E(e^{sX}) = 1 - \frac{1}{y} + \frac{1}{y2\pi i}\oint_{|z(v)|=r>b} e^{sv}\frac{1}{z(v)}\,dv. \tag{1.9}$$

Since $G_I(z) \to 0$ as $|z| \to \infty$, for any $\delta \in (0,1)$ we have for all r sufficiently large

$$(1-\delta)|z| \leqq |v| \leqq (1+\delta)|z| \tag{1.10}$$

along the contour.

To derive $F_{y,y'}$, $0 < y' < 1$, we apply (1.9) to the case when Y^{-1} has density $f_{y'}$. Using the identity

$$\int_c^d \frac{\sqrt{(x-c)(d-x)}}{x}\,dx = \frac{\pi}{2}(\sqrt{d} - \sqrt{c})^2 \tag{1.11}$$

valid for $0 \leqq c < d$, it is straightforward to show, first for z real, $z > (1-\sqrt{y'})^{-2}$, and therefore for all $z \in \mathbb{C} - [(1+\sqrt{y'})^{-2}, (1-\sqrt{y'})^{-2}]$,

$$G_I(z) = \frac{1}{2\pi y'}\int_{(1-\sqrt{y'})^2}^{(1+\sqrt{y'})^2} \frac{1}{(1-xz)x}\sqrt{\left(x - (1-\sqrt{y'})^2\right)\left((1+\sqrt{y'})^2 - x\right)}\,dx \tag{1.12}$$

$$= \frac{1 - z(1-y') + (1-y')\sqrt{\left(z - (1+\sqrt{y'})^{-2}\right)\left(z - (1-\sqrt{y'})^{-2}\right)}}{2y'z}$$

where we will interpret all square roots of the form

$$\sqrt{(z-a_1)(z-a_2)}, \qquad a_1, a_2 \in \mathbb{R}, \quad a_1 < a_2 \tag{1.13}$$

to be positive on (a_2, ∞) and to vary continuously off this interval. Notice then, that the square root will be negative for $z \in (-\infty, a_1)$.

Solving for z in (1.8) we find

$$z=\frac{(2y'/y+(1-y'))v+1-y\pm\sqrt{(v(1-y')+(1-y))^2-4v}}{2(y'/y+1-y')} \tag{1.14}$$

$$=\frac{(2y'/y+(1-y'))v+1-y\pm(1-y')\sqrt{(v-b_1)(v-b_2)}}{2(y'/y+1-y')}$$

where

$$b_1=\left(\frac{1-\sqrt{1-(1-y)(1-y')}}{1-y'}\right)^2, \qquad b_2=\left(\frac{1+\sqrt{1-(1-y)(1-y')}}{1-y'}\right)^2.$$

Notice in (1.14) if the plus sign in front of the square root is used we would have $z\sim v$ for v large, whereas if the minus sign is used, then $z\sim(y'/y)/((y'/y)+(1-y'))$. Therefore, for r in (1.9) sufficiently large (1.8) is invertible along $|z|=r$ and we have

$$z(v)=\frac{(2y'/y+(1-y'))v+1-y+(1-y')\sqrt{(v-b_1)(v-b_2)}}{2(y'/y+1-y')} \tag{1.15}$$

and

$$\frac{1}{z(v)}=\frac{(2y'/y+(1-y'))v+1-y-(1-y')\sqrt{(v-b_1)(v-b_2)}}{2v(vy'/y+1)}. \tag{1.16}$$

Integrating $e^{sv}/z(v)$ along contours as in Fig. 1 when $y\neq 1$, and letting the two horizontal lines approach the real axis, we get (noting the discontinuity of the square root across $[b_1,b_2]$)

$$\begin{aligned} E(e^{sX})=1-\frac{1}{y}+\frac{1}{y2\pi i}\oint_{|z+y/y'|=r_1<y/y'} e^{sv}\frac{1}{z(v)}\,dv \\ +\frac{1}{y2\pi i}\oint_{|z|=r_2<\min(y/y',b_1)} e^{sv}\frac{1}{z(v)}\,dv \\ +\frac{1}{2\pi}\int_{b_1}^{b_2} e^{sx}\frac{(1-y')\sqrt{(x-b_1)(b_2-x)}}{x(xy'+y)}\,dx. \end{aligned} \tag{1.17}$$

For $y=1$ the limiting inner contour should not encompass the origin, and we will get (1.17) except the second integral will not appear.

We see that when $v=-y/y'$, the numerator of $1/z(v)$ is zero. Therefore the first integral in (1.17) vanishes. When $v=0$ the numerator of $1/z(v)$ is $2(1-y)$ when $0<y\leqq 1$, and is zero when $y>1$. Therefore, the term involving the second integral in (1.17) is $(1/y-1)I_{(0,1]}^{(y)}$, where I_A is the indicator function on the set A.

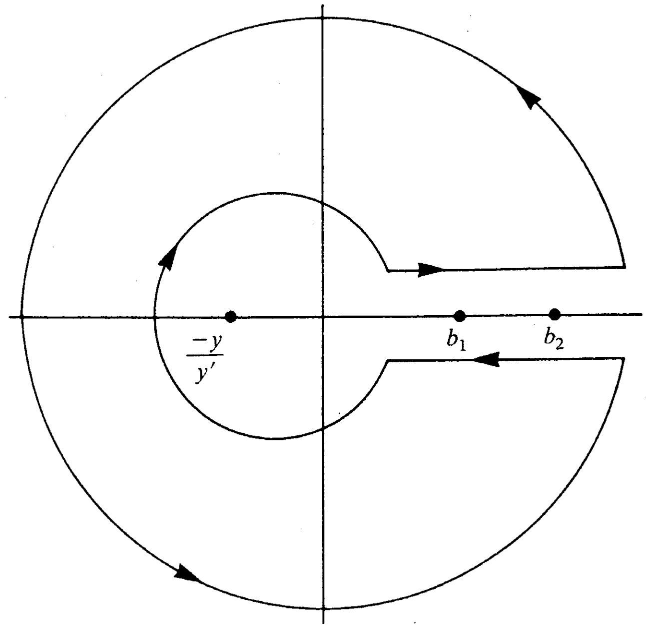

FIG. 1.

We therefore have

$$E(e^{sX}) = \left(1 - \frac{1}{y}\right) I_{(1,\infty)}^{(y)} + \int_{-\infty}^{\infty} e^{sx} f_{y,y'}^{(x)}\,dx. \tag{1.18}$$

Using the fact that $F_{y,y'}$ is a proper probability d.f. we conclude that (1.18) for $s = it$, $t \in \mathbb{R}$, is the characteristic function of the random variable X with d.f. $F_{y,y'}$, so that the d.f. of X must be $F_{y,y'}$.

REFERENCES

[1] U. GRENANDER AND J. W. SILVERSTEIN, *Spectral analysis of networks with random topologies*, SIAM J. Appl. Math, 32 (1977), pp. 499–519.

[2] D. JONSSON, *Some limit theorems for the eigenvalues of a sample covariance matrix*, J. Multivariate Anal., 12 (1982), pp. 1–38.

[3] E. C. TITCHMARSH, *The Theory of Functions*, 2nd ed., Oxford Univ. Press, London, 1939.

[4] K. W. WACHTER, *The strong limits of random matrix spectra for sample matrices of independent elements*, Ann. Probab., 6 (1978), pp. 1–18.

[5] Y. Q. YIN, Z. D. BAI AND P. R. KRISHNAIAH, *Limiting behavior of the eigenvalues of a multivariate F matrix*, J. Multivariate Anal., 13 (1983), pp. 508–516.

[6] Y. Q. YIN AND P. R. KRISHNAIAH, *A limit theorem for the eigenvalues of product of two random matrices*, J. Multivariate Anal., 13 (1983), pp. 489–507.