

THE EIGENVALUES OF RANDOM SYMMETRIC MATRICES

by

Z. FÜREDI and J. KOMLÓS

Mathematical Institute of the Hungarian Academy of Sciences
Budapest, Hungary H-1053

Received 13 December 1980

Let $A=(a_{ij})$ be an $n \times n$ matrix whose entries for $i \neq j$ are independent random variables and $a_{ji}=a_{ij}$. Suppose that every a_{ij} is bounded and for every $i > j$ we have $Ea_{ij}=\mu$, $D^2a_{ij}=\sigma^2$ and $Ea_{ii}=\nu$.

E. P. Wigner determined the asymptotic behavior of the eigenvalues of A (semi-circle law). In particular, for any $c > 2\sigma$ with probability $1 - o(1)$ all eigenvalues except for at most $o(n)$ lie in the interval $I = (-c\sqrt{n}, c\sqrt{n})$.

We show that with probability $1 - o(1)$ all eigenvalues belong to the above interval I if $\mu=0$, while in case $\mu > 0$ only the largest eigenvalue λ_1 is outside I , and

$$\lambda_1 = \frac{\sum_{i,j} a_{ij}}{n} + \frac{\sigma^2}{\mu} + O\left(\frac{1}{\sqrt{n}}\right)$$

i.e. λ_1 asymptotically has a normal distribution with expectation $(n-1)\mu + \nu + (\sigma^2/\mu)$ and variance $2\sigma^2$ (bounded variance!).

1. Introduction

E. P. Wigner published in 1955 his famous semi-circle law for the distribution of eigenvalues of random symmetric matrices (used in quantum mechanics). Here we recall the following generalization due to L. Arnold [1] (see also U. Grenander [3]):

Let $A=(a_{ij})$, $1 \leq i, j \leq n$, be an $n \times n$ symmetric matrix where the entries a_{ij} , $i \neq j$, are independent real-valued random variables. Furthermore, the a_{ij} with $i > j$ are required to have the same distribution function F , while all the a_{ii} possess the same distribution G . Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be the eigenvalues of A (all real by the symmetry of A). We also use the notation $\lambda_i(A)$. Further, $W_{A,n}$ or simply W_n denotes their empirical distribution function, i.e.

$$W_{A,n}(x) =: (\text{number of eigenvalues } \leq x)/n.$$

One is interested in the limiting behavior of the sequence W_n of random distribution functions as $n \rightarrow \infty$.

AMS subject classification (1980): 15 A 52, 15 A 18; 05 C 50

Semi-circle law. Assume that $\int |x|^k dF < \infty$, $\int |x|^k dG < \infty$, $k=1, 2, \dots$ and set $D^2 a_{ij} = \text{Var } a_{ij} = \sigma^2$. Then

$$(1) \quad \lim_n W_n(x2\sigma\sqrt{n}) = W(x)$$

in probability, where W is an absolutely continuous distribution function with density (semi-circle!)

$$W(x) = \begin{cases} \frac{2}{\pi}(1-x^2)^{1/2} & \text{for } |x| \leq 1 \\ 0 & \text{for } |x| > 1 \end{cases}$$

Since (1) gives only a limit distribution, it does not describe the behavior of the largest eigenvalues. This latter is done in the paper of F. Juhász [4]:

Theorem (Juhász [4]). We use the notations $\int |x|^k dF = \mu_k$, $\int |x|^k dG = \nu_k$, $\mu_1 = \mu$, $\nu_1 = \nu$. Under the conditions of the semi-circle law, in case $\mu > 0$ we have with probability approaching 1

$$(2) \quad \lambda_1(A) = \mu n + o(n).$$

For the other eigenvalues we have for any $\varepsilon > 0$

$$(3) \quad \max_{i \geq 2} |\lambda_i(A)| = o(n^{1/2+\varepsilon}).$$

If $\mu = 0$ then for any $\varepsilon > 0$

$$(4) \quad \max_{1 \leq i \leq n} |\lambda_i(A)| = o(n^{1/2+\varepsilon}).$$

The aim of the present paper is to sharpen the statements (2), (3), (4).

2. The results

Theorem 1. Let a_{ij} , $i \leq j$, be independent (not necessarily identically distributed) random variables bounded with a common bound K . Assume that for $i > j$, the a_{ij} have a common expectation μ and variance σ^2 , further that $Ea_{ii} = \nu$. Define a_{ij} for $i < j$ by $a_{ij} = a_{ji}$. (The numbers K , μ , σ^2 , ν will be kept fixed as n will tend to infinity.)

If $\mu > 0$ then the distribution of the largest eigenvalue of the random symmetric matrix $A = (a_{ij})$ can be approximated in order $1/\sqrt{n}$ by a normal distribution of expectation

$$(5) \quad (n-1)\mu + \nu + \sigma^2/\mu$$

and variance $2\sigma^2$. Further, with probability tending to 1,

$$(6) \quad \max_{i \geq 2} |\lambda_i(A)| < 2\sigma\sqrt{n} + O(n^{1/3} \log n).$$

That $\max_{i \geq 2} |\lambda_i(A)|$ cannot be much smaller than $2\sigma\sqrt{n}$, is guaranteed by the semi-circle law.

Note that the largest eigenvalue of the deterministic matrix $a_{ij} = \mu$ for $i \neq j$, $a_{ii} = \nu$, is equal to $(n-1)\mu + \nu$ (and all the other eigenvalues equal to $\nu - \mu$). Thus the fluctuation of the entries of the matrix changes λ_1 only with a normal random variable $N(\sigma^2/\mu, 2\sigma^2)$ of expectation and variance not growing with n .

Theorem 2. *Under the conditions of Theorem 1, in case $\mu=0$ we have*

$$(7) \quad \max_{1 \leq i \leq n} |\lambda_i(A)| = 2\sigma \sqrt{n} + O(n^{1/3} \log n) \quad (\text{with prob. tending to } 1).$$

Thus we see that though the semi-circle law allows $o(n)$ eigenvalues to be larger than $(2 + \varepsilon)\sigma \sqrt{n}$ (in fact they could be arbitrarily large), there is at most one large eigenvalue (in case $\mu \neq 0$), and then this large eigenvalue has a bounded variance (around the expectation $(n-1)\mu + v + \sigma^2/\mu$).

The theorems will be proved in the reverse order (7), (6), (5). The determination of the distribution of λ_1 will be based on two observations. The first one is that, as proved by Juhász, the largest eigenvalue is a great deal larger than the others, the second one is that the vector $\mathbf{1}=(1, 1, \dots, 1)$ is “nearly” an eigenvector corresponding to λ_1 . Thus, starting with the initial value $\mathbf{v}_0 = \mathbf{1}$, practically any numerical iteration method yields the eigenvalue λ_1 and the corresponding eigenvector in two-three steps with a high accuracy. Here we are going to use the von Mises iteration that will give the approximations $\sum S_i/n$ and $\sum S_i^2/\sum S_i$ for λ_1 in the first two steps (while for the eigenvector the approximations are $\mathbf{1}$ and \mathbf{S}), where S_i denotes the sum of the elements in the i -th row of A , ε id \mathbf{S} is the vector (S_1, \dots, S_n) .

The error of the first approximation is bounded and that of the second is of order $1/n$, more precisely we will have

$$(8) \quad P(|\lambda_1 - \sum S_i^2/\sum S_i| > x/n) < c_1/x^2 + 1/n.$$

The quantity $\sum S_i^2/\sum S_i$ is easily seen to have an asymptotically normal distribution with parameters of (5) (see Lemma 3). It will also be seen that the simpler quantity

$$\sum S_i/n + \sigma^2/\mu$$

approximates λ_1 in order $1/\sqrt{n}$.

3. Proof of Theorem 2

3.1. Just as the earlier papers, we will use the method of Wigner and calculate the moments of the eigenvalues. It is well-known that

$$\sum_{i=1}^n \lambda_i^k = \text{trace } A^k.$$

Thus

$$\begin{aligned} E_{n,k} &=: E \sum_{i=1}^n \lambda_i^k = E \text{trace } A^k = E \sum_{i_0=1}^n \sum_{i_1=1}^n \dots \sum_{i_k=1}^n a_{i_0 i_1} a_{i_1 i_2} \dots a_{i_{k-1} i_k} = \\ &= \sum_{i_0=1}^n \dots \sum_{i_k=1}^n E a_{i_0 i_1} a_{i_1 i_2} \dots a_{i_{k-1} i_k} = \\ &= \sum_{p=1}^{k+1} \left(\sum_{\substack{i_0=1 \\ |(i_0, i_1, \dots, i_k)|=p}}^n \dots \sum_{i_k=1}^n E a_{i_0 i_1} a_{i_1 i_2} \dots a_{i_{k-1} i_k} \right) = \sum_{p=1}^{k+1} E_{n,k,p} \end{aligned}$$

Now we are going to estimate $E_{n,k}$. Since $\mu=0$, if some a_{ij} ($i \neq j$) has multiplicity 1 in the product $a_{i_0 i_1} a_{i_1 i_2} \dots a_{i_{k-1} i_k}$ then the expectation of this product is zero. Thus it is sufficient to count those products in which every a_{ij} , $i \neq j$, occurs with multiplicity not equal to 1. Hence $E_{n,k,p}=0$ for $p > (k/2)+1$.

3.2. Let k be even. We show that

$$(9a) \quad E_{n,k,(k/2)+1} = \frac{1}{(k/2)+1} \binom{k}{k/2} n(n-1) \dots (n-(k/2)) \sigma^k,$$

and

$$(9b) \quad \sum_{p \leq k/2} E_{n,k,p} = O(k^6/n) E_{n,k,(k/2)+1}.$$

Consider the complete graph on vertex set $\{1, 2, \dots, n\}$ with a loop at every vertex. We assign to the edge (i, j) the random variable $a_{ij}=a_{ji}$, and to the walk (i_0, i_1, \dots, i_k) the product of the variables assigned to the edges of the walk. (We use the word “walk” here rather than “path” for we allow that vertices and edges occur in the sequence more than once.) Let us estimate, for given p , the sum of the expectations of all products with p different vertices (i.e. $E_{n,k,p}$). To this we need estimating the number of all walks of length k with p different vertices in which every edge (that is not a loop) occurs at least twice (perhaps once as (a, b) and once as (b, a)). We are going to assign a code to such a walk and count the number of possible codes.

Let us choose first the p vertices in order of occurrence in the sequence $I=(i_0, i_1, \dots, i_k)$. This can be done in $n(n-1) \dots (n-p+1)$ ways. Let us start now from i_0 and define a code consisting of the symbols “+”, “-” and “ (u, v) ”, where $1 \leq u, v \leq p$, as follows. We define a spanning rooted tree T (having of course $p-1$ edges):

$$(i_{\alpha-1}, i_{\alpha}) \in T \quad \text{if} \quad i_{\alpha} \notin \{i_0, i_1, \dots, i_{\alpha-1}\},$$

i.e. to the vertex $i_{\alpha} \in I$, $i_{\alpha} \neq i_0$, that edge $(i_{\alpha-1}, i_{\alpha}) \in T$ which first leads to i_{α} on the walk. Now during the walk we mark an edge e by “+” if $e \in T$ and we use this edge for the first time (e.g. the walk always starts with a + unless it starts with a loop); we mark the edge e with a “-” if $e \in T$ and we use the edge for the second time (regardless of direction), and finally we mark the (yet unmarked) edge $e=(x, u)$ by “ (u, v) ” if either $e \notin T$ or we have used e at least twice before; for defining v , we start from u and keep on going on the walk and the first time we use an edge (y, z) that does not belong to T , we take $v=y$, i.e. v is the place where we leave the edges of T . Since any two vertices u, v are connected on the tree T in a unique way, the mark (u, v) makes it well defined how we proceed when decoding and see a “-” mark.

We use altogether $p-1$ “+”, the same number of “-”, and $k-2p+2$ other marks, further the number of “+” marks is always at least as large in any first section of the code as that of the “-” marks. It can be checked easily that different walks will have different code-sequences and in case $p=(k/2)+1$ all codes (+ and - only in this case) belong to walks. The number of such \pm sequences is well-known to be equal to $\frac{1}{p} \binom{2p-2}{p-1}$ (see e.g. [5], Problem 1.33), thus the number of code sequences is at most

$$\binom{k}{k-2p+2} p^{2(k-2p+2)} \frac{1}{p} \binom{2p-2}{p-1}.$$

The expectation of a particular product $a_{i_0 i_1} \dots a_{i_{k-1} i_k}$ can be estimated

$$|E a_{i_0 i_1} \dots a_{i_{k-1} i_k}| \cong \sigma^{2p-2} K^{k-2p+2}$$

(since $E|a_{ij}|^k \cong K^{k-2}\sigma^2$ for $k \geq 3, i \neq j$), and hence

$$|E_{n,k,p}| \cong \sigma^{2p-2} K^{k-2p+2} n(n-1) \dots (n-p+1) \binom{k}{k-2p+2} p^{2(k-2p+2)} \frac{1}{p} \binom{2p-2}{p-1} =: S_{n,k,p}.$$

Further $E_{n,k,(k/2)+1}$ satisfies (9a). Now we have $S_{n,k,p-1} < S_{n,k,p} (K^2 k^6 / (4\sigma^2 n))$ whence

$$\sum_{p \leq k/2} |E_{n,k,p}| < E_{n,k,(k/2)+1} 2K^2 k^6 / (4\sigma^2 n)$$

if only $K^2 k^6 / (4\sigma^2 n) < 1/2$, e.g. for $k < (\sigma/K)^{1/3} n^{1/6}$.

3.3. Using Markov's inequality we get from (9a) and (9b)

$$\begin{aligned} P(\max |\lambda| > 2\sigma \sqrt{n} + v) &= P(\max |\lambda|^k > (2\sigma \sqrt{n} + v)^k) \cong \\ &\cong (2\sigma \sqrt{n} + v)^{-k} E \max |\lambda|^k < (2\sigma \sqrt{n} + v)^{-k} n^{(k/2)+1} 2^k \sigma^k = \\ &= \sqrt{n} \left(1 - \frac{v}{2\sigma \sqrt{n} + v} \right)^k < \sqrt{n} e^{-kv / (2\sigma \sqrt{n} + v)}. \end{aligned}$$

Set $k = (\sigma/K)^{1/3} n^{1/6}$ and $v = 50Kn^{1/3} \log n$, then

$$(10a) \quad P(\max |\lambda| > 2\sigma \sqrt{n} + 50Kn^{1/3} \log n) < 1/n^{10} \quad \text{for } n > n_0.$$

Now, for $n > n_0, E \sum \lambda_i^k > \frac{1}{2} E_{n,k,(k/2)+1}$, thus we get similarly

$$(10b) \quad P(\max |\lambda| < 2\sigma \sqrt{n} - 50Kn^{1/3} \log n) < 1/n^{10} \quad \text{for } n > n_0.$$

4. The proof of (6)

Lemma 1. *If $A = (a_{ij})$ is an $n \times n$ real symmetric matrix, and $B = A - tJ$ (where J is the matrix with all 1 entries) then*

$$\lambda_2(A) \cong \lambda_1(B).$$

Proof. $\lambda_1(A) = \max_{\|x\|=1} xAx$ and by the Courant—Fisher theorem (cf. [2])

$$(11) \quad \lambda_2(A) = \min_v \max_{\substack{(x,v)=0 \\ \|x\|=1}} xAx,$$

thus

$$\lambda_2(A) \cong \max_{\substack{(x,1)=0 \\ \|x\|=1}} xAx = \max_{\substack{(x,1)=0 \\ \|x\|=1}} x(B+tJ)x = \max_{\substack{(x,1)=0 \\ \|x\|=1}} xBx \cong \lambda_1(B),$$

since $(x, 1) = 0$ implies $Jx = 0$. ■

Similarly, for the smallest eigenvalue $\lambda_{-\infty}(A)$ of A we get

Lemma 2. *If $A=(a_{ij})$ is a real symmetric matrix, and $B=A-tJ$, $t>0$, then*

$$\lambda_{-\infty}(A) \cong \lambda_{-\infty}(B).$$

Proof. For $t>0$ the matrix tJ is positive definite (i.e. $\mathbf{x}tJ\mathbf{x} \cong 0$ for all $\mathbf{x} \in \mathbf{R}^n$), hence

$$\lambda_{-\infty}(A) = \min_{\|\mathbf{x}\|=1} \mathbf{x}A\mathbf{x} \cong \min_{\|\mathbf{x}\|=1} \mathbf{x}B\mathbf{x} + \min_{\|\mathbf{x}\|=1} \mathbf{x}tJ\mathbf{x} \cong \min_{\|\mathbf{x}\|=1} \mathbf{x}B\mathbf{x} = \lambda_{-\infty}(B). \blacksquare$$

Thus, (6) follows from Theorem 2 on substituting $t=\mu$. \blacksquare

5. The investigation of λ_1

5.1. Let us split the vector $\mathbf{1}$ into an eigenvector \mathbf{v} of λ_1 and a component orthogonal to \mathbf{v} :

$$(12) \quad \mathbf{1} = \mathbf{v} + \mathbf{r}, \quad (\mathbf{v}, \mathbf{r}) = 0, \quad A\mathbf{v} = \lambda_1\mathbf{v}.$$

First we will show that the “remaining term” \mathbf{r} is very small (short) with large probability ($\|\mathbf{r}\|^2 < 4\sigma^2/\mu^2$ cf. (19)).

Let S_i denote the sum of elements of A in the i -th row, $\mathbf{S}=(S_1, \dots, S_n)$ and $ES_i = n\mu - \mu + \nu =: L$. Applying the linear operation A to (12) we get

$$(13) \quad \mathbf{S} = A\mathbf{1} = A\mathbf{v} + A\mathbf{r} = \lambda_1\mathbf{v} + A\mathbf{r}.$$

Subtract $L\mathbf{1}$ from both sides of (13)

$$\mathbf{S} - L\mathbf{1} = (\lambda_1 - L)\mathbf{v} + (A\mathbf{r} - L\mathbf{r}).$$

Since both \mathbf{r} and $A\mathbf{r}$ (and all their linear combinations) are orthogonal to \mathbf{v} , applying the theorem of Pythagoras we find

$$(14) \quad \sum_{i=1}^n (S_i - L)^2 = \|\mathbf{S} - L\mathbf{1}\|^2 = (\lambda_1 - L)^2 \|\mathbf{v}\|^2 + \|A\mathbf{r} - L\mathbf{r}\|^2.$$

We will need a good approximation for the left-hand side of (14).

Lemma 3. *Under the conditions and notations of Theorem 1 we have*

$$(15) \quad P\left(\left|\sum_i (S_i - L)^2 - n^2\sigma^2\right| > 2\sigma^2 n^{3/2}x\right) < 1/x^2,$$

$$(16) \quad P\left(\left|\sum_i \sum_j a_{ij}(S_i - L)(S_j - L)\right| > n^2(K^3 + x)\right) < 1/x^2,$$

$$(17) \quad E\left(\sum S_i/n\right) = (n-1)\mu + \nu = L, \quad D^2\left(\sum S_i/n\right) = 2\sigma^2 + (\sigma_v^2 - 2\sigma^2)/n,$$

where σ_v^2 is the common variance of the variables a_{ii} . According to the central limit theorem, $\sum S_i/n$ is actually very nearly normal.

$$(18) \quad P\left(\left|\frac{\sum S_i^2}{\sum S_i} - \frac{\sum S_i}{n} - \frac{\sigma^2}{\mu}\right| > 3\sigma^2 x/\sqrt{n}\right) < \frac{1}{x^2}.$$

Lemma 3 will be proved in § 6 by elementary probability means. Returning to (14), the relation $(\mathbf{r}, \mathbf{v})=0$, the Courant—Fisher theorem (11) and Theorem 2 (formula (10a)) imply that $\|\mathbf{Ar}\| \cong \lambda_2(A)\|\mathbf{r}\| < 3\sigma\sqrt{n}\|\mathbf{r}\|$ with very large probability. Thus $\mathbf{Ar}-L\mathbf{r}$ is at least $(L-3\sigma\sqrt{n})\|\mathbf{r}\|$ in length, whence

$$(19) \quad \|\mathbf{r}\|^2 < \sum (S_i-L)^2/(L-\lambda_2)^2 < 2n^2\sigma^2/(\mu n-3\sigma\sqrt{n})^2 < 4\sigma^2/\mu^2$$

with probability $\cong 1-\frac{1}{n}$.

5.2. Now we apply the von Mises iteration method (cf. [6]) for determining the eigenvalue λ_1 .

$$(20) \quad \sum S_i^2/\sum S_i = (\mathbf{S}, \mathbf{S})/(\mathbf{1}, \mathbf{S}) = \|\mathbf{A1}\|^2/(\mathbf{1A1}) = \|\lambda_1\mathbf{v} + \mathbf{Ar}\|^2/(\mathbf{v} + \mathbf{r}, \lambda_1\mathbf{v} + \mathbf{Ar}) = \\ = (\lambda_1^2\|\mathbf{v}\|^2 + \|\mathbf{Ar}\|^2)/(\lambda_1\|\mathbf{v}\|^2 + \mathbf{rAr}) = \lambda_1 + (\|\mathbf{Ar}\|^2 - \lambda_1\mathbf{rAr})/\sum S_i.$$

The remaining term on the right of (20) will be estimated using (17) and (19). Since $\|\mathbf{Ar}\|^2 < 9\sigma^2n\|\mathbf{r}\|^2$ (with probability $> 1-\frac{1}{n}$), we get from (19)

$$(21) \quad \|\mathbf{Ar}\|^2 < (4\sigma^2/\mu^2)9\sigma^2n < (50\sigma^4/\mu^2)n \\ \left(\text{w. pr. } 1-\frac{1}{n}\right).$$

Now $|\mathbf{rAr}| \cong \|\mathbf{r}\|\|\mathbf{Ar}\|$ together with (19) and (21) give

$$(22) \quad |\mathbf{rAr}| < (20\sigma^3/\mu^2)\sqrt{n} \\ \left(\text{w. pr. } 1-\frac{1}{n}\right).$$

λ_1 can always be estimated from above by $\max_i \sum_j |a_{ij}| < Kn$, thus (21), (22) and (17) show that the remaining term on the right of (20) is of the order $O(1/\sqrt{n})$, more precisely

$$(23) \quad P\left(\left|\frac{\sum S_i^2}{\sum S_i} - \lambda_1\right| > c/\sqrt{n}\right) < 1/n \\ (\text{where } c < 50K\sigma^3/\mu^3).$$

5.3. Actually the above remaining term is of order $O(1/n)$, which gives (8). This will be proved by improving on (22).

$$(24) \quad \mathbf{rAr} \text{ is bounded in probability (cf. (24a)).}$$

Let us expand the sum on the left of (16).

$$(25) \quad \sum_i \sum_j a_{ij}(S_i-L)(S_j-L) = (\mathbf{S}-L\mathbf{1}, A(\mathbf{S}-L\mathbf{1})) = (\mathbf{A1}-L\mathbf{1}, A(\mathbf{A1}-L\mathbf{1})) = \\ = ((\lambda_1-L)\mathbf{v} + \mathbf{Ar} - L\mathbf{r}, \lambda_1(\lambda_1-L)\mathbf{v} + A^2\mathbf{r} - L\mathbf{Ar}) = \\ = \lambda_1(\lambda_1-L)^2\|\mathbf{v}\|^2 + \mathbf{rA}^3\mathbf{r} - 2L\|\mathbf{Ar}\|^2 + L^2\mathbf{rAr}.$$

We show that the left hand side of (25), the first three terms on its right and L^2 are all of the order $O(n^2)$ (with large probability), thus \mathbf{rAr} is bounded in probability, i.e.

$$(24a) \quad P(|\mathbf{rAr}| > C_2 + x) < C_3/x^2 + O(1/n).$$

Indeed, the left hand side of (25) can be estimated by (16). In the first term of the right hand side $|\lambda_1| \leq Kn$, $\|\mathbf{v}\|^2 \leq n$ and the boundedness of $\lambda_1 - L$ follows from (18) and (23). Further, $|\mathbf{rA}^3\mathbf{r}| \leq |\lambda_2|^3 \|\mathbf{r}\|^2 < (3\sigma \sqrt{n})^3 (4\sigma^2/\mu^2) = O(n^{3/2})$ according to (19) and Theorem 2. Finally, by (21),

$$2L\|\mathbf{Ar}\|^2 < 2L(50\sigma^4/\mu^2)n = O(n^2).$$

Thus we established (24a), and hence, through (20), also (8).

6. The proof of Lemma 3

6.1. Denote $b_{ij} = a_{ij} - Ea_{ij}$ and Σ_{15} and Σ_{16} the sums in (15) resp. (16). Since $\Sigma_{15} = \sum_i (S_i - L)^2 = \sum_i (\sum_j b_{ij})^2$, thus

$$E\Sigma_{15} = \sum_i \sum_j \sum_k Eb_{ij}b_{ik} = \sum_{i \neq j} \sum_k Eb_{ij}b_{ik} + \sum_i \sum_k Eb_{ik}^2 = (n^2 - n)\sigma^2 + n\sigma_v^2.$$

Further,

$$\begin{aligned} D^2\Sigma_{15} &= E\left[\sum_i \sum_j \sum_k b_{ik}b_{jk} - (n^2 - n)\sigma^2 - n\sigma_v^2\right] = \\ &= E\sum_i \sum_j \sum_k \sum_I \sum_J \sum_K b_{ik}b_{jk}b_{IK}b_{JK} - ((n^2 - n)\sigma^2 + n\sigma_v^2)^2. \end{aligned}$$

Now the expectation of $b_{ik}b_{jk}b_{IK}b_{JK}$ is non-zero only if every factor in it has a multiplicity ≥ 2 . Thus

$$\begin{aligned} D^2\Sigma_{15} &= E\sum_i \sum_k \sum_I \sum_K b_{ik}^2 b_{IK}^2 - ((n^2 - n)\sigma^2 + n\sigma_v^2)^2 + \\ &+ 2E\sum_k \sum_{\substack{i \\ i \neq j}} \sum_j b_{ik}^2 b_{jk}^2 - \sum_i \sum_k b_{ik}^4 = O(\sigma^4 n^3). \end{aligned}$$

(15) follows from Chebyshev's inequality.

6.2. (16) will be proved similarly.

$$E\Sigma_{16} = \sum_i \sum_j a_{ij} \sum_k b_{ik}b_{jk} = \sum_i \sum_j \sum_k Ea_{ij}b_{ik}b_{jk} = n^2 v\sigma^2 + nEb_{ii}^3,$$

and

$$(26) \quad D^2\Sigma_{16} = \sum_i \sum_j \sum_k \sum_I \sum_J \sum_K Ea_{ij}b_{ik}b_{jk}a_{IJ}b_{IK}b_{JK} - (n^2 v\sigma^2 + nEb_{ii}^3)^2.$$

The expectation of such a 6-product is non-zero only if there are more coincidences among the indices, i.e. if $|\{i, j, k, I, J, K\}| \leq 4$. Thus, among the n^6 terms in (26) only n^4 is non-zero, hence

$$D^2\Sigma_{16} = O(n^4).$$

(Actually, it is $O(n^3)$, but we do not need that.)

6.3. The proof of (17) is trivial. For (18) we use the following identity.

$$(27) \quad \frac{\sum S_i^2}{\sum S_i} - \frac{\sum S_i}{n} = \frac{\sum (S_i - L)^2}{\sum S_i} - \frac{((\sum S_i/n) - L)^2}{\sum S_i/n}.$$

The second term on the right of (27) is, by (17), $O(1/n)$. The first term has a numerator $n^2\sigma^2 + O(n^{3/2})$ according to (15), while its denominator is $\mu n^2 + O(n)$, hence the ratio equals $\sigma^2/\mu + O(1/\sqrt{n})$. ■

References

- [1] L. ARNOLD, On the asymptotic distribution of the eigenvalues of random matrices, *J. Math. Analysis and Appl.* **20** (1967), 262—268.
- [2] F. R. GRANTMACHER, *Applications of the theory of matrices*, Intersciences, 1959.
- [3] U. GRENANDER, *Probabilities on algebraic structures*, Almquist and Wiksell, Stockholm 1963, 178—180.
- [4] F. JUHÁSZ, On the spectrum of a random graph, in: *Algebraic methods in graph theory* (Lovász et al., eds), *Coll. Math. Soc. J. Bolyai* **25**, North-Holland 1981, 313—316.
- [5] L. LOVÁSZ, *Combinatorial problems and exercises*, Akadémiai Kiadó—North-Holland, Budapest—Amsterdam, 1979.
- [6] A. RALSTON, *A first course in numerical analysis*, McGraw-Hill, 1965.
- [7] A. RÉNYI, *Foundations of probability*, Holden-Day, San Francisco, 1970.
- [8] E. P. WIGNER, Characteristic vectors of bordered matrices with infinite dimensions, *Ann. Math.* **62** (1955), 548—564.
- [9] E. P. WIGNER, On the distribution of the roots of certain symmetric matrices, *Ann. Math.* **67** (1958), 325—327.