Distribution Of Subdominant Eigenvalues of Random Matrices

G. Goldberg
Programming Recourses Company
Hartford, Connecticut 06105-1990

P. Okunev
Department of Mathematics
University of Connecticut
Storrs, Connecticut 06269–3009

M. Neumann
Department of Mathematics
University of Connecticut
Storrs, Connecticut 06269–3009

H. Schneider
Department of Mathematics
University of Wisconsin
Madison, Wisconsin 53706

Abstract

We mainly investigate the behavior of the subdominant eigenvalue of matrices $B = (b_{i,j}) \in \mathbb{R}^{n,n}$ whose entries are independent random variables with an expectation $E(b_{i,j}) = 1/n$ and with a variance $\mathrm{Var}(b_{i,j}) \leq c/n^2$ for some constant $c \geq 0$. For such matrices we show that for large n, the subdominant eigenvalue is, with great probability, in a small neighborhood of 0. Actually, we also show that for large n, the spectral radius of such matrices is, with great probability, in a small neighborhood of 1.

1 INTRODUCTION AND MAIN RESULT

Let $K \in \mathbb{R}^{n,n}$, the space of all real $n \times n$ matrices, and denote the spectral radius of K by $\rho(K)$. Let $(\lambda_1, \ldots, \lambda_n)$ be any arrangement of the eigenvalues of K in which $|\lambda_1| = \rho(K)$ Then a *subdominant eigenvalue* of K is any eigenvalue μ of K for which

$$|\mu| = \max_{2 \leq i \leq n} |\lambda_i|.$$

In linear iterative methods in which the powers of the iteration matrix converge, but to a nonzero limit, so that necessarily the spectral radius of the iteration matrix is 1, it is well known the *magnitude of a subdominant eigenvalue determines the asymptotic rate of convergence of the process*, see, for instance, Berman and Plemmons [3, p.199]. An important example of an application of such iterative methods occurs in the problem of finding the stationary distribution vector of a finite homogeneous Markov chain by iteration. Because it is this type of an application that has served in part as the motivation of our present study, we now describe this application in more detail. Suppose that $P = (p_{i,j})$ is a (row stochastic) transition matrix for a finite ergodic homogeneous Markov process on n states and let $\gamma(P)$ be the magnitude of a subdominant eigenvalue of P. Let e be the n–vector of all 1's and let v be the stationary distribution vector for the chain, in which case $v^T P = v^T$ and $v^T e = 1$. In Seneta [16, p.9], it is shown that if $\gamma(P) \neq 0$, then, as $k \to \infty$,

$$P^k \;=\; ev^T + O\left(k^s \gamma^k(T)\right),$$

where s is one less than the largest multiplicity of any subdominant eigenvalue of P.

To avoid having to compute $\gamma(P)$, various estimates have been developed in the literature. One estimate, due to Ostrowski [14] is that

$$\gamma(P) \;\leq\; \frac{M-m}{M+m},$$

where $M = \max_{1 \leq i,j \leq n} p_{i,j}$ and $m = \min_{1 \leq i,j \leq n} p_{i,j}$. A second estimate is that

$$\gamma(P) \;\leq\; \frac{1}{2} \max_{1 \leq i,j \leq n} \sum_{s=1}^{n} |p_{i,s} - p_{j,s}|. \tag{1.1}$$

This estimate was found by Bauer, Deutsch, and Zenger [2]. Actually, the right hand side of (1.1) is also a special case of a *coefficient of ergodicity* of P which can be found in Dobrushin [5, p.335] and which was used and studied by many others, see, for example Paz [15] and Seneta [17]. Specifically, let ν be a norm on $\mathbb{R}^n$. Then the *coefficient of ergodicity of P with respect to ν* is given by

$$\nu(P) \;=\; \max_{\nu(x)=1,\ x^T e=1} \nu(x^T P).$$

It is known, see Tan [19], that $\nu(P) \geq \gamma(P)$ and that, in the case when ν is the 1–vector norm, $\nu(P)$ is equal to the right hand side of (1.1).

As a more immediate motivation for the numerical experimentation which we carried out and which lead subsequently to the investigation in this paper we mention the problem of the *weak ergodicity* of an infinite product of of stochastic matrices, see Neumann and Schneider [13]. This is the case when the infinite product of stochastic matrices may not converge, but approaches a rank 1 matrix (see also Seneta [16, Definition 3.3]).

For background material on the statistical concepts used in the paper see Feller [9] and for background material concerning nonnegative matrices and applications to Markov chains see Berman and Plemmons [3] and Campbell and Meyer [4].

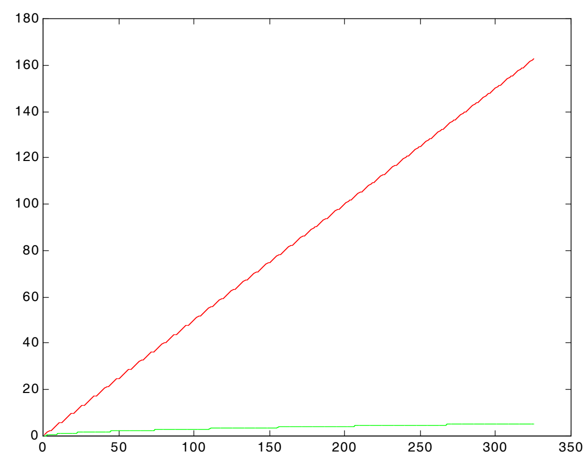
In this paper we shall mainly prove results concerning the distribution of the subdominant eigenvalues of $n \times n$ matrices $B = (b_{i,j})$ whose entries are independent random variables from *any distribution*, provided that the entries have an expectation $E(b_{i,j}) = 1/n$ and a variance bounded by c/n^2, for some constant $c \geq 0$. Our results here were motivated by numerical experiments which we carried out before hand and in which the elements of the matrices B were mutually independent and uniformly distributed in the interval $[0,1]$ and by unpublished results of Friedman [8] from 1983 concerning the distribution of the singular values of such matrices. Friedman stated that the singular values of B other than $\|B\|_2$ are uniformly distributed in the interval $[0, \|B\|_2/\sqrt{n}\ ]$. We comment that our approach to the proof of our main result, Theorem 1.1, is quite different from the approach taken in Friedman notes and that our results are not implied by his since, in general and even for stochastic matrices, the second largest singular value does not necessarily dominate the subdominant eigenvalues of a matrix.

We comment that another, different, direction of investigation of the behavior of eigenvalues of random matrices has lead to the so called *empirical spectral distribution* and its consequence, the so called *circular law*. Briefly they are as follow. Consider the $n \times n$ matrix Ξ_n whose entries $\xi_{k,j} = (1/\sqrt{n})x_{k,j}$, where the $x_{k,j}$, $k,j \geq 1$, form an infinite double array of independent, randomly distributed, complex random variables of mean 0 and variance 1. One then uses the eigenvalues $\lambda_1, \ldots, \lambda_n$ of Ξ_n to construct the two dimensional empirical distribution given by:

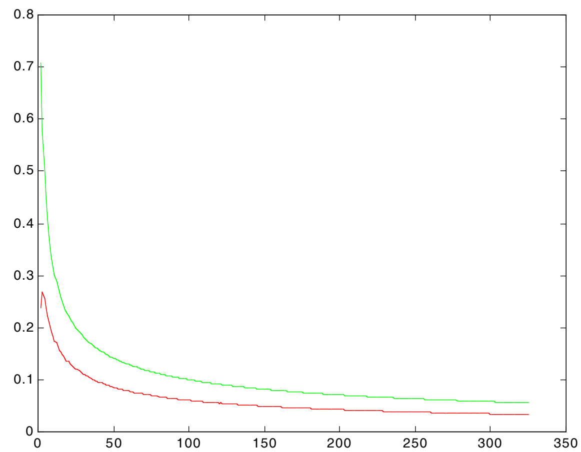
$$\mu_n = \frac{1}{n}\# \left\{ i \leq n \mid \Re(\lambda_k) \leq x, \Im(\lambda_k) \leq y \right\}.$$

In Bai [1] the following result (the circular law) is proved: *Suppose that the* *entries of X have a finite sixth moment and that the joint distribution of the real and imaginary part of the entries has a bounded density. Then, with probability* 1*, the empirical distribution $\mu_n(x, y)$ tends to the uniform distribution over the unit disc in two–dimensional space.* The paper [1] and the working manuscript [7] by Edelman describe applications of large scale random matrices in physics to quantum mechanics and in other disciplines. We also refer the reader to the list [6] of some 200 papers on random matrices and their applications compiled by Edelman which is available on the web.

Let us begin by presenting some of our numerical results. We used the *random* generator facility in MATLAB which supplies a random number from $(0, 1)$ to generate $n \times n$ matrices of sizes $n = 1, \ldots, 325$. For each n in this range, we generated 300 examples and averaged the values of the subdominant eigenvalue as well as the spectral radius. Our first illustration gives a plot of both the average subdominant eigenvalue (the dotted line) and the average of the spectral radius (the more solid line) as functions of n:



The graph clearly shows that we may expect the subdominant eigenvalue to be very small compared to the size of the matrix. Even more interesting is the next graph:



In this plot the bottom curve is the ratio of the subdominant eigenvalue to the spectral radius in the graph above. The curve at the top is the plot of the function $1/\sqrt{n}$. This paper will be devoted to proving **some** of the phenomena which we observe in these graphs.

Let $M = (m_{i,j})$ be an $n \times n$ matrix with mutually independent elements uniformly distributed in the interval $[0, 1]$. One of the goals of this paper will be to investigate the distribution of the subdominant eigenvalues of the matrix

$$B = (b_{i,j}) \;=\; \frac{M}{n/2}.$$

Using Feller [9, p.5], we can easily find that the expected of each entry of B is equal to $1/n$, while the variance of each entry of B is equal to $1/(3n^2)$ Let us now rewrite the entries of B as

$$b_{i,j} \;=\; \frac{1}{n} + a_{i,j}, \quad 1 \leq i, j \leq n.$$

Then, obviously, $E(a_{i,j}) = 0$ and $\mathrm{Var}(a_{i,j}) = \frac{1}{3n^2}$. *Note that the $a_{i,j}$'s are not necessarily nonnegative. Furthermore, as the elements of M are mutually independent, so are the elements of the matrix $A = (a_{i,j})$.*

We are now ready to state the main result of this paper. *Note that our result does not actually require the restriction that the entries of B come specifically from a uniform distribution nor do they require that the entries of B are, themselves, nonnegative!*

THEOREM 1.1 *Let $0 < \epsilon < 1$ and $0 \leq p < 1$. Suppose that $B = \left(\frac{1}{n} + a_{i,j}\right)$ is an $n \times n$ matrix whose elements are independent random variables with $E(a_{i,j}) = 0$ and $\mathrm{Var}(a_{i,j}) \leq \frac{c}{n^2}$, for some nonnegative constant c, $i, j = 1, \ldots, n$. The the following hold:*

(i) *There is an integer $N(\epsilon, p)$ such that for any $n \geq N(\epsilon, p)$, with probability of at least p, $n-1$ of the eigenvalues of B are in an open disc of radius ϵ centered at the origin.*

(ii) *There is an integer $N_1(\epsilon, p)$ such that for any $n \geq N_1(\epsilon, p)$, $P(\mid \rho(B) - 1 \mid < \epsilon) \geq p$, where $\rho(B)$ is the spectral radius of B.*

We shall devote the next section to the proof of the Theorem 1.1. The conclusion of the theorem is a consequence of several results which are of independent interest. A main idea in the proof is to split the characteristic polynomial $p_B(\lambda)$ of B into two parts: the *principal part* which equals $\lambda^n - \lambda^{n-1}$ and which is also shown earlier to be the expectation of $p_B(\lambda)$ and the *remainder* $g_B(\lambda) = p_B(\lambda) - (\lambda^n - \lambda^{n-1})$. We then use *(i)* the reverse case of Chebyshev's inequality (which says that if X is a random variable, then $P(|X| < \epsilon) \geq 1 - E(X^2)/\epsilon^2$, see, for example, Manoukian [11, p.11, (iv)–(v)]), *(ii)* Rouché's theorem (which says that if f and h are analytic functions in a domain containing the track and the interior of a closed Jordan contour γ and $|h(z)| < |f(z)|$ on γ, then f and $f + h$ have same number of zeros inside γ, see, for example, Tall [18, p.38]), and *(iii)* a sequence of estimations on the expected values of squares of sums of determinants to show that as $n \to \infty$, with great probability, the characteristic polynomial of B has in any disc of radius $\epsilon \neq 1$ as many roots as the polynomial $\lambda^n - \lambda^{n-1}$. From this it follows that for n large enough, with great probability, all the eigenvalues of B with the exception of spectral radius are in a small neighborhood of 0, whereas the spectral radius of B is in a small neighborhood of 1.

2 PROOF OF THEOREM 1.1

As mentioned in the introduction, the proof of the Theorem 1.1 is a consequence of auxiliary results, some of which are of interest in their own right. As several of these results will require the same set of assumptions, let us formulate the common assumptions now and refer to them in the section when they are required simultaneously:

ASSUMPTIONS 2.1 *The entries of the real matrix $A = (a_{i,j})$ are independent random variables such that $E(a_{i,j}) = 0$ and such that $E\left(a_{i,j}^2\right) \leq \frac{c}{n^2}$.*

We begin with the following simple lemma concerning the expected value of the determinant. *No originality* is claimed for this lemma and it could be deduced from works of Mehta [12] and Girko [10], but as its latter part serves as a motivation for the splitting of the characteristic polynomial of B into a sum of the principal part and the remainder, which are concepts mentioned in the introduction and which will be precisely defined later, it is presented here for the sake of completeness.

LEMMA 2.2 *Let $B \in \mathrm{I\!R}^{\mathrm{n,n}}$ be matrix whose elements are independent random variables with finite expectation. Then*

$$E\left(\det(B)\right) \;=\; \det\left(E(B)\right).$$

In particular, if $B = (b_{i,j}) \in \mathrm{I\!R}^{\mathrm{n,n}}$ is a matrix whose elements are independent random variables with $E(b_{i,j}) = \frac{1}{n}$, then the expectation of the characteristic polynomial is $\lambda^n - \lambda^{n-1}$.

<u>**Proof**</u>: Using the independence of the entries of B we have the following sequence of equalities:

$$E(\det(B)) \;=\; E\left(\sum_{\sigma\in\mathcal{S}_n} \mathrm{sign}(\sigma) b_{1,\sigma(1)}\dots b_{n,\sigma(n)}\right)$$

$$=\; \sum_{\sigma\in S_n} E\left(\mathrm{sign}(\sigma) b_{1,\sigma(1)}\dots b_{n,\sigma(n)}\right)$$

$$=\; \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) E(b_{1,\sigma(1)})\dots E\left(b_{m,\sigma(n)}\right) \;=\; \det\left(E(B)\right).$$

The second part of the lemma follows as now $E(B)$ is a rank 1 matrix whose only nonzero eigenvalue is its constant row sum, namely, 1. □

In some of our estimates on the maximum of the remainder of the characteristic polynomial takes on the boundary of discs of radius ϵ we shall also require the following estimate:

LEMMA 2.3 *Let A be a $k \times k$ matrix whose elements are independent random variables such that $E\left(a_{i,j}\right) = 0$ and $E\left(a_{i,j}^2\right) \le \frac{c}{n^2}$. Then*

$$E\left(\det{}^2(A)\right) \le \frac{c^k k!}{n^{2k}}.$$

Proof: We can write that

$$
\begin{aligned}
&E((\det(A))^2)\\
&= E\left(\left(\sum_{\sigma\in S_k}\operatorname{sign}(\sigma)a_{1,\sigma(1)}\dots a_{k,\sigma(k)}\right)\left(\sum_{\tau\in S_k}\operatorname{sign}(\tau)a_{1\tau(1)}\dots a_{k\tau(k)}\right)\right)\\
&= E\left(\sum_{\sigma\in S_k,\tau\in S_k}\operatorname{sign}(\sigma)a_{1,\sigma(1)}\dots a_{k,\sigma(k)}\operatorname{sign}(\tau)a_{1,\tau(1)}\dots a_{k,\tau(k)}\right)\\
&= E\left(\sum_{\sigma\in S_k,\tau\in S_k}\operatorname{sign}(\sigma)\operatorname{sign}(\tau)\left(a_{1,\sigma(1)}a_{1,\tau(1)}\right)\dots\left(a_{k,\sigma(k)}a_{k,\tau(k)}\right)\right)\\
&= \sum_{\sigma\in S_k,\tau\in S_k}E(\operatorname{sign}(\sigma)\operatorname{sign}(\tau)\left(a_{1,\sigma(1)}a_{1,\tau(1)}\right)\dots\left(a_{k,\sigma(k)}a_{k,\tau(k)}\right)\\
&= \sum_{\sigma\in S_k,\tau\in S_k}\operatorname{sign}(\sigma)\operatorname{sign}(\tau)E\left(a_{1\sigma,(1)}a_{1,\tau(1)}\right)\dots E\left(a_{k,\sigma(k)}a_{k,\tau(k)}\right).
\end{aligned}
\tag{2.1}
$$

The last line in (2.1) follows because all the expressions appearing in parenthesis in the line preceding it are mutually independent since they are made up from the elements of different rows of A. Now, if for some i, $\sigma(i)\neq\tau(i)$ then because elements of A are independent we have that

$$E\left(a_{i,\sigma(i)}a_{i,\tau(i)}\right) \;=\; E\left(a_{i,\sigma(i)}\right)E\left(a_{i,\tau(i)}\right) \;=\; 0$$

and therefore

$$\operatorname{sign}(\sigma)\operatorname{sign}(\tau)E\left(a_{1,\sigma(1)}a_{1,\tau(1)}\right)\dots E\left(a_{k,\sigma(k)}a_{k,\tau(k)}\right) \;=\; 0.$$

From this and equation (2.1) it follows that

$$E\left(\det(A)^2\right) \;=\; \sum_{\sigma\in S_k}\operatorname{sign}(\sigma)^2E\left(a^2_{1,\sigma(1)}\right)\dots E\left(a^2_{k,\sigma(k)}\right).$$

Finally, since $E\left(a^2_{i,j}\right)\leq\frac{c}{n^2}$, we have that

$$E\left(\det{}^2(A)\right)\leq\frac{c^kk!}{n^{2k}}.$$

This concludes the proof. □

A further auxiliary result which we shall need is the following lemma:

LEMMA 2.4 *Let A be an $n \times n$ matrix whose elements are independent random variables such that $E(a_{i,j}) = 0$. Let X be an $\ell \times \ell$ submatrix of A and let Y be a $k \times k$ submatrix of A. If $X \neq Y$, then*

$$E(\det(X)\det(Y)) \;=\; 0.$$

__**Proof**__: Since $X \neq Y$ their respective elements cannot come from exactly the same rows and exactly the same columns of M. Suppose that X contains elements of ith row of A and Y contains no elements of that row. Now $\det(X)$ is a sum of $\ell!$ numbers, each of which is up to a sign to a product of ℓ elements of X. Similarly, $\det(Y)$ is a sum of $k!$ numbers, each of which is up to a sign a product of k elements of Y. Thus $\det(X)\det(Y)$ is a sum of $\ell! \times k!$ numbers each being equal, up to a sign, a product of ℓ elements of X and k elements of Y. To complete the proof of the lemma we only need to show that expectation of every such product is 0. Let

$$\prod_{p=1}^{\ell} a_{j_p,s_p} \prod_{t=1}^{k} a_{q_t,r_t}$$

be such a product. This product contains exactly one element of i–th row of A. Suppose, without loss of generality, that a_{j_1,s_1} is an element of the i–th row of A. Then a_{j_1,s_1} is independent from

$$\prod_{p=2}^{\ell} a_{j_p,s_p} \prod_{t=1}^{k} a_{q_t,r_t}$$

(since elements of A are mutually independent) and therefore

$$E\left(\prod_{p=1}^{\ell} a_{j_p,s_p} \prod_{t=1}^{k} a_{q_t,r_t}\right)$$

$$= \underbrace{E\left(a_{j_1,s_1}\right)}_{=\,0} E\left(\prod_{p=2}^{\ell} a_{j_p,s_p} \prod_{t=1}^{k} a_{q_t,r_t}\right) \;=\; 0.$$

□

For our analysis of the behavior of the remainder of the characteristic polynomial, the following definition will be helpful:

DEFINITION 2.5 *Let $\mathcal{S}_k$ be the set of all subsets of $\{1, \ldots, n\}$ of cardinality k. Suppose that $A = (a_{i,j}) \in \mathbb{R}^{n,n}$. For $L \in \mathcal{S}_k$, let G_L be the $n \times n$*

matrix defined by

$$g_{i,j} \;=\; \begin{cases} -a_{i,j}, & \textit{if } j \in L \\ -\frac{1}{n}, & \textit{if } j \notin L \textit{ and } i \neq j \\ \lambda - \frac{1}{n}, & \textit{if } j \notin L \textit{ and } i = j. \end{cases} \tag{2.2}$$

Now let $L \in \mathcal{S}_k$ be fixed. Then there are exactly $k(n-k)$ sets H in $\mathcal{S}_k$ such that $|L \cap H| = k-1$. Denote these sets by L_i, $i = 1 \ldots k(n-k)$.

In our next lemma we shall use standard notation from matrix theory in which if $U \in \mathrm{I\!R}^{\mathrm{m,m}}$ and L and L' are subsequences of strictly increasing integers from between 1 and m, then $U[L, L']$ is the submatrix of U whose rows and columns are determined by L and L'.

LEMMA 2.6 *Suppose that* $A = (a_{i,j}) \in \mathrm{I\!R}^{\mathrm{n,n}}$ *and let* $L \in \mathcal{S}_k$ *and* G_L *be as given in Definition* 2.5. *Then*

$$\det(G_L) \;=\; \lambda^{n-k-1}\left(\frac{n-k}{n} - \lambda\right)\xi_L + \frac{1}{n}\sum_{i=1}^{k(n-k)} \lambda^{n-k-1}\xi_{L_i},$$

where ξ_L *is up to a sign* $\det(A[L, L])$ *and* ξ_{L_i} *is up to a sign* $\det(A[L_i, L])$.

Proof: We first notice that from the definition of G_L in (2.2) it follows that

$$\det(G_L[L^c, L^c]) \;=\; \lambda^{n-k-1}\left(\lambda - \frac{n-k}{n}\right)$$

and that $\det(G_L[L_i^c, L^c])$ is up to a sign equal to $(1/n)\lambda^{n-k-1}$. Also for any $H \in \mathcal{S}_k$ such that $|L \cap H| < k-1$, the matrix $G_L[H^c, L^c]$ has two rows with all elements equal to $-\frac{1}{n}$. Therefore,

$$\det(G_L([H^c, L^c])) \;=\; 0$$

for any such H. From definition of G_L it follows that $\det(G_L[L, L])$ is equal up to a sign to $\det(A[L, L])$ and $\det(G_L[L_i, L])$ is equal up to a sign to $\det(A[L_i, L])$. The lemma now follows from the Laplace expansion of the determinant. □

Let $B = \left(\frac{1}{n} + a_{i,j}\right)$ be an $n \times n$ matrix and suppose that $p_B(\lambda) = \det(\lambda I - B)$ is the characteristic polynomial of B. We now want to separate $p_B(\lambda)$ into its *principal part* which we define as

$$\lambda^n - \lambda^{n-1}$$

and its *remainder* which we define as

$$g_B(\lambda) \;=\; p_B(\lambda) - (\lambda^n - \lambda^{n-1}) \tag{2.3}$$

Since $\det(\cdot)$ is a semilinear function in the columns of matrix, we see that

$$p_B(\lambda) \;=\; \det(\lambda I - B) \;=\; \sum_{k=0}^{n} \sum_{L \in \mathcal{S}_k} \det(G_L)\,.$$

But then, as

$$\lambda^n - \lambda^{n-1} \;=\; \sum_{L \in \mathcal{S}_0} \det(G_L)\,,$$

we see that

$$g_B(\lambda) \;=\; \sum_{k=1}^{n} \sum_{L \in \mathcal{S}_k} \det(G_L)\,.$$

Applying Lemma 2.6 we get that:

$$g_B(\lambda) \;=\; \sum_{k=1}^{n} \sum_{L \in \mathcal{S}_k} \left[\lambda^{n-k-1} \left(\frac{n-k}{n} - \lambda \right) \xi_L + \frac{1}{n} \sum_{i=1}^{k(n-k)} \lambda^{n-k-1} \xi_{L_i} \right].$$

Thus, since $|\mathcal{S}_k| = \binom{n}{k}$, we have proved that:

LEMMA 2.7 *Let* $B = \left(\frac{1}{n} + a_{i,j}\right)$. *Then*

$$g_B(\lambda) \;=\; \sum_{k=1}^{n} \left[B_k \left(\frac{n-k}{n} - \lambda \right) + C_k \right] \lambda^{n-k-1},$$

where

$$B_k \;=\; \sum_{i=1}^{\binom{n}{k}} Y_i \tag{2.4}$$

and

$$C_k \;=\; \frac{1}{n} \sum_{j=1}^{\binom{n}{k} k(n-k)} X_j \tag{2.5}$$

with Y_i's and X_j's being up to a sign determinants of different $k \times k$ sub-matrices of the matrix $A = (a_{i,j})$

In the next few results we estimate the expected values of the squares of the coefficients B_k's and C_k's given in (2.4) and (2.5), respectively, and of other related quantities all of which we shall require to estimate $E(g_B(\lambda)$ on certain circles in Lemma 2.13.

LEMMA 2.8 *Let* $B = \left(\frac{1}{n} + a_{i,j}\right) \in \mathbb{R}^{n,n}$*, where the entries of* $A = (a_{i,j}) \in \mathbb{R}^{n,n}$ *satisfy the requirements of Assumption* 2.1*. Let* B_k *and* C_k *be as given in* (2.4) *and* (2.5)*. Then*

$$E\left(B_k^2\right) \;\le\; \frac{c^k}{n^k}$$

and

$$E\left(C_k^2\right) \;\le\; \frac{c^k}{n^k}.$$

Proof: Let us compute the C_k's (the computation of B_k's being similar, but simpler). Now

$$C_k \;=\; \frac{1}{n}\sum_{j=1}^{\binom{n}{k}k(n-k)} X_j,$$

with X_j's being up to a sign the determinants of different $k \times k$ submatrices of the matrix $B = (-a_{i,j})$. By Lemma 2.4

$$E(X_iX_j) = 0$$

whenever $i \neq j$, therefore

$$E\left(C_k^2\right) \;=\; E\left(\left(\sum_{j=1}^{\binom{n}{k}k(n-k)} \frac{1}{n}X_j\right)^2\right) \;=\; \sum_{j=1}^{\binom{n}{k}k(n-k)} E\left(\left(\frac{1}{n}X_j\right)^2\right).$$

By Lemma 2.3

$$E\left(X_j^2\right) \le \frac{c^k k!}{n^{2k}}.$$

Therefore

$$\sum_{j=1}^{\binom{n}{k}k(n-k)} E\left(\frac{1}{n}X_j\right)^2 \le \sum_{j=1}^{\binom{n}{k}k(n-k)} \frac{1}{n^2}\frac{c^k k!}{n^{2k}}.$$

We have thus shown that

$$E\left(C_k^2\right) \le \sum_{j=1}^{\binom{n}{k}k(n-k)} \frac{1}{n^2}\frac{c^k k!}{n^{2k}} \le \binom{n}{k}k(n-k)\frac{1}{n^2}\frac{c^k k!}{n^{2k}}.$$

Since $k(n-k) \le n^2$ and

$$\binom{n}{k} \;\le\; \frac{n^k}{k!}$$

we have that

$$E\left(C_k^2\right) \ \leq \ \frac{c^k}{n^k}.$$

□

LEMMA 2.9 *Let* $B = \left(\frac{1}{n} + a_{i,j}\right) \in \mathrm{I\!R}^{\mathrm{n,n}}$, *where the entries* $A = (a_{i,j}) \in \mathrm{I\!R}^{\mathrm{n,n}}$ *satisfy the requirements of Assumption* 2.1. *Let* B_k *and* C_k *be as given in* (2.4) *and* (2.5). *Then*

$$E\left(\mid B_i \mid\mid B_j \mid\right) \ \leq \ \frac{c^{(i+j)/2}}{n^{(i+j)/2}},$$

$$E\left(\mid B_i \mid\mid C_j \mid\right) \ \leq \ \frac{c^{(i+j)/2}}{n^{(i+j)/2}},$$

and

$$E\left(\mid C_i \mid\mid C_j \mid\right) \ \leq \ \frac{c^{(i+j)/2}}{n^{(i+j)/2}}.$$

Proof: The lemma immediately follows from the lemma above and the Cauchy–Schwartz inequality. □

LEMMA 2.10 *Let* $B = \left(\frac{1}{n} + a_{i,j}\right) \in \mathrm{I\!R}^{\mathrm{n,n}}$, *where the entries of* $A = (a_{i,j}) \in \mathrm{I\!R}^{\mathrm{n,n}}$ *satisfy the requirements of Assumption* 2.1. *Let* B_k *and* C_k *be as given in* (2.4) *and* (2.5). *Furthermore, let*

$$D_k \ = \ \sum_{i+j=k} |B_i||B_j| + \sum_{i+j=k} |B_i||C_j| + \sum_{i+j=k} |C_i||B_j| + \sum_{i+j=k} |C_i||C_j|.$$

Then

$$E\left(D_k\right) \ \leq \ \frac{4kc^{k/2}}{n^{k/2}}.$$

Proof: The result follows from the fact that each of the four summations above runs over at the most $(k-1)$ pairs of indices i,j and by Lemma 2.9. □

LEMMA 2.11 *Let* $B = \left(\frac{1}{n} + a_{i,j}\right) \in \mathrm{I\!R}^{\mathrm{n,n}}$ *and let* $g_B(\lambda)$ *be the remainder of the characteristic polynomial of* B *as defined in* (2.3). *Let* $0 \leq \epsilon \leq 1$. *Then*

$$\max_{|\lambda|=\epsilon} |g_B(\lambda)| \ \leq \ 2\sum_{k=1}^{n} \left(|B_k| + |C_k|\right) \epsilon^{n-k-1}.$$

Proof: Lemma 2.7 we know that

$$g_B(\lambda) \;=\; \sum_{k=1}^{n} \left[B_k \left(\frac{n-k}{n} - \lambda\right) + C_k\right] \lambda^{n-k-1}.$$

This and the fact that $\left|\frac{n-k}{n} - \lambda\right| \leq 2$ yield the conclusion of the lemma. □

LEMMA 2.12 *Let* $B = \left(\frac{1}{n} + a_{i,j}\right) \in \mathrm{I\!R}^{\mathrm{n,n}}$ *and let* $g_B(\lambda)$, B_k, *and* C_k *be as given in* (2.3), (2.4), *and* (2.5), *respectively. Then*

$$\left(\max_{|\lambda| \,=\, \epsilon} |g_B(\lambda)|\right)^2 \;\leq\; 4 \sum_{k=2}^{2n-2} \epsilon^{2n-2-k} D_k.$$

Proof: The result immediately follows from the definition of D_k in Lemma 2.10 and by Lemma 2.11. □

LEMMA 2.13 *Let* $B = \left(\frac{1}{n} + a_{i,j}\right) \in \mathrm{I\!R}^{\mathrm{n,n}}$, *where the entries of* $A = (a_{i,j}) \in \mathrm{I\!R}^{\mathrm{n,n}}$ *satisfy the requirements of Assumption* 2.1. *Let* $g_B(\lambda)$ *be the remainder of the characteristic polynomial of* B *as defined in* (2.3). *Then*

$$E\left(\left(\max_{|\lambda| \,=\, \epsilon} |g_B(\lambda)|\right)^2\right) \;\leq\; 4 \sum_{k=2}^{2n-2} \epsilon^{2n-2-k} \frac{4kc^{k/2}}{n^{k/2}}.$$

Proof: The lemma follows from Lemmas 2.12 and 2.10. □

LEMMA 2.14 *Let* $B = \left(\frac{1}{n} + a_{i,j}\right) \in \mathrm{I\!R}^{\mathrm{n,n}}$, *where the entries of* $A = (a_{i,j}) \in \mathrm{I\!R}^{\mathrm{n,n}}$ *satisfy the requirements of Assumption* 2.1. *Let* $g_B(\lambda)$ *be the remainder of the characteristic polynomial of* B *as defined in* (2.3). *Let* $0 < \epsilon \leq 1$, $\alpha > c^{1/2}$, *and* $\epsilon n^{1/2} \geq \alpha$. *Then*

$$E\left(\left(\max_{|\lambda| \,=\, \epsilon} |g_B(\lambda)|\right)^2\right) \;\leq\; \frac{4\epsilon^{2n-4} G(\alpha)}{n},$$

where $G(\alpha)$ *is a constant which depends on* α *only.*

Proof: Follows from Lemma 2.13 with $G(\alpha)$ equal to the sum of convergent series

$$\sum_{k=2}^{\infty} \frac{4kc^{k/2}}{\alpha^{k-2}}.$$

□

LEMMA 2.15 *Let* $B = \left(\frac{1}{n} + a_{i,j}\right) \in \mathbb{R}^{n,n}$*, where the entries of* $A = (a_{i,j})\mathbb{R}^{n,n}$ *satisfy the requirements of Assumption 2.1. Let* $g_B(\lambda)$ *be the remainder of the characteristic polynomial of* B *as defined in (2.3). Let* $0 < \epsilon < 1$ *be fixed. Then the probability that for all* λ *such that* $|\lambda| = \epsilon$ *the absolute value of the remainder* $|g_B(\lambda)|$ *is strictly less than the absolute value of the principal part* $|\lambda^n - \lambda^{n-1}|$ *of the characteristic polynomial tends to 1 as n tends to infinity.*

Proof: Let us fix $\alpha > 1$ and suppose that $\epsilon n^{1/2} \geq \alpha$. Notice that $|\lambda^n - \lambda^{n-1}| > \frac{1}{2}\epsilon^{n-1}(1-\epsilon)$ when $|\lambda| = \epsilon$. Therefore,

$$P\left(\max_{|\lambda|=\epsilon} |g_B(\lambda)| \;<\; \min_{|\lambda|=\epsilon} \left|\lambda^n - \lambda^{n-1}\right|\right) \;\geq\; P\left(\max_{|\lambda|=\epsilon} |g_B(\lambda)| < \frac{1}{2}\epsilon^{n-1}(1-\epsilon)\right).$$

But then, by (the reverse case of) Markov's inequality,

$$P\left(\max_{|\lambda|=\epsilon} \;|g_B(\lambda)| \;<\; \tfrac{1}{2}\epsilon^{n-1}(1-\epsilon)\right) \;\geq\; 1 - \frac{E\left(\left(\max_{|\lambda|=\epsilon} \;|g_B(\lambda)|\right)^2\right)}{\left[\frac{1}{2}\epsilon^{n-1}(1-\epsilon)\right]^2}$$

$$\geq\; 1 - \frac{4\epsilon^{2n-4}G(\alpha)}{\left[\frac{1}{2}\epsilon^{n-1}(1-\epsilon)\right]^2 n} \;=\; 1 - \frac{4G(\alpha)}{\frac{1}{4}\epsilon^2(1-\epsilon)^2 n},$$

where the last inequality follows from Lemma 2.14. Our result follows from the inequality above. □

Proof of Theorem 1.1(i): Fix an $0 < \epsilon < 1$. Then the principal part of the characteristic polynomial has exactly $n-1$ roots inside the open ϵ neighborhood of 0. Rouché's theorem tells us that whenever the absolute value of the principal part of the characteristic polynomial $|\lambda^n - \lambda^{n-1}|$ is greater than the absolute value of the remainder $|g_B(\lambda)|$ on the circle of radius ϵ in the complex plain then the characteristic polynomial has exactly as many roots inside the ϵ–ball as the principal part. In our case it has exactly $n-1$ roots inside the ϵ–ball. The previous lemma shows that the probability of the above condition being true tends to 1 as n tends to infinity. This concludes the proof of the main result.

REMARK 2.16 An implication of Theorem 1.1 is, of course, that, under the assumptions of the theorem, the subdominant eigenvalues of the random matrices B tend, in probability, to 0 as n tends to infinity.

REMARK 2.17 We can estimate the rate of convergence of the subdominant eigenvalues of B to 0. Let $m = \max_{2\le i\le n}|\lambda_i|$. For $0 \le p < 1$ and each fixed positive integer n, let $\epsilon(p,n) \ge 0$ be *the smallest number such that*

$$P\left(m \le \epsilon(p,n)\right) \ \ge \ p.$$

From the rightmost expressions in the proof of Lemma 2.15, it follows that for each $p \in [0,1)$, there is a constant $C(p)$ such that

$$\epsilon(p,n) \le \frac{C(p)}{\sqrt{n}}.$$

We call the interval $[0, \epsilon(p,n)]$ *a confidence interval.* By this we mean that with probability or confidence of at least p, the number $m = \max_{2\le i\le n}|\lambda_i|$ is in this interval. An outcome of Theorem 1.1 is that for every $p \in [0,1)$, the right–end of the confidence interval tends to zero as some constant over the square root of n as n tends to infinity.

REMARK 2.18 The considerations used to prove Theorem 1.1(i) can be slightly modified to prove a similar result foe matrices $B = (b_{i,j}) \in \mathrm{IR}^{\mathrm{n,n}}$ whose entries are given by $\frac{1}{n} + a_{i,j}$, where the $a_{i,j}$'s are independent random variables with $E(a_{i,j}) = 0$, but now with $\mathrm{Var}(a_{i,j}) = \frac{c}{n^{1+\delta}}$, where $\delta > 0$.

Thus far we have proved that for any $0 < \epsilon < 1$ and for sufficiently large n, with great probability the characteristic polynomial of B has as many roots in the open ball of radius ϵ about the origin as the polynomial $\lambda^n - \lambda^{n-1}$. The same arguments, with minor modifications, can now be used to show that for any $\epsilon > 1$ and for sufficiently large n, with great probability the characteristic polynomial of B has as many roots in the open ball of radius ϵ about the origin as the polynomial $\lambda^n - \lambda^{n-1}$. It follows that for any $\delta > 0$ and for n large enough, with great probability the spectral radius of B is within δ of 1. This establishes then the validity of part (ii) of Theorem 1.1.

REMARK 2.19 Just as in the Remark 2.17, we can define $\tilde{\epsilon}(p,n) \geq 0$ to be the smallest number such that

$$P\ (\ |1-|\lambda_1|\ |\ \leq \tilde{\epsilon}(p,n)\ )\ \geq\ p.$$

and derive for each $p \in [0,1)$, a constant $C(p)$ such that

$$\tilde{\epsilon}(p,n) \leq \frac{C(p)}{\sqrt{n}}.$$

This leads again to a confidence interval $[0, \tilde{\epsilon}(p,n)]$ such that with probability or confidence of at least p, the number $|1-|\lambda_1||$ is in the interval. The right end point of the confidence interval tends to 0 as some constant over the square root of n as n tends to infinity.

ACKNOWLEDGEMENT: The authors would like to thank Professor Alan Edelman of MIT and the University of California at Berkeley for his very kind help and advice.

References

[1] Z. D. Bai. Circular law. *Anal. of Probability*, 25:494–529, 1997.

[2] F. L. Bauer, Eckart Deutsch, and J. Stoer. Abschätzungen für die Eigenwerte positive linearer Operatoren. *Lin. Alg. Appl.*, 2:275–301, 1969.

[3] A. Berman and R. J. Plemmons. Nonnegative Matrices in the Mathematical Sciences. *SIAM*, Philadelphia, 1994.

[4] S. L. Campbell and C. D. Meyer, Jr. Generalized Inverses of Linear Transformations. *Dover Publications*, New York, 1991.

[5] R. L. Dobrushin. Central limit theorems for nonstationary Markov chains. *Theory of Probab. Appl.*, 1: 65–79 (Part I), 329 –383 (Part 2), 1956.

[6] A. Edelman. Random Eigenvalue Bibliography. http://www. math.berkeley.edu/˜ edelman, MSRI, Univ. Calif., Berkeley, 1999.

[7] A. Edelman. Random Eigenvalues. *Work in progress*, MSRI, Berkeley 1999.

[8] Y. Friedman. On singular numbers of a matrix with random entries. *Unpublished notes*, 1983.

[9] W. Feller. An Introduction to Probability Theory and its Applications, Vol. II. *John Wiley*, New York, 1966.

[10] V. L. Girko. An Introduction to Statistical Analysis of Random Arrays. *VSP International Science Pub.*, Utrecht, 1988.

[11] E. B. Manoukian. Modern Concepts and Theorems of Mathematical Statistics. *Springer Verlag Series in Statistics*, New–York, 1985.

[12] M. L. Mehta. Random Matrices. *Academic Press*, New–York, 1991.

[13] M. Neumann and H. Schneider. Partial norms and the convergence of general products of matrices. *Lin. Alg. Appl.*, 287:307–314, 1999.

[14] A. M. Ostrowsi. On positive matrices. *Math. Anal.*, 150:276–284, 1963.

[15] A. Paz. Introduction to Probabilistic Automata. *Academic Press*, New–York, 1971.

[16] E. Seneta. Non–negative Matrices and Markov Chains. Second Edition. *Springer Series in Statistics*, Springer–Verlag, New York, 1980.

[17] E. Seneta. Coefficients of ergodicity:structure and applications. *Av. Appl. Prob.*, 11:576–590, 1979.

[18] D. O. Tall. Functions of a Complex Variable, Vo.II. *Routledge & Kegan Paul Ltd.*, London, 1970.

[19] C. P. Tan. Coefficients of ergodicity with respect to vector norms. *J. Appl. Probab.*, 20:277–287, 1983.