Ramanujan J (2019) 48:131–149
https://doi.org/10.1007/s11139-017-9930-5

Statistics and characterization of matrices by determinant and trace

Emre Alkan[1] · **Ekin Sıla Yörük**[2]

Received: 23 May 2017 / Accepted: 1 June 2017 / Published online: 16 August 2017

Abstract Answering a question of Erdös, Komlós proved in 1968 that almost all $n \times n$ Bernoulli matrices are nonsingular as $n \to \infty$. In this paper, we offer a new perspective on the question of Erdös by studying $n \times n$ matrices with prime number entries in an almost all sense. Precisely, it is shown that, as $x \to \infty$, the probability of randomly choosing a nonsingular $n \times n$ matrix among all $n \times n$ matrices with prime number entries that are $\leq x$ is 1. If A is a unitary matrix, then it is well known that $|\det A| = 1$. However, the converse is far from being true. As a remedy of this defect, we search for necessary and sufficient conditions for being a unitary matrix by teaming up determinant with trace. In this way, we are led to simple characterizations of unitary matrices in the set of normal matrices. The question of which nonsingular commuting complex matrices with real eigenvalues have the same characteristic polynomial is formulated via determinant and trace conditions. Finally, through a study of eigenvectors, we obtain new characterizations of Hermitian and normal matrices. Our approach to proving these results benefits from a modular interpretation of nonsingularity and the spectral theorem for normal operators together with equality cases of classical inequalities such as the arithmetic–geometric mean inequality and the Cauchy–Schwarz inequality.

Keywords Matrices with prime number entries · Prime numbers in progressions · Almost all · Determinant · Trace · Unitary matrix · Hermitian matrix · Normal matrix

✉ Emre Alkan
ealkan@ku.edu.tr

Ekin Sıla Yörük
eyoruk13@ku.edu.tr

[1] Department of Mathematics, Koç University, Rumelifeneri Yolu, 34450 Sarıyer, Istanbul, Turkey

[2] Department of Physics, Koç University, Rumelifeneri Yolu, 34450 Sarıyer, Istanbul, Turkey

Mathematics Subject Classification 11N13 · 15B36 · 15A57

1 Introduction

Let A be a square matrix of size n with complex entries. Two important quantities that are associated with A are the determinant and trace, denoted as $\det A$ and $\mathrm{tr}(A)$, respectively. If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then

$$\det A = \prod_i \lambda_i \quad \text{and} \quad \mathrm{tr}(A) = \sum_i \lambda_i,$$

where both of them are symmetric functions of the λ_i. Despite the well-known fact that determinant and trace are both invariants under similarity transformations, the converse almost never holds. One of the motivations of this paper is to repair this defect for specific classes of matrices and further obtain characterizations of them in terms of statements involving only determinant and trace. To quote another well-known fact, note that the vanishing of determinant characterizes the singularity of a matrix. In this regard, singular matrices are expected to be statistically rare among all matrices. For this expectation to be meaningful, we need to put constraints on the entries of the matrices under consideration. A natural and simple setting would be to focus on $n \times n$ matrices with nonnegative integer entries that are $\leq x$ (here we are assuming that $x \geq 1$). Answering a question of Erdös, Komlós [12,13] proved in this setting that the probability of randomly choosing a singular matrix tends to 0 as $n \to \infty$. In particular, for the first interesting case when $x = 1$, matrices with 0, 1 entries are called Bernoulli matrices. An arithmetic question on unimodular matrices was nicely treated by Dănescu et al. [5]. There are alternative ways to quantify the rareness of singular matrices. For fixed n, one could as well look at matrices whose entries are $\leq x$ and belong to a thin subset of positive integers, such as the set of prime numbers, with the hope of showing that the singular ones among them are still rare as $x \to \infty$. This would then give a new direction on the question of Erdös. Our main result below confirms the desired statistical expectation in an almost all sense among matrices all of whose entries belong to the set of prime numbers. We should mention that almost all type results are common in number theory. Recall that a property $\mathcal{P}$ holds for almost all positive integers, if

$$\lim_{x\to\infty} \frac{\mathcal{P}(x)}{x} = 1,$$

where $\mathcal{P}(x)$ is the number of integers that are $\leq x$ and have the property $\mathcal{P}$. A classical result of Hardy and Ramanujan [10] states that almost all positive integers n have about $\log\log n$ prime factors. A celebrated theorem of Erdös and Kac [8] extends this phenomenon to the values of a wide class of arithmetic functions (for a characterization of additive arithmetic functions with continuous limiting distributions, see [7]). A special case of their striking discovery shows that the number of prime factors of a positive integer n behaves like a Gaussian normal distribution with mean and variance both equal to $\log\log n$. We can now state the main result.

Theorem 1 *Asymptotically, almost all $n \times n$ matrices with prime number entries are nonsingular. Precisely, if $M_n(x)$ is the number of nonsingular $n \times n$ matrices with prime number entries that are $\leq x$ and $\pi(x)$ is the number of prime numbers that are $\leq x$, then*

$$\lim_{x \to \infty} \frac{M_n(x)}{\pi(x)^{n^2}} = 1.$$

Consequently, as $x \to \infty$, the probability of randomly choosing a nonsingular $n \times n$ matrix among all $n \times n$ matrices with prime number entries that are $\leq x$ is 1.

Some remarks on Theorem 1 are now in order. First by the prime number theorem (see Chap. 18 of [6])

$$\pi(x) \sim \frac{x}{\log x} \sim li(x) = \int_2^x \frac{1}{\log t} \mathrm{d}t$$

as $x \to \infty$. Thus Theorem 1 can be rephrased as

$$M_n(x) \sim \left(\frac{x}{\log x} \right)^{n^2}$$

as $x \to \infty$. Recall that a set of positive integers $\mathcal{A}$ has asymptotic density 0 if

$$\lim_{x \to \infty} \frac{\mathcal{A}(x)}{x} = 0,$$

where $\mathcal{A}(x)$ is the number of integers in $\mathcal{A}$ that are $\leq x$. Note that the set of prime numbers has asymptotic density 0 and more generally, a set with asymptotic density 0 can be viewed as a thin subset of positive integers. It would be an interesting task to obtain analogs of Theorem 1 for matrices with entries coming from other thin subsets of integers. It turns out that, by adapting our method of proof of Theorem 1, such a task can be undertaken if the set of integers, where the entries of the matrices under consideration belong to, is sufficiently well distributed over arithmetic progressions. In particular, the prime number theorem for arithmetic progressions (see Chap. 20 of [6]) is needed in the proof of Theorem 1, thereby exploiting the fact that prime numbers are asymptotically equally distributed among arithmetic progressions which admit infinitely many of them.

Matrices possessing more structure such as symmetry are ubiquitous throughout linear algebra (and most of the rest of mathematics as well). These include representations arising from unitary, Hermitian, skew-Hermitian, and normal operators which indeed constitute a foundation for the mathematical formulation of quantum mechanics (see Chap. 3 of [9]). As a result of the impact of such operators on applied problems, alternative characterizations of them were obtained and used (see Chap. 7 of [14]). Recall that A is real unitary precisely when

$$AA^t = I,$$

where A^t is the transpose of A. An obvious necessary condition for being real unitary is then

$$|\det A| = 1.$$

However, this condition is far from being a sufficient condition. An elegant necessary and sufficient condition can be given by teaming determinant with trace. Although it may still not be clear which matrix should be associated with the trace, one is able to infer a slightly stronger statement. For the notation, we let $\langle\ .\ ,\ .\ \rangle$ be a positive definite bilinear form (or scalar product) on $\mathbb{R}^n$.

Theorem 2 *Let A be a matrix of size n with real entries. Then A is similar (with respect to an orthonormal basis arising from $\langle\ .\ ,\ .\ \rangle$) to a block matrix B consisting of only 2×2 and 1×1 blocks of the forms*

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad [1], \quad [-1]$$

on the main diagonal (so that other entries of B are 0) if and only if

$$|\det A| = 1 \ \text{ and } \ \operatorname{tr}(AA^t) = n.$$

Our proof of Theorem 2 rests on the relation between determinant and trace via the arithmetic–geometric mean inequality,

$$\frac{1}{n}\sum_{i=1}^{n} x_i \geq \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$$

for nonnegative real numbers $x_1, \ldots, x_n$, equality being possible only when all x_i are the same. For a complex unitary matrix A satisfying

$$AA^* = I,$$

where A^* is the conjugate transpose of A, a similar characterization can be given as a bonus by making the necessary modifications in Theorem 2. This time we let $\langle\ .\ ,\ .\ \rangle$ be a positive definite Hermitian form on $\mathbb{C}^n$.

Theorem 3 *Let A be a matrix of size n with complex entries. Then A is similar (with respect to an orthonormal basis arising from $\langle\ .\ ,\ .\ \rangle$) to a diagonal matrix B consisting of entries of the form $e^{i\theta}$ on the main diagonal if and only if*

$$|\det A| = 1 \ \text{ and } \ \operatorname{tr}(AA^*) = n.$$

The conditions of being Hermitian and normal for a matrix A are given by

$$A = A^* \text{ and } AA^* = A^*A,$$

respectively. Consequently, unitary and Hermitian matrices are normal, though Hermitian and normal matrices do not possess group structure, a key property shared by unitary matrices. Furthermore, unitary matrices are characterized by preserving unit vectors (see p. 189 of [14]). The following result keeps the essence of Theorems 2 and 3 by providing a new characterization of complex unitary matrices when they are viewed as members of the collection of normal matrices.

Corollary 1 *Assume that A is a normal matrix of size n with complex eigenvalues $\lambda_1, \ldots, \lambda_n$. Then A is complex unitary if and only if*

$$|\det A| = 1 \text{ and } \sum_i |\lambda_i| = n.$$

Indeed it is possible to formulate the unitary condition solely as an equation involving determinant and trace.

Theorem 4 *Let N be a nonsingular complex matrix of size n. Then there exists a positive number c such that cN is unitary if and only if the equality*

$$\frac{\operatorname{tr}(NN^*)}{n} = |\det N|^{\frac{2}{n}} \tag{1}$$

holds.

Assume for the moment that A and B are similar matrices. Then $B = P^{-1}AP$ holds for some nonsingular matrix P. We certainly have then $\det A = \det B$, $\operatorname{tr}(A) = \operatorname{tr}(B)$, and A, B have the same characteristic polynomial so do the same eigenvalues counting multiplicity. Again, in general, converses of these implications are not true. A better measure of similarity might go through the comparison of Jordan normal forms of A and B (see Chap. 11 of [14]). Observe that the computation of determinant and trace uses data only from the main diagonal of the Jordan normal form but not from the actual blocks appearing in the Jordan normal form. Therefore, we cannot hope to deduce the similarity of matrices in an obvious way from determinant and trace conditions. Despite this, it is possible to salvage the conclusion that two matrices have the same eigenvalues counting multiplicity out of a combination of determinant and trace conditions.

Theorem 5 *Let A and B be nonsingular commuting complex matrices with real eigenvalues. Then A and B have the same characteristic polynomial if and only if*

$$\det A = \det B \text{ and } (\operatorname{tr}(AB))^2 = \operatorname{tr}(A^2)\operatorname{tr}(B^2). \tag{2}$$

It is interesting to note how the trace condition in (2) mimics the equality case of the Cauchy–Schwarz inequality. The set of eigenvalues of an operator is called the spectrum of that operator so that in the case of Theorem 5, A and B would also have the same spectrum if the hypotheses are satisfied. Let us have a brief digression to generate plenty of examples of nonsingular commuting matrices with real eigenvalues

as in Theorem 5. To this end, first start with two Jordan normal forms, say J_1 and J_2, subject to the following conditions. If J_1 has a block of the form $\lambda I + N$, where N is the nilpotent part consisting of 1's above the main diagonal, then J_2 either has a block of the same form, namely has a block of the form $\mu I + N$, or it has a diagonal block of the form μI having the same size. Note that $\lambda I + N$ commutes with $\mu I + N$ and μI. Consequently, assuming also that the corresponding blocks of J_1 and J_2 appear at the same places on the main diagonal, we see that $J_1 J_2 = J_2 J_1$. Next take any nonsingular matrix P. Then let $A = P^{-1} J_1 P$ and $B = P^{-1} J_2 P$ so that A and B commute as J_1 and J_2 commute. Finally, the eigenvalues of A, B are the λ's and μ's so we may choose them to be nonzero real numbers. This guarantees that A and B are nonsingular as well. Having generated examples, we may assume that A, B are nonsingular commuting matrices with real eigenvalues satisfying (2). Thus Theorem 5 applies and shows that A and B have the same eigenvalues with the same multiplicities. We finish this digression by a discussion of some extra conditions imposed on A and B which would force them to be similar. In addition to the above conditions, assume that the geometric multiplicity of each eigenvalue of A, B is the same. This means that for each common eigenvalue λ,

$$\dim \operatorname{Ker}(A - \lambda I) = \dim \operatorname{Ker}(B - \lambda I)$$

holds, where $\dim Ker(A - \lambda I)$ and $\dim Ker(B - \lambda I)$ denote the dimensions of the kernel of the corresponding operators. Moreover, if $(t - \lambda)^{m_1}$ and $(t - \lambda)^{m_2}$ are the factors, corresponding to λ, in the minimal polynomials of A and B, respectively, then in either of the cases when

$$\max(m_1, m_2) \leq 2$$

for all λ or

$$m_1 = m_2 \geq m - 3$$

for all λ, where m is the multiplicity of λ (so that $(t - \lambda)^m$ is the factor corresponding to λ in the common characteristic polynomial of A, B), one can infer that A, B are similar. To justify this first assume that $\max(m_1, m_2) \leq 2$. Consider the Jordan normal forms of A and B. All blocks involving λ must have size at most 2 in both of the Jordan normal forms of A, B. As the geometric multiplicities are the same, the total number of 1's appearing in all of the blocks involving λ must be the same (observe that the total number of 1's is equal to the multiplicity of λ minus the geometric multiplicity of λ) in both of the Jordan normal forms of A, B. This means that the number of 2×2 blocks should be the same in both of the Jordan normal forms of A, B for every common eigenvalue λ. Therefore, the Jordan normal forms of A, B are similar and it follows that A, B are similar. For the other case when $m_1 = m_2 \geq m - 3$, if there is a block of size $m - 3$ corresponding to λ in both of the Jordan normal forms of A, B, then the remaining part can decompose as a single 3×3 or one 2×2 and one 1×1 or three 1×1. Again the total number of 1's must be the same, and the blocks in the Jordan normal forms of A, B are identical except possibly their places on the main

diagonal. Thus A and B are forced to be similar. If there is a block of size $> m-3$ corresponding to λ in both of the Jordan normal forms of A, B, then it can be shown as above that A and B are again similar.

The spectral theorem holds for normal matrices (see p. 227 of [14]) which makes them amenable to diagonalization with respect to an orthonormal basis consisting of eigenvectors. A characterization of normal operators was defined by Hoffman and Taussky [11]. In this connection, it is possible to give a simple and elegant treatment of Hermitian and normal matrices separately by focusing on their eigenvectors. The case of hermitian matrices is handled in the next theorem.

Theorem 6 *Assume that A is a complex matrix with real eigenvalues. Then A is Hermitian if and only if A and A^* have the same eigenvectors.*

For normal matrices there is an analogous criterion in terms of eigenvectors.

Theorem 7 *Let A be a complex matrix. Then A is normal if and only if AA^* and A^*A have the same eigenvectors corresponding to every common eigenvalue of AA^* and A^*A.*

2 A modular interpretation of nonsingularity

Let $A = (a_{ij})$ be a given $n \times n$ matrix with positive integer entries. For any prime number p, let

$$A_p = (\overline{a_{ij}})$$

be the $n \times n$ matrix obtained from A by reducing entries of A modulo p. Thus A_p can be regarded as a matrix over the field

$$\mathbb{Z}_p = \{\overline{0}, \overline{1}, \ldots, \overline{p-1}\}$$

of congruence classes modulo p. The following characterization of the nonsingularity of A will be one of the key ingredients in the proof of Theorem 1.

Theorem 8 (Modular interpretation of nonsingularity) *Let A be a matrix of size n with positive integer entries. Then A is nonsingular if and only if A_p is nonsingular for some prime number p.*

To prove Theorem 8, first let $A = (a_{ij})$. Then by the well-known determinant formula,

$$\det A = \sum_{\sigma \in S_n} (-1)^{\epsilon(\sigma)} a_{1,\sigma(1)} \ldots a_{n,\sigma(n)},$$

where σ ranges over all permutations in the symmetric group S_n and $\epsilon(\sigma)$ gives the parity of the permutation which is 0 or 1 according to when σ is an even or odd

permutation, respectively. Now $\det A$ is an integer and if $\overline{\det A} \in \mathbb{Z}_p$ denotes the reduction of $\det A$ modulo p, then we see from the above formula that

$$\overline{\det A} = \sum_{\sigma \in S_n} (-1)^{\epsilon(\sigma)} \overline{a_{1,\sigma(1)}} \dots \overline{a_{n,\sigma(n)}} = \det A_p$$

as $A_p = (\overline{a_{ij}})$. If A_p is nonsingular for some prime p, then $\det A_p \neq \overline{0}$ and this forces $\det A \neq 0$. Therefore, A has to be nonsingular. Conversely, if A is nonsingular, then $\det A$ is a nonzero integer. Let p be a prime number not dividing $\det A$. Then $\overline{\det A} \neq \overline{0}$ in $\mathbb{Z}_p$ and this forces $\det A_p \neq \overline{0}$. Thus A_p is nonsingular for some prime number p. This completes the verification of Theorem 8.

It is worth remarking that Theorem 8 builds on the simple idea of reading entries modulo a prime number. Reading entries modulo prime numbers is a fruitful idea in other areas of mathematics such as the local to global principles in the theory of Diophantine equations which motivated most of the developments in algebraic number theory and class field theory (see Chaps. 2, 5, 6 of [15] and first part of [17] for the Hasse–Minkowski theorem). This idea also led to the theory of supersingular primes by studying elliptic curves p-adically. For results on the Fourier coefficients of cusp forms associated with elliptic curves using the distribution of supersingular primes, see [2,3] which form part of the first author's doctoral thesis [1] (also see [4]). Thus, inspired by these applications, there is good motivation to name this section and Theorem 8 as above.

In conclusion, to show that such a matrix A as in Theorem 8 is nonsingular, it is enough to find a prime number p and show that A_p is nonsingular over $\mathbb{Z}_p$. It makes sense, of course, to search for the smallest prime number p that does the job. Let us therefore call the smallest prime number p such that A_p is nonsingular, the complexity of A. Note that the complexity of the 3×3 matrix

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 15 & 3 & 12 \\ 25 & 50 & 101 \end{bmatrix}$$

is 5 as A_2 and A_3 are singular but A_5 is not as $\det A_5 = \overline{1}$ in $\mathbb{Z}_5$. It is also clear from the proof of Theorem 8 that if A is nonsingular, then the complexity of A is the smallest prime number not dividing $\det A$. Although our modular interpretation of nonsingularity is theoretically interesting, verifying that a given matrix A is nonsingular with the help of this principle can be arbitrarily difficult. This is due to the fact that the complexity is unbounded. Indeed the unboundedness even holds for matrices having prime number entries.

Theorem 9 *Let N be a positive integer and $n \geq 2$. Then there exist infinitely many $n \times n$ matrices with prime number entries whose complexity are all $> N$.*

Observe that Theorem 9 is not true when $n = 1$ as the complexity of a 1×1 matrix $[p]$, where p is an odd prime, is always 2. We may now give the proof of Theorem 9. Let p_{k+1} be the least prime that is $> N$ and put

$$P = \prod_{i=1}^{k} p_i,$$

where $2 = p_1 < 3 = p_2 < \cdots < p_i < \ldots$ is the sequence of primes in increasing order. Then consider the progression $Pm + 1$, where m ranges over positive integers. By Dirichlet's theorem, there are infinitely many prime numbers belonging to this progression. We prove by induction over $n \geq 2$ that there exist infinitely many $n \times n$ matrices with prime number entries in the progression $Pm + 1$ whose complexity is all $> N$. For the base case of the induction, let $A = (a_{ij})$ be a 2×2 matrix with prime number entries in the progression $Pm + 1$. Then note that

$$A_{p_i} = \begin{bmatrix} \overline{1} & \overline{1} \\ \overline{1} & \overline{1} \end{bmatrix}$$

and A_{p_i} is singular for all $1 \leq i \leq k$. Moreover, we have $\det A = a_{11}a_{22} - a_{21}a_{12}$ and keeping a_{12}, a_{21}, a_{22} fixed, we may vary a_{11} along the progression $Pm + 1$ to guarantee that $\det A \neq 0$. Thus for each such choice of a_{11}, the corresponding A is nonsingular and its complexity is $\geq p_{k+1} > N$. Clearly, there are infinitely many such A as there are infinitely many choices for a_{11}. This settles the base case. For the inductive step, let $A = (a_{ij})$ be an $n \times n$ matrix with prime number entries in the progression $Pm + 1$. As above, all entries of A_{p_i} are $\overline{1}$ and A_{p_i} is singular for $1 \leq i \leq k$. Let A_{11} be the $(n-1) \times (n-1)$ minor obtained from A by deleting the first row and the first column of A. By the inductive hypothesis, we may assume that $\det A_{11} \neq 0$. Again we may keep all a_{ij} fixed when $i \neq 1$, $j \neq 1$, and vary a_{11} in the progression $Pm + 1$. We also have

$$\det A = a_{11} \det A_{11} + \sum_{j=2}^{n} a_{1j} \det A_{1j},$$

where A_{1j} is the $(n-1) \times (n-1)$ minor obtained from A by deleting the first row and the jth column of A. As $\det A_{11} \neq 0$ and

$$\sum_{j=2}^{n} a_{1j} \det A_{1j}$$

is fixed, we may vary a_{11} to guarantee that $\det A \neq 0$. Therefore, A is nonsingular and the complexity of A is $\geq p_{k+1} > N$. Lastly, there exist infinitely many such A since there are infinitely many choices for a_{11} as a prime number in the progression $Pm + 1$. This completes the induction and the proof of Theorem 9.

3 Proof of Theorem 1

The claim trivially holds when $n = 1$ so we may assume for the rest of the argument that $n \geq 2$. First note that the number of all $n \times n$ matrices with prime number entries

that are $\leq x$ is $\pi(x)^{n^2}$. Since $M_n(x) \leq \pi(x)^{n^2}$, we see that

$$\limsup_{x\to\infty} \frac{M_n(x)}{\pi(x)^{n^2}} \leq 1. \tag{3}$$

Let p be a prime number and let $B = (\overline{b_{ij}})$ be a nonsingular matrix over $\mathbb{Z}_p$ such that all entries of B are in $\{\overline{1}, \ldots, \overline{p-1}\}$. An important step in the proof is to use the modular interpretation of nonsingularity as in Theorem 8. To this end, consider any matrix $A = (a_{ij})$ with prime number entries that are $\leq x$ and satisfying $A_p = B$, where the matrix A_p is defined as in Sect. 2. As B is assumed to be nonsingular over $\mathbb{Z}_p$, Theorem 8 tells us that A is nonsingular. Clearly, each entry of A is a prime number $\leq x$ that lies in a progression with common difference p. Precisely,

$$a_{ij} \equiv b_{ij} \pmod{p}$$

and $\overline{b_{ij}} \neq \overline{0}$ for all i, j. Therefore, by the prime number theorem for arithmetic progressions (see Chap. 20 of [6]), it follows that each entry of A can be chosen in

$$(1 + o(1))\frac{\pi(x)}{p-1}$$

many ways as $x \to \infty$. Consequently, the number of such nonsingular matrices A corresponding to B, by reducing entries of A modulo p, is

$$(1 + o(1))\left(\frac{\pi(x)}{p-1}\right)^{n^2} \tag{4}$$

as $x \to \infty$. Next let $N(p)$ be the number of nonsingular matrices over $\mathbb{Z}_p$ with entries in $\{\overline{1}, \ldots, \overline{p-1}\}$. As the A's corresponding to different such B's have to be different, we infer from (4) that

$$M_n(x) \geq (1 + o(1))\left(\frac{\pi(x)}{p-1}\right)^{n^2} N(p) \tag{5}$$

as $x \to \infty$. Now (5) further gives that

$$\liminf_{x\to\infty} \frac{M_n(x)}{\pi(x)^{n^2}} \geq \frac{N(p)}{(p-1)^{n^2}}. \tag{6}$$

To complete the proof, we need to find a good lower bound for $N(p)$. Note that each such matrix B as above may be viewed as an invertible linear map belonging to the general linear group $GL_n(\mathbb{Z}_p)$. Let $\{v_1, \ldots, v_n\}$ be a basis for $\mathbb{Z}_p^n$ over $\mathbb{Z}_p$. It suffices to find a good lower bound for the number invertible linear maps from $\mathbb{Z}_p^n$ to $\mathbb{Z}_p^n$, where the allowed coefficients are in $\{\overline{1}, \ldots, \overline{p-1}\}$. Let $f : \mathbb{Z}_p^n \to \mathbb{Z}_p^n$ be such a linear map.

Then f is determined uniquely by $f(v_1), \ldots, f(v_n)$. Starting with $f(v_1)$, say

$$f(v_1) = \sum_{i=1}^{n} \overline{x_i} v_i$$

with $\overline{x_i} \in \{\overline{1}, \ldots, \overline{p-1}\}$. Clearly, $f(v_1) \neq 0$ and $f(v_1)$ can be chosen in $(p-1)^n$ different ways. In general, for $2 \leq j \leq n$, $f(v_j)$ cannot be a $\mathbb{Z}_p$ combination of $f(v_1), \ldots, f(v_{j-1})$. Observe that

$$\left|\{\overline{\alpha_1} f(v_1) + \cdots + \overline{\alpha_{j-1}} f(v_{j-1}) : \overline{\alpha_1}, \ldots, \overline{\alpha_{j-1}} \in \mathbb{Z}_p\}\right| = p^{j-1}$$

as $\{v_1, \ldots, v_{j-1}\}$ are already chosen to be linearly independent. Therefore, the number of choices for $f(v_j)$, so as to make f invertible, is at least

$$(p-1)^n - p^{j-1}.$$

Let us remark here that this is far from being an exact count for the number of choices of $f(v_j)$ since some of the combinations of $f(v_1), \ldots, f(v_{j-1})$ in the form

$$\overline{\alpha_1} f(v_1) + \cdots + \overline{\alpha_{j-1}} f(v_{j-1})$$

with $\overline{\alpha_1}, \ldots, \overline{\alpha_{j-1}} \in \mathbb{Z}_p$ can be written as a combination of $v_1, \ldots, v_n$, where some of the coefficients are $\overline{0}$. It follows that the number of such invertible linear maps is at least

$$(p-1)^n \prod_{j=2}^{n} \left((p-1)^n - p^{j-1}\right)$$

which amounts to the lower bound

$$N(p) \geq (p-1)^n \prod_{j=2}^{n} \left((p-1)^n - p^{j-1}\right). \tag{7}$$

It is worth pointing out that (7) is nontrivial only when p is large enough in terms of n. Combining (6) with (7), we deduce that

$$\liminf_{x \to \infty} \frac{M_n(x)}{\pi(x)^{n^2}} \geq (p-1)^{n-n^2} \prod_{j=2}^{n} \left((p-1)^n - p^{j-1}\right). \tag{8}$$

We also have

$$(p-1)^{n-n^2} \prod_{j=2}^{n} \left((p-1)^n - p^{j-1}\right) = \prod_{j=2}^{n} \left[1 - \frac{1}{p^{n-j+1}\left(1-\frac{1}{p}\right)^n}\right], \tag{9}$$

where the product over j on the right hand side of (9) is nonempty since $n \geq 2$ is assumed. Moreover, $n - j + 1 \geq 1$ holds for all $2 \leq j \leq n$ and using the fact that p can be arbitrarily large, one easily sees that

$$\lim_{p\to\infty} \prod_{j=2}^{n} \left[1 - \frac{1}{p^{n-j+1}\left(1 - \frac{1}{p}\right)^{n}} \right] = 1. \tag{10}$$

As a result of (8)–(10), we have

$$\liminf_{x\to\infty} \frac{M_n(x)}{\pi(x)^{n^2}} = 1 \tag{11}$$

Finally assembling (3) with (11), one completes the proof of Theorem 1.

4 Proof of Theorem 2

First assume that A is similar to such a matrix B with respect to an orthonormal basis of $\mathbb{R}^n$. Then there exists a real unitary matrix P satisfying $P^{-1}AP = B$ and $P^{-1} = P^t$. It is easy to see that

$$\det B = \pm \prod_{i=1}^{k} (\cos^2\theta_i + \sin^2\theta_i) = \pm 1, \tag{12}$$

for some $k \leq [n/2]$, where k is the number of 2×2 blocks appearing on the main diagonal of B. Note that $\det A = \det B$ and $|\det A| = 1$ follows from (12). Moreover, we have $\mathrm{tr}(BB^t) = n$. Using the facts that P is real unitary and

$$P^t A^t (P^{-1})^t = B^t,$$

we infer that

$$P^{-1}AA^tP = P^{-1}AA^t(P^{-1})^t = P^{-1}APP^tA^t(P^{-1})^t = BB^t. \tag{13}$$

Thus by (13), AA^t is similar to BB^t and we obtain that $\mathrm{tr}(AA^t) = \mathrm{tr}(BB^t) = n$. This completes the proof of the necessity part of the claim. For the sufficiency part of the claim, assume that $|\det A| = 1$ and $\mathrm{tr}(AA^t) = n$. Note that AA^t is a symmetric matrix. Let λ be an eigenvalue of AA^t. It is well known that eigenvalues of symmetric matrices are real so that λ is real. Moreover, if $v \neq 0$ is an eigenvector for λ, then using $AA^tv = \lambda v$, one gets

$$\left\langle AA^tv, v \right\rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle . \tag{14}$$

Also

$$\left\langle AA^tv, v \right\rangle = \left\langle A^tv, A^tv \right\rangle \geq 0 \tag{15}$$

holds as $\langle\ .\ ,\ .\ \rangle$ is positive definite. Clearly, AA^t is nonsingular and $\langle v, v\rangle > 0$. Then comparison of (14) and (15) tells us that all eigenvalues of AA^t are positive real numbers. Let $\lambda_1, \ldots, \lambda_n$ be these eigenvalues. But we know that

$$n = \mathrm{tr}(AA^t) = \sum_{i=1}^{n} \lambda_i \tag{16}$$

and as A is real matrix with $|\det A| = 1$, we also have that

$$1 = (\det A)^2 = \det AA^t = \prod_{i=1}^{n} \lambda_i. \tag{17}$$

Combining (16) and (17), one sees that

$$\frac{1}{n}\sum_{i=1}^{n} \lambda_i = \left(\prod_{i=1}^{n} \lambda_i\right)^{\frac{1}{n}}, \tag{18}$$

where $\lambda_1, \ldots, \lambda_n$ are positive real numbers. (18) represents the equality case of the arithmetic–geometric mean inequality and it is well known that this can happen only when

$$\lambda_1 = \cdots = \lambda_n = \lambda \tag{19}$$

for some $\lambda > 0$. From (17) and (19), one infers that $\lambda^n = 1$ and this forces $\lambda = 1$. As there exists a basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$ consisting of eigenvectors of AA^t (see Theorem 4.3 on p. 219 of [14]) and all eigenvalues of AA^t are 1, AA^t fixes all of the basis elements. This shows that $AA^t = I$. Therefore, A is real unitary and A has to be similar to the desired block matrix B with respect to an orthonormal basis arising from $\langle\ .\ ,\ .\ \rangle$ (see Theorem 6.4 on p. 230 of [14]). This completes the proof of Theorem 2.

5 Proof of Theorem 3

For the necessity part of the claim, note that

$$\det B = e^{i\sum \theta_k} \text{ and } \mathrm{tr}(BB^*) = n, \tag{20}$$

where $e^{i\theta_k}$, $1 \le k \le n$ are the diagonal entries of B. Moreover, there exists a complex unitary matrix U such that

$$U^{-1}AU = B \text{ and } U^{-1} = U^*. \tag{21}$$

First from (20) and (21), $|\det A| = |\det B| = 1$ follows. Similarly as in the proof of Theorem 2, we can show, using (20) and (21), that $\mathrm{tr}(AA^*) = \mathrm{tr}(BB^*) = n$. The sufficiency part of the claim can be shown similarly as in Theorem 2 by noting that AA^* is Hermitian and has positive real eigenvalues. Thus the equality case of

the arithmetic–geometric mean inequality is again applicable. In this way, one may deduce that A is complex unitary. Finally, using Theorem 6.2 on p. 228 of [14], we see that A has to be similar to such a matrix B with respect to an orthonormal basis arising from the hermitian product $\langle\ .\ ,\ .\ \rangle$. This completes the proof of Theorem 3.

6 Proof of Corollary 1

Assume that A is complex unitary with eigenvalues $\lambda_1, \ldots, \lambda_n$. Since $\det A^* = \overline{\det A}$ (here of course $\overline{\det A}$ is the complex conjugate of $\det A$, unlike its meaning in Sect. 2), we have $|\det A| = 1$. Moreover, all eigenvalues of A are of the form $e^{i\theta}$ for some real number θ so that

$$\sum_{i=1}^{n} |\lambda_i| = n$$

holds. This completes the necessity part of the claim. For the sufficiency part of the claim, assume that

$$|\det A| = 1 \text{ and } \sum_{i=1}^{n} |\lambda_i| = n. \tag{22}$$

Since A is a normal matrix, by the spectral theorem (see p. 227 of [14]), there exists a complex unitary matrix U such that

$$U^{-1} A U = B, \tag{23}$$

where B is a diagonal matrix consisting of eigenvalues of A on the main diagonal. From (22), we see that

$$\prod_{i=1}^{n} |\lambda_i| = |\det A| = 1$$

and consequently that

$$\frac{1}{n} \sum_{i=1}^{n} |\lambda_i| = 1 = \left(\prod_{i=1}^{n} |\lambda_i| \right)^{\frac{1}{n}}. \tag{24}$$

By the equality case of the arithmetic–geometric mean inequality, (24) implies that $|\lambda_i| = 1$ for all i. Therefore, the diagonal entries of B are of the form $e^{i\theta}$ and B is complex unitary. As complex unitary matrices form a group under multiplication and U is complex unitary, one obtains from (23) that $A = UBU^{-1}$ is complex unitary as well. This completes the proof of Corollary 1.

7 Proof of Theorem 4

First assume that cN is unitary for some positive number c. Then

$$cNcN^* = c^2 NN^* = I \tag{25}$$

and it follows from (25) that

$$\det NN^* = \frac{1}{c^{2n}} \text{ and } \operatorname{tr}(NN^*) = \frac{n}{c^2}. \tag{26}$$

But we also have $\det NN^* = \det N\overline{\det N} = |\det N|^2$ so that

$$|\det N|^{\frac{2}{n}} = \frac{1}{c^2} \tag{27}$$

follows from (26). Combining (26) and (27), (1) is obtained. Conversely, assume that (1) holds. Since N is nonsingular, NN^* is a positive definite Hermitian matrix. Therefore, by the spectral theorem (see Theorem 5.3 on p. 226 of [14]), NN^* is similar to a diagonal matrix consisting of the positive eigenvalues of NN^* on the main diagonal. Using this, we see that

$$|\det N|^2 = \det NN^* = \prod_{i=1}^{n} \lambda_i \text{ and } \operatorname{tr}(NN^*) = \sum_{i=1}^{n} \lambda_i, \tag{28}$$

where $\lambda_1, \ldots, \lambda_n$ are the positive eigenvalues of NN^*. As a result of (28), (1) can be written in the form

$$\frac{1}{n}\sum_{i=1}^{n} \lambda_i = \left(\prod_{i=1}^{n} \lambda_i\right)^{\frac{1}{n}}. \tag{29}$$

Once again, (29) gives the equality case of the arithmetic–geometric mean inequality and one concludes that $\lambda_i = \lambda > 0$ for all i. Finally, NN^* is similar to the diagonal matrix λI and we get that

$$\left(\frac{1}{\sqrt{\lambda}} N\right)\left(\frac{1}{\sqrt{\lambda}} N^*\right) = I. \tag{30}$$

Consequently from (30), cN is unitary with $c = \frac{1}{\sqrt{\lambda}} > 0$ and this completes the proof of Theorem 4.

8 Proof of Theorem 5

As A and B are commuting complex matrices, we may use a special case of a classical theorem of Frobenius on the simultaneous triangulation of an arbitrary set of commuting matrices by unitary similarity. For a nice treatment and proof of this result based

on representation theory, the reader is referred to a paper of Newman [16]. Precisely, by Theorem 1 of [16], there exists a unitary matrix S such that

$$U_1 = S^*AS \text{ and } U_2 = S^*BS \tag{31}$$

are both upper triangular matrices. Therefore, A and B can be simultaneously put into triangular form with respect to an orthonormal basis of $\mathbb{C}^n$. Using (31), one obtains that

$$\operatorname{tr}(U_1U_2) = \operatorname{tr}(S^*ASS^*BS) = \operatorname{tr}(S^*ABS) = \operatorname{tr}(AB). \tag{32}$$

Again from (31), we also have

$$\operatorname{tr}(U_1^2) = \operatorname{tr}(S^*A^2S) = \operatorname{tr}(A^2) \text{ and } \operatorname{tr}(U_2^2) = \operatorname{tr}(S^*B^2S) = \operatorname{tr}(B^2). \tag{33}$$

Assume now that (2) holds for A and B. Then combining (32) with (33), we see that

$$(\operatorname{tr}(U_1U_2))^2 = \operatorname{tr}(U_1^2)\operatorname{tr}(U_2^2). \tag{34}$$

Observe that U_1 and U_2 are upper triangular matrices whose main diagonals consist of eigenvalues of A and B, respectively. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A and let $\mu_1, \ldots, \mu_n$ be the eigenvalues of B. By assumption, λ_i and μ_i are all nonzero real numbers. Moreover, U_1U_2 is also an upper triangular matrix whose main diagonal consists of the numbers $\lambda_i\mu_i$. It follows that

$$(\operatorname{tr}(U_1U_2))^2 = \left(\sum_{i=1}^{n} \lambda_i\mu_i\right)^2. \tag{35}$$

By similar reasoning, one further infers that

$$\operatorname{tr}(U_1^2) = \sum_{i=1}^{n} \lambda_i^2 \text{ and } \operatorname{tr}(U_2^2) = \sum_{i=1}^{n} \mu_i^2. \tag{36}$$

Assembling (34)–(36), we deduce that

$$\left(\sum_{i=1}^{n} \lambda_i\mu_i\right)^2 = \left(\sum_{i=1}^{n} \lambda_i^2\right)\left(\sum_{i=1}^{n} \mu_i^2\right). \tag{37}$$

Clearly, (37) represents the equality case of the Cauchy–Schwarz inequality. Then it is well known that

$$\mu_i = c\lambda_i \tag{38}$$

holds for all i, where

$$c = \left(\frac{\sum_{i=1}^{n} \mu_i^2}{\sum_{i=1}^{n} \lambda_i^2}\right)^{\frac{1}{2}} > 0. \tag{39}$$

From (2), we have

$$\prod_{i=1}^{n} \lambda_i = \det A = \det B = \prod_{i=1}^{n} \mu_i. \tag{40}$$

Consequently from (38)–(40) and the fact that A and B are nonsingular, one obtains $c^n = 1$ and $c = 1$. Therefore, A and B have the same eigenvalues, counting multiplicity, so have the same characteristic polynomial. Conversely, if A and B have the same characteristic polynomial so have the same eigenvalues, counting multiplicity, say $\lambda_1, \ldots, \lambda_n$, then we have

$$\det A = \prod_{i=1}^{n} \lambda_i = \det B. \tag{41}$$

Moreover, by (32) and (33),

$$(\operatorname{tr}(AB))^2 = (\operatorname{tr}(U_1 U_2))^2 = \left(\sum_{i=1}^{n} \lambda_i^2 \right)^2 = \operatorname{tr}(U_1^2) \operatorname{tr}(U_2^2) = \operatorname{tr}(A^2) \operatorname{tr}(B^2) \tag{42}$$

holds. From (41) and (42), one completes the proof of Theorem 5.

9 Proof of Theorem 6

The necessity part of the claim is obvious. For the sufficiency part, assume that A and A^* have the same eigenvectors. Let us show by induction on the size of the matrices that A and A^* are simultaneously diagonalizable. The base case of the induction clearly holds for 1×1 matrices A and A^*. Thus for the inductive step, we may assume that A and A^* are $n \times n$ matrices with $n > 1$. Since A is a complex matrix, it has an eigenvalue, say λ. Then $Av_1 = \lambda v_1$ holds for some unit vector v_1. Let

$$W = \{cv_1 : c \in \mathbb{C}\}$$

be the complex space spanned by v_1. Also let $\langle \, . \, , \, . \, \rangle$ be the standard positive definite Hermitian product on $\mathbb{C}^n$ so that for any two $v = (a_1, \ldots, a_n)$, $w = (b_1, \ldots, b_n)$ in $\mathbb{C}^n$, we have

$$\langle v, w \rangle = \sum_{i=1}^{n} a_i \overline{b_i},$$

where $\overline{b_i}$ is the complex conjugate of b_i. In this setting, one gets by the Gram–Schmidt orthogonalization process that

$$W \oplus W^{\perp} = \mathbb{C}^n, \tag{43}$$

where $W^{\perp}$ is the orthogonal complement of W in $\mathbb{C}^n$ with respect to the standard Hermitian product. Note that

$$\dim_{\mathbb{C}} W^{\perp} = n - 1 \tag{44}$$

follows from (43). Let us see that $W^{\perp}$ is stable under A. To this end, take any $w \in W^{\perp}$. Then $\langle cv_1, w\rangle = 0$ for all $c \in \mathbb{C}$. As A and A^* have the same eigenvectors, v_1 is an eigenvector of A^* and $A^*v_1 = \lambda_1 v_1$ holds for some eigenvalue λ_1 of A^*. Then we observe that

$$\langle cv_1, Aw\rangle = \left\langle cA^*v_1, w\right\rangle = \langle c\lambda_1 v_1, w\rangle = 0. \tag{45}$$

Therefore, from (45), $Aw \in W^{\perp}$ follows and $W^{\perp}$ is stable under A. Similarly, one can show that $W^{\perp}$ is also stable under A^*. By (44), the inductive hypothesis holds for $W^{\perp}$ and A and A^* are simultaneously diagonalizable over $W^{\perp}$. This means that there exists an orthonormal basis $\mathcal{B}$ of $W^{\perp}$ such that A and A^* are both diagonal with respect to $\mathcal{B}$. But then A and A^* are both diagonal over $\mathbb{C}^n$ with respect to $\mathcal{B} \cup \{v_1\}$. Next let v be any basis element in $\mathcal{B} \cup \{v_1\}$. Then note that since all eigenvalues of A are real,

$$Av = \mu_1 v \text{ and } A^*v = \mu_2 v \tag{46}$$

holds with real μ_1. Using (46) and the properties of the Hermitian product, one may deduce that

$$\mu_2 \langle v, v\rangle = \left\langle A^*v, v\right\rangle = \langle v, Av\rangle = \overline{\mu_1}\, \langle v, v\rangle = \mu_1 \langle v, v\rangle . \tag{47}$$

Since $\langle v, v\rangle > 0$, we obtain from (46) and (47) that $\mu_1 = \mu_2$ and consequently that

$$Av = A^*v$$

for all $v \in \mathcal{B} \cup \{v_1\}$. Thus $A = A^*$ and A is hermitian. This completes the proof of Theorem 6.

10 Proof of Theorem 7

The necessity part of the claim is obvious. For the sufficiency part, assume that AA^* and A^*A have the same eigenvectors corresponding to every common eigenvalue of AA^* and A^*A. Note that AA^* is a Hermitian matrix, so by the spectral theorem (see Theorem 5.3 on p. 226 of [14]) it is diagonalizable. It follows that there exists a basis $\mathcal{B}$ of $\mathbb{C}^n$ consisting of the eigenvectors of AA^*. Let v be any basis element of $\mathcal{B}$. Then v is an eigenvector of AA^* so that

$$AA^*v = \lambda v$$

for some real number λ. Note that AA^* and A^*A have the same eigenvalues so that λ is also an eigenvalue of A^*A. Thus by assumption, v is also an eigenvector of A^*A

corresponding to the common eigenvalue λ. It follows that

$$A^*Av = \lambda v = AA^*v.$$

As $v \in \mathcal{B}$ is arbitrary, we conclude that $AA^* = A^*A$ and A is therefore normal. This completes the proof of Theorem 7.

References

1. Alkan, E.: Multiplicative number theory with applications to modular forms and enumeration of groups. PhD thesis, University of Wisconsin, Madison (2003)
2. Alkan, E.: Nonvanishing of Fourier coefficients of modular forms. Proc. Am. Math. Soc. **131**, 1673–1680 (2003)
3. Alkan, E.: On the sizes of gaps in the Fourier expansion of modular forms. Can. J. Math. **57**, 449–470 (2005)
4. Alkan, E., Zaharescu, A.: On the gaps in the Fourier expansion of cusp forms. Ramanujan J. **16**, 41–52 (2008)
5. Dănescu, A., Vâjâitu, V., Zaharescu, A.: Unimodular matrices whose entries are squares of those of a unimodular matrix. Rev. Roum. Math. Pures Appl. **46**, 419–430 (2001)
6. Davenport, H.: Multiplicative Number Theory. Graduate Texts in Mathematics, vol. 74, 3rd edn. Springer, New York (2000)
7. Erdös, P., Wintner, A.: Additive arithmetical functions and statistical independence. Am. J. Math. **61**, 713–721 (1939)
8. Erdös, P., Kac, M.: The Gaussian law of errors in the theory of additive number theoretic functions. Am. J. Math. **62**, 738–742 (1940)
9. Griffiths, D.J.: Introduction to Quantum Mechanics, 2nd edn. Prentice Hall, Pearson (2005)
10. Hardy, G.H., Ramanujan, S.: The normal order of prime factors of a number n. Quart. J. Math. **48**, 76–92 (1917)
11. Hoffman, A.J., Taussky, O.: A characterization of normal matrices. J. Res. Nat. Bur. Stand. **52**, 17–19 (1954)
12. Komlós, J.: On the determinant of (0, 1) matrices. Stud. Sci. Math. Hung. **2**, 7–21 (1967)
13. Komlós, J.: On the determinant of random matrices. Stud. Sci. Math. Hung. **3**, 387–399 (1968)
14. Lang, S.: Linear Algebra. Undergraduate Texts in Mathematics, 3rd edn. Springer, New York (1987)
15. Neukirch, J.: Algebraic Number Theory. Grundlehren der mathematischen Wissenschaften (A Series of Comprehensive Studies in Mathematics), vol. 322. Springer, Berlin (1999)
16. Newman, M.: Two classical theorems on commuting matrices. J. Res. Nat. Bur. Stand. B **71B**, 69–71 (1967)
17. Serre, J.P.: A Course in Arithmetic. Graduate Texts in Mathematics, vol. 7. Springer, New York (1973)