ON THE HILBERT MATRIX[1]

TOSIO KATO

The present paper is concerned with the existence[2] of the eigenvalue π of the Hilbert matrix $A=((i+k-1)^{-1})$ or $A=((i+k)^{-1})$, i, $k=1, 2, 3, \cdots$. It is well known that,[3] considered as a linear operator in the Hilbert space l^2 of vectors with finite square sum of components, A is symmetric, positive-definite and bounded, the upper bound being equal to π. It is further known that[4] π is not an eigenvalue of A thus defined. However, the question has remained open whether there exists any eigenvector (not belonging to the Hilbert space) with the eigenvalue π of the matrix A.

In what follows we shall show that there exists such an eigenvector x and that x may be chosen positive. Further we shall show that x is logarithmically convex in the sense that $[x(i+1)]^2 \leq x(i) \cdot x(i+2)$.

Actually we shall establish these results for a rather wide class of matrices containing the Hilbert matrix as a special case. Our method is quite simple and elementary: we consider the dominant eigenvectors[5] of the $n\times n$ segments A_n of A and show that the ith components of these eigenvectors form (when properly normalized), for each fixed i, a monotone converging sequence; the limiting vector thus obtained being shown to be the required eigenvector of A.

These results may be of some interest in view of various numerical work[6] done recently on the segments A_n of the Hilbert matrix. Actually the present investigation was suggested by a table[7] of the dominant eigenvectors of A_n.

1. In what follows we consider matrices $A=(a(i, k))$ which may be finite or infinite, square or rectangular. In any case we assume that the indices i, k take on positive integral values starting with 1, that is, $i=1, 2, \cdots, m(A)$ and $k=1, 2, \cdots, n(A)$, where $m(A)$ and $n(A)$, which may be finite or infinite, denote respectively the number of

Received by the editors March 31, 1956.

[1] This paper was prepared in part under a National Bureau of Standards contract with The American University sponsored by the Office of Naval Research.

[2] This gives a solution to a research problem raised by Taussky [9].

[3] See Schur [7], Magnus [4; 5], Taussky [8], Hardy, Littlewood and Polya [3],

[4] See Magnus [5], Taussky [8].

[5] See paragraph 2 below.

[6] For instance, see Fairthorne and Miller [2], Savage and Lukacs [6], Todd [10].

[7] Fairthorne and Miller [2]. The writer is also indebted to Professor Forsythe for communicating his interesting numerical results regarding these eigenvectors.

the rows and columns of A. Also we consider column-vectors $x=(x(i))$ as special cases of matrices with only one column.

DEFINITION 1.1. A matrix $A=(a(i, k))$ is said to be a P-matrix if (1) A is positive (that is, all $a(i, k)$ are positive) and (2) all minor determinants of second order

$$\begin{vmatrix} a(i, k) & a(i, k+1) \\ a(i+1, k) & a(i+1, k+1) \end{vmatrix}$$

composed of four neighboring elements are non-negative.

It is convenient to regard any positive vector as a P-matrix.

Actually the restriction in (2) above that the four elements of the minor determinant be neighboring is superfluous. In fact, it follows from (2) that

$$\frac{a(i, k+1)}{a(i, k)} \leqq \frac{a(i+1, k+1)}{a(i+1, k)} \leqq \frac{a(i+2, k+1)}{a(i+2, k)} \leqq \cdots \leqq \frac{a(i+p, k+1)}{a(i+p, k)}$$

for $p>0$, and hence further that

$$\frac{a(i+p, k)}{a(i, k)} \leqq \frac{a(i+p, k+1)}{a(i, k+1)} \leqq \cdots \leqq \frac{a(i+p, k+q)}{a(i, k+q)} \tag{1.1}$$

for $p>0$, $q>0$. This shows that all minor determinants of second order are non-negative.

DEFINITION 1.2. Let $A=(a(i, k))$ and $B=(b(i, k))$ be two positive matrices. We shall write $A \ll B$ if (1) the size of A is not larger than that of B (that is, $m(A) \leqq m(B)$, $n(A) \leqq n(B)$) and (2) the ratio $c(i, k)=b(i, k)/a(i, k)$ is a monotone nondecreasing function of i and k (that is, $c(i, k) \leqq c(i+1, k)$, $c(i, k) \leqq c(i, k+1)$) as long as it is defined (that is, for $1 \leqq i \leqq m(A)$, $1 \leqq k \leqq n(A)$).

The condition (2) may also be expressed as

$$\frac{a(i+p, k+q)}{a(i, k)} \leqq \frac{b(i+p, k+q)}{b(i, k)}, \qquad p \geqq 0, q \geqq 0. \tag{1.2}$$

The special case in which A and B reduce to vectors $x=(x(i))$ and $y=(y(i))$ is particularly important. Thus we write $x \ll y$ whenever x, y are positive and $y(i)/x(i)$ is nondecreasing with i. In this case (1.2) becomes

$$\frac{x(i+p)}{x(i)} \leqq \frac{y(i+p)}{y(i)}, \qquad p \geqq 0. \tag{1.3}$$

LEMMA 1.1. *Let A, B be two finite, rectangular, positive matrices such that $A \ll B$, and let x, y be two positive vectors such that $x \ll y$. Furthermore, let the size of these matrices and vectors be such that the products Ax, By are defined. Then $Ax \ll By$ provided B is a P-matrix.*

PROOF. We have only to prove that $D(i) \geqq 0$ where

$$D(i) = \begin{vmatrix} (Ax)(i) & (By)(i) \\ (Ax)(i+1) & (By)(i+1) \end{vmatrix}$$
$$= \begin{vmatrix} \sum_j a(i, j)x(j) & \sum_k b(i, k)y(k) \\ \sum_j a(i+1, j)x(j) & \sum_k b(i+1, k)y(k) \end{vmatrix}$$
$$= \sum_j \sum_k \begin{vmatrix} a(i, j) & b(i, k) \\ a(i+1, j) & b(i+1, k) \end{vmatrix} x(j)y(k),$$

the indices j and k running from 1 to $n(A)$ and from 1 to $n(B)$ respectively.

The terms with $j=k$ on the right are non-negative by (1.2). Furthermore, the terms with $j \leqq n(A) < k$ are, if any, non-negative since

$$\frac{a(i+1, j)}{a(i, j)} \leqq \frac{b(i+1, j)}{b(i, j)} \leqq \frac{b(i+1, k)}{b(i, k)} \tag{1.4}$$

by (1.2) and (1.1).

The remaining terms can be arranged in pairs such as

$$\begin{aligned} D(i, j, k) = & \begin{vmatrix} a(i, j) & b(i, k) \\ a(i+1, j) & b(i+1, k) \end{vmatrix} x(j)y(k) \\ & + \begin{vmatrix} a(i, k) & b(i, j) \\ a(i+1, k) & b(i+1, j) \end{vmatrix} x(k)y(j), \end{aligned} \tag{1.5}$$

where $1 \leqq j < k \leqq n(A)$. We shall show that $D(i, j, k) \geqq 0$ so that $D(i) \geqq 0$ follows.

We have

$$\begin{aligned} D(i, j, k) = & \left[\frac{b(i+1, k)}{b(i, k)} - \frac{a(i+1, j)}{a(i, j)}\right] a(i, j)b(i, k)x(j)y(k) \\ & + \left[\frac{b(i+1, j)}{b(i, j)} - \frac{a(i+1, k)}{a(i, k)}\right] a(i, k)b(i, j)x(k)y(j). \end{aligned}$$

The expression in the first [ ] on the right is non-negative, since (1.4) is valid here too. Moreover, the other factors of the first term satisfy the inequalities

$$a(i, j)b(i, k) \geqq a(i, k)b(i, j),$$
$$x(j)y(k) \geqq x(k)y(j)$$

since $A \ll B$ and $x \ll y$ respectively (see (1.2) and (1.3)). We have therefore

$$D(i, j, k) \geqq \left[\frac{b(i+1, k)}{b(i, k)} - \frac{a(i+1, j)}{a(i, j)} + \frac{b(i+1, j)}{b(i, j)} - \frac{a(i+1, k)}{a(i, k)}\right] \cdot a(i, k)b(i, j)x(k)y(j).$$

But the combination of the first and the fourth terms in [ ] on the right is non-negative since $A \ll B$, as well as the combination of the second and third terms. This gives the desired result that $D(i, j, k) \geqq 0$.

2. Let A be a finite, positive, square matrix.[8] Then there is a positive eigenvalue λ, called the dominant eigenvalue, of A which is larger in absolute value than any other eigenvalues of A. There is only one linearly independent eigenvector x of A corresponding to the dominant eigenvalue, and x can be taken positive. In what follows we shall call x the dominant eigenvector of A, when x is normalized in such a way that $x(1)=1$.

The dominant eigenvector can be constructed by means of the so-called iteration method. Let x^0 be an arbitrary positive vector with the length $n(A)$ normalized by $x^0(1)=1$, and let a sequence of vectors x^r be determined successively by $x^r = \text{const. } Ax^{r-1}$, $x^r(1)=1$. Then $\lim_{r\to\infty} x^r = x$ exists and coincides with the dominant eigenvector of A.

LEMMA 2.1. *Let A, B be two positive, finite, square matrices such that $A \ll B$ and let B be a P-matrix. Let x, y be the dominant eigenvectors of A, B respectively. Then we have $x \ll y$. In particular* $0 < x(i) \leqq y(i)$ *for* $i=1, 2, \cdots, n$ *where* $n=n(A)$.

PROOF. Let x^0 and y^0 be the vectors with all components unity and with lengths equal to $n(A)$ and $n(B)$ respectively. Let us apply the iteration method described above to A and B, starting with these initial vectors x^0 and y^0 respectively. Thus we get two sequences x^r and y^r of positive vectors. Since we have obviously $x^0 \ll y^0$ (see Definition 1.2), successive application of Lemma 1.1 shows that $x^r \ll y^r$ holds for all $r=0, 1, \cdots$, for the relation $u \ll v$ is preserved when u or v is multiplied by a positive scalar. But since $\lim x^r = x$ and $\lim y^r = y$, we obtain $x \ll y$. The relation $x(i) \leqq y(i)$ follows from (1.3) and

[8] As regards the properties of positive matrices used here, see for instance Wielandt [11], where other references may also be found.

the normalization condition $x(1)=y(1)=1$.

3. We are now in a position to give our main theorem.

THEOREM I. *Let $A=(a(i,k))$, $i, k=1, 2, \cdots$, be an infinite P-matrix. Let λ_n and $x_n=(x_n(i))$ $(x_n(1)=1)$ be the dominant eigenvalue and eigenvector of the $n\times n$ segment A_n of A. Then the sequence $\{\lambda_n\}$ is increasing,*[9] *and the sequence $\{x_n\}$ is nondecreasing in the sense that $m<n$ implies $x_m \ll x_n$. In particular the sequences $\{x_n(i)\}$ with fixed i are nondecreasing. If $\lim_{n\to\infty}\lambda_n=\lambda$ is finite,*[10] *then λ is an eigenvalue*[11] *of A, and there exists a positive eigenvector x such that $Ax=\lambda x$, $x_n\ll x$, $x_n(i)\leqq x(i)$, $\lim_{n\to\infty} x_n(i)=x(i)$, $x(1)=1$.*

PROOF. It is obvious that all segments A_n are P-matrices and that $m<n$ implies $A_m\ll A_n$. It follows from Lemma 2.1 that $x_m\ll x_n$, hence in particular $0<x_m(i)\leqq x_n(i)$, for $i\leqq m<n$. For a fixed i, $\{x_n(i)\}$ is therefore a nondecreasing sequence of positive numbers. This proves the first part of Theorem I.

Suppose now that the increasing sequence $\{\lambda_n\}$ of positive numbers is bounded and let $\lim_{n\to\infty}\lambda_n=\lambda$. Then

$$\sum_{k=1}^{n} a(1,k)x_n(k) = \lambda_n x_n(1) = \lambda_n < \lambda,$$

so that $x_n(i)\leqq\lambda a(1,i)^{-1}$, $i\leqq n$. Thus the nondecreasing sequence $\{x_n(i)\}$ with a fixed i is also bounded, so that the limit $x(i) =\lim_{n\to\infty} x_n(i)$ exists for each $i=1, 2, \cdots$. Obviously we have

$$0 < x_n(i) \leqq x(i), \qquad x(1) = 1. \tag{3.1}$$

We shall now show that the infinite vector $x=(x(i))$ is an eigenvector of A for the eigenvalue λ.

We first note that for $m<n$

$$\sum_{k=1}^{m} a(i,k)x_n(k) < \sum_{k=1}^{n} a(i,k)x_n(k) = \lambda_n x_n(i), \qquad i \leqq n.$$

Let $n\to\infty$ for fixed m and i. Since $\lim \lambda_n=\lambda$ and $\lim x_n(i)=x(i)$, we obtain $\sum_{k=1}^{m} a(i,k)x(k)\leqq\lambda x(i)$. Since this is true for all m, we obtain

$$\sum_{k=1}^{\infty} a(i,k)x(k) \leqq \lambda x(i), \qquad i = 1, 2, \cdots, \tag{3.2}$$

[9] This is well known and is a simple consequence of an inclusion theorem given in Collatz [1]. Cf. also Wielandt [11].

[10] This is the case if, for instance, A is a symmetric, bounded matrix in the sense of Hilbert. Then λ is precisely the upper bound of A.

[11] It is not clear whether λ is in any way distinguished among the eigenvalues of A. In particular it is doubtful that λ is the largest eigenvalue of A.

the convergence of the infinite series on the left being established.

On the other hand, we have

$$\sum_{k=1}^{n} a(i, k)x(k) \geqq \sum_{k=1}^{n} a(i, k)x_n(k) = \lambda_n x_n(i), \qquad i \leqq n,$$

by (3.1). Letting $n \to \infty$, we obtain

$$\sum_{k=1}^{\infty} a(i, k)x(k) \geqq \lambda x(i), \qquad i = 1, 2, \cdots . \tag{3.3}$$

The two opposite inequalities (3.2) and (3.3) give the desired relation

$$\sum_{k=1}^{\infty} a(i, k)x(k) = \lambda x(i), \qquad i = 1, 2, \cdots .$$

4. Under some additional conditions on the matrix A we can get further information on the eigenvector x of Theorem I.

DEFINITION 4.1. A positive matrix $A = (a(i, k))$ is said to be column-wise logarithmically convex if $a(i, k)a(i+2, k) \geqq [a(i+1, k)]^2$ holds whenever the expressions are significant. In particular a positive vector $x = (x(i))$ is said to be logarithmically convex if $x(i)x(i+2) \geqq [x(i+1)]^2$.

The property of a positive matrix being column-wise logarithmically convex is closely connected with the relation $\ll$ introduced by Definition 1.2. To see this, we introduce two $(m-1)\times m$ matrices $U_m = (u_m(i, k))$ and $V_m = (v_m(i, k))$ defined by $u_m(i, i) = 1$, $v_m(i, i+1) = 1$, $i = 1, 2, \cdots, m-1$, all other elements being equal to zero. For any $m \times n$ matrix A, both U_mA and V_mA are $(m-1)\times n$ matrices: U_mA is obtained from A simply by omitting the last row, while V_mA is obtained by omitting the first row of A and renumbering the remaining rows. The relationship stated above is now given by the following lemma.

LEMMA 4.1. *Let A be an $m \times n$ P-matrix. Then A is column-wise logarithmically convex if and only if $U_mA \ll V_mA$.*

PROOF. By Definition 1.2 the property $U_mA \ll V_mA$ is equivalent to the condition that

$$\frac{a(i+1, k)}{a(i, k)} \leqq \frac{a(i+2, k)}{a(i+1, k)}, \qquad \frac{a(i+1, k)}{a(i, k)} \leqq \frac{a(i+1, k+1)}{a(i, k+1)}.$$

But the second of these inequalities is satisfied by the assumption that A is a P-matrix, while the first is equivalent to the condition that A be column-wise logarithmically convex.

LEMMA 4.2. *Let A be an $m \times n$ P-matrix column-wise logarithmically convex, where $m, n < \infty$. Then, for any positive vector x of length n, the vector Ax is logarithmically convex.*

PROOF. By Lemma 4.1 we have $U_m A \ll V_m A$. Since $x \ll x$ is trivially satisfied and since $V_m A$ is a P-matrix (with A), it follows from Lemma 1.1 that $U_m Ax \ll V_m Ax$. But as the positive vector Ax may be regarded as a P-matrix, we see from Lemma 4.1, applied to the vector Ax instead of to A, that the vector Ax is logarithmically convex.

LEMMA 4.3. *Let A be a finite, square P-matrix column-wise logarithmically convex. Then its dominant eigenvector x is logarithmically convex.*

PROOF. This is an immediate consequence of Lemma 4.2 and the relation $\lambda x = Ax$, $\lambda > 0$.

These lemmas lead to the following theorem.

THEOREM II. *In Theorem I let A be column-wise logarithmically convex and let* $\lim_{n\to\infty} \lambda_n = \lambda$ *be finite. Then the eigenvector x is logarithmically convex.*

PROOF. By Lemma 4.3 each eigenvector x_n of Theorem I is logarithmically convex, so that the same is true with their limit x.

5. The preceding results can be applied to the Hilbert matrix. Let us consider the generalized Hilbert matrices A_θ with the elements

$$a_\theta(i, k) = (i + k + \theta)^{-1}, \qquad i, k = 1, 2, 3, \cdots,$$

where θ is a real number and we assume in what follows that $\theta > -2$. Then it is easily verified that A_θ are P-matrices column-wise logarithmically convex, and that $A_{\theta'} \ll A_{\theta''}$ for $\theta' < \theta''$. Also it is known that A_θ are non-negative-definite, bounded matrices in the sense of Hilbert, the precise upper bounds M_θ being given by[12]

$$M_\theta = \pi/\sin \pi\theta \quad \text{for} \quad -2 < \theta \leqq -3/2,$$
$$M_\theta = \pi \quad \text{for} \quad \theta \geqq -3/2.$$

This implies[13] that the sequence of the dominant eigenvalues of the $n \times n$ segments $A_{\theta,n}$ of A_θ is bounded and has the limit M_θ for $n \to \infty$. Thus Theorems I and II show that A_θ has an eigenvalue equal to M_θ with a positive eigenvector x_θ with $x_\theta(1) = 1$. This eigenvector x_θ has the following properties.

[12] See Schur [7] and Magnus [4].

[13] See footnote 10.

(1) x_θ is positive, logarithmically convex and $\lim_{i\to\infty} x_\theta(i)=0$.

(2) $\theta'<\theta''$ implies $x_{\theta'}\ll x_{\theta''}$, in particular $x_{\theta'}(i)\leqq x_{\theta''}(i)$ for all i. Thus for larger θ, $x_\theta(i)$ is more slowly converging to zero for $i\to\infty$.

(3) For $\theta\geqq -1$, the square sum of the components of x_θ is infinite.[14]

The first part of (1) is an immediate consequence of Theorems I and II. The logarithmic convexity implies that $x_\theta(i)$ tends to a finite or infinite limit for $i\to\infty$. That the limit must be zero follows from the convergence of the series expressing the components of the left-hand side of $A_\theta x_\theta = M_\theta x_\theta$. To prove (2), we consider the dominant eigenvectors $x_{\theta,n}$ of the segments $A_{\theta,n}$ of A_θ. Then $\theta'<\theta''$ implies $A_{\theta',n}\ll A_{\theta'',n}$, hence $x_{\theta',n}\ll x_{\theta'',n}$ by Lemma 2.1, and the limiting procedure $n\to\infty$ gives $x_{\theta'}\ll x_{\theta''}$. Property (3) is known[15] for $\theta=-1$, and the result (2) shows that it is also true for $\theta>-1$.

The above results are still unsatisfactory in many respects. These questions are still open: Is x_θ the only linearly independent eigenvector to the eigenvalue M_θ? Are there other eigenvectors of A_θ, in particular, are there eigenvalues of A_θ larger than M_θ with or without positive eigenvectors? The writer wishes to discuss some of these questions on another occasion.

REFERENCES

1. L. Collatz, *Einschliessungssatz für die charakteristischen Zahlen von Matrizen*, Math. Zeit. vol. 48 (1942) pp. 221–226.

2. R. A. Fairthorne and J. C. P. Miller, *Mathematical Tables and Other Aids to Computation* vol. 3 (1948–1949) pp. 399–400.

3. G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Chapter IX, Cambridge, 1952.

4. W. Magnus, *Ueber einige beschränkte Matrizen*, Archiv der Mathematik vol. 2 (1949–1950) pp. 405–412.

5. ———, *On the spectrum of Hilbert's matrix*, Amer. J. Math. vol. 72 (1950) pp. 699–704.

6. I. R. Savage and E. Lukacs, *Tables of inverses of finite segments of the Hilbert matrix, contributions to the solutions of systems of linear equations and the determination of eigenvalues*, National Bureau of Standards Applied Mathematics Series vol. 39 (1950) pp. 105–108.

7. I. Schur, *Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, J. Reine Angew. Math. vol. 140 (1911) pp. 1–28.

8. O. Taussky, *A remark concerning the characteristic roots of the finite segments of the Hilbert matrix*, Quart. J. Math. Oxford Ser. vol. 20 (1949) pp. 80–83.

9. ———, Bull. Amer. Math. Soc. Research problem 60-3-12.

[14] Probably the square sum is finite for $-2<\theta<-3/2$, for it is quite plausible that $x_\theta=y_\theta$, where $y_\theta=\{y_\theta(i)\}=\{\Gamma(\theta+1+i)/(i-1)!\Gamma(\theta+2)\}$ is an eigenvector of A_θ with the eigenvalue M_θ and $\sum_i y_\theta(i)^2<\infty$ for $-2<\theta<-3/2$. (That y_θ is an eigenvector of A_θ is not positively stated but is essentially contained in Magnus [**4**].)

[15] Magnus [**5**].

10. J. Todd, *The conditions of the finite segments of the Hilbert matrix, contributions to the solutions of systems of linear equations and the determination of eigenvalues,* National Bureau of Standards Applied Mathematics Series vol. 39 (1950) pp. 109–116.

11. H. Wielandt, *Unzerlegbare, nicht negative Matrizen,* Math. Zeit. vol. 52 (1950) pp. 642–648.

AMERICAN UNIVERSITY AND
UNIVERSITY OF TOKYO

THE GIBBS PHENOMENON FOR BOREL MEANS

LEE LORCH

1. Statement of result.

We prove here the following

THEOREM. *Let $B_x(t)$ denote the xth Borel exponential or integral mean of the Fourier series*

$$\sum_{n=1}^{\infty} \frac{\sin nt}{n}. \tag{1}$$

Then, for given T, $0 \leqq T \leqq \infty$,

$$\lim_{x\to\infty} B_x(t_x) = \int_0^T \frac{\sin v}{v}\, dv, \tag{2}$$

where[1]

$$t_x \to 0+ \quad \textit{and} \quad xt_x \to T. \tag{3}$$

Thus, the Borel means display the same Gibbs phenomenon and have the same Gibbs ratio as classic convergence, even achieving this ratio for the same value, π, of the parameter T. Except for the last assertion, the same is true (as O. Szász has shown [5; 6]) of the generalized Euler means E_r, $0<r<1$, all of which are equivalent to the Borel summation method for Fourier series and whose Lebesgue

Presented to the Society, April 14, 1956; received by the editors January 27, 1956 and, in revised form, February 20, 1956.

[1] The assumption that $t_x \to 0+$ is redundant except when T is infinite. The more restrictive condition that $nt_n^2 \to 0$, which, again, is needed only when $T=\infty$, is imposed by O. Szász in his first discussion of the corresponding problem for generalized Euler means [5]. The analogous restriction here would also simplify the technical details of the proof, as shown in §3.