Extremal numbers of positive entries of imprimitive nonnegative matrices

Xingzhi Zhan[1]
Department of Mathematics, East China Normal University
Shanghai 200062, China
e-mail: zhan@math.ecnu.edu.cn

Abstract. We determine the maximum and minimum numbers of positive entries of imprimitive nonnegative matrices with a given imprimitivity index. One application of the results is to estimate the imprimitivity index by the number of positive entries. The proofs involve the study of a cyclic quadratic form. This completes a research initiated by M. Lewin in 1990.

Key words. nonnegative matrix, imprimitive matrix, quadratic form

AMS subject classifications. 15A48, 05D99

1. Introduction

A square nonnegative matrix A is said to be primitive if A^p is a positive matrix for some positive integer p; otherwise A is called imprimitive. A primitive matrix is necessarily irreducible. The imprimitivity index k of A is the number of eigenvalues of A whose moduli are equal to the spectral radius of A. A is primitive if $k = 1$, and imprimitive if $k > 1$.

Let A be an $n \times n$ irreducible nonnegative matrix with imprimitivity index k. Then A is permutation similar to

$$\begin{pmatrix} 0 & A_1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & A_2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & A_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & A_{k-1} \\ A_k & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{1}$$

[1]The author's research was supported by the NSFC grant 10571060.

where the zero blocks along the diagonal are square (see [1, p.32] or [2, p.71]). The matrix in (1) is called the canonical form of A.

In [3] M. Lewin proved that if an irreducible nonnegative matrix has more positive entries than zero entries then it is primitive. R.A. Brualdi (see reference 2 in [4]) observed that this result follows immediately from the canonical form (1). Let $\sigma(A)$ denote the number of positive entries of a nonnegative matrix A. For a real number x, denote by $\lfloor x \rfloor$ the largest integer not exceeding x. Suppose A is an $n \times n$ irreducible nonnegative matrix with imprimitivity index k. M. Lewin [4] proved that if $k \leq 4$ then $\sigma(A) \leq \lfloor n^2/k \rfloor$ and this bound is sharp, and if $k \geq 5$ then $\sigma(A) \leq \lfloor n^2/4 \rfloor$. The case $n = 7$ and $k = 5$ shows that the bound $\lfloor n^2/4 \rfloor$ is not sharp. In fact by the canonical form (1) it is easy to see that if a nonnegative matrix A is 7×7 and of imprimitivity index 5 then $\sigma(A) \leq 10$, while $\lfloor 7^2/4 \rfloor = 12$.

The purpose of this note is to determine the sharp upper and lower bounds for $\sigma(A)$, and consider related applications. One interesting thing here is that the rule for the maximum $\sigma(A)$ changes when k turns from 4 to 5.

2. Main results

Let A be an $n \times n$ irreducible nonnegative matrix of imprimitivity index k whose canonical form is the matrix in (1). Suppose the orders of the zero blocks on the diagonal of (1) are $x_1, x_2, \ldots, x_k$ respectively. Since the positive entries of A are contained in the submatrices A_j, $j = 1, \ldots, k$,

$$\sigma(A) \leq x_1 x_2 + x_2 x_3 + \cdots + x_{k-1} x_k + x_k x_1. \tag{2}$$

Thus to determine the maximum $\sigma(A)$ it suffices to determine the maximum value of the cyclic quadratic form

$$f(x_1, x_2, \cdots, x_k) = \sum_{i=1}^{k} x_i x_{i+1} \tag{3}$$

under the condition that $x_1, \ldots, x_k$ are positive integers and $\sum_{i=1}^{k} x_i = n$. Here $x_{k+1} = x_1$ and throughout we read the subscripts of x_i modulo k. We will also determine the minimum value of this quadratic form under the same restriction on its variables.

Theorem 1. *Given positive integers $n \geq k \geq 2$, let $x_1, \ldots, x_k$ be positive integers with $\sum_{i=1}^{k} x_i = n$. Then*

$$2n - k \leq \sum_{i=1}^{k} x_i x_{i+1}. \tag{4}$$

For any n and k this lower bound can be attained. If $k \leq 4$ then

$$\sum_{i=1}^{k} x_i x_{i+1} \leq \lfloor n^2/k \rfloor. \tag{5}$$

If $k \geq 5$ then

$$\sum_{i=1}^{k} x_i x_{i+1} \leq 2n - k + \lfloor (n-k)^2/4 \rfloor. \tag{6}$$

For any n and k the upper bounds in (5) and (6) can be attained.

Proof. Our strategy is to change successively a pair of variables without changing their sum to make the value of the quadratic form smaller or bigger. We first prove inequality (4). If $k = 2$, then

$$2x_1x_2 = 2x_1(n - x_1) \geq 2n - 2.$$

Let $k = 3$. Use the notation (3). There exists some i such that $x_i \leq x_{i+1}$. Since the value of f is invariant under cyclic permutations of x_1, x_2, x_3, we may suppose $x_1 \leq x_2$. Then

$$\begin{aligned} f(x_1 - d, x_2 + d, x_3) &= f(x_1, x_2, x_3) - d(x_2 - x_1) - d^2 \\ &\leq f(x_1, x_2, x_3) \end{aligned}$$

for any nonnegative number d. Taking $d = x_1 - 1$ we get

$$f(1, x_2 + x_1 - 1, x_3) \leq f(x_1, x_2, x_3).$$

Thus the minimum value of f can be attained at some (x_1, x_2, x_3) with $x_1 = 1$. Then we have

$$\begin{aligned} f(1, x_2, x_3) = x_2 + x_3 + x_2x_3 &= n - 1 + x_2x_3 \\ &\geq n - 1 + n - 2 \\ &= 2n - 3. \end{aligned}$$

Let $k = 4$. Since $f(x_1, x_2, x_3, x_4) = f(x_1 - d, x_2, x_3 + d, x_4)$ for any number d, we may suppose $x_1 = 1$. Then $x_2 + x_3 + x_4 = n - 1$ and

$$\begin{aligned} f(1, x_2, x_3, x_4) = x_2 + x_2x_3 + x_3x_4 + x_4 &= (x_2 + x_4)(x_3 + 1) \\ &= (n - 1 - x_3)(x_3 + 1) \\ &\geq 2n - 4. \end{aligned}$$

Now let $k \geq 5$. We use induction on k. Suppose inequality (4) holds for $k - 1$. Since

$$f(x_1, x_2 + d, x_3, x_4 - d, x_5, \cdots, x_k) - f(x_1, x_2, \cdots, x_k) = d(x_1 - x_5),$$

by choosing $d = 1 - x_2$ if $x_1 > x_5$ and choosing $d = x_4 - 1$ if $x_1 \leq x_5$ we see that the minimum value of f is attained at some $(x_1, \ldots, x_k)$ with $x_2 = 1$ or $x_4 = 1$. By cyclic permutations of the variables it suffices to consider the case $x_1 = 1$. We have

$$f(1, x_2 + x_4 - 1, x_3, 1, x_5, \cdots, x_k) \leq f(1, x_2, x_3, x_4, \cdots, x_k).$$

Thus we need consider only $f(1, x_2, x_3, 1, x_5, \cdots, x_k)$. But

$$f(1, x_2 + x_3 - 1, 1, 1, x_5, \cdots, x_k) \leq f(1, x_2, x_3, 1, x_5, \cdots, x_k)$$

if $x_3 < x_2$ and

$$f(1, 1, x_2 + x_3 - 1, 1, x_5, \cdots, x_k) \leq f(1, x_2, x_3, 1, x_5, \cdots, x_k)$$

if $x_3 \geq x_2$. Therefore in any case the minimum value of f can be attained when two consecutive variables are equal to 1. By cyclic permutations if necessary we may suppose $x_1 = x_2 = 1$. Then by the inductive hypothesis we have

$$\begin{aligned} f(1, 1, x_3, x_4, \cdots, x_k) &= 1 + f(1, x_3, x_4, \cdots, x_k) \\ &\geq 1 + 2(n - 1) - (k - 1) \\ &= 2n - k. \end{aligned}$$

This completes the proof of inequality (4). For any n and k the lower bound can be attained at $x_1 = \cdots = x_{k-1} = 1$, $x_k = n - k + 1$.

Inequality (5) and its sharpness are proved in [4]. We sketch the proof here. When $k \leq 4$, for any positive real numbers $x_1, \ldots, x_k$ with $\sum_{i=1}^{k} x_i = n$ it is not hard to show $\sum_{i=1}^{k} x_i x_{i+1} \leq n^2/k$. Thus if $x_1, \ldots, x_k$ are positive integers we have (5). By considering the division $n = km + r$ with $0 \leq r < k$ we can show that $\lfloor n^2/k \rfloor$ can be attained for each of $k = 2, 3, 4$.

Next we suppose $k \geq 5$ and prove (6). It is easy to verify that

$$f(\cdots, x_{i-2}, x_{i-1}, x_i, x_{i+1}, x_{i+2}, \cdots) \leq f(\cdots, x_{i-2}, 1, x_i, x_{i+1} + x_{i-1} - 1, x_{i+2}, \cdots) \quad (7)$$

if $x_{i-2} \leq x_{i+2}$ and

$$f(\cdots, x_{i-2}, x_{i-1}, x_i, x_{i+1}, x_{i+2}, \cdots) \leq f(\cdots, x_{i-2}, x_{i-1} + x_{i+1} - 1, x_i, 1, x_{i+2}, \cdots) \quad (8)$$

if $x_{i-2} > x_{i+2}$. Applying (7) or (8) with $i = 3$ we see that the maximum value of f can be attained when $x_2 = 1$ or $x_4 = 1$. By cyclic permutations of the variables we may suppose $x_1 = 1$. Now using (7) repeatedly we have

$$\begin{aligned} f(1, x_2, \cdots, x_k) &\leq f(1, 1, x_3, x_4 + x_2 - 1, x_5, \cdots, x_k) \\ &\leq f(1, 1, 1, x_4 + x_2 - 1, x_5 + x_3 - 1, \cdots, x_k) \\ &\leq \cdots \\ &\leq f(1, 1, \cdots, 1, a, b) \\ &= n - 1 + ab \end{aligned}$$

where a and b are positive integers and $a + b = n - k + 2$. Then

$$ab \leq \left(\frac{a+b}{2}\right)^2 = \frac{(n-k+2)^2}{4}.$$

But ab is an integer, so

$$f(1, x_2, \cdots, x_k) \leq n - 1 + ab \leq n - 1 + \lfloor (n-k+2)^2/4 \rfloor = 2n - k + \lfloor (n-k)^2/4 \rfloor.$$

This proves (6). The upper bound in (6) can be attained at $x_1 = \cdots = x_{k-2} = 1$, $x_{k-1} = x_k = (n-k+2)/2$ when $n - k$ is even and at $x_1 = \cdots = x_{k-2} = 1$, $x_{k-1} = (n-k+1)/2$, $x_k = (n-k+3)/2$ when $n - k$ is odd. □

Theorem 2. *Let $\Gamma(n, k)$ be the set of $n \times n$ irreducible nonnegative matrices with imprimitivity index k. Then*

$$\max\{\sigma(A) | A \in \Gamma(n, k)\} = \begin{cases} \lfloor n^2/k \rfloor & \text{if } 1 \leq k \leq 4, \\ 2n - k + \lfloor (n-k)^2/4 \rfloor & \text{if } k \geq 5. \end{cases}$$

Proof. The case $k = 1$ is trivial. Let $k \geq 2$. These upper bounds follow from (2) and Theorem 1. On the other hand, let A_i be the $x_i \times x_{i+1}$ matrix with all entries equal to 1, where $x_1, \ldots, x_k$ are positive integers such that $\sum_{i=1}^k x_i x_{i+1}$ attains its upper bound in (5) or (6) according as $k \leq 4$ or $k \geq 5$. Then the matrix in (1) is in $\Gamma(n, k)$ and has the maximum number of positive entries, which is just the bound in Theorem 2. $\square$

We denote by $\mathrm{ind}(A)$ the imprimitivity index of an irreducible nonnegative matrix A. One application of Theorem 2 is to estimate $\mathrm{ind}(A)$ in terms of $\sigma(A)$.

Corollary 3. *Let A be an $n \times n$ irreducible nonnegative matrix whose number of positive entries is $\sigma(A)$. If $n^2 < 4\sigma(A)$ then*

$$\mathrm{ind}(A) \leq \lfloor n^2/\sigma(A) \rfloor.$$

If $n^2 \geq 4\sigma(A)$ then

$$\mathrm{ind}(A) \leq \lfloor n + 2 - 2\sqrt{\sigma(A) - n + 1} \rfloor.$$

Proof. Denote

$$g(n,k) = \begin{cases} \lfloor n^2/k \rfloor & \text{if} \quad 1 \leq k \leq 4, \\ 2n - k + \lfloor (n-k)^2/4 \rfloor & \text{if} \quad k \geq 5. \end{cases}$$

It is not hard to verify that for a given n, $g(n, k)$ is strictly decreasing in k for $1 \leq k \leq n$. By Theorem 2 and the strict monotonicity of $g(n, k)$ in k, we deduce that if $\sigma(A) > g(n, k)$ then $\mathrm{ind}(A) \leq k - 1$.

When $n^2 < 4\sigma(A)$, we have $k_1 \equiv \lfloor n^2/\sigma(A) \rfloor + 1 \leq 4$ and

$$\sigma(A) > \lfloor n^2/k_1 \rfloor.$$

Thus $\mathrm{ind}(A) \leq k_1 - 1 = \lfloor n^2/\sigma(A) \rfloor$. Since A is irreducible, $\sigma(A) \geq n$. If $\sigma(A) = n$, the conclusion holds trivially. When $n^2 \geq 4\sigma(A)$ and $\sigma(A) \geq n + 1$, we have

$$n \geq k_2 \equiv \lfloor n + 2 - 2\sqrt{\sigma(A) - n + 1} \rfloor + 1 \geq 5$$

and

$$\sigma(A) > 2n - k_2 + \lfloor (n-k_2)^2/4 \rfloor.$$

Thus $\mathrm{ind}(A) \le k_2 - 1 = \lfloor n + 2 - 2\sqrt{\sigma(A) - n + 1} \rfloor$. This completes the proof. □

We remark that Corollary 3 includes Lewin's result that if $\sigma(A) > n^2/2$ then $\mathrm{ind}(A) = 1$, i.e., A is primitive.

The matrix of the quadratic form f in (3) is the $k \times k$ matrix $G_k/2$ where

$$G_k = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

By the way we point out two interesting properties of G_k:
(i) $\det G_{2m-1} = 2$ and $\det G_{2m} = 2((-1)^m - 1)$. Thus if $4|k$, G_k is singular and otherwise G_k is nonsingular.
(ii) When $k \ge 3$ the maximum eigenvalue of G_k is always 2.

Finally we determine the minimum number of positive entries of an irreducible nonnegative matrix with a given imprimitivity index.

Theorem 4. *Let $\Gamma(n,k)$ be the set of $n \times n$ irreducible nonnegative matrices with imprimitivity index k. Then*

$$\min\{\sigma(A) | A \in \Gamma(n,k)\} = \begin{cases} n+1 & \text{if } \; k < n, \\ n & \text{if } \; k = n. \end{cases}$$

Proof. Let $A \in \Gamma(n,k)$. Since A is irreducible, A cannot have zero rows or zero columns. Thus $\sigma(A) \ge n$. The case $k = 1$, i.e., when A is a primitive matrix, is easy and well-known: $\min\sigma(A) = n + 1$. Now suppose $k \geq 2$ and consider the matrix

$$A = \begin{pmatrix} 0 & A_1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & A_2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & A_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & A_{k-1} \\ A_k & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{9}$$

which is in its canonical form. If $k = n$, each A_i is forced to be 1×1 and A with each $A_i = (1)$ satisfies $A \in \Gamma(n, n)$ and $\sigma(A) = n$.

Next we assume $2 \leq k \leq n - 1$. An $n \times n$ irreducible nonnegative matrix with exact n positive entries has the same sign pattern as that of a permutation matrix, and hence has imprimitivity index n. Therefore

$$\min\{\sigma(A) | A \in \Gamma(n, k)\} \geq n + 1$$

for $k < n$. It remains to exhibit an $A \in \Gamma(n, k)$ which attains this lower bound.

First we consider the case $k|n$. Let $n = km$ with $m \geq 2$. Denote by I_t the identity matrix of order t. Define A to be the matrix in (9) with $A_1 = A_2 = \cdots = A_{k-1} = I_m$ and with A_k being the matrix of order m whose only $m + 1$ nonzero entries are $A_k(1, 2) = A_k(2, 3) = \cdots = A_k(m - 1, m) = A_k(m, 1) = A_k(m, 2) = 1$. Then the digraph of A is strongly connected, and hence A is irreducible. We have

$$A^k = \mathrm{diag}(A_k, A_k, \cdots, A_k).$$

Since A_k is a primitive matrix (the Wielandt matrix), $\mathrm{ind}(A_k) = 1$. Hence $\mathrm{ind}(A) = k$. So $A \in \Gamma(n, k)$ and $\sigma(A) = n + 1$.

Then we consider the case $k \nmid n$. Let $n = km + r$ with $1 \leq r < k$. Denote $e_m = (0, \cdots, 0, 1)^T \in \mathbb{R}^m$. Define A to be the matrix in (9) with $A_1 = A_2 = \cdots = A_{r-1} = I_{m+1}$, $A_{r+1} = \cdots = A_{k-1} = I_m$,

$$A_r = \begin{pmatrix} I_m \\ e_m^T \end{pmatrix}, \quad A_k = (e_m, I_m).$$

The digraph of A with vertices $1, 2, \ldots, n$ is represented by the following arcs:

$$1 \to (m+1)+1 \to \cdots \to r(m+1)+1 \to r(m+1)+m+1 \to \cdots \to (k-1)m+r+1 \to 2$$

$$2 \to (m{+}1){+}2 \to \cdots \to r(m{+}1){+}2 \to r(m{+}1){+}m{+}2 \to \cdots \to (k{-}1)m{+}r{+}2 \to 3$$

$$\cdots \quad \cdots \quad \cdots$$

$$m \to \cdots \to r(m+1)-1 \to r(m+1)+m \to r(m+1)+2m \to \cdots \to km+r \to 1$$

$$m+1 \to \cdots \to r(m+1) \to r(m+1)+m$$

$$km+r \to m+1.$$

This digraph contains two closed paths going through all the vertices. Thus it is strongly connected and A is irreducible. We also have

$$A^k = \text{diag}(B_1, B_2, \cdots, B_k)$$

where $B_1 = A_1A_2\cdots A_k$, $B_2 = A_2A_3\cdots A_kA_1$, $\ldots$, $B_k = A_kA_1\cdots A_{k-1}$. Obviously all the $B_j, j = 1, \ldots, k$, have the same nonzero eigenvalues. Observe that

$$B_k = A_kA_r$$

is irreducible and has a positive trace. So B_k is primitive and hence $\text{ind}(B_k) = 1$. Thus each $B_j, j = 1, \ldots, k$ has exactly one eigenvalue with modulus equal to the spectral radius, and consequently $\text{ind}(A) = k$. We conclude that $A \in \Gamma(n,k)$ and evidently $\sigma(A) = n + 1$. This completes the proof. □

Acknowledgment. The author is grateful to Dr. Li Qiu for many delightful discussions. This work was done while the author was visiting the Hong Kong University of Science and Technology. He thanks HKUST for its hospitality and support.

References

1. R.B. Bapat and T.E.S. Raghavan, Nonnegative Matrices and Applications, Cambridge University Press, 1997.

2. R.A. Brualdi and H.J. Ryser, Combinatorial Matrix Theory, Cambridge University Press, 1991.

3. M. Lewin, On the primitivity of a nonnegative matrix with many entries, Ars Combinatoria, 29C(1990), 41-47.

4. M. Lewin, A family of inequalities and the sparsity of imprimitive matrices, Ars Combinatoria, 36(1993), 27-31.