

## Nonnegative roots of the unit matrix

***Citation for published version (APA):***

Steutel, F. W., & Harn, van, K. (1985). *Nonnegative roots of the unit matrix*. (Eindhoven University of Technology : Dept of Mathematics : memorandum; Vol. 8504). Technische Hogeschool Eindhoven.

***Document status and date:***

Published: 01/01/1985

***Document Version:***

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

***Please check the document version of this publication:***

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

***General rights***

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

***Take down policy***

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

702602

TECHNISCHE HOGESCHOOL EINDHOVEN

Onderafdeling der Wiskunde en Informatica

Memorandum 1985-04

March 1985

NONNEGATIVE ROOTS OF THE UNIT MATRIX

by

F.W. Steutel and K. van Harn

University of Technology

Department of Mathematics and Computing Science

PO Box 513, Eindhoven

The Netherlands

## NONNEGATIVE ROOTS OF THE UNIT MATRIX

by

F.W. Steutel and K. van Harn

Abstract. The full solution is given to the equation  $A^n = I_k$ , where  $A$  is a nonnegative  $k \times k$  matrix,  $I_k$  is the  $k \times k$  unit matrix, and  $k$  and  $n \in \mathbb{N}$  are fixed.

### 1. Introduction and summary

If  $S$  is a  $k \times k$  stochastic matrix, i.e. if the entries  $s_{ij}$  of  $S$  satisfy

$$s_{ij} \geq 0 \quad (i, j = 1, \dots, k), \quad \sum_{j=1}^k s_{ij} = 1 \quad (i = 1, \dots, k),$$

then any  $k$ -variate probability generating function (pgf)  $G$  is transformed into a pgf  $G_S$  by

$$G_S(z) = G(Sz) \quad (z = (z_1, \dots, z_k)' \in [0, 1]^k).$$

If  $G_S = G$  and if  $S$  is nonsingular, then also  $G_{S^{-1}} = G$ , and it is of interest to know when  $S^{-1}$  is stochastic as well.

In Section 2 it is shown in an elementary way that this is the case if and only if  $S$  is a permutation matrix.

This result can be rephrased as follows: the stochastic matrices  $S$  for which

$$S^n = I \text{ for some } n \in \mathbb{N},$$

are exactly the permutation matrices. Both results are extended to general nonnegative matrices. In Section 3, the second result is made more precise by indicating all nonnegative solutions of the equation  $A^n = I$ , for a fixed  $n$ . For this we need a basic property of nonnegative matrices, as given e.g. in [2]. For a discussion on general roots of the unit matrix we refer to [3].

## 2. Permutation matrices

Let  $\mathcal{P}_k$  denote the group of  $k \times k$  permutation matrices, i.e. matrices representing a linear transformation of  $\mathbb{R}^k$  that only permutes coordinates (so the entries  $p_{ij}$  of  $P \in \mathcal{P}_k$  satisfy  $p_{ij} = 1$  if  $\pi(i) = j$ , and  $= 0$  otherwise, where  $\pi$  is a permutation of the index set  $\{1, \dots, k\}$ ). Note that  $\# \mathcal{P}_k = k!$  and that  $P^{-1} = P'$  for  $P \in \mathcal{P}_k$ .

### Theorem 1

A nonsingular  $k \times k$  stochastic matrix  $S$  has a stochastic inverse if and only if  $S$  is a permutation matrix ( $S \in \mathcal{P}_k$ ).

### Proof

Let  $S$  and  $S^{-1}$  both be stochastic, and let  $\|x\| := \max_i |x_i|$  for  $x = (x_1, \dots, x_k)' \in \mathbb{R}^k$ . Then  $\|Sx\| \leq \|x\|$  and  $\|S^{-1}x\| \leq \|x\|$  ( $x \in \mathbb{R}^k$ ); hence

$$\|x\| = \|S^{-1}(Sx)\| \leq \|Sx\| \leq \|x\| \quad (x \in \mathbb{R}^k),$$

and so

$$\|Sx\| = \|x\| \quad (x \in \mathbb{R}^k).$$

Taking the  $j$ -th unit vector for  $x$  we see that  $\max_i s_{ij} = 1$  ( $j=1, \dots, k$ ), i.e. each column of  $S$  contains (at least) a 1. As  $S$  is stochastic, it follows that  $S$  is a permutation matrix. The converse is trivial. ■

Remark

From the proof it is clear that we may replace (twice) "stochastic" in Theorem 1 by "sub-stochastic" (i.e. row-sums  $\leq 1$ ); see also Lemma 1 (ii).

Now let  $\tilde{\mathcal{P}}_k$  denote the group of (nonnegative) matrices  $A$  of the form  $A = RP$  where  $P \in \mathcal{P}_k$  and  $R = \text{diag}(r_1, \dots, r_k)$  with  $r_i > 0$  for all  $i$  (so  $A$  is obtained from a permutation matrix by replacing the 1's by arbitrary positive numbers). Before extending Theorem 1 to general nonnegative matrices (i.e. matrices with nonnegative entries) we state the following lemma without its simple proof.

Lemma 1

Let  $A$  be nonsingular. Then:

- (i)  $A$  has row-sums 1 if and only if  $A^{-1}$  has row-sums 1;
- (ii) If  $A$  and  $A^{-1}$  are both sub-stochastic, then  $A$  and  $A^{-1}$  are stochastic.

Theorem 2

A nonsingular nonnegative  $k \times k$  matrix  $A$  has a nonnegative inverse if and only if  $A \in \tilde{\mathcal{P}}_k$ .

Proof

Let  $A$  and  $A^{-1}$  both be nonnegative, and define  $r_i := \sum_{j=1}^k a_{ij}$  for  $i=1, \dots, k$ . As  $A$  is nonsingular, the  $r_i$  are positive. Define  $R := \text{diag}(r_1, \dots, r_k)$ , then  $S := R^{-1}A$  is stochastic, and since  $S^{-1} = A^{-1}R$  is nonnegative,  $S^{-1}$  is stochastic as well (cf. Lemma 1 (i)). From Theorem 1 it now follows that  $S \in \mathcal{P}_k$ , and hence  $A \in \tilde{\mathcal{P}}_k$ . The converse is again trivial. ■

3. Roots of the unit matrix

It is well known and easily verified that a permutation matrix  $P$  satisfies  $P^n = I$  for some  $n \in \mathbb{N}$ . Since a stochastic matrix  $S$  satisfying  $S^n = I$  has a stochastic inverse, Theorem 1 immediately yields the following result (the dimension of unit matrices is indicated by a subscript).

Theorem 3

A stochastic matrix  $S$  satisfies  $S^n = I_k$  for some  $n \in \mathbb{N}$  if and only if  $S$  is a permutation matrix ( $S \in \mathcal{P}_k$ ).

Similarly, from Theorem 2 it follows that a nonnegative matrix  $A$  satisfying  $A^n = I_k$  for some  $n \in \mathbb{N}$ , is in  $\tilde{\mathcal{P}}_k$ , where, because of  $|\det A| = 1$ , the positive entries  $r_1, \dots, r_k$  of  $A$  satisfy  $\prod_{i=1}^k r_i = 1$ . The converse of this is *not* true in general: take for  $A$  a  $2 \times 2$  diagonal matrix with entries 2 and  $\frac{1}{2}$ .

By using a simple result on the structure of nonnegative matrices, it is not hard to determine the  $\tilde{\mathcal{P}}_k$ -matrices that do satisfy  $A^n = I_k$  for some  $n \in \mathbb{N}$ . We will do so by solving the equation  $A^n = I_k$  for nonnegative  $A$  and a *fixed*  $n$  (and fixed  $k$ ).

Consider, to this end, a general nonnegative  $k \times k$  matrix  $A$ , and for  $n \in \mathbb{N}$  denote the entries of  $A^n$  by  $a_{ij}^{(n)}$ . The elements of the index set  $\{1, \dots, k\}$  of  $A$  are called the *indices* of  $A$ ; in the case of stochastic matrices they are called *states*: the matrix then describes the transitions of a Markov chain. Further, an index  $i$  is called *essential* if  $i$  leads to some  $j$  (i.e.  $a_{ij}^{(n)} > 0$  for some  $n \in \mathbb{N}$ ) and any such  $j$  leads (back) to  $i$ ; in this case  $i$  and  $j$  are said to *communicate*. Now it is easily seen that the index set of an  $A$  having only essential indices can be partitioned into classes  $C_1, \dots, C_m$  that are *self-communicating*, i.e. all the indices belonging to one class communicate but cannot lead to an index outside the class. From this it follows that by (possibly) relabelling the indices one can put  $A$  into a block-diagonal form, or, more precisely, that there is a  $P \in \mathcal{P}_k$  such that  $B := PAP'$  is a block-diagonal matrix having as many blocks as there are classes, and where the size of the  $j$ -th block  $B_j$  is given by  $\# C_j$  ( $B_j$  is essentially the restriction of  $A$  to  $C_j \times C_j$ ). For more information on the partitioning of nonnegative matrices we refer to [2], p. 15.

We are now ready to prove the main result of this note.

Theorem 4

Let  $n \in \mathbb{N}$  be fixed. Then the nonnegative solutions  $A$  of the equation

$$(1) \quad A^n = I_k$$

are given by the  $\tilde{\mathcal{P}}_k$ -matrices for which each self-communicating class  $C$  has the following properties:  $\# C$  divides  $n$ , and the product of the positive entries in the restriction of  $A$  to  $C \times C$  (i.e. in the block corresponding to  $C$ ) is one.

Proof

Since we already observed (using Theorem 2) that only  $\tilde{\mathcal{P}}_k$ -matrices can be solutions of (1) we may suppose  $A$  to be in  $\tilde{\mathcal{P}}_k$ . Such an  $A$  has only essential indices, and hence the index set can be partitioned in self-communicating classes. Further, we can find a  $P \in \mathcal{P}_k$  such that  $B := PAP'$  is a block-diagonal matrix with the following property: if  $C$  is a self-communicating class with  $\#C = v$ , say, then the block  $B_1$ , say, corresponding to  $C$  has the form

$$B_1 = \begin{bmatrix} 0 & r_1 & \dots & \dots & \dots & 0 \\ \vdots & & & & & \vdots \\ \vdots & & r_2 & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & & & & & r_{v-1} \\ r_v & 0 & \dots & \dots & \dots & 0 \end{bmatrix},$$

where  $r_1, \dots, r_v$  are the positive entries in the restriction of  $A$  to  $C \times C$  (which is indeed a  $\tilde{\mathcal{P}}_v$ -matrix). Finally, note that the smallest  $\ell \in \mathbb{N}$  for which  $B_1^\ell = c I_v$  for some  $c > 0$  is  $\ell = v$ , in which case necessarily

$$c = \prod_{i=1}^v r_i.$$

Now, if  $A^n = I_k$ , then also  $B^n = I_k$ , and hence  $B_1^v = I_v$ , from which in view of the observations above we conclude that  $v$  divides  $n$  and that  $\prod_{i=1}^v r_i = 1$ .

Conversely, if  $C$  has these properties, and similarly for the other classes (if any), then the  $n$ -th power of each block of  $B$  is the identity; hence

$$B^n = I_k, \text{ and so } A^n = I_k. \quad \blacksquare$$

Example

The  $\tilde{\mathcal{P}}_5$ -matrix A given by

$$A = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{3} \\ 2 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 \end{bmatrix}$$

is a root of the  $5 \times 5$  unit matrix:  $A^6 = I_5$ .

Specializing Theorem 4 for stochastic matrices gives the following improvement of Theorem 3.

Corollary

Let  $n \in \mathbb{N}$  be fixed. Then the *stochastic* solutions S of the equation  $S^n = I_k$  are given by the permutation matrices for which each self-communicating class C has the property that  $\#C$  divides  $n$ .

This corollary, and hence Theorem 3, can also be proved by using properties of eigenvalues of stochastic matrices (see e.g. [1]), all of which have modulus 1 if  $S^n = I$ . Further we note that the cyclically moving sub-classes for the Markov chain associated with a solution  $P \in \mathcal{P}_k$  all consist of one state, and that each class of communicating states corresponds to a cycle in the permutation associated with P; hence these cycles have lengths that divide  $n$ .

The number of stochastic solutions of  $A^n = I_k$  is, of course, bounded by  $k!$  Clearly, it depends on the divisibility properties of  $n$ . For  $A^7 = I_5$  the only

negative solution is  $A = I_5$ , whereas  $A^6 = I_5$  has many solutions, even for stochastic  $A$ .

Acknowledgements.

The problem considered in this note arose from joint work of the authors with S.J. Wolfe, University of Delaware. The present proof of Theorem 1 was suggested by F. van Schagen, Free University, Amsterdam.

References

- [1] Fritz, F.J., Hupert, B. and Willems, W. Stochastische Matrizen. Springer-Verlag, Berlin (1979).
- [2] Seneta, E. Nonnegative matrices and Markov-chains, 2<sup>nd</sup> ed. Springer-Verlag, Berlin (1980).
- [3] Wedderburn, J.H.M. Lectures on matrices, Dover, New York (1934).