# Optimization of the Spectral Radius of a Product for Nonnegative Matrices

Jonathan Axtell * Lixing Han † Daniel Hershkowitz ‡ Michael Neumann §

Nung–Sing Sze ¶

May 28, 2007

**Abstract**

Let $A$ be an $n \times n$ irreducible nonnegative matrix. We show that over the set $\Omega_n$ of all $n \times n$ doubly stochastic matrices $S$, the multiplicative spectral radius $\rho(SA)$ attains a minimum and a maximum at a permutation matrix. For the case when $A$ is a symmetric nonnegative matrix, a by-product of our technique of proof yields a result allowing us to show that $\rho(S_1A) \geq \rho(S_2A)$, when $S_1$ and $S_2$ are two symmetric matrices such that both $S_1A$ and $S_2A$ are nonnegative matrices and $S_1 - S_2$ is a positive semidefinite matrix. This result has several corollaries. One corollary is that $\rho(S_1A) \geq \rho(S_2A)$, when $S_1 = (1/n)J$ and $S_2 = (1/(n-1))(J-I)$, where $J$ is the matrix of all one's. A second corollary is a comparison theorem for weak regular splittings of two monotone matrices.

**AMS classification:** 15A48, 15A18

**Keywords:** nonnegative matrix, spectral radius, stochastic matrix

## 1 Introduction

Let $A$ be an $n \times n$ irreducible nonnegative matrix. The problem of optimizing the spectral radius of the sum $A + X$, where $X$ runs through the $n \times n$ matrices of Frobenius norm 1 or through all nonnegative diagonal matrices of a fixed trace has been considered by several researchers, see, for example Han, Neumann, and Tsatsomeros [7], Hershkowitz, Huang, Neumann, and Schneider [8], and Johnson, Loewy, Olesky, and van den Driessche [9]. In this paper, we study the problem of optimizing the spectral radius of the product $SA$, where $S$

*Department of Mathematics, University of Connecticut, Storrs, CT 06269-3009 (axtell@math.uconn.edu)

†Department of Mathematics, University of Michigan, Flint, MI 48502 (lxhan@umflint.edu)

‡Department of Mathematics, Technion, Haifa 32000, Israel (hershkow@technion.ac.il)

§Department of Mathematics, University of Connecticut, Storrs, CT 06269–3009 (neumann@math.uconn.edu). Research supported in part by NSA Grant No. 06G–232

¶Department of Mathematics, University of Connecticut, Storrs, CT 06269–3009 (sze@math.uconn.edu)

runs through the set of all $n \times n$ doubly stochastic matrices.

Denote by $\Omega_n$ and $\mathcal{P}_n$ the sets of the $n \times n$ doubly stochastic matrices and the $n \times n$ permutation matrices, respectively. Recall the well known result of Birkhoff, see [2], that $\Omega_n$ is the convex hull of $\mathcal{P}_n$, the extreme points in $\Omega_n$.

Since $\Omega_n$ is a closed and bounded set, the extremal values, both minimal and maximal, of $\rho(SA)$, where $\rho(\cdot)$ denotes the spectral radius of a matrix, are attained on $\Omega_n$. In the main result of this paper, cf. Theorem 2.1 in Section 2, we show that the extremal values are always attained on $\mathcal{P}_n$. From convex analysis we know that the extremum values of every convex function defined on $\Omega_n$ are attainable at the extremal points of $\Omega_n$. However, in general the function $S \mapsto \rho(SA)$ is not a convex function on $\Omega_n$. Notice that as $\rho(XY) = \rho(YX)$ for any matrices $X$ and $Y$, the optimization of $\rho(SAT)$, where $S$ and $T$ run through the set of all doubly stochastic matrices, is also solved. We go on to provide a variation of Theorem 2.1. For example, in Theorem 2.3, we consider the case where the optimization of $\rho(SA)$ is taken over all $n \times n$ doubly stochastic matrices $S$ which are a direct sum of $k$ doubly stochastic matrices of sizes $n_1, \cdots, n_k$, with $n_1 + \ldots + n_k = n$.

A by-product of an intermediate step in the proof of Theorem 2.1 leads us to consider the special case when $A$ is symmetric. Suppose that $S_1$ and $S_2$ are two symmetric matrices such that the difference $S_1 - S_2$ is a positive semidefinite matrix and such that $S_1A$ and $S_2A$ are nonnegative matrices. In Theorem 3.1 of Section 3 we show that

$$\rho(S_2A) \;\leq\; \rho(S_1A).$$

Suppose now that $\hat{J}_n$ is the $n \times n$ matrix of all 1's and that $J_n = (1/n)\hat{J}_n$. Set $K_n = (1/(n-1))(\hat{J}_n - I_n)$. Then on letting $S_1 = J_n$ and $S_2 = K_n$, we see that $S_1 - S_2 = \frac{1}{n-1}I_n - \frac{1}{n-1}J_n$ which is positive semidefinite. We thus obtain the corollary that for any $n \times n$ symmetric irreducible nonnegative $A$,

$$\rho(K_nA) \;\leq\; \rho(J_nA).$$

Theorem 3.1 has application to *comparison theorems for nonnegative iteration matrices.* Recall that a splitting of an $n \times n$ matrix $B$ into $B = M - N$ is called *regular* if $N \geq 0$, $M$ is invertible, and $M^{-1} \geq 0$. A celebrated comparison result due to Varga [14] states that if $B = M_1 - N_1 = M_2 - N_2$ are two regular splittings of $B$ such the $N_1 \geq N_2$, then

$$\rho(M_2^{-1}N_2) \;\leq\; \rho(M_1^{-1}N_1).$$

Since Varga's comparison theorem for regular splittings was published, many papers have appeared in the literature in which various relaxations of the conditions for a splitting to be regular have been considered. For example, Ortega and Rheinboldt [13] introduced the notation of a *weak regular splitting* in which we require that $M$ is invertible, $M^{-1} \geq 0$, and $M^{-1}N \geq 0$, while comparison theorems for weak regular splittings have been developed in Csordas and Varga [3], Elsner [4], Elsner, Frommer, Nabben, Schneider, and Szyld [5], Neumann and Miller [10], and Neumann and Plemmons [11]. Using Theorem 3.1, we are able to mix conditions involving symmetry of matrices with conditions involving nonnegativity of matrices to obtain comparison theorems for *weak regular splittings of matrices.*

Before we proceed to the development of the results of this paper, let us mention the results in two papers which are of some relevance to the present results and which may interest the reader. In the 1968 paper [1], Brualdi and Wielandt show that a matrix $A \in R^{n,n}$ is stochastic if and only if for every permutation matrix $P$, $\rho(PA) = 1$. The second paper of interest is [6] by Friedland, Hemasinha, Schneider, Stuart, and Weaver. Let $A \in \mathbb{R}^{n,n}$ be a nonnegative and irreducible matrix with $\rho(A) < 1$ so that $(I - A)^{-1}$ exists and is a positive matrix. They consider the question of when the Perron vector of $A$ and the vector of the row sums of $(I - A)^{-1}$ share the same grading, namely, that both vectors can be simultaneously permuted to vectors whose entries are nonincreasing.

# 2 The Extremal Problem $\rho(SA)$ Over the Doubly Stochastic Matrices

Let $A$ be an $n \times n$ nonnegative and irreducible matrix and, as before, let $\Omega_n$ be the set of all the $n \times n$ nonnegative doubly stochastic matrices. In the main result of this section we consider the problem of the extremal values of $\rho(SA)$ as $S$ varies over $\Omega_n$.

Recall that $\Omega_n$ is a closed and bounded set and that the spectral radius function $\rho(\cdot)$ is continuous on $\mathbb{R}^{n,n}$. Hence, for any arbitrary but fixed $n \times n$ nonnegative and irreducible matrix $A \in \mathbb{R}^{n,n}$, $\rho(SA)$, viewed as a function on $\Omega_n$, attains its bounds on $\Omega_n$. Indeed, since, by Birkhoff expansion, if $S \in \Omega_n$, then $S = \sum_{i=1}^m a_i P_i$, for some permutation matrices $P_1, \ldots, P_m$ and nonnegative numbers $a_1, \ldots a_m$, such that $\sum_{i=1}^m a_i = 1$, we can write that:

$$\rho(SA) \;\le\; \|SA\|_2 \;=\; \|\sum_{i=1}^m a_i P_i A\|_2 \;\le\; \sum_{i=1}^m a_i \|P_i A\|_2 \;=\; \sum_{i=1}^m a_i \|A\|_2 \;=\; \|A\|_2.$$

The question is thus whether we can identify where in $\Omega_n$ does $\rho(SA)$ attains its bounds. It is known that the specral radius function is not convex and hence, a priori, we do not know whether $\rho(SA)$ attains its bounds at the extreme points of $\Omega_n$, namely, in $\mathcal{P}_n$, the set of all the $n \times n$ permutation matrices. Therefore our main result is somewhat surprising.

**Theorem 2.1** *Let $A \in \mathbb{R}^{n,n}$ be a nonnegative and irreducible matrix. Then there are permutation matrices $P^*$ and $Q^*$ such that*

$$\rho(P^* A) \;=\; \min_{S \in \Omega_n} \rho(SA) \quad \textit{and} \quad \rho(Q^* A) \;=\; \max_{S \in \Omega_n} \rho(SA). \tag{2.1}$$

We will now proceed with some preliminaries to the proof of Theorem 2.1. First, a key idea is to introduce a function of two arguments, both matrices. Let $S_1$ and $S_2$ be two distinct matrices such that $S_1A$ and $S_2A$ are irreducible nonnegative matrices. Define the map $f_{S_1,S_2}$ by

$$f_{S_1,S_2}(\alpha) \;=\; \rho((\alpha S_1 + (1-\alpha)S_2)A), \quad \alpha \in [0,1]. \tag{2.2}$$

Suppose, next, that $x_\alpha$ and $y_\alpha$ are positive right and left Perron vectors of the (irreducible nonnegative) matrix $(\alpha S_1 + (1-\alpha)S_2)A$ normalized in some fixed manner. Then

$$\begin{aligned}(\beta-\alpha)y_\beta^t(S_1-S_2)Ax_\alpha &= y_\beta^t[(\beta S_1+(1-\beta)S_2)A-(\alpha S_1+(1-\alpha)S_2)A]x_\alpha \\ &= f_{S_1,S_2}(\beta)y_\beta^t x_\alpha - f_{S_1,S_2}(\alpha)y_\beta^t x_\alpha.\end{aligned}$$

It follows that

$$\frac{f_{S_1,S_2}(\beta)-f_{S_1,S_2}(\alpha)}{\beta-\alpha} \;=\; \frac{1}{y_\beta^t x_\alpha} y_\beta^t(S_1-S_2)Ax_\alpha.$$

Note that $y_\beta \to y_\alpha$, as $\beta \to \alpha$, and so

$$f'_{S_1,S_2}(\alpha) \;=\; \lim_{\beta\to\alpha}\frac{f_{S_1,S_2}(\beta)-f_{S_1,S_2}(\alpha)}{\beta-\alpha} \;=\; \frac{1}{y_\alpha^t x_\alpha} y_\alpha^t(S_1-S_2)Ax_\alpha. \tag{2.3}$$

Here, $f'_{S_1,S_2}(0)$ and $f'_{S_1,S_2}(1)$ are defined to be the corresponding usual one-sided limits.

We are now ready to present the following useful lemma:

**Lemma 2.2** *Suppose $A$, $S_1$ and $S_2$ are matrices in $\mathbb{R}^{n,n}$ such that both $S_1A$ and $S_2A$ are nonnegative and irreducible and* $\mathrm{rank}\,(S_1-S_2)=1$. *Then the map $f_{S_1,S_2}$ defined by (2.2) is either a strictly monotone function or a constant function on $[0,1]$. Furthermore, if $x$ and $y$ are the right and left Perron vectors of $S_2A$, then:*

(a) *$f_{S_1,S_2}$ is strictly increasing if $y^t(S_1-S_2)Ax \;>\; 0$.*

(b) *$f_{S_1,S_2}$ is strictly decreasing if $y^t(S_1-S_2)Ax \;<\; 0$.*

(c) *$f_{S_1,S_2}$ is a constant function if $y^t(S_1-S_2)Ax \;=\; 0$.*

*Proof.* Suppose the map $f_{S_1,S_2}$ is not strictly monotone on $[0,1]$. Then the map must have some local extremum in $(0,1)$, say at $0<\beta<1$. By (2.3),

$$0 \;=\; f'_{S_1,S_2}(\beta) \;=\; \frac{1}{y_\beta^t x_\beta} y_\beta^t(S_1-S_2)Ax_\beta.$$

Hence, $y_\beta^t(S_1-S_2)Ax_\beta=0$. Since $\mathrm{rank}\,(S_1-S_2)=1$, we have that either $y_\beta^t(S_1-S_2)A=0$ or $(S_1-S_2)Ax_\beta=0$.

If $y_\beta^t(S_1 - S_2)A = 0$, then for any $\alpha \in [0,1]$, we have

$$\begin{aligned} y_\beta^t(\alpha S_1 + (1-\alpha)S_2)A &= y_\beta^t S_2 A + \alpha y_\beta^t (S_1 - S_2)A = y_\beta^t S_2 A + \beta y_\beta^t (S_1 - S_2) A \\ &= y_\beta^t(\beta S_1 + (1-\beta)S_2)A = \rho((\beta S_1 + (1-\beta)S_2)A) y_\beta^t. \end{aligned}$$

Thus, $(\alpha S_1 + (1-\alpha)S_2)A$ and $(\beta S_1 + (1-\beta)S_2)A$ have the same spectral radius. That is, $f(\alpha) = f(\beta)$ for all $\alpha \in [0,1]$. The same result holds if $(S_1 - S_2)Ax_\beta = 0$. In both cases, the map $f_{S_1,S_2}$ is a constant function. The second part of the lemma can be easily verified by considering $f'_{S_1,S_2}(0)$. □

We are now ready present our proof of Theorem 2.1.

**Proof of Theorem 2.1.** We shall prove here only the left equality in (2.1), that is that the minimum of $\rho(SA)$ over $\Omega_n$ is attained at a permutation matrix as the proof that the maximum of $\rho(SA)$ over $\Omega_n$ is also attained at a permutation matrix can been proved along similar lines.

Suppose that $S^* \in \Omega_n$ is a matrix such that

$$\rho(S^*A) = \min_{S \in \Omega_n} \rho(SA).$$

We claim that if $S^*$ has exactly $q \geq 0$ entries equal one, then we can construct another matrix $S^\dagger \in \Omega_n$ such that $S^\dagger$ has at least $q+1$ entries equal one and $\rho(S^\dagger A) = \rho(S^* A)$. Thus, inductively, we can construct a matrix $P^*$ in $\Omega_n$ having $n$ entries equal one, which is in fact a permutation matrix, such that $\rho(P^*A) = \rho(S^*A)$. Our result will then follow.

To prove our claim, suppose $S^* = (s_{i,j})$ has exactly $q$ entries equal one. Then there are permutation matrices $P$ and $Q$ in $\mathcal{P}_n$ such that

$$PS^*Q = \begin{bmatrix} S_1^* & 0 \\ 0 & I_q \end{bmatrix},$$

for some $S_1^* \in \Omega_p$, with $p + q = n$. Without loss of generality, we may assume that $P = Q = I_n$. Otherwise, we can replace $S^*$ and $A$ by $PS^*Q$ and $Q^tAP^t$, respectively. Note that all entries of $S_1^*$, or equivalently, all $s_{i,j}$ with $1 \leq i, j \leq p$, must be less than one.

Let $x$ and $y = (y_1, \ldots, y_n)^t$ be right and left Perron vectors of $S^*A$, respectively, and set $w = (w_1, \ldots, w_n)^t = Ax$. We can further assume that

$$y_1 \geq y_2 \geq \cdots \geq y_p \quad \text{and} \quad w_1 \leq w_2 \leq \cdots \leq w_p. \tag{2.4}$$

This follows since we can further replace $S^*$, $A$, $x$, and $y$ by $PS^*Q^t$, $QAP^t$, $Px$, and $Py$, respectively, in which both $P$ and $Q$ have the form $R \oplus I_q$ in $\mathcal{P}_n$. Now let

$$u = (s_{1,1} - 1, s_{2,1}, \ldots, s_{n,1})^t = (s_{1,1} - 1, s_{2,1}, \ldots, s_{p,1}, 0, \ldots, 0)^t$$

and

$$v \;=\; (s_{1,1}-1, s_{1,2}, \ldots, s_{1,n})^t \;=\; (s_{1,1}-1, s_{1,2}, \ldots, s_{1,p}, 0, \ldots, 0)^t$$

and define the matrix

$$S^\dagger \;=\; (s^\dagger_{i,j}) \;=\; S^* + (1-s_{1,1})^{-1} u v^t.$$

Now $S^\dagger$ has the form $\begin{bmatrix} S_1^\dagger & 0 \\ 0 & I_q \end{bmatrix}$ with

$$S_1^\dagger \;=\; \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,p} \\ s_{2,1} & & & \\ \vdots & & s_{ij} & \\ s_{p,1} & & & \end{bmatrix} + \begin{bmatrix} 1-s_{1,1} & -s_{1,2} & \cdots & -s_{1,p} \\ -s_{2,1} & & & \\ \vdots & & \frac{s_{i,1}s_{1,j}}{1-s_{1,1}} & \\ -s_{p,1} & & & \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & s_{i,j}+\frac{s_{i,1}s_{1,j}}{1-s_{1,1}} & \\ 0 & & & \end{bmatrix},$$

so that $S^\dagger$ is nonnegative with at least $q+1$ entries equal 1. Furthermore, as all the row and column sums of $uv^t$ equal zero, the row and columns sums of $S^\dagger$ coincide, respectively, with those of $S^*$. Hence $S^\dagger$ is a doubly stochastic matrix.

To complete the proof it remains to be shown that $\rho(S^\dagger A) = \rho(S^* A)$. As $y$ and $w$ satisfy (2.4) we have that

$$y^t u \;=\; \sum_{i=1}^{n} s_{i,1} y_i - y_1 \;=\; \sum_{i=1}^{p} s_{i,1} y_i - y_1 \;\le\; \sum_{i=1}^{p} s_{i,1} y_1 - y_1 \;=\; y_1 - y_1 \;=\; 0$$

and

$$v^t w \;=\; \sum_{j=1}^{n} s_{1,j} w_j - w_1 \;=\; \sum_{j=1}^{p} s_{1,j} w_j - w_1 \ge \sum_{j=1}^{p} s_{1,j} w_1 - w_1 \;=\; w_1 - w_1 = 0.$$

Hence $(y^t u)(v^t w) \le 0$. Now as

$$\operatorname{rank}(S^\dagger - S^*) \;=\; \operatorname{rank}(uv^t) \;=\; 1 \quad \text{and} \quad y^t(S^\dagger - S^*)Ax \;=\; (1-s_{1,1})^{-1} y^t u v^t w \le 0,$$

by Lemma 2.2, the map $f_{S^\dagger,S^*}$ is either a strictly decreasing function or a constant function. But $f_{S^\dagger,S^*}$ cannot be strictly decreasing as $f_{S^\dagger,S^*}(0) = \rho(S^*A) \le \rho(S^\dagger A) = f_{S^\dagger,S^*}(1)$. Thus, we must have that

$$\rho(S^\dagger A) \;=\; f_{S^\dagger,S^*}(1) \;=\; f_{S^\dagger,S^*}(0) \;=\; \rho(S^* A).$$

□

A careful consideration of the proof of Theorem 2.1 shows that the proof actually works for a more general result. Specifically, we can verify that for any irreducible nonnegative matrix $A \in \mathbb{R}^{n,n}$, there is a $k \times k$ permutation $P^*$ such that

$$\rho((P^* \oplus I_{n-k})A) \;=\; \min_{S \in \Omega_k} \rho((S \oplus I_{n-k})A).$$

If we now replace $A$ by $(I_k \oplus T)A$ for some $(n-k) \times (n-k)$ doubly stochastic matrix $T$, then we obtain that

$$\rho((P^* \oplus T)A) \;=\; \min_{S \in \Omega_k} \rho((S \oplus T)A),$$

for some permutation $P^*$ in $\mathcal{P}_k$. In other words,

$$\rho((P^* \oplus T)A) \;\leq\; \rho((S \oplus T)A), \quad \text{for all } S \in \Omega_k \tag{2.5}$$

From the above developments it is readily seen that we can extend Theorem 2.1 as follows:

**Theorem 2.3** *For any $n \times n$ irreducible nonnegative $A$ and positive integers $n_1, \ldots, n_k$ with $n_1 + \cdots + n_k = n$, there exist $P_i^* \in \mathcal{P}_{n_i}$ for $i = 1, \ldots, k$, such that*

$$\rho\left((P_1^* \oplus \cdots \oplus P_k^*)A\right) \;=\; \min_{(S_1,\ldots,S_k) \in \Omega_{n_1} \times \cdots \times \Omega_{n_k}} \rho\left((S_1 \oplus \cdots \oplus S_k)A\right). \tag{2.6}$$

*Similarly, there exist $Q_i^* \in \mathcal{P}_{n_i}$ for $i = 1, \ldots, k$, such that*

$$\rho\left((Q_1^* \oplus \cdots \oplus Q_k^*)A\right) \;=\; \max_{(S_1,\ldots,S_k) \in \Omega_{n_1} \times \cdots \times \Omega_{n_k}} \rho\left((S_1 \oplus \cdots \oplus S_k)A\right). \tag{2.7}$$

*Proof.* Again we shall only prove here (2.6), the part of our theorem which is concerned with minimization, as the proof of (2.7) follows along similar lines.

Suppose $(S_1^*, \ldots, S_k^*) \in \Omega_{n_1} \times \cdots \times \Omega_{n_k}$ satisfies

$$\rho((S_i^* \oplus \cdots \oplus S_k^*)A) \;=\; \min_{(S_1,\ldots,S_k) \in \Omega_{n_1} \times \cdots \times \Omega_{n_k}} \rho\left((S_1 \oplus \cdots \oplus S_k)A\right) \tag{2.8}$$

By (2.5) with $T = S_2^* \oplus \cdots \oplus S_k^*$, there is $P_1^* \in \mathcal{P}_{n_1}$ such that

$$\rho((S_1^* \oplus S_2^* \oplus \cdots \oplus S_k^*)A) \;\geq\; \rho((P_1^* \oplus S_2^* \oplus \cdots \oplus S_k^*)A).$$

We can now proceed by an inductive argument. Suppose that $(P_1^*, \ldots, P_i^*) \in S_{n_1} \times \cdots \times S_{n_i}$ already exists. We apply (2.5) to the $(i+1)$–th diagonal block. Then just as above, there is $P_{i+1}^* \in \mathcal{P}_{n_{i+1}}$ such that

$$\rho((P_1^* \oplus \cdots \oplus P_i^* \oplus S_{i+1}^* \oplus S_{i+2}^* \oplus \cdots \oplus S_k^*)A) \;\geq\; \rho((P_1^* \oplus \cdots \oplus P_i^* \oplus P_{i+1}^* \oplus S_{i+2}^* \oplus \cdots \oplus S_k^*)A).$$

In conclusion , we have found $(P_1^*, \ldots, P_k^*) \in \mathcal{P}_{n_1} \times \cdots \times \mathcal{P}_{n_k}$ such that

$$\rho((S_1^* \oplus S_2^* \oplus \cdots \oplus S_k^*)A) \;\geq\; \rho((P_1^* \oplus S_2^* \oplus \cdots \oplus S_k^*)A) \;\geq\; \cdots \geq \rho((P_1^* \oplus \cdots \oplus P_k^*)A).$$

But then (2.8) shows that the above inequalities are indeed equalities and we are done. □

**Example 2.4** We now present an example to show that for an $n \times n$ symmetric nonnegative and irreducible matrix $A$, the extremal value of the spectral radius $\rho(SA)$, as $S$ varies over the $n \times n$ doubly stochastic matrices $\Omega_n$, can be attained at more than one extreme point of the convex set $\Omega_n$ of all stochastic matrices, but not necessarily on the interior of the line joining these points. For that purpose let

$$A = \begin{bmatrix} 0.9712 & 1.745 & 0.9995 & 1.153 & 0.5299 \\ 1.745 & 0.3860 & 1.857 & 1.849 & 0.9953 \\ 0.9995 & 1.857 & 1.627 & 1.268 & 1.099 \\ 1.153 & 1.849 & 1.268 & 1.189 & 1.308 \\ 0.5299 & 0.9953 & 1.099 & 1.308 & 0.8793 \end{bmatrix}.$$

Then for the permutation matrices

$$P_1 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad P_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix},$$

we find that the spectral radii of $\rho(P_1A)$ and $\rho(P_2A)$ is minimum and equals 6.0376, but for $S = (P_1 + P_2)/2$, $\rho(SA) = 6.0384$.

We next provide a necessary condition for a doubly stochastic matrix $S^*$ to be an extremum for $\rho(SA)$ as $S$ varies over $\Omega_n$.

**Theorem 2.5** *Let $A$ be an $n \times n$ irreducible nonnegative matrix. Suppose that $S^*$ is a matrix in $\Omega_n$ such that*

$$\rho(S^*A) = \min_{S \in \Omega_n} \rho(SA). \tag{2.9}$$

*Then*

$$y^t S^* A x = \min_{S \in \Omega_n} y^t S A x = \min_{P \in \mathcal{P}_n} y^t P A x, \tag{2.10}$$

*where $x$ and $y$ are the right and left Perron vectors of $S^*A$. Similarly, if $S^\dagger$ is a matrix in $\Omega_n$ such that*

$$\rho(S^\dagger A) = \max_{S \in \Omega_n} \rho(SA),$$

*then*

$$\tilde{y}^t S^\dagger A \tilde{x} = \max_{S \in \Omega_n} \tilde{y}^t S A \tilde{x} = \max_{P \in \mathcal{P}_n} \tilde{y}^t P A \tilde{x}, \tag{2.11}$$

*where $\tilde{x}$ and $\tilde{y}$ are the right and left Perron vectors of $S^\dagger A$.*

*Proof.* We note that the second equality in (2.10) always holds by virtue of Birkhoff's theorem which says that every doubly stochastic matrix is a linear combination of permutation matrices.

Suppose now that $S^* \in \Omega_n$ satisfies (2.9). Fix an element $S \in \Omega_n$ and consider the map $f_{S,S^*}$ defined in (2.2). Then $f_{S,S^*}$ attains its minimum at $\alpha = 0$. Hence, by (2.3), we have that:

$$\frac{1}{y^t x} y^t (S - S^*) A x \;=\; f'_{S,S^*}(0) \geq 0,$$

where $x$ and $y$ are right and left Perron vectors of $S^* A$. Thus, $y^t S A x \geq y^t S^* A x$ and hence the first equality in (2.10) is also satisfied.

The proof of (2.11) follows along similar lines. □

Note that in Theorem 2.5, if we write $S^* = \sum_{i=1}^m a_i P_i^*$ with permutation matrices $P_i^*$ and $\sum_{i=1}^m a_i = 1$ in which all the $a_i$'s are positive, then we can further deduce that

$$y^t P_i^* A x \;=\; \min_{P \in \mathcal{P}_n} y^t P A x, \quad \text{for all } i = 1, \ldots, k.$$

In particular, if the entries of each of the vectors $y$ and $Ax$ are mutually distinct, then there exists a unique $Q$ in $\mathcal{P}_n$ such that

$$y^t Q A x \;=\; \min_{P \in \mathcal{P}_n} y^t P A x.$$

Thus we have the following corollary:

**Corollary 2.6** *Let $A \in \mathbb{R}^{n,n}$ be a nonnegative and irreducible matrix. Suppose that $S^* \in \Omega_n$ satisfies (2.9) and that $x$ and $y$ are right and left Perron vectors of $S^* A$. If entries of each of the vectors $y$ and $Ax$ are mutually distinct, then $S^*$ is a permutation matrix.*

We comment that condition (2.10) in Theorem 2.5 is not a sufficient condition for $S^*$ to be a minimum as we show in the following example:

**Example 2.7** Let $A = \begin{bmatrix} 4 & 0 & 5 \\ 1 & 6 & 1 \\ 2 & 1 & 8 \end{bmatrix}$ and $Q = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. Then $\rho(QA) = 9.1394$ and the corresponding right and left Perron vectors are

$$x \;=\; (0.3163, 0.3781, 0.3056)^t \quad \text{and} \quad y \;=\; (0.2938, 0.2166, 0.4896)^t.$$

As $QAx = (2.8905, 3.4556, 2.7932)^t$, we see that $y^t Q A x = \min_{P \in \mathcal{P}_n} y^t P A x$. However, on taking $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ we have that $\rho(PA) = 9.0466 < 9.1384 = \rho(QA)$.

# 3 The Case of the Nonnegative $A \in \mathbb{R}^{n,n}$ Being Symmetric

The proof of Theorem 2.1 suggests the development of a result for a symmetric matrix $A \in \mathbb{R}^{n,n}$ which allows comparison under appropriate assumptions of the spectral radii of two matrices: $S_1A$ and $S_2A$.

**Theorem 3.1** *Suppose $A$, $S_1$ and $S_2$ are $n \times n$ symmetric matrices such that both $S_1A$ and $S_2A$ are irreducible and nonnegative. If $S_1 - S_2$ is positive semidefinite, then the map $\alpha \mapsto \rho(\alpha S_1 + (1-\alpha)S_2)A)$ is an increasing function. In particular,*

$$\rho(S_2A) \;\leq\; \rho(S_1A).$$

*Proof.* As both $S_1A$ and $S_2A$ are irreducible and nonnegative matrices, it follows that the matrix $(\alpha S_1 + (1-\alpha)S_2)A$ is an irreducible nonnegative matrix, for each $\alpha \in [0,1]$.

Suppose next that $x_\alpha$ is a right Perron vector of $(\alpha S_1 + (1-\alpha)S_2)A$. Then:

$$\begin{aligned}
[(\alpha S_1 + (1-\alpha)S_2)A]^t A x_\alpha \;&=\; A^t(\alpha S_1^t + (1-\alpha)S_2^t)Ax_\alpha \;=\; A(\alpha S_1 + (1-\alpha)S_2)Ax_\alpha \\
&=\; A\left(\rho(\alpha S_1 + (1-\alpha)S_2)x_\alpha\right) \;=\; \rho(\alpha S_1 + (1-\alpha)S_2)A)Ax_\alpha,
\end{aligned}$$

which implies that $Ax_\alpha$ is a left eigenvector corresponding to $\rho(\alpha S_1 + (1-\alpha)S_2)A)$. Let $r_\alpha$ be the sum of the entries of $Ax_\alpha$ and take $y_\alpha = \frac{1}{r_\alpha}Ax_\alpha$. Then $y_\alpha$ is the left Perron vector of $(\alpha S_1 + (1-\alpha)S_2)A$. Now by (2.3),

$$f'_{S_1,S_2}(\alpha) \;=\; \frac{1}{y_\alpha^t x_\alpha} y_\alpha^t (S_1 - S_2)Ax_\alpha \;=\; \frac{1}{(Ax_\alpha)^t x_\alpha}(Ax_\alpha)^t(S_1 - S_2)Ax_\alpha \geq 0.$$

Hence the map is an increasing function as claimed. □

An interesting corollary to Theorem 3.1 is the following:

**Corollary 3.2** *Suppose that $\hat{J}$ is the $n \times n$ matrix of all 1's. Set $J_n = (1/n)\hat{J}_n$ and $K_n = (1/(n-1))(\hat{J}_n - I_n)$. Let $A \in \mathbb{R}^{n,n}$ be a symmetric irreducible nonnegative matrix. Then:*

$$\rho(K_nA) \;\leq\; \rho(J_nA),$$

*where $\rho(J_nA)$ equals the average of the column sums of $A$.*

*Proof.* Set $S_1 = J_n$ and $S_2 = K_n$. Then

$$S_1 - S_2 \;=\; J_n - K_n \;=\; \frac{1}{n}\hat{J}_n - \frac{1}{n-1}\hat{J}_n + \frac{1}{n-1}I_n \;=\; \frac{1}{n-1}I_n - \frac{1}{n(n-1)}\hat{J}_n.$$

Now it is easily determined that the distinct eigenvalues of $(1/(n-1))I - (1/(n(n-1)))\hat{J}_n$ are 0 and $1/(n-1)$. Hence the matrix $S_1 - S_2$ is positive semidefinite and the result follows from Theorem 3.1. □

A second consequence of Theorem 3.1 is to the iterative method for solving linear systems. Given the linear system of equations $Bx = c$, with $B \in \mathbb{R}^{n,n}$, one way to solve the system is by an indirect method, namely, via an iteration scheme. One begins by spliting $B$ into $B = M - N$, with $M$ nonsingular. Then, starting from an arbitrary initial vector $x_0 \in \mathbb{R}^n$, one carries out the iteration $x_i = M^{-1}Nx_{i-1} + M^{-1}c$ and it is well known that the iteration scheme will converge to the unique solution to the system if and only if the spectral radius of the *iteration matrix* $M^{-1}N$ satisfies that $\rho(M^{-1}N) < 1$. Furthermore, $\rho(M^{-1}N)$ determines the asymptotic rate of convergence of the scheme and hence the interest in numerical analysis in being able to compare the rate of convergence of different schemes for solving the same or even two linear systems, see [3, 4, 5, 10, 11, 12, 13, 14, 15].

Recall that according to Ortega and Rheinboldt [13], a splitting of $B \in \mathbb{R}^{n,n}$ into $B = M - N$ is called a *weak regular splitting* if $M$ is invertible, $M^{-1} \geq 0$, and $M^{-1}N \geq 0$. We are now ready to state the second corollary to Theorem 3.1.

**Corollary 3.3** *Let $N \in \mathbb{R}^{n,n}$ be a symmetric nonnegative matrix. Suppose that $B_1 = M_1 - N$ and $B_2 = M_2 - N$ are two weak regular splittings of the matrices $B_1$ and $B_2$, respectively. If $M_1^{-1} - M_2^{-1}$ is positive semidefinite and both $M_1^{-1}N$ and $M_2^{-1}N$ are irreducible, then*

$$\rho(M_2^{-1}N) \;\leq\; \rho(M_1^{-1}N).$$

*Proof.* Set $A = N$, $S_1 = M_1^{-1}$, and $S_2 = M_2^{-1}$. The result now follows directly from Theorem 3.1. □

# References


[1] R. A. Brualdi and H. W. Wielandt, A spectral characterization of stochastic matrices, Lin. Alg. Appl. 1:65–71, 1968.

[2] G. Birkhoff, Tres observaciones sobre el algebra linea, Univ. Nac. Rev. Ser. A 5:147–150, 1946.

[3] G. Csordas and R. S. Varga, Comparisons of regular splittings of matrices, Numer. Math. 44:23–35, 1984.

[4] L. Elsner, Comparisons of weak regular splittings and multisplitting methods, Numer. Math. 56:283–289, 1989.

[5] L. Elsner, A. Frommer, R. Nabben, H. Schneider, and D. Szyld, Conditions for strict inequality in comparisons of spectral raddi of splittings of different matrices, Lin. Alg. Appl. 363:65–80, 2003.

[6] S. Friedland, R. Hemasinha, H. Schneider, J. Stuart, and J. Weaver, Row sums and inverse row sums for nonnegative matrices, SIAM J. Matrix Anal. and Appl. 15: 1157–1166, 1994.

[7] L. Han, M. Neumann and M. Tsatsomeros, Spectral radii of fixed Frobenius norm perturbations of nonnegative matrices, SIAM J. Matrix Anal. Appl. 21:79–92, 1999.

[8] D. Hershkowitz, W. Huang, M. Neumann and H. Schneider, Minimization of norms and the spectral radius of a sum of nonnegative matrices under diagonal equivalence, Linear Algebra Appl. 241-243:431–453, 1996.

[9] C. Johnson, R. Loewy, D. Olesky, and P. van den Driessche, Maximizing the spectral radius of fixed trace diagonal perturbations of nonnegative matrices, Linear Algebra Appl. 241-243:635–654, 1996.

[10] V. Miller and M. Neumann, A note on comparison theorems for nonnegative matrices, Numer. Math. 47:427–434, 1985.

[11] M. Neumann and R. J. Plemmons, Convergence of parallel multisplitting iterative methods for M–matrices. Lin. Alg. Appl. 88/89: 559–573, 1987.

[12] J.M. Ortega and R.J. Plemmons, Extensions of the Ostrowski-Reich theorem for SOR iterations, Lin. Alg. Appl. 28:177–191, 1979

[13] J. M. Ortega and W. C. Rheinboldt, Monotone iterations for nonlinear equations with application to Gauss–Seidel methods, SIAM J. Numer. Anal. 4:171–190, 1967.

[14] R. S. Varga, Matrix Iterative Analysis, Prentice Hall Englewood Cliffs, N.J., 1962.

[15] D. M. Young, Iterative Solutions of Large Linear Systems, Academic Press, N.Y., 1971.