2009

Calculation of Equilibrants for Semipositive Matrices

Zheng Tong
College of William & Mary - Arts & Sciences

Follow this and additional works at: https://scholarworks.wm.edu/etd

Part of the Mathematics Commons

Recommended Citation
https://dx.doi.org/doi:10.21220/s2-0h15-s322

This Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
Calculation of Equilibrants for Semipositive Matrices

Zheng Tong
Beijing, China

Master of Science, Wright State University, 2001
Bachelor of Medicine, Capital Institute of Medical Science, 1991

A Thesis presented to the Graduate Faculty
of the College of William and Mary in Candidacy for the Degree of
Master of Science

Department of Applied Science

The College of William and Mary
May, 2009
This Thesis is submitted in partial fulfillment of the requirements for the degree of Master of Science

Zheng Tong

Approved by the Committee, April 03, 2009

Charles Johnson
Committee Chair
Class of 1961 Professor Charles Johnson, Mathematics
The College of William and Mary

Gunter Luepke
Associate Professor Gunter Luepke, Applied Science
The College of William and Mary

Jianjun Paul Tian
Assistant Professor Jianjun Paul Tian, Mathematics
The College of William and Mary
For square, semipositive matrices $A$, there is an $x > 0$ such that $Ax > 0$, two (nonnegative) equilibrants $e(A)$ and $E(A)$ are defined as:

\[
e(A) = \inf_{x > 0, Ax > 0} \prod_{i=1}^{n} (Ax)_i = \prod_{i=1}^{n} \left( \inf_{x > 0, Ax > 0} x_i \right) = \prod_{i=1}^{n} \left( \inf_{x > 0, Ax > 0} x_i \right)
\]

and

\[
E(A) = \sup_{x > 0, Ax > 0} \prod_{i=1}^{n} (Ax)_i = \prod_{i=1}^{n} \left( \sup_{x > 0, Ax > 0} x_i \right) = \prod_{i=1}^{n} \left( \sup_{x > 0, Ax > 0} x_i \right)
\]

Our primary goal is to develop theory from which each may be calculated. To this end, the collection of semipositive matrices is partitioned into three subclasses for each equilibrant: (1) nonnegative matrices and (2) those that have some negative entries.

We break the nonnegative matrices down further into (1i) those that are positive diagonally equivalent to DS-matrices and (1ii) those that are not. A connection to those matrices that are scalable to doubly stochastic matrices is made. In the case of DS-scalable matrices, the “inf” in the definition of $e(A)$ is attained. For the invertible SP matrices, for which inverse is DS-scalable, the “sup” in the definition of $E(A)$ is attained. Some consequence of our results are given. In the process, a certain matrix/vector equation: $x^{(-1)} = A^T (Ax)^{(-1)}$, that is related to scalability of a matrix to one with line sums 1 is derived and discussed.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedication</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Elementary Properties of the Equilibrants</td>
<td>4</td>
</tr>
<tr>
<td>Calculating $e(A)$</td>
<td>8</td>
</tr>
<tr>
<td>Calculating $E(A)$</td>
<td>14</td>
</tr>
<tr>
<td>Observations and Special Cases</td>
<td>18</td>
</tr>
<tr>
<td>When $e(A)$ or $E(A)$ is Attained</td>
<td>22</td>
</tr>
<tr>
<td>References</td>
<td>26</td>
</tr>
<tr>
<td>Appendix A</td>
<td>27</td>
</tr>
<tr>
<td>Appendix B</td>
<td>36</td>
</tr>
</tbody>
</table>
DEDICATION

I dedicate this work to my parents who live in distant China, love me, miss me and support me in everything I do.
ACKNOWLEDGMENTS

I would like to thank professor Charles Johnson for his wisdom, kind, never ending patience and guidance during my time at the College of William and Mary. I would also like to thank my committee members professor Gunter Luepke and professor Jianjun Paul Tian for their support, also like to thank so many people for their help.
INTRODUCTION

A real matrix $A$ is called semipositive (SP) if there is a positive vector $x$ such that $Ax$ is positive. We use $>$, $\geq$ to denote entry-wise inequalities for vectors and matrices, as well as scalars. Semipositivity makes sense for rectangular matrices, but we restrict our attention to square matrices here. We are interested in two natural notions that are well-defined for SP matrices $A \in M_n(R)$ and that, after [H], we call equilibrants:

$$
eq(A) = \inf_{x > 0, Ax > 0} \frac{\prod_{i=1}^{n} (Ax)_i}{\prod_{i=1}^{n} x_i},$$

and

$$E(A) = \sup_{x > 0, Ax > 0} \frac{\prod_{i=1}^{n} (Ax)_i}{\prod_{i=1}^{n} x_i}.$$ 

In [H], only $e(A)$ was defined and slightly differently: Hoffman’s equilibrant was the $n$-th root of our $e(A)$, but there is no real theoretical difference. Also, in [H] only entry-wise nonnegative matrices were considered. Various natural questions about $e(A)$ were left unanswered in [H], and, in the meantime, there seems to have been little progress on theoretical issues about $e(A)$. For completeness, if there are positive vectors $x$ such that $Ax \geq 0$, but no positive vectors $x$ such that $Ax > 0$, we define $e(A) = 0$.

We became interested in $E(A)$ because of questions about scaling M-matrices [HJ, pp112-133] to diagonal dominance. It is known that if $A$ is an M-
matrix, then there are positive diagonal similarities, $D^{-1}AD$, of $A$ that are (row) diagonally dominant, i.e. that have positive row sums, and positive diagonal equivalences, $DAE$, that are row and column diagonally dominant. Of course, M-matrices are SP. (Analogous remarks maybe made about real H-matrices with positive diagonal entries [HJ, pp123-125].)

Generally, matrix calculations are more accurate in the presence of diagonal dominance (We were interested in the LU factorization of M-matrices), so that (approximate) optimizing of diagonal dominance seems natural. One natural measure of latent diagonal dominance is the "sup" of the product of row sums of $D^{-1}AD$ among positive diagonal matrices $D$ such that $D^{-1}AD$ is row diagonally dominant (has positive row sums in the case of M-matrices). A simple calculation shows that this "sup" is just $E(A)$, a simpler version.

Our interest here lay in considering these two equilibrants together and in developing their properties. Each has natural, but different, motivations. Our emphasis is upon how $e(A)$ and $E(A)$ may be calculated. By this we mean the evaluation of $e(A)$ and $E(A)$ as opposed to numerical aspects of their computation. For this reason, we are interested in when the "inf" ("sup") in the definition of $e(A)$ ($E(A)$) is a "min" ("max"). It turns out that these are related. When the "inf" is a "min" ("sup" is a "max"), we can either determine it directly or determine it by solving a related problem in which the "inf" is a "min" ("sup" is a "max"). When the "inf" is a "min" ("sup" is a "max"), a solution is typically interior, in which case critical points of the objective are relevant. These critical points are solutions of an intriguing matrix/vector equation that we derive.
In the process, we give other interesting properties of $e(A)$ and $E(A)$.

First we mention some elementary properties of $e(A)$ and $E(A)$ that we use.
ELEMENTRY PROPERTIES OF THE EQUILIBRANTS

Here we list several properties of $e(A)$ and $E(A)$ that are easily proven.

It is clear from the definitions that for each SP matrix $A \in M_n(R)$,

(1a) \[ 0 \leq e(A) < \infty \]

(1b) \[ 0 < E(A) \leq \infty \]

Of course, the first equality can occur only when the "inf" is not a "min" because if the "inf" is a "min", 0 can be actually attained, $e(A) = 0$, from the definition we will see $\inf_{x > 0, A \geq 0} \prod_{i=1}^{n} (Ax)_i = 0$. This come into conflict with $A$ is SP. And the second equality can occur only when the "sup" is not a "max" because $\infty$ never can be actually attained. Also, it is clear that, whether or not the "inf" ("sup") is attained, we have for any SP matrix $B \in M_n(R)$ and any $A \in M_n(R)$, $A \geq B$,

(2a) \[ e(A) \geq e(B) \]

(2b) \[ E(A) \geq E(B) \]

Of course, B is SP and $A \geq B$ implies A is SP, but equality in (2a), need not imply equality between the matrices, unless the "inf" is a "min" ("sup" is a "max") in the definition of both $e(A)$ and $e(B)$ ($E(A)$ and $E(B)$).

Since positive diagonal scaling will be important (and clearly preserves SP), it should be observed that positive diagonal scaling has a simple effect upon $e(A)$ and $E(A)$.
For an SP matrix $A$ and positive diagonal matrices $D$ and $F, D = \text{diag}(d_1, d_2, \cdots, d_m), F = \text{diag}(f_1, f_2, \cdots, f_m)$, note that $DAF$ remains SP; we then have

\begin{align*}
(3a) \quad e(DAF) &= (\det DF)e(A) \\
(3b) \quad E(DAF) &= (\det DF)E(A) \\
(3c) \quad e(D^{-1}AD) &= e(A) \\
(3d) \quad E(D^{-1}AD) &= E(A).
\end{align*}

Proof of (3a): From the definition,

\[ e(DAF) = \inf_{x>0, Ax>0} \prod_{i=1}^{n} (DAFx)_i = \inf_{x>0, Ax>0} \prod_{i=1}^{n} \left( \prod_{j=1}^{m} (d_{ij}f_{ij}(Ax))_j \right) = \inf_{x>0, Ax>0} \prod_{i=1}^{n} \left( \prod_{j=1}^{m} (d_{ij}f_{ij}(Ax))_j \right) \]

\[ = \inf_{x>0, Ax>0} \prod_{i=1}^{n} (d_{ii}f_{ii}(Ax)) = \det(DF) \inf_{x>0, Ax>0} \prod_{i=1}^{n} (A) = \det(DF)e(A). \]

This completes the proof. □

The (3b), (3c) and (3d) are also can be proved by same logic.

A nonnegative matrix $A \in M_\mathbb{R}$ with the property that all its row (column) sums are +1 is said to be a row (column) stochastic matrix because each row (column) may be thought of as a discrete probability distribution on a sample space with $n$ points. If all its row and column sums are one, it is said to be a doubly stochastic matrix (DS-matrix) [HJ, pp40]. If a nonnegative matrix can be scaled to be a (doubly) stochastic matrix by a positive diagonal matrix, it is called "DS-scalable".
Now, when an SP matrix $A$ is reducible, it is straightforward to break down the calculation of $e(A)$. If

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

is SP, then $A_{22}$ is SP; if $A_{11}$ is also SP, then

$$(4a) \quad e(A) = e(A_{11})e(A_{22})$$

Of course, if $A_{12} = 0$, $A_{11}$ must be SP. If $A_{12} = 0$, then the “inf” is a “min” in the definition of $e(A)$ if and only if this is so for both $A_{11}$ and $A_{22}$; moreover, if the “inf” is a “min”, the “min” cannot be uniquely attained. If $A_{12} \neq 0$, the “inf” cannot be a “min”. If $A_{11}$ is not SP (but $A$ is), we shall see that $e(A) = 0$. The issue of $E(A)$ is more subtle, unless $A_{12} = 0$. In that event, we have

$$(4b) \quad E(A) = E(A_{11})E(A_{22})$$

If $A_{12} \neq 0$, the above equality need not hold. An inequality may occur in either direction.

Permutation equivalence changes neither semipositivity, nor the value of $e(A)$ and $E(A)$. This is obvious for left permutation; on the right, the same feasible vectors are input to $Ax$, just in a different order. If $A \in M_n(R)$ is semipositive and $P$ and $Q$ are permutation matrices, then $PAQ$ is SP and

$$(5a) \quad e(PAQ) = e(A)$$

$$(5b) \quad E(PAQ) = E(A).$$
Finally, note that if $A$ is invertible,

\begin{equation}
A \text{ is SP if and only if } A^{-1} \text{ is SP.}
\end{equation}
CALCULATING $e(A)$

Now, we consider $e(A)$; it is useful to break down the possible SP argument $A$ into:

(1) nonnegative matrices and

(2) those that have some negative entries.

We break the nonnegative matrices down further into

(1i) those that are positive diagonally equivalent to DS-matrices and

(1ii) those that are not.

Recall that any nonnegative matrix $A$ is permutation equivalent to one of the form

$$
\begin{bmatrix}
A_{11} & A_{12} & \cdots & A_{1k} \\
0 & \ddots & \cdots & \cdots \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & A_{kk}
\end{bmatrix}
$$

(7)

in which each $A_{ij}$ is either irreducible via permutation equivalence or a one-by-one 0 matrix. If no diagonal block in the form (7) is 0, we call $A$ nondegenerate, otherwise it is degenerate. Also, it is well known [BPS] and easy to see that $A$ is DS-scalable if and only if $A_{ij} = 0$ for $i < j$ and $A$ is nondegenerate. We first consider those nonnegative matrices that are DS-scalable.

**Theorem 1.** Let $A \in M_n(R)$ satisfy $A \geq 0$ and $A = DBE$, with $B$ doubly stochastic and $D, E$ positive diagonal matrices. Then $e(A) = \det DE$, and the “inf”
in the definition of $e(A)$ is attained, i.e. is a “min”. Moreover, the vector that attains $e(A)$ is unique up to positive scalar multiples if and only if $A$ is irreducible under permutation equivalence.

Proof: By the (3a), $e(DBE) = \det(DE)e(B)$. Thus, it suffices to assume that $A$ is doubly stochastic and $D = E = I$. For the first claim, we show that (a) $e(A) \geq 1$ and (b) that 1 is attained. For (a), it was shown in [JK, theorem 1] that for $A$ doubly stochastic and $x > 0$, $\prod_{i=1}^{n} (Ax)_i \geq \prod_{i=1}^{n} x_i$, which implies that $e(A) \geq 1$.

For (b), consider the vector $x = e = (1,1,1,...,1)^T$. Then $\prod_{i=1}^{n} (Ax)_i = \prod_{i=1}^{n} x_i = 1$.

This means that $e(A) = 1$ and that the "inf" is a "min".

If $A$ is irreducible under permutation equivalence, it also follows from [JK, theorem 1] that equality is attained in $\prod_{i=1}^{n} (Ax)_i \geq \prod_{i=1}^{n} x_i$ for a positive vector $x$ if and only if is a constant vector, i.e. a positive multiple of e. If $A$ is reducible under permutation equivalence, there exist permutation matrices $P$ and $Q$ such that

$$PAQ = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

with $A_1 \in M_{n_1}$ and $A_2 \in M_{n_2}$ and with $A_1$ and $A_2$ DS. Then, any vector

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \text{ with } x_1 \in R^{n_1} \text{ and } x_2 \in R^{n_2}, x_1 > 0 \text{ and constant, } x_2 > 0 \text{ and}$$
independently constant, satisfies \( \prod_{i=1}^{n} (Ax_i) = \prod_{i=1}^{n} x_i \), so that \( x \) need not itself be a constant vector. This completes the proof. \( \square \)

In the event that \( A \geq 0 \), but is not DS-scalable, \( A \) must be reducible via permutation equivalence, and we may assume that \( A \) is in the form (7) with \( k > 1 \) and some \( A_{ij} \neq 0 \), \( i < j \). For such \( A \geq 0 \), we define \( \hat{A} \) to be that matrix that agrees with \( A \), except that all blocks \( A_{ij} \), \( i < j \) are replaced by 0. As a result \( \hat{A} \geq 0 \) is DS-scalable if \( A \) was nondegenerate. If \( A \) was degenerate, then \( \hat{A} \) is not SP, but \( e(\hat{A}) = e(A) = 0 \). If \( A \) is not in the normal form (7), by \( \hat{A} \) we mean the matrix obtained from a normal form of \( A \) and then permuted back to be comparable to \( A \). We then have

**Theorem 2.** Suppose \( A \in M_n \) and \( A \geq 0 \) and that \( A \) is not DS-scalable. Then if \( A \) is nondegenerate, we have \( e(A) = e(\hat{A}) > 0 \) and the “inf” in the definition of \( e(A) \) is not attained. If \( A \geq 0 \) is degenerate, then \( e(A) = 0 \).

Proof: If \( A \in M_n \), \( A \geq 0 \), then \( A \) is permutation equivalent to a matrix in form (7) in which each \( A_{ii} \in M_{n_i} \) is irreducible under permutation equivalence or \( A_{ii} = 0 \in M_1 \), \( i = 1,2,\ldots,k \).

Since \( A \) is not DS-scalable, at least one super-diagonal block in \( A \) must be zero.

suppose \( A \) is nondegenerate; then
\[
\hat{A} = \begin{bmatrix}
A_{11} & 0 & \ldots & 0 \\
0 & \ldots & \ldots \\
\ldots & \ldots & 0 \\
0 & \ldots & 0 & A_{kk}
\end{bmatrix}
\]

\(A \geq \hat{A}\), so, by (2a), \(e(A) \geq e(\hat{A})\), which is positive because of nondegeneracy. By (4a), \(e(\hat{A}) = e(A_{11})e(A_{22})\ldots e(A_{kk})\).

Now, consider \(e(A)\) and let
\[
D(\varepsilon) = \begin{bmatrix}
I_1 & 0 & 0 & 0 \\
0 & \varepsilon I_2 & 0 & 0 \\
0 & 0 & \ldots & 0 \\
0 & 0 & 0 & \varepsilon^{k-1}I_k
\end{bmatrix}
\]
in which \(I_j\) is an identity matrix with the same number of rows as \(A_{jj}\).

Then, for \(A_\varepsilon = D(\varepsilon)^{-1}AD(\varepsilon)\), \(e(A) = e(A_\varepsilon)\), by (3c). Since \(A_\varepsilon \to \hat{A}\) as \(\varepsilon \to 0\), \(e(A) = e(\hat{A})\). The "inf" is not attained, as, for each positive vector \(x\), \(\hat{A}x > 0\) and
\[
\prod_{i=1}^{n}(Ax)_i > \prod_{i=1}^{n}(\hat{A}x)_i.
\]

If \(A\) is degenerate, then some \(A_{ij} = 0\), and \(e(\hat{A}) = e(A_{11})e(A_{22})\ldots e(A_{kk}) = 0\), and, \(e(A) = e(\hat{A}) = 0\). This completes the proof. \(\square\)

Example 1. If
\[
A = \begin{bmatrix}
2 & 1 & 2 \\
0 & 1 & 2 \\
0 & 12 & 2
\end{bmatrix}
\]
then by theorem 2, $e(A) = e\left(\hat{A}\right)$ for
\[
\hat{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 12 & 2 \end{bmatrix}.
\]

Now, $\hat{A}$ is DS-scalable by virtue of its pattern, and, in particular,
\[
\hat{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 12 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 3 & 3 \\ 0 & 3 & 3 \end{bmatrix},
\]
in which the second factor is doubly stochastic. It then follows from theorem 1 and 3(a) that $e(\hat{A}) = 36$ and from theorem 2 that $e(A) = 36$.

Finally, if $A \in M_n(R)$ is SP but $A$ is not $\geq 0$, we have

**Theorem 3.** If $A \in M_n(R)$ is SP, and has some negative entries, then $e(A) = 0$.

**Proof:** Let
\[
A = \begin{bmatrix} \cdot & A_1 & \cdot \\ \cdot & A_2 & \cdot \\ \cdot & A_n & \cdot \end{bmatrix},
\]
a partition of $A$ by rows. Suppose there are some negative entries in row $A_i$. The matrix $A'$ is obtained by replacing any negative entries in the other rows with 0's. Then $A' \geq A$, so, $e(A') \geq e(A)$.
Pick a vector $x > 0$, $x \perp A_j$, then $A_j x = 0$. Let $j$ be the index of a positive entry of $A_i$, $e_j$ the $j^{th}$ unit vector, $t > 0$, and define $x' = x + te_j$. Then,

$$A_i x' = 0 + t(A_i)_j > 0,$$

and $A' x' > 0$. Now, $\prod_{i=1}^n (A' x')_i > 0$ for any $t > 0$, but, as $t \to 0$,

$$\prod_{i=1}^n (A' x')_i \to \prod_{i=1}^n (A' x)_i = 0.$$ This implies that $e(A') = \inf_{x > 0} \frac{\prod_{i=1}^n (A' x')_i}{\prod_{i=1}^n x'_i} = 0$. Since $e(A) \leq e(A')$ and $e(A)$ cannot be negative, we must have $e(A) = 0$. This completes the proof. □
CALCULATING $E(A)$

In order to understand how to obtain $E(A)$ for an SP $A$, it is again useful to partition the possible $A$'s into subclasses. If $A$ is invertible, those subclasses depend upon the nature of $A^{-1}$, relative to the partition used in the prior section, and we reduce the calculation of $E(A)$ to the calculation of $e(A^{-1})$ as given in the prior section. If $A$ is singular, we will see that $E(A) = \infty$. First, it is convenient to have a general lemma relating $e(A)$ and $E(A^{-1})$.

**Lemma 4.** Suppose that $A \in M_n(\mathbb{R})$ and that $BA = I$. Then, $A$ is SP if and only if $B$ is SP, and, if $A$ is SP, $e(A) > 0$ if only if $E(B) < \infty$. Moreover, if $A$ is SP and $e(A) > 0$, then $e(A)E(B) = 1$, and the “inf” in the definition of $e(A)$ is a “min” if and only if the “sup” in the definition of $E(B)$ is a “max”.

Proof: If $x \in M_n(\mathbb{R})$, then $BAx = lx = x$, and $Ax = y$ if and only if $By = x$. Now, if $A$ is SP, we may choose $x > 0$, so that $y > 0$ and $B$ is SP because $By = x$. If $B$ is SP, the argument $A$ is SP is the same, reversing the roles of $y$ and $x$.

If $A$ is SP and $e(A) > 0$,

$$e(A) = \inf_{x > 0,Ax > 0} \prod_{i=1}^{n} (Ax)_i = \inf_{y > 0, Ay > 0} \prod_{i=1}^{n} y_i = \left( \sup_{y > 0,A^{-1}y > 0} \prod_{i=1}^{n} (A^{-1}y)_i \right)^{-1} = E(A^{-1})^{-1} = E(B)^{-1}.$$ 

So, $e(A)E(B) = 1$, and the “inf” in the definition of $e(A)$ is a “min” if and only if the “sup” in the definition of $E(B)$ is a “max”. 

14
This means, in particular that for $A$ SP, $e(A) > 0$ if and only if $E(B) < \infty$.

This completes the proof. □

Now, if $A$ is invertible and $A^{-1}$ is DS-scalable, we have our first observation about how $E(A)$ maybe calculated.

**Theorem 5.** Suppose $A \in M_n(R)$ is invertible. If $A^{-1} \geq 0$, and $A^{-1}$ is DS-scalable. Then the “sup” in definition of $E(A)$ is a “max” and $E(A) = e(A^{-1})^{-1}$.

Proof: By theorem 1, because $A^{-1}$ is DS-scalable, so $e(A^{-1})$ is attained, and is a “min”. By lemma 4, $E(A)$ is also attained, and is a “max”. Also $e(A^{-1})E(A) = 1$, so that, $E(A) = e(A^{-1})^{-1}$. This completes the proof. □

**Theorem 6.** Suppose that $A \in M_n(R)$ is SP and invertible. If $A^{-1} \geq 0$ but it is not DS-scalable, then the “sup” in the definition of $E(A)$ is not attained, but $E(A) = e(A^{-1})^{-1} = e(A^{-1})^{-1}$.

Proof: Because $A^{-1} \geq 0$ but it is not DS-scalable, by theorems 1 and 2, the $e(A^{-1})$ cannot attained, “inf” is not “min”, and by lemma 4, the “sup” in the definition of $E(A)$ is not attained, also by lemma 4, $E(A) = e(A^{-1})^{-1} = e(A^{-1})^{-1}$.

This completes the proof. □

**Example 2.** The inverse of matrix $A$ in computation example 1 is

$$A^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 2 & 1 \\ 0 & -1 & 11 \\ 0 & 6 & \frac{1}{11} \\ 11 & 11 & 1 \end{bmatrix}.$$
\[
\hat{A}^{-1} = \begin{bmatrix}
2 & 0 & 0 \\
0 & 1 & 2 \\
0 & 12 & 2
\end{bmatrix}^{-1} = \left( \begin{bmatrix}
2 & 0 & 0 \\
0 & 3 & 0 \\
0 & 6 & 0
\end{bmatrix} \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 2 \\
0 & 3 & 3
\end{bmatrix} \right)^{-1} = \begin{bmatrix}
1 & 0 & 0 \\
0 & -1 & 2 \\
0 & 2 & -1
\end{bmatrix} \begin{bmatrix}
\frac{1}{2} & 0 & 0 \\
0 & \frac{1}{3} & 0 \\
0 & 0 & \frac{1}{6}
\end{bmatrix},
\]

in which the first factor is DS. It then follows from Lemma 4 and theorem 5 that

\[E(\hat{A}^{-1}) = \frac{1}{36}\]

and from theorem 5 that \[E(A^{-1}) = \frac{1}{36}\].

**Theorem 7.** Suppose \(A \in M_n(R)\) is SP and invertible. If \(A^{-1}\) has some negative entries, then \(E(A)\) is infinite.

Proof: Theorems 3 shows \(e(A^{-1}) = 0\) if \(A^{-1}\) has some negative entries, and lemma 4 shows \(E(A) = e(A^{-1})^{-1}\). So, \(E(A)\) is infinite.

This completes the proof. \(\square\)

**Theorem 8.** If \(A \in M_n(R)\) is SP but not invertible, then \(E(A)\) is infinite.

Proof: We identify a sequence of vector \(z\) such that \(z > 0, \, Az > 0\), but

\[
\prod_{i=1}^{n}(Az_i) \to 0, \text{ while } \prod_{i=1}^{n}(z_i) \text{ is bounded below. This means that } \frac{\prod_{i=1}^{n}(Az_i)}{\prod_{i=1}^{n}(z_i)} \text{ is not bounded above and that } E(A) = \infty.
\]

Suppose that \(x \neq 0\) and \(Ax = 0\); the existence of such an \(x\) being guaranteed by the singularity of \(A\). Suppose also that \(z > 0, \, Az > 0\); the existence of such a \(z\) being guaranteed by the semipositivity of \(A\). WLOG, (perhaps by replacing \(x\) with \(-x\)), we may suppose that \(z > 0\) is such that

16
$z + tx \geq 0$ and $z + tx$ is not greater than 0. Let $T$ be the set of indices $i$ for which $(z + tx)_i = 0$. Now, the finiteness of $t$ guarantees that $P = \max \prod_{i \in T} (z + sx)_i$ is positive but finite.

Now, consider the set of vectors $z + sx$, $0 \leq s \leq t$, and let $s$ approach $t$. Then $\prod_{i=1}^n (z + sx)_i = P \prod_{i \in T} (z + sx)_i$. Since $\prod_{i \in T} (z + sx)_i \to 0$ as $s \to t$, the first product approaches 0 as $s \to t$.

However $A(z + sx) = Az > 0$ and $z + sx > 0$ for all $0 \leq s \leq t$, giving the promised set of vectors and completing the proof. □
OBSERVATIONS AND SPECIAL CASES

Before we show some special cases, we summarize the theorems which proved in sections of "Calculating $e(A)$" and "Calculating $E(A)$".

Calculating $e(A)$:
- **DS-scalable:** $e(A) = (\det DB)^{\frac{1}{n}}$, $DAE$ is DS
- **Nonnegative**
  - Not DS-scalable: "inf" is greater than or equal to 0, but not attained.
- **SP**
  - Some negative entries: "inf" is equal to 0, but not attained.

Calculating $E(A)$:
- $A^{-1}$ is nonnegative
  - **DS-scalable:** $E(A) = e(A^{-1})^{-1}$
- **invertible**
  - Not DS-scalable: "sup" is but not attained, $E(A) = e(A^{-1})^{-1}$.
  - $A^{-1}$ has some negative entries: $E(A)$ is infinite.
- **SP**
  - Not invertible: $E(A)$ is infinite.

For $\phi \subset \alpha \subset \{1, 2, \ldots, n\}$, $A[\alpha]$ is the principal submatrix of $A \in M_n(R)$ lying in rows and columns $\alpha$, while $A(\alpha)$ is the principal submatrix resulting from deletion of rows and columns $\alpha$. If $A \in M_n(R)$ is nonnegative, it is clear that
$A[\alpha]$ and $A(\alpha)$ being SP implies that $A$ is SP. In this event, a calculation shows that

$$e(A) \geq e(A[\alpha])e(A(\alpha)).$$

This, of course, occurs when $A > 0$. Note that every proper submatrix of $A$ is SP if and only if $A > 0$.

Examples. In general, $A$ being SP does not mean that $A[\alpha]$ or $A(\alpha)$ need be SP, as shown by the example

$$A_1 = \begin{bmatrix} -2 & 4 \\ 5 & -1 \end{bmatrix},$$

nor do $A[\alpha]$ and $A(\alpha)$ being SP mean that inequality (8) holds, as shown by the example

each of $A[\alpha], A(\alpha)$ is SP unless the matrix is, e.g., positive. (8) is not for all SP matrices. For example:

$$A_2 = \begin{bmatrix} 4 & -2 \\ -1 & 5 \end{bmatrix}.$$

As $e(A_2) = 0$, while $e(A_2[1]) = 4, e(A_2[2]) = 5$.

By choice of the vector $x > 0$, we may make the value of $\prod_{i=1}^{n} (Ax)_i$ as large as possible, when $A$ is positive. Thus,

$$E(A) = \infty \text{ for } A > 0.$$
If $A \geq 0$ and DS-scalable, and $D$ and $E$ are positive diagonal matrices such that $DAE$ is a DS-matrix, we conclude from theorem 1 that

(10) \[ e(A) = (\det DE)^{-1}. \]

If $A$ is a permutation matrix,

(11) \[ e(A) = E(A) = 1. \]

A monomial matrix is a square matrix having one and only one nonzero entry per row and column. Thus, a monomial matrix is simply the product of a permutation matrix and a nonsingular diagonal matrix. The inverse of an nonsingular, nonnegative matrix is not usually nonnegative; the only exception is a nonnegative monomial matrix[6]. For a monomial matrix $\prod_{i=1}^{n} (Ax)_i$, is just the product of the nonzero entries of $A$ times the entries of $x$ and a monomial matrix is SP if and only if it is nonnegative.

(12) \[ e(A) = E(A) = \text{the product of positive entries of } A. \]

If $A \geq 0$ is SP and not monomial, note that, generalizing (9), we have

(9') \[ E(A) = \infty. \]

If the inverse of $A$ is DS-scalable and $DA^{-1}E$ is a DS-matrix, from theorem 1 and lemma 4, we have

(13) \[ E(A) = \det DE. \]

If $A$ is a DS-matrix, but not a permutation matrix, then

(14) \[ E(A) = \infty. \]
Theorem 9. If $A \in M_n(R)$ is SP, the following are equivalent:

(i) $E(A) = e(A)$;

(ii) $E(A) < \infty$, $e(A) > 0$;

and

(iii) $A \geq 0$ is monomial.

Proof:

(i) $\Rightarrow$ (ii):

Suppose $E(A) = e(A)$. Because of 1(a), $E(A) \neq \infty$ and because of 1(b), $e(A) \neq 0$, so that both are finite and positive.

(ii) $\Rightarrow$ (iii):

It then follows from theorem 8 that $A$ is invertible. Since $e(A) > 0$, $A \geq 0$, and since $E(A) < \infty$, $A^{-1} \geq 0$ (because of lemma 4). For $A \geq 0$ and $A^{-1} \geq 0$, it is known [F], that $A$ is a nonnegative monomial matrix.

(iii) $\Rightarrow$ (i):

It follows from (12), that if $A$ is a nonnegative monomial matrix, $E(A) = e(A)$. □

Theorem 9 is, of course, a strong converse to observation (12), and we use it and other observations to better understand the vectors for which $e(A)$ and $E(A)$ are attained, as well as scalability, in the next section.
WHEN $e(A)$ OR $E(A)$ IS ATTAINED

We know from section 3 (4) that $e(A)(E(A))$ is attained by a vector $x > 0$ precisely when $A(A^{-1})$ is a nonnegative matrix that is DS-scalable. Furthermore, in all non-attainable cases, the value of $e(A)$ or $E(A)$ may be deduced from results of sections 3 or 4 or be reduced to the DS-scalable case. If $A^{-1}$ is DS-scalable, then $A$ may be scalable by positive diagonal matrices to achieve row and column sums 1, and it can happen that non-SP matrices may be scaled by invertible (not necessarily positive) diagonal matrices to achieve row and column sums 1. Finally, note from the definition ($Ax = y, x, y > 0$) that $A$ is SP if and only if it may be positive diagonally scaled ($D^{-1}_x ADD_x$) to have row sums 1. Here, as throughout, we use the notation that $D_x$ is the diagonal matrix whose diagonal entries are the entries of $x$, so that $D_x e = x$, in which we use $e$ (without confusion) to be the vector of 1's of appropriate size.

If for an SP matrix $A$, either $e(A)$ or $E(A)$ is attained, the vector $x > 0$ is a critical point of the Lagrangian

$$L(x, \lambda) = \prod_{i=1}^{n} (Ax)_i - \lambda \left( \prod_{i=1}^{n} x_i - 1 \right)$$

for the obvious associated optimization problem. According to theorem 9, not both $e(A)$ and $E(A)$ can be attained for the same matrix, except for the simple case in which $A$ is monomial. Henceforth, we consider critical points of this Lagrangian.
Differentiation reveals (see appendix A) that a (necessarily) totally nonzero vector $x$ being a critical point means that

$$\left[ \prod_{i=1}^{n} (Ax)_{i} \right] A^{T} (Ax)^{(-1)} = \lambda x^{(-1)},$$

in which we use, as throughout, that $\gamma^{(-1)}$ denotes the entry-wise inverse of a totally nonzero vector $\gamma$.

Multiplication on the left by $x^{T}$ implies that $\lambda = \prod_{i=1}^{n} (Ax)_{i}$. Since this product can be neither 0 nor $\infty$, division yields the matrix/vector (nonlinear) equation

$$A^{T} (Ax)^{(-1)} = x^{(-1)}. \quad (15)$$

Solution of this equation then becomes of interest. In particular if $e(A)$ or $E(A)$ is attained, the attaining vector will be a solution. In all other cases in which $e(A)$ is positive or $E(A)$ is finite, evaluating $e(A)$ or $E(A)$ reduces to the attained case. In addition solutions to equation (15) generally relate to scaling to achieve row and column sums 1.

If $x$ is a totally nonzero solution to (15), with $(Ax)$ totally nonzero, we may calculate:

$$D_{Ax}^{-1}AD_{x}e = D_{Ax}^{-1}Ax = e$$

and

$$e^{T} D_{Ax}^{-1}AD_{x} = (Ax)^{(-1)} AD_{x} = \left[ A^{T} (Ax)^{(-1)} \right]^{T} D_{x} = x^{(-1)^{T}} D_{x} = e.$$ 

This means that the scaled matrix has row and column sums 1, $D_{Ax}^{-1}$ is positive diagonal matrix that diagonal entries are $(Ax)_{i}^{-1}$. 

23
On the other hand, if $F$ and $E$ are invertible diagonal matrices such that

$$FAE$$

has row and column sums 1, let $x = Ee$. Then $FAEe = e$ implies $FAx = e$ or that $Ax$ is totally nonzero and $F = D_x^{-1}$. We may then calculate

$$e^T = e^T FAE = (Ax)^{-1}r^T AE,$$

which means

$$(Ax)^{-1}r^T A = e^T E^{-1} = x^{-1}r^T.$$ 

Transposing, we obtain equation (15) for this choice of $x$. 

We record the above ideas as

**Theorem 10.** There exist invertible diagonal matrices $F$ and $E$ such that $FAE$ has row and column sums 1 if and only if the equation

$$x^{-1}r = A^T (Ax)^{-1}$$

has a totally nonzero solution $x$ such that $Ax$ is totally nonzero. In this event we may take

$$E = D_x \ (or \ x = Ee)$$

and

$$F = D_{Ax}^{-1} \ (or \ Fe = (Ax)^{-1}).$$

The matrix $A$ is DS-scalable if and only if $A \geq 0$ and there is a solution $x > 0$.

We close by noting several things. Solvability of equation (15) characterizes some sort of scaling to constant (1) line sums. We know of no
effective characterization of matrices for which (15) has a solution (Here, one could consider an arbitrary field), nor the best algorithm to solve (15) when it has a solution. There is always the ambiguity of a scale factor in $x$, but, even aside from this, there may be multiple solutions, and we do not know when "Uniqueness" occurs. It appears that all logical possibilities occur for possible solutions: of course, a positive solution for a nonnegative matrix, but also a mixed sign solution for a nonnegative matrix and a positive solution for a mixed sign matrix, as well as a mixed sign solution for a mixed sign matrix. Nonunique solution may occur for a nonnegative matrix, and a mixed sign solution may produce a scaling to a nonnegative matrix.
REFERENCES


APPENDIX A

HOW TO OBTAIN THE MATRIX/VECTOR FORMULA

\[ x^{(-1)} = A^T (Ax)^{(-1)} \]

Suppose matrix \( A \) is SP, it can be scaled to DS-matrix. Let

\[ A = [a_{ij}], \quad D = \text{diag}(x_1, x_2, \ldots, x_n), \quad D^{-1} = \text{diag}(\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}), \quad x_i > 0, \prod_{i=1}^{n} x_i = 1. \]

Then

\[ f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} D^{-1} A D e = (a_{11} + a_{12} \frac{x_2}{x_1} + \ldots + a_{1n} \frac{x_n}{x_1}) \]

\[ (a_{21} \frac{x_1}{x_2} + a_{22} + \ldots + a_{nn} \frac{x_n}{x_2}) \ldots \ldots \]

\[ (a_{n1} \frac{x_1}{x_n} + a_{n2} \frac{x_2}{x_n} + \ldots + a_{nn} \frac{x_n}{x_n}) \]

\[ = \frac{1}{x_1 x_2 \ldots x_n} \left( a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \right) \left( a_{21} x_1 x_2 + a_{22} x_2 + \ldots + a_{2n} x_n \right) \ldots \ldots \]

\[ (a_{n1} x_1 + a_{n2} x_2 + \ldots + a_{nn} x_n) \]

\[ = (Ax)_1 (Ax)_2 \ldots (Ax)_n \quad \text{[use } (Ax)_i \text{ represent } \sum_{j=1}^{n} a_{ij} x_j \text{].} \]
set $\phi(x_1, x_2, \ldots, x_n) = x_1 x_2 \ldots x_n - 1$, the Lagrange function is

$$F(x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n) - \lambda \phi(x_1, x_2, \ldots, x_n) = (Ax)_1 (Ax)_2 \ldots (Ax)_n - \lambda (x_1 x_2 \ldots x_n - 1),$$

$\lambda$ is Lagrange Multiplier.

Once the value of $\lambda$ is determined, back to the original number of variables and so can go on to find the minimum (and/or maximum) value of $f(x_1, \ldots, x_n)$.

Take derivative respect to $x_i$ and let them equal to zero: $\frac{\partial F}{\partial x_i} = 0$

$$\frac{\partial F}{\partial x_1} = 0 \Rightarrow a_1 (Ax)_2 \ldots (Ax)_n + (Ax)_1 a_2 (Ax)_3 \ldots (Ax)_n + (Ax)_1 (Ax)_2 a_3 (Ax)_4 \ldots (Ax)_n + \ldots + (Ax)_1 \ldots (Ax)_{n-1} a_n - \lambda x_1 x_2 \ldots x_n = 0$$

$$\frac{\partial F}{\partial x_2} = 0 \Rightarrow a_2 (Ax)_2 \ldots (Ax)_n + (Ax)_1 a_2 (Ax)_3 \ldots (Ax)_n + (Ax)_1 (Ax)_2 a_3 (Ax)_4 \ldots (Ax)_n + \ldots + (Ax)_1 \ldots (Ax)_{n-1} a_n - \lambda x_1 x_2 \ldots x_n = 0$$

$$\frac{\partial F}{\partial x_3} = 0 \Rightarrow a_3 (Ax)_2 \ldots (Ax)_n + (Ax)_1 a_2 (Ax)_3 \ldots (Ax)_n + (Ax)_1 (Ax)_2 a_3 (Ax)_4 \ldots (Ax)_n + \ldots + (Ax)_1 \ldots (Ax)_{n-1} a_n - \lambda x_1 x_2 x_3 \ldots x_n = 0$$

$$\frac{\partial F}{\partial x_4} = 0 \Rightarrow a_4 (Ax)_2 \ldots (Ax)_n + (Ax)_1 a_2 (Ax)_3 \ldots (Ax)_n + (Ax)_1 (Ax)_2 a_3 (Ax)_4 \ldots (Ax)_n + \ldots + (Ax)_1 \ldots (Ax)_{n-1} a_n - \lambda x_1 x_2 x_3 x_4 \ldots x_n = 0$$

......
\[
\frac{\partial F}{\partial x_n} = 0 \implies \\
a_{1n}(Ax)^2 + (Ax)a_{2n}(Ax)^3 + \ldots + (Ax)_{n-1}a_{nn}(Ax)^n = 0
\]

\[
\frac{\partial F}{\partial \lambda} = 0 \implies \\
x_1x_2\ldots x_n - 1 = 0
\]

Sum of all \( \frac{\partial F}{\partial x_i} = 0 \), get:

\[
\begin{bmatrix}
a_{11} & a_{21} & a_{31} & \ldots & a_{n1} \\
a_{12} & a_{22} & a_{32} & \ldots & \ldots \\
a_{13} & a_{23} & a_{33} & \ldots & \ldots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_{1n} & \ldots & \ldots & \ldots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
\prod_{i=1}^{n} (Ax)_i \\
\prod_{i=1}^{n} (Ax)_i \\
\prod_{i=1}^{n} (Ax)_i \\
\prod_{i=1}^{n} (Ax)_i
\end{bmatrix}
- \lambda
\begin{bmatrix}
\prod_{i=1}^{n} x_i \\
\prod_{i=1}^{n} x_i \\
\prod_{i=1}^{n} x_i \\
\prod_{i=1}^{n} x_i
\end{bmatrix}
= 0,
\]

where \((Ax)_i = \sum_{j=1}^{n} a_{ij}x_j\).

that is:
Note: 

\[
\begin{bmatrix}
  a_{11} & a_{21} & a_{31} & \cdots & a_{n1} \\
  a_{12} & a_{22} & a_{32} & \cdots & a_{n2} \\
  a_{13} & a_{23} & a_{33} & \cdots & a_{n3} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn}
\end{bmatrix}
\]

\[= A^T\]

Change the entries in 

\[
\begin{bmatrix}
  \prod_{i=1}^{n} (Ax)_i \\
  \prod_{i=1}^{n} (Ax)_i \\
  \prod_{i=1}^{n} (Ax)_i \\
  \vdots \\
  \prod_{i=1}^{n} (Ax)_i
\end{bmatrix}
\]

and \(\lambda\) 

\[
\begin{bmatrix}
  \prod_{i=1}^{n} x_i \\
  \prod_{i=1}^{n} x_i \\
  \prod_{i=1}^{n} x_i \\
  \vdots \\
  \prod_{i=1}^{n} x_i
\end{bmatrix}
\]

to the product of when 

\[i=1, \ldots, n\]
Multiply \[ \frac{1}{(Ax)_1} \quad 0 \]
at left side
\[
\begin{bmatrix}
\frac{1}{x_1} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \frac{1}{x_n}
\end{bmatrix}
\]
at right side, it doesn't change the equality.

i.e. \[ A^T \]
\[
\begin{bmatrix}
\prod_{i=1}^{n} (Ax)_i \\
\prod_{i=1}^{n} (Ax)_i \\
\prod_{i=1}^{n} (Ax)_i
\end{bmatrix}
\]

\[
\begin{bmatrix}
\frac{1}{x_1} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \frac{1}{x_n}
\end{bmatrix}
\]
is equal to \[ \lambda \]
\[
\begin{bmatrix}
\prod_{i=1}^{n} x_i \\
\prod_{i=1}^{n} x_i \\
\prod_{i=1}^{n} x_i
\end{bmatrix}
\]
équation (1.1)
Then, equation (1.1) becomes: 

$$\prod_{i=1}^{n} (A_i x_i) = \prod_{i=1}^{n} (A_x) = \prod_{i=1}^{n} \lambda x_i = \prod_{i=1}^{n} \lambda x_i$$ 

Then, equation (1.1) becomes to:
\[
\begin{pmatrix}
\frac{1}{(Ax)_1} & 0 \\
\vdots & \ddots \\
0 & \cdots & \frac{1}{(Ax)_n}
\end{pmatrix}
\]

\[
A^T \begin{bmatrix}
\prod_{i=1}^n (Ax)_i \\
\end{bmatrix} = \lambda \begin{bmatrix}
\prod_{i=1}^n x_i \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 \\
1 \\
\vdots \\
1 \\
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
\]

\[
\prod_{i=1}^n (Ax)_i \quad \text{and} \quad \prod_{i=1}^n x_i \quad \text{are scalars.}
\]

Multiplying both sides on the left by \(e^T D_e\), gives:

\[
\begin{pmatrix}
\frac{1}{(Ax)_1} & 0 \\
\vdots & \ddots \\
0 & \cdots & \frac{1}{(Ax)_n}
\end{pmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} A^T
\begin{bmatrix}
\prod_{i=1}^n (Ax)_i \\
\end{bmatrix} = \lambda \begin{bmatrix}
\prod_{i=1}^n x_i \\
\end{bmatrix}
\]

\[
(Ax)_1 (Ax)_2 \ldots (Ax)_n e = \lambda x_1 \ldots x_n n
\]

-----equation (1.2)

Set \(P_e = (Ax)_1 (Ax)_2 \ldots (Ax)_n = \prod_{i=1}^n (Ax)_i\), and because \(x_1 \ldots x_n = 1\), so,

equation (1.2) become to:

33
\[ x^T A^T (Ax)^{(-1)} P_A e = \lambda n \]

\[ P_A (Ax)^T (Ax)^{(-1)} e = \lambda n \]

\[ P_A e^T e = \lambda n \]

\[ P_A = \lambda \]

The scalar \( \lambda = P_A = (Ax)_1 (Ax)_2 \ldots (Ax)_n = \prod_{i=1}^n (Ax)_i . \)

Now, value of scalar \( \lambda \) is determined, substitute it to the equation (1.1), find the critical points of \( x \), and so can go on to find the minimum (and/or maximum) value of \( f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n D^{-1} AD e \).

\[
\begin{bmatrix}
1 \\
x_1 \\
\vdots \\
\vdots \\
\vdots \\
x_n
\end{bmatrix}
= A^T
\begin{bmatrix}
\frac{1}{(Ax)_1} & \ldots & 0 \\
\frac{1}{(Ax)_2} & \ldots & 0 \\
\frac{1}{(Ax)_3} & \ldots & 0 \\
\vdots & \ddots & \vdots \\
0 & \ldots & \frac{1}{(Ax)_n}
\end{bmatrix}
\theta
\]

That means the value of critical point \( \frac{1}{x_i} \) are the row sum of

\[
\begin{bmatrix}
\frac{1}{(Ax)_1} & \ldots & 0 \\
\frac{1}{(Ax)_2} & \ldots & 0 \\
\vdots & \ddots & \vdots \\
0 & \ldots & \frac{1}{(Ax)_n}
\end{bmatrix}
A^T
\]

34
Then,
\[ \frac{1}{x_i} = \sum_{j=1}^{n} \frac{a_{ji}}{(Ax)_j} = \sum_{j=1}^{n} \frac{a_{ji}}{\sum_{j=1}^{n} (a_{ij}x_j)_{i=1,...,n}} \]

are the critical points for finding the minimum (and/or maximum) value of

\[ \prod_{i=1}^{n} D^{-1}ADe \], which is minimum (and/or maximum) value of \( \prod_{i=1}^{n} (Ax)_i \), where

\[ A = [a_{ij}], \quad (Ax)_i = \sum_{j=1}^{n} (a_{ij}x_j)_{i=1,...,n} \] And, \( \min_{x>0} \prod_{i=1}^{n} (Ax)_i = e(A) \), \( \max_{x>0} \prod_{j=1}^{n} (Ax)_j = E(A) \),

so, these critical points also are those of finding \( e(A) \) (and/or \( E(A) \)).

Now, we obtain the matrix/vector formula: it is \( x^{(-1)} = A^T (Ax)^{(-1)} \).
APPENDIX B

SOME EXAMPLES FOR MATRIX/VECTOR FORMULA

\[ x^{(-1)} = A^T (Ax)^{(-1)} \]

Example 1. Let matrix

\[
A = \begin{bmatrix}
2 & 2 & 1 \\
3 & 15 & 5 \\
1 & 3 & 1 \\
2 & 10 & 5 \\
1 & 1 & 2 \\
2 & 10 & 5
\end{bmatrix},
\]

vector

\[
x = \begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}, \quad \prod_{i=1}^{3} x_i = 1.
\]

By theorem 1, and use the formulas which derived for achieving the \( e(A) \),

\[ x^{(-1)} = A^T (Ax)^{(-1)}, \]

we have:

\[
\begin{bmatrix}
1 \\
x_1 \\
1 \\
x_2 \\
x_1 \\
x_3
\end{bmatrix} = \begin{bmatrix}
2 & 1 & 2 \\
2 & 2 & 1 \\
2 & 10 & 10 \\
5 & 5 & 5
\end{bmatrix} \begin{bmatrix}
1 \\
\frac{2-x_1}{3} + \frac{2-x_2}{15} + \frac{1-x_3}{5} \\
\frac{2-x_1}{3} + \frac{2-x_2}{15} + \frac{1-x_3}{5} \\
\frac{2-x_1}{2} + \frac{1-x_2}{10} + \frac{1-x_3}{5} \\
\frac{2-x_1}{2} + \frac{1-x_2}{10} + \frac{1-x_3}{5} \\
\frac{2-x_1}{2} + \frac{1-x_2}{10} + \frac{1-x_3}{5}
\end{bmatrix},
\]

solve this system, then the values of critical points are attained:

\[
x = [0.523292 \ 1.73393 \ 1.10211]^T,
\]
\[ Ax = \begin{bmatrix}
0.796438 \\
1.00225 \\
0.875882
\end{bmatrix}, \]

\[
\prod_{i=1}^{3} x_i = 0.523292 \times 1.73393 \times 1.10211 = 1.000001.
\]

Let

\[ D_{Ax} = diag((Ax)_1, (Ax)_2, (Ax)_3) \]

\[ = diag(0.796438, 1.00225, 0.875882) \]

\[ = \begin{bmatrix}
0.796438 & 0 & 0 \\
0 & 1.00225 & 0 \\
0 & 0 & 0.875882
\end{bmatrix}. \]

Then

\[ D_{Ax}^{-1} = \begin{bmatrix}
1.255591 & 0 & 0 \\
0 & 0.997755 & 0 \\
0 & 0 & 1.141706
\end{bmatrix}. \]

\[ D_{x}^{-1} = diag\left(\frac{1}{x_1}, \frac{1}{x_2}, \frac{1}{x_3}\right) \]

\[ = diag\left(\frac{1}{0.523292}, \frac{1}{1.73393}, \frac{1}{1.10211}\right) \]

\[ = \begin{bmatrix}
1.910979 & 0 & 0 \\
0 & 0.567246 & 0 \\
0 & 0 & 0.9073504
\end{bmatrix}. \]
Then

\[
D_x = \begin{bmatrix}
0.523292 & 0 & 0 \\
0 & 1.73393 & 0 \\
0 & 0 & 1.10211
\end{bmatrix}.
\]

\[
D_{Ax}^{-1}AD_x = \begin{bmatrix}
0.4380270 & 0.2902808 & 0.2767598 \\
0.2610586 & 0.5190112 & 0.2199272 \\
0.2987229 & 0.1979639 & 0.5033144
\end{bmatrix} = B.
\]

\[
A = D_{Ax}BD_x^{-1},
\]

which \( B \) is DS-matrix, three row sums are 0.9978085, 1.007256 and 1.000001, three column sums are 1.005068, 0.999997 and 1.000001. It means \( A \) is DS-scalable.

\[
\det(D_{Ax}D_x^{-1}) = \det(D_{Ax}^{-1}D_x)^{-1} = 0.699154
\]

\[
e(A) = \prod_{i=1}^{3} (Ax)_i = 0.796438 \times 1.00225 \times 0.875882
\]

\[
= 0.699154 = \det(D_{Ax}D_x^{-1}) = \det(D_{Ax}^{-1}D_x)^{-1}.
\]

This result does agree with theorem 1.

Example 2. Let

\[
A = \begin{bmatrix}
3 & -1 & -1 \\
-3 & 5 & -1 \\
-3 & 0 & 4
\end{bmatrix},
\]

38
then

\[
A^{-1} = \begin{bmatrix}
2 & 2 & 1 \\
3 & 15 & 5 \\
1 & 3 & 1 \\
2 & 10 & 5 \\
2 & 10 & 5
\end{bmatrix},
\]

\(A^{-1}\) is DS-scalable, it has been shown in the example 1, \(e(A^{-1}) = 0.6991543\).

By theorem 5, and use the formula obtained which for finding the \(E(A)\) and the same procedure what has been used in Example 1, we have the critical points for achieving the \(E(A)\),

\[x = \begin{bmatrix}
0.900978 & 1.12546 & 0.986178
\end{bmatrix}^T,
\]

and

\[Ax = \begin{bmatrix}
0.591294 & 1.938185 & 1.241786
\end{bmatrix}^T.
\]

\[E(A) = \prod_{i=1}^{3} (Ax_i) = 1.42313
\]

Use the same logic used in the example 1, let

\[D_{Ax} = diag((Ax)_i), D_{Ax}^{-1} = diag\left(\frac{1}{(Ax)_i}\right), D_{x}^{-1} = diag\left(\frac{1}{x_i}\right), \]

\[D_{x} = diag(x_i),
\]

\[D_{Ax}^{-1}AD_{x} = \begin{bmatrix}
4.571211 & -1.903382 & -1.667827 \\
-1.394568 & 2.903382 & -0.5088144 \\
-2.176664 & 0.000000 & 3.1766644
\end{bmatrix} = B,
\]

it is DS-matrix,
\[ A = D_{Ax}BD_x^{-1}, \]

and

\[ E(A) = \det(D_{Ax}D_x^{-1}) = \det(D_{Ax}^{-1}D_x) = 1.42313219. \]

From the example for theorem 1, we know

\[ e(A^{-1}) = 0.6991543, \]

\[ E(A) \cdot e(A^{-1}) = 0.99498899, \]

it is almost 1.

This result does agree with theorem 5.