Modifying the Power Method in Max Algebra

Ludwig Elsner [*]
Fakultät für Mathematik
Universität Bielefeld
Postfach 100 131
D-33501 Bielefeld, GERMANY

P. van den Driessche [†]
Department of Mathematics and Statistics
University of Victoria
Victoria, B.C.
V8W 3P4, CANADA

November 9, 1999

Abstract

In the max algebra system, the eigenequation for an n-by-n irreducible nonnegative matrix $A = [a_{ij}]$ is $A \otimes x = \mu(A)x$. Here $(A \otimes x)_i = \max_j a_{ij}x_j$ and $\mu(A)$ is the maximum circuit geometric mean. The complexity of the power method given in [5] to compute $\mu(A)$ and x is considered. Under some assumptions on the critical matrix, it is shown that the algorithm may have time complexity $O(n^4)$. A modified power method, based on Karp's formula, is presented. For this new algorithm, with no assumptions on the critical matrix, $\mu(A)$ and x can be computed in $O(n^3)$ time. Furthermore, this algorithm can be used to compute all linearly independent eigenvectors corresponding to $\mu(A)$.

[*]elsner@mathematik.uni-bielefeld.de

[†]Research partially supported by SFB 343, Diskrete Strukturen in der Mathematik, and an NSERC research grant. pvdd@math.uvic.ca

1 Introduction

In [5] a power method algorithm is given to compute the eigenvalue $\mu(A)$ and eigenvector x of an n-by-n irreducible nonnegative matrix $A = [a_{ij}]$ in the max algebra system. The eigenequation of this system is $A \otimes x = \mu(A)x$, where $(A \otimes x)_i = \max_j a_{ij}x_j$ for $i = 1, \ldots, n$. Letting $\mathcal{G}(A)$ be the directed graph corresponding to A, i.e., $\mathcal{G}(A) = (V, E)$ where $V = \{1, \ldots, n\}$ and $E = \{(i, j) : a_{ij} > 0\}$, the eigenvalue $\mu(A)$ is the maximum circuit geometric mean in $\mathcal{G}(A)$. Our aim here is to find the complexity of the power method given in [5] and to present a modified algorithm, based on Karp's formula, that is in general more efficient.

Following the notation in [1] and [5], the *critical matrix* $A^{\mathcal{C}} = [a^{\mathcal{C}}_{ij}]$ is formed from the principal submatrix of A on rows and columns containing a *critical circuit* (i.e., a circuit with product equal to $\mu(A)$) by setting $a^{\mathcal{C}}_{ij} = a_{ij}$ if (i, j) lies on a critical circuit, and $a^{\mathcal{C}}_{ij} = 0$ elsewhere. Vertices on a critical circuit are called *critical vertices*. The critical graph $\mathcal{G}(A^{\mathcal{C}})$ has vertex set $V^{\mathcal{C}} = \{\text{critical vertices}\}$; thus $V^{\mathcal{C}} \subseteq V$.

In the max algebra, $(A \otimes B)_{ik} = \max_j a_{ij}b_{jk}$ for compatible matrices A, B; the notation $A^2_{\otimes}$ means $A \otimes A$ and similarly for higher powers. An identity matrix is denoted by I, a matrix with every entry 1 by J, and a vector with every entry 1 by e, thus $J = ee^T$; dimensions are specified by subscripts when necessary. Inequalities are componentwise, thus $A \leq J_n$ means that $a_{ij} \leq 1$ for all $i, j = 1, \ldots, n$.

In Section 2 we consider the power method algorithm given in [5, Algorithm 3.4]. We assume that $V = V^{\mathcal{C}}$ and that $A^{\mathcal{C}}$ is the direct sum of primitive matrices. Recall that matrix $B \geq 0$ is *primitive* if $B^h > 0$ for some integer h. Under these assumptions, we show that $A_\infty := \lim_{k\to\infty} A^k_{\otimes}/(\mu(A))^k = A^p_{\otimes}/(\mu(A))^p$, where a bound for p depends only on the structure of $A^{\mathcal{C}}$, and $p \leq n^2$. When A is scaled so that $\mu(A) = 1$ and $A \leq J_n$, then we give the entries of A_∞. Even under the above assumptions on A, the power method may have the time complexity $O(n^4)$.

In Section 3 we give a modified power method (Algorithm 3.4) that follows in spirit Karp's formula [8] for $\mu(A)$. With this algorithm, which assumes only that A is nonnegative and irreducible, $\mu(A)$ and an eigenvector can be computed in $O(n^3)$ time. We conclude by showing, in Section 4, that if $\mathcal{G}(A^{\mathcal{C}})$ has r strongly connected components, then Algorithm 3.4 can be used to compute the r linearly independent eigenvectors corresponding to $\mu(A)$.

2 A bound for the power method

In [5, Theorem 2.2] we used tools from nonnegative matrix theory to prove that if $A^{\mathcal{C}}$ is the direct sum of primitive matrices, then $\lim_{k\to\infty} A_{\otimes}^{k}/\mu(A)^{k}$, denoted by A_∞, exists, and $A_{\otimes}^{k} = \mu(A)^{k} A_\infty$ for k large enough. This led to a power method algorithm for finding $\mu(A)$ and an eigenvector. This algorithm is based on the following result, see [5, Theorem 3.3].

Theorem 2.1 *Let $A \geq 0$ be irreducible, and consider the sequence $x(t+1) = A \otimes x(t), t = 0, \ldots,$ where $x(0) > 0$ is given. If the critical matrix $A^{\mathcal{C}}$ is the direct sum of primitive matrices, then $x(t)$ is an eigenvector of A for some finite t_0 (and hence for all $t \geq t_0$).*

In [5], we left unresolved the question of a bound for t_0, which is equivalent to the number of steps in the algorithm (Algorithm 3.4 in [5]). The condition for convergence of the powers $A_{\otimes}^{k}$ given above is well known in the max plus-algebra (see, e.g., [1, Section 3.7.4]). But as noted there, the convergence can be very slow.

Example 2.2 (cf. [1, Example 3.110]) Let

$$A = \begin{bmatrix} 1 & 1 \\ e^{-1} & e^{-1/m} \end{bmatrix},$$

where m is a large positive integer. Here $V = \{1, 2\}$ and $V^{\mathcal{C}} = \{1\}$. Only after $k = m$ steps

$$A_{\otimes}^{k} = A_\infty = \begin{bmatrix} 1 & 1 \\ e^{-1} & e^{-1} \end{bmatrix}$$

with $\mu(A) = 1$. Thus an arbitrary large power is needed to give convergence, and this power depends on the magnitude of the entry a_{22}.

The above example shows that in order to obtain a bound on t_0 that is independent of the entries of A, we must restrict attention to the case $V = V^{\mathcal{C}}$. To obtain a bound in this case, we introduce some needed notation.

For a connected directed graph $G = (V, E)$ and $i, j \in V$ let $d(i, j) = \min\{l :$ there exists a path of length l from i to $j\}$. Then

$$d(G) = \max\{d(i, j) : i, j \in V, i \neq j\}$$

is the *diameter* of G. When G consists of a single vertex, set $d(G) = 0$. When A is a nonnegative primitive matrix, then the *exponent* of A is defined as $\exp(A) = \min\{k : A^k > 0\}$. It is convenient to define $\exp(G) = \exp(A(G))$, where $A(G)$ is the adjacency matrix of G; see, e.g., [4, Chapter 3].

To formulate our result, we make the following assumptions on the n-by-n irreducible matrix $A \geq 0$.

(A1) $V^{\mathcal{C}} = V = \{1, \ldots, n\}$.

(A2) The matrix $A^{\mathcal{C}}$ is the direct sum of r primitive matrices, i.e.

$$A^{\mathcal{C}} = A^{\mathcal{C}}_{11} \dotplus A^{\mathcal{C}}_{22} \dotplus \ldots \dotplus A^{\mathcal{C}}_{rr}.$$

Assumption (A2) leads to further notation. For $r \geq 2$ the vertex set V is partitioned by the matrices $A^{\mathcal{C}}_{ss}$, that is $V = V_1 \cup V_2 \cup \ldots \cup V_r$, where $A^{\mathcal{C}}_{ss} = A^{\mathcal{C}}[V_s]$, the principal submatrix of $A^{\mathcal{C}}$ with rows and columns in V_s. Also $|V_s| = n_s$ with $\sum_{s=1}^{r} n_s = n$. This leads to a block decomposition $A = [A_{st}]$, $s, t = 1, \ldots, r$, of A. For each such block, let $A_{st} \leq b_{st} e_s e_t^T$ with b_{st} minimal.

Assume now that A is scaled so that $\mu(A) = 1$ and $A \leq J_n$. Then $b_{ss} = 1$ and $b_{st} \leq 1$ for $s \neq t$. Define the r-by-r matrix B by $B = [b_{st}]$ and

$$B_\infty = \lim_{k \to \infty} B^k_\otimes = B^q_\otimes = [\beta_{st}] \tag{2.1}$$

with $\beta_{ss} = 1$. The existence of this limit is ensured by the result of Lemma 2.1 in [5], and $q \leq r - 1$ since any path of length r must contain a cycle, and all cycles have weight ≤ 1.

For each matrix $A^{\mathcal{C}}_{ss}$, let $G_s = \mathcal{G}(A^{\mathcal{C}}_{ss})$, with $d_s = d(G_s)$ and $\sigma_s = \exp(G_s)$ for $s = 1, \ldots, r$. Finally we define

$$p = r - 1 + \sum_{h=1}^{r} d_h + \max\{\sigma_h - d_h : d_h \geq 2\}. \tag{2.2}$$

The existence of $A_\infty := \lim_{k \to \infty} A^k_\otimes / \mu(A)^k = A^w_\otimes / \mu(A)^w$ for some w is in Theorem 2.2 of [5] under Assumption (A2). Under the additional assumption (A1), the following result shows that p is an upper bound on w. Note that the value of the bound p depends only on the structure of $A^{\mathcal{C}}$.

Theorem 2.3 *Let $A \geq 0$ be an n-by-n irreducible matrix satisfying assumptions (A1) and (A2) with block decomposition $A = [A_{st}]$ for $s, t = 1, \ldots, r$. Then $A_\infty := \lim_{k \to \infty} A^k_\otimes / \mu(A)^k = A^p_\otimes / \mu(A)^p$ where p is given by (2.2).*

If, in addition, A is scaled to have $\mu(A) = 1$ and $A \leq J_n$, then

$$A_\infty = [\beta_{st} e_s e_t^T], \tag{2.3}$$

where for $s, t = 1, \ldots, r$ the numbers β_{st} are defined by (2.1).

Proof. Assume that $\mu(A) = 1$ and $A \leq J_n$. Observe that in this case $A^{\mathcal{C}}$ is a (0,1)-matrix. To find the entries of A_∞, we show that

(i) $A^l_\otimes \leq [\beta_{st} e_s e_t^T]$ for all l,

(ii) $A^l_\otimes \geq [\beta_{st} e_s e_t^T]$ for all $l \geq p$.

Then (i) and (ii) together imply, under these assumptions that $A^p_\otimes = A_\infty = [\beta_{st} e_s e_t^T]$, as required in (2.3). For a general A, a scaling can be applied to give $\mu(A) = 1$. A diagonal similarity can then be applied to achieve $A \leq J_n$; see, e.g., [6], or the similarity displayed in our Algorithm 3.4 in Section 3.

To prove (i), let $C = [c_{ij}] = A^l_\otimes$ for a given l. We need to show that $c_{ij} \leq \beta_{st}$ for any pair $i \in V_s, j \in V_t$. For such a pair, $c_{ij} = \prod_{g=0}^{l-1} a_{i_g i_{g+1}}$ with $i_0 = i, i_l = j$ for a suitable sequence $i_0, \ldots, i_l$. Let $i_g \in V_{y_g}, g = 0, \ldots, l$. Then $c_{ij} \leq \prod_{g=0}^{l-1} b_{y_g y_{g+1}}$, implying $c_{ij} \leq \beta_{st}$, by the definition of β_{st} in (2.1) and by the fact that the powers of B form a monotonically increasing sequence. This proves (i).

For given $i \in V_s, j \in V_t$ we prove (ii) by constructing a path in A from i to j of length p with path product β_{st} in the following way. In the construction we twice use the following simple observation. Given any two vertices $a, b \in V_s$ and any number $l \geq \sigma_s$, we find a path within $A^{\mathcal{C}}_{ss}$ of length l connecting a to b.

For $s = t$, as $p \geq \sigma_s$, such a path of length p exists. Its path product is 1, because all nonzero entries in $A^{\mathcal{C}}_{ss}$ are 1 and $1 = \beta_{ss}$. For $r \geq 2$, if $s \neq t$, let $\beta_{st} = b_{s_0 s_1} b_{s_1 s_2} \cdots b_{s_v s_{v+1}}$, where $s_0 = s, s_{v+1} = t$. The length of this path in B is $v + 1$, hence $v + 1 \leq q \leq r - 1$. Each term is of the form $b_{s_g s_{g+1}} = a_{m_g \tilde{m}_{g+1}}$, for $g = 0, \ldots, v$, where $m_g \in V_{s_g}$ and $\tilde{m}_{g+1} \in V_{s_{g+1}}$. In addition set $\tilde{m}_0 = i, m_{v+1} = j$. The path connects the following vertices in V: $i = \tilde{m}_0 \to m_0 \to \tilde{m}_1 \to m_1 \to \cdots \to m_v \to \tilde{m}_{v+1} \to m_{v+1} = j$. Here each pair $(\tilde{m}_g, m_g)$ for $g = 0, \ldots, v + 1$ can be connected within $A^{\mathcal{C}}_{s_g s_g}$ by a path of length $\leq d_{s_g}$, (it is possible that $\tilde{m}_g = m_g$, so there is no connection necessary), while $(m_g, \tilde{m}_{g+1})$ is an edge in $\mathcal{G}(\mathcal{A})$. This path has path product β_{st} and length $\hat{p} \leq r - 1 + \sum_{h=1}^{r} d_h \leq p$. We need to extend the length of this path to p. If on this path at least one $d_h = 0$, then we can use this diagonal entry with weight one to extend the path to any length without changing the path product. If on the path all $d_h \geq 1$, then we extend the length of the path in the first block $A^{\mathcal{C}}_{ss}$. Let $\tilde{p}$ be the length of the path from m_0 to j, thus $\tilde{p} \leq r - 1 + \sum_{h=1}^{r} d_h - d_s$. Note that $p - \tilde{p} \geq \max\{\sigma_h - d_h : d_h \geq 1\} + d_s \geq \sigma_s$. By the observation above we find a path of length $p - \tilde{p}$ in $A^{\mathcal{C}}_{ss}$ from i to m_0. Adding the path above with length $\tilde{p}$ from m_0 to j gives the required path from i to j of length p and product β_{st}. This path can be extended in the same way to any length $\geq p$. If $d_h = 1$ then $\sigma_h = 1$, thus we can in (2.2) restrict the max to those blocks with $d_h \geq 2$. This completes the proof of (ii). □

We now consider some special cases.

Corollary 2.4 *Let A satisfy all the assumptions of Theorem 2.3, and in addition assume that $A^{\mathcal{C}}_{ss} = e_s e_s^t$ for all $s = 1, \ldots, r$. Then $A_\infty = A^p_\otimes$ with $p \leq r - 1 + p_1 \leq n - 1$, where p_1 is the number of blocks that have $n_s \geq 2$.*

Proof. If all $A^{\mathcal{C}}_{ss} = e_s e_s^T$, then $d_s = 1$ if $n_s \geq 2$, $d_s = 0$ if $n_s = 1$, and all $\sigma_s = 1$. Thus from (2.2), $p \leq r - 1 + p_1$. This bound is always $\leq n - 1$, since the case when $r = n, p_1 = 0$ gives the maximum value of $r - 1 + p_1$. Theorem 2.3 now gives the result. □

In particular, if $n_s = 1$ for all s, then $r = n$ and Corollary 2.4 gives $p \leq n - 1$. This bound $n - 1$ is tight for A tridiagonal with $a_{ii} = 1, a_{ij} = a < 1$ for $|i - j| = 1$ and $a_{ij} = 0$ otherwise. The bound in Corollary 2.4 can also be tight when $n_s \geq 2$, as the following example shows.

Example 2.5 Let

$$A = \begin{bmatrix} 1 & 1 & a & 0 \\ 1 & 1 & 0 & 0 \\ a & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

with $a < 1$. Then $A^2_\otimes \neq A_\infty$, $A^3_\otimes = A_\infty$, and Corollary 2.4 gives $p \leq 3$.

Corollary 2.6 *Let A satisfy the assumptions (A1) and (A2) of Theorem 2.3 and $\mu(A) = 1$. Then $A_\infty = A^p_\otimes$, where $p \leq n^2 - 1$.*

Proof. ¿From Theorem 2.3 and (2.2), $p \leq n - 1 + n + (n^2 - 2n + 2) - 2 = n^2 - 1$, since $r \leq n$, $\sum_{h=1}^{r} d_h \leq n$, and $n^2 - 2n + 2$ is Wielandt's bound for the exponent; see e.g., [4, Theorem 3.5.6]. □

Thus, in the case when A has maximum exponent (i.e., $A^{\mathcal{C}}$ is a Wielandt matrix), $O(n^2)$ steps may be needed to compute A_∞. Accordingly the power method may need $O(n^2)$ matrix-vector multiplications. Thus the time for the power method discussed in [3] or [5] may be $O(n^4)$. In the next section we give a modified power method, that is one order better, and is independent of the structure of $A^{\mathcal{C}}$.

We conclude this section by noting that if the assumption (A2) is dropped, then the convergence of the powers of A is not guaranteed. In general we have only

$$A^{\mathcal{C}} = A_{11}^{\mathcal{C}} \dot{+} A_{22}^{\mathcal{C}} \dot{+} \dots \dot{+} A_{rr}^{\mathcal{C}},$$

with each $A_{ii}^{\mathcal{C}}$ irreducible with index of imprimitivity s_i. As shown in [5, Theorem 2.3], there is periodic behaviour with period $d = l.c.m.(s_1, \dots, s_r)$, and $(A_\otimes^d)^{\mathcal{C}}$ is equal to the direct sum of $\sum_{i=1}^r s_i$ primitive matrices. If $V = V^{\mathcal{C}}$, then Theorem 2.3 can be applied to $A_\otimes^d$ to give a bound on the number of steps of the power method given in Algorithm 3.4 in [5].

3 A modified power method

We begin by proving a generalization of Karp's formula ([8], see also [1, Theorem 2.19]) for $\mu(A)$.

Theorem 3.1 *Let $A \geq 0$ be an n-by-n irreducible matrix and $x_0 \geq 0, x_0 \neq 0$ be an n-vector. Define $x_i = A \otimes x_{i-1}, i = 1, \dots, n$, with $x_i = [x_{1i}, \dots, x_{ni}]^T$ for $i = 0, \dots, n$. Then the eigenvalue $\mu(A)$ is given by*

$$\mu(A) = \max_{i=1,\dots,n} \min_{k=0,\dots,n-1} (\frac{x_{in}}{x_{ik}})^{1/(n-k)}. \tag{3.1}$$

Proof. If A is replaced by tA, then both sides of (3.1) are multiplied by t. Thus we can assume that $\mu(A) = 1$. Every entry in $A_\otimes^n$ is a product of n entries of A, and must therefore contain a cycle (which has product of at most one) with remaining terms that are not greater than a term in a previous power. Hence

$$I \oplus A \oplus A_\otimes^2 \oplus \dots \oplus A_\otimes^{n-1} \geq A_\otimes^n. \tag{3.2}$$

Notice that (3.2) also holds for $\mu(A) \leq 1$. Let u_i be the n-vector with 1 in position i and 0 everywhere else. Then from (3.2)

$$u_i^T \otimes (I \oplus A \oplus A_\otimes^2 \oplus \dots \oplus A_\otimes^{n-1}) \otimes x_0 \geq u_i^T \otimes A_\otimes^n \otimes x_0,$$

equivalently

$$\max_{k=0,\dots,n-1} x_{ik} \geq x_{in}.$$

As A is irreducible the left hand side is positive, hence

$$\frac{x_{in}}{\max_{k=0,\dots,n-1} x_{ik}} = \min_{k=0,\dots,n-1} \frac{x_{in}}{x_{ik}} \leq 1.$$

So we have shown that in (3.1) the inequality $\geq$ holds.

For equality, we need to show that there exists i such that $x_{in} = \max\{x_{ik} : k = 0, \dots, n-1\}$. Assume that no such i exists, that is

$$x_{in} < \max_{k=0,\dots,n-1} x_{ik}$$

for all i, equivalently

$$x_n < x_0 \oplus x_1 \oplus \cdots \oplus x_{n-1}.$$

Using the definitions of x_i, this is

$$A_{\otimes}^{n} \otimes x_0 < (I \oplus A \oplus A_{\otimes}^{2} \oplus \cdots \oplus A_{\otimes}^{n-1}) \otimes x_0,$$

contradicting (3.2). $\square$

Note that in computing $\mu(A)$ by (3.1), $0/0$ must be regarded as 0. Observe that we have used the irreducibility of A only to ensure the positivity of $\max_k x_{ik}$. This holds also in the reducible case if we assume that $x_0 > 0$. Hence Theorem 3.1 also holds in the general case, if we require x_0 to be positive. Note that if $x_0 = u_j$, then the formula (3.1) reduces to Karp's formula. For $x_0 = e$ and A replaced by A^T, (3.1) can be found also in [10, Theorem 4] and [11], where no irreducibility is required.

To find an eigenvector, we introduce some notation, see ([1, Section 1.2.1]). Let A be an n-by-n nonnegative matrix with $\mu(A) \leq 1$. Define

$$A^{+} = A \oplus A_{\otimes}^{2} \oplus \cdots \oplus A_{\otimes}^{n},$$

$$A^{*} = I \oplus A \oplus A_{\otimes}^{2} \oplus \cdots \oplus A_{\otimes}^{n-1}.$$

Then $A \otimes A^{*} = A^{+}$ and by (3.2) $A_{\otimes}^{n} \leq A^{*}$, which implies that

$$I \oplus A \otimes A^{*} = A^{*} \tag{3.3}$$

and $A^{+} \leq A^{*}$. These relations are used in the following lemma.

Lemma 3.2 *Let $A \geq 0$ be irreducible with $\mu(A) \leq 1$, $x \geq 0, x \neq 0$, and $z = A^* \otimes x$. Then $z > 0$. If $D = diag(z_i)$ then $D^{-1}AD \leq J_n$, i.e. each entry has magnitude less than or equal to 1.*

Proof. $z > 0$ follows from the irreducibility of A. Also

$$A \otimes z = A \otimes A^* \otimes x = A^+ \otimes x \leq A^* \otimes x = z.$$

This inequality means that $a_{ik}z_k \leq z_i$ for all i, k; that is $a_{ik}z_k/z_i \leq 1$. $\square$

The following result is proved in the max-plus algebra in [1, Theorem 3.17], so we omit the simple proof.

Lemma 3.3 *Let $A \geq 0$ be irreducible with $\mu(A) \leq 1$, and $x \geq 0$. Then a solution of*

$$y = A \otimes y \oplus x$$

is given by $y = A^* \otimes x$. If $\mu(A) < 1$, then this solution is unique.

We now give the steps of an algorithm to compute $\mu(A)$ and an eigenvector, followed by a theorem that proves the validity of the algorithm and shows that it stops in $O(n^3)$ time.

Algorithm 3.4 *For $A \geq 0$ an n-by-n irreducible matrix and $x_0 \geq 0, x_0 \neq 0$, follow the steps 1-6 below.*

1. *Calculate $x_i = A \otimes x_{i-1}$ for $i = 1, \ldots, n$.*
2. *Denoting $x_k = [x_{1k}, \ldots, x_{nk}]^T$, find*
$$\mu(A) = \max_{i=1,\ldots,n} \min_{k=0,\ldots,n-1} (\frac{x_{in}}{x_{ik}})^{1/(n-k)}.$$
3. *Calculate*
$$y = x_0 \oplus \frac{1}{\mu(A)} x_1 \oplus \cdots \oplus \frac{1}{\mu(A)^{n-1}} x_{n-1}.$$
If $A \otimes y = \mu(A)y$ then set $x = y$; otherwise proceed further.
4. *Form $\tilde{A} = 1/\mu(A) D_y^{-1} A D_y$, where $D_y = diag(y_i)$.*

5. *Find a principal submatrix of $\tilde{A}$, which by a permutation can be $\tilde{A}_{11}$ of dimension t-by-t, such that $\tilde{A}_{11} \otimes e_t = e_t, t < n$. Hence (by permutation, if necessary)*

$$\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix},$$

where $\tilde{A}_{21}$ is $(n-t)$-by-t.

6. *Form the $(n-t)$-vector $z_0 = \tilde{A}_{21} \otimes e_t$ and calculate*

$$z_i = z_{i-1} \oplus \tilde{A}_{22} \otimes z_{i-1}, i = 1, \ldots, n-t-1.$$

Set $z = z_{n-t-1}$ and form $x = D_y \begin{bmatrix} e_t \\ z \end{bmatrix}$.

Remark Step 5 can be achieved by a simple algorithm. Let $I_0 = \{1, \ldots, n\}$ and for $v \geq 1$, $I_v = \{i \in I_{v-1} : (\tilde{A}[I_{v-1}] \otimes e)_i = 1\}$. If $I_v = I_{v+1}$, then stop and set $\tilde{A}_{11} = \tilde{A}[I_v]$. Since $I_v \supseteq I_{v+1}$, and all I_v contain a critical circuit, this algorithm stops after at most $n-1$ steps. This algorithm produces the largest t at step 5.

Theorem 3.5 *Let $A \geq 0$ be an n-by-n irreducible matrix. Then $\mu(A)$ and x produced by Algorithm 3.4 satisfy $A \otimes x = \mu(A)x$.*

Proof. The fact that $\mu(A)$ as produced by steps 1 and 2 is an eigenvalue of A follows from Theorem 3.1. Since A is irreducible, the vector y computed in step 3 is positive. By Lemma 3.2 applied to the matrix $A/\mu(A)$ and $x = x_0$, every entry of $\tilde{A}$ is less than or equal to 1. Also $\mu(\tilde{A}) = 1$. The vector z calculated in step 6 is $z = \tilde{A}_{22}^* \otimes \tilde{A}_{21} \otimes e_t$, hence it satisfies (by Lemma 3.3) the equation $z = \tilde{A}_{21} \otimes e_t \oplus \tilde{A}_{22} \otimes z$. Also $\tilde{A}_{11} \otimes e_t \oplus \tilde{A}_{12} \otimes z = e_t$, which gives

$$\tilde{A} \otimes \begin{bmatrix} e_t \\ z \end{bmatrix} = \begin{bmatrix} e_t \\ z \end{bmatrix}.$$

It follows that $A \otimes x = \mu(A)x$. □

.

The time complexity of Algorithm 3.4 is $O(n^3)$, or more exactly $O(nm)$, where m is the number of nonzero entries of A. Here steps 1 and 6 need at most nm operations each, while steps 2 - 5 each take $O(n^2)$ operations. This algorithm is thus an order of magnitude better than the worst case of the power method analyzed in Section 2. Also note that Algorithm 3.4 is valid for any nonnegative irreducible matrix, whereas Theorem 2.3 requires the additional assumptions (A1) and (A2). In fact under the assumption (A1), namely $V = V^{\mathcal{C}}$, Algorithm 3.4 stops after step 3. The difference between the two algorithms is illustrated by the following example.

Example 3.6 Let

$$A = \begin{bmatrix} a & a & 1 & 1 \\ 1 & a & a & a \\ a & 1 & a & a \\ a & a & 1 & a \end{bmatrix}$$

with $a < 1$. Then $A^{\mathcal{C}}$ is a Wielandt matrix with $n = 4$, satisfying (A1) and (A2) with $r = 1$. As predicted in Section 2, since $\exp(A^{\mathcal{C}}) = 10$, it takes 10 steps to compute $A_{\otimes}^{10} \otimes x_0 = A_{\infty} \otimes x_0$ when $x_0 = [0, 0, 0, 1]^T$. Applying Algorithm 3.4 with the same initial vector x_0, gives $x_1 = [1, a, a, a]^T, x_2 = [a, 1, a, a]^T, x_3 = [a, a, 1, a]^T$, and $y = [1, 1, 1, 1]^T$ in step 3, which is an eigenvector. Thus only 3 matrix-vector multiplications are required.

4 Finding all eigenvectors

If $\mathcal{G}(A^{\mathcal{C}})$ has $r \geq 2$ strongly connected components, then A has r linearly independent eigenvectors corresponding to $\mu(A)$, see, e.g., [2, Theorem 5]. A set of vectors is linearly independent (in the max algebra) if no vector is a nonnegative linear combination of the others. We show in the following that by starting with different choices of $\tilde{A}_{11}$ in step 5 of Algorithm 3.4, r linearly independent eigenvectors are computed. It suffices to consider only the algorithm starting from step 5, so we may assume that $A = \tilde{A}$, $\mu(A) = 1$ and $A \leq J_n$.

Theorem 4.1 *Let $A \geq 0$ be an n-by-n irreducible matrix with $\mu(A) = 1, A \leq J_n$, and its critical matrix*

$$A^{\mathcal{C}} = A_{11}^{\mathcal{C}} \dot{+} A_{22}^{\mathcal{C}} \dot{+} \ldots \dot{+} A_{rr}^{\mathcal{C}}, \tag{4.1}$$

with each $A_{ii}^{\mathcal{C}}$ irreducible. If in step 5 of Algorithm 3.4, A_{11} is taken as $A[V_s]$ for $s = 1, \ldots, r$, then the algorithm yields r linearly independent eigenvectors of A. Each eigenvector computed in this way is a column of A^* and A^+.

Proof. We start from step 5 of Algorithm 3.4 with

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

where we take $A_{11} = A[V_s]$. Note that $A_{11} \otimes e_t = e_t$, where $t = |V_s|$. It is easy to see by using $A \le J_n$ that in A^* the left upper block is J_t. Hence

$$A^* = \begin{bmatrix} J_t & X \\ Z & Y \end{bmatrix},$$

where X, Y, Z have all entries ≤ 1. From (3.3) for the (2,1)-block

$$Z = A_{21} \otimes J_t \oplus A_{22} \otimes Z,$$

since I does not contribute to this block. Each column z of Z satisfies the same equation $z = A_{21} \otimes e_t \oplus A_{22} \otimes z$. If $\mu(A_{22}) < 1$, then by Lemma 3.3 the solution is unique, and hence coincides with the solution $z = A_{22} \otimes A_{21} \otimes e_t$ computed by the algorithm. For $\mu(A_{22}) = 1$ the same result is obtained by a continuity argument. In particular, all columns of A^* belonging to the same V_s are identical. Thus the eigenvector calculated is a column of A^* belonging to a critical vertex. As this column has diagonal entry equal to 1, it is also a column of A^+. By [2, Theorem 5] the set of r eigenvectors generated in this way is linearly independent. $\square$

We have thus shown that Algorithm 3.4 can also be used to find all eigenvectors. The complexity of this procedure seems however not to be optimal. In addition, the decomposition (4.1) has to be determined; for an algorithm see, e.g., [7, pp. 18-20]. An alternative to calculate the maximal set of linearly independent eigenvectors as in Theorem 4.1 uses the Floyd-Warshall procedure, see, e.g., [7, p. 58] or [9, Theorem 1], and is of complexity $O(n^3)$. For completeness we give a short description.

For $A = A_0 \ge 0$ with $\mu(A) = 1$ calculate

$$A_{k+1} = A_k \oplus (A_k \otimes u_{k+1})(u_{k+1}^T \otimes A_k), \qquad k = 0, \ldots, n-1. \tag{4.2}$$

$(A_k)_{ij}$ is the maximum path product of all paths connecting i and j where the inner vertices are in $\{1, \ldots, k\}$. These paths are of length $\le k+1$. Then obviously $A_n = A^+$, and it is clear from (4.2) that we need n^3 multiplications and comparisons. Now if we start with a matrix A satisfying also $A \le J_n$, then according to the considerations above the columns corresponding to indices $i \in V^C$, characterized by $(A_n)_{ii} = 1$, are eigenvectors and columns corresponding to indices in the same connected component of $V^{\mathcal{C}}$ are equal. Hence by at most $n^2(n-1)/2$ comparisons, the maximal number of r linearly independent eigenvectors is found.

Acknowledgement: We thank Hans Schneider for discussions and for pointing out several connections to previous results.

References

[1] F.L. Baccelli, G. Cohen, G.J. Olsder, and J.-P. Quadrat. *Synchronization and Linearity: An Algebra for Discrete Event Systems*, Wiley (1992).

[2] R.B. Bapat, *A Max Version of the Perron-Frobenius Theorem*, *Linear Algebra Appl.*, **275-276** (1998), 3-18

[3] J.G. Braker and G.-J. Olsder, *The Power Method in Max Algebra*, *Linear Algebra Appl.*, **182** (1993), 67-89.

[4] R.A. Brualdi and H.J. Ryser, *Combinatorial Matrix Theory*, Cambridge University Press (1991).

[5] L. Elsner, P. van den Driessche. *On the Power Method in Max Algebra*, *Linear Algebra Appl.*, to appear.

[6] G.M. Engel and H. Schneider, *Diagonal Similarity and Equivalence for Matrices over Groups with O*, *Czechoslovak Math. J.*, **25** (1975), 389-403.

[7] M. Gondran and M. Minoux, *Graphs and Algorithms*, Wiley (1984).

[8] R. M. Karp, *A Characterization of the Minimum Cycle Mean in a Digraph*, Discrete Mathematics, **23** (1978), 309-311.

[9] G.-J. Olsder, K. Roos and R.-J. van Egmond, *An efficient algorithm for critical circuits and finite eigenvectors in the max-plus algebra*, *Linear Algebra Appl.*, **295** (1999), 231-240.

[10] H. Schneider and M. H. Schneider *An l_∞ Balancing of a Weighted Directed Graph OR-Group Report Series 88-1*, Johns Hopkins University, Baltimore, Maryland (1987).

[11] H. Schneider and M. H. Schneider *Max-Balancing Weighted Directed Graphs and Matrix-Scaling, Mathematics of Operation Research*, **16** (1991), 208-222.