Chapter Two

Spectral Theory

This chapter is devoted to spectral theory of matrices over the max-plus semiring. In Section 2.1 we will study the relation between graphs and matrices over the max-plus semiring. The basic observation is that any square matrix can be translated into a weighted graph (to be defined shortly) and that products and powers of matrices over the max-plus semiring have entries with a nice graph-theoretical interpretation. This interpretation will be further studied in Section 2.2. The key result will be that, under mild conditions, a square matrix over the max-plus semiring possesses a unique eigenvalue that equals the maximal average weight of circuits in the associated graph. This result plays a crucial role in numerical analysis of systems that can be modeled by max-plus recurrence relations. The chapter is concluded with Section 2.3 in which some elementary results on certain max-plus linear equations are presented.

2.1 MATRICES AND GRAPHS

A *directed graph* $\mathcal{G}$ is a pair $(\mathcal{N}, \mathcal{D})$, where $\mathcal{N}$ is a finite set of elements called *nodes* (or *vertices*) and $\mathcal{D} \subset \mathcal{N} \times \mathcal{N}$ is a set of ordered pairs of nodes called *arcs* (or *edges*). The word *ordered* means that the arcs (i, j) and (j, i) will be distinguished. If $(i, j) \in \mathcal{D}$, then we say that $\mathcal{G}$ contains an arc from i to j, and the arc (i, j) is called an *incoming arc* at j and an *outgoing arc* at i. Suppose that $(i, j) \in \mathcal{D}$, but $(j, i) \notin \mathcal{D}$; then an arc from i to j exists but there isn't an arc from j to i. This distinction in the direction of an arc explains the name *directed graph*. A directed graph is also called a *digraph* in the literature. A directed graph is called *weighted* if a weight $w(i, j) \in \mathbb{R}$ is associated with any arc $(i, j) \in \mathcal{D}$. From now on we will deal exclusively with weighted directed graphs and will refer to them as "graphs" for simplicity.

To any $n \times n$ matrix A over $\mathbb{R}_{\max}$ a graph can be associated, called the *communication graph* of A. The graph will be denoted by $\mathcal{G}(A)$. The set of nodes of the graph is given by $\mathcal{N}(A) = \underline{n}$, and a pair $(i, j) \in \underline{n} \times \underline{n}$ is an arc of the graph if $a_{ji} \neq \varepsilon$ (this is not a typo!), i.e., in symbols $(i, j) \in \mathcal{D}(A) \Leftrightarrow a_{ji} \neq \varepsilon$, where $\mathcal{D}(A)$ denotes the set of arcs of the graph.

For any two nodes i, j, a sequence of arcs $p = ((i_k, j_k) \in \mathcal{D}(A) : k \in \underline{m})$, such that $i = i_1$, $j_k = i_{k+1}$, for $k < m$, and $j_m = j$ is called a *path* from i to j. The path is then said to consist of the nodes $i = i_1, i_2, \ldots, i_m, j_m = j$ and to have *length* m. The latter will be denoted as $|p|_1 = m$. Further, if $i = j$, then the path is called a *circuit*. A circuit $p = ((i_1, i_2), (i_2, i_3), \ldots, (i_m, i_1))$ is called *elementary* if, restricted to the circuit, each of its nodes has only one incoming and one outgoing arc or, more formally, if nodes i_k and i_l are different for $k \neq l$. A circuit consisting of just one arc, from a node to itself, is also called a *self-loop*.

The set of all paths from i to j of length $m \geq 1$ is denoted by $P(i,j;m)$. For an arc (i,j) in $\mathcal{G}(A)$, the weight of (i,j) is given by a_{ji} (again, this is not a typo!), and the *weight* of a path in $\mathcal{G}(A)$ is defined by the sum of the weights of all arcs constituting the path. More formally, for $p = ((i_1,i_2),(i_2,i_3),\ldots,(i_m,i_{m+1})) \in P(i,j;m)$ with $i = i_1$ and $j = i_{m+1}$, define the weight of p, denoted by $|p|_{\mathrm{w}}$, through

$$|p|_{\mathrm{w}} = \bigotimes_{k=1}^{m} a_{i_{k+1}i_k}.$$

Note that in conventional notation $|p|_{\mathrm{w}} = \sum_{k=1}^{m} a_{i_{k+1}i_k}$. The *average weight* of a path p is given by $|p|_{\mathrm{w}}/|p|_{\mathrm{l}}$. For circuits the notions of weight, length, and average weight are defined similarly as for paths. Also, the phrase *circuit mean* is used instead of the phrase *average circuit weight*.

Paths in $\mathcal{G}(A)$ can be combined in order to construct a new path. For example, let $p = ((i_1,i_2),(i_2,i_3))$ and $q = ((i_3,i_4),(i_4,i_5))$ be two paths in $\mathcal{G}(A)$. Then,

$$p \circ q = \big((i_1,i_2),(i_2,i_3),(i_3,i_4),(i_4,i_5)\big)$$

is a path in $\mathcal{G}(A)$ as well. The operation $\circ$ is called the *concatenation of paths*. Clearly, the operation is not commutative, even when both $p \circ q$ and $q \circ p$ are defined.

Example 2.1.1 *Let*

$$A = \begin{pmatrix} \varepsilon & 15 & \varepsilon \\ \varepsilon & \varepsilon & 14 \\ 10 & \varepsilon & 12 \end{pmatrix}.$$

The communication graph of A is shown in Figure 2.1. The graph $\mathcal{G}(A)$ has node

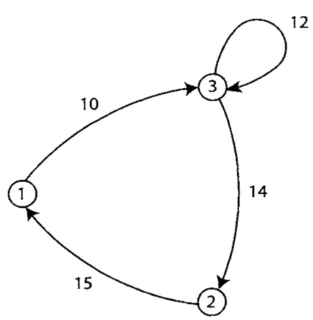

Figure 2.1: The communication graph of matrix A in Example 2.1.1.

set $\mathcal{N}(A) = \{1,2,3\}$ and arc set $\mathcal{D}(A) = \{(1,3),(3,2),(2,1),(3,3)\}$. Specifically, $\mathcal{G}(A)$ consists of two elementary circuits, namely, $\rho = ((1,3),(3,2),(2,1))$ and $\theta = (3,3)$. The weight of ρ is given by

$$|\rho|_{\mathrm{w}} = a_{12} + a_{23} + a_{31} = 39,$$

and the length of ρ equals $|\rho|_1 = 3$. Circuit θ has weight $|\theta|_w = a_{33} = 12$ and is of length 1.

The communication graph $\mathcal{G}(A)$ and powers of A are closely related to each other. As will be proved in the next theorem, the element $[A^{\otimes k}]_{ji}$ yields the maximal weight of a path of length k from node i to node j, provided that such a path exists.

THEOREM 2.1 *Let $A \in \mathbb{R}_{\max}^{n \times n}$. It holds for all $k \geq 1$ that*

$$\left[A^{\otimes k}\right]_{ji} = \max \{ |p|_w : p \in P(i,j;k) \},$$

where $[A^{\otimes k}]_{ji} = \varepsilon$ in the case where $P(i,j;k)$ is empty, i.e., when no path of length k from i to j exists in $\mathcal{G}(A)$.

Proof. The proof is done by induction. Let (i,j) be an arbitrary element of $\underline{n} \times \underline{n}$. For $k = 1$, paths in $P(i,j;k)$ consist only of one arc whose weight is by definition $[A]_{ji}$. If $[A]_{ji} = \varepsilon$, then there is no arc (i,j) in $\mathcal{G}(A)$ and $P(i,j;1) = \emptyset$.

Suppose now that the theorem holds true for k. Consider $p \in P(i,j;k+1)$ and assume that there exists at least one path in $P(i,j;k+1)$. Such a path, say, p can be split up into a subpath of length k running from i to some node l, and a path consisting of one arc from l to j or, more formally,

$$p = \hat{p} \circ (l,j) \quad \text{with} \quad \hat{p} \in P(i,l;k).$$

The maximal weight of any path in $P(i,j;k+1)$ can thus be obtained from

$$\max_{l \in \underline{n}} \left([A]_{jl} + \max \left\{ |\hat{p}|_w : \hat{p} \in P(i,l;k) \right\} \right). \tag{2.1}$$

In accordance with the induction hypothesis, it holds that

$$\max \left\{ |\hat{p}|_w : \hat{p} \in P(i,l;k) \right\} = \left[A^{\otimes k}\right]_{li},$$

and the expression for the maximal weight of a path from i to j of length $(k+1)$ in (2.1) reads

$$\begin{aligned}
\max_{l \in \underline{n}} \left(a_{jl} + \left[A^{\otimes k}\right]_{li} \right) &= \bigoplus_{l=1}^{n} a_{jl} \otimes \left[A^{\otimes k}\right]_{li} \\
&= \left[A \otimes A^{\otimes k}\right]_{ji} = \left[A^{\otimes (k+1)}\right]_{ji}.
\end{aligned}$$

Now turn to the case in which $P(i,j;k+1) = \emptyset$; i.e., there exists no path of length $k+1$ from i to j. Obviously, this implies that for any node l it holds true that either there exists no path of length k from i to l or there exists no arc from l to j (or both). Hence, $P(i,j;k+1) = \emptyset$ implies that for any l at least one of the values a_{jl} and $\left[A^{\otimes k}\right]_{li}$ equals ε. As a consequence,

$$\left[A^{\otimes (k+1)}\right]_{ji} = \varepsilon,$$

which completes the proof of the theorem. $\square$

The above theorem is illustrated with the following example.

Example 2.1.2 *Consider a railway network with a set of stations $\underline{n}$, with $n \geq 1$. The railway system can be mapped on a graph $\mathcal{G} = (\mathcal{N}, \mathcal{D})$ in the following way. Set $\mathcal{N} = \underline{n}$, and for $i, j \in \underline{n}$, let (i,j) be an arc in $\mathcal{D}$ if there is a direct connection from station i to station j in the railway network. To arc $(i,j) \in \mathcal{D}$, associate weight a_{ji}, where a_{ji} denotes the travel time from i to j. Suppose one is interested in the maximal time to travel from a particular station i to a particular station j in m steps. A direct way of computing this number would be to check all paths of length m from i to j in $\mathcal{G}$ and to compare their weights. An alternative way for computing the maximal travel time is as follows. Let A be the $n \times n$ matrix whose elements represent the weights of arcs in $\mathcal{G}$, where $a_{ji} = \varepsilon$ if $(i,j) \notin \mathcal{D}$. Then the maximal weight of a path from i to j of length m is given by $[A^{\otimes m}]_{ji}$; i.e., the maximal weight is given by the element (j,i) of the mth power of A.*

For $A \in \mathbb{R}_{\max}^{n \times n}$, let

$$A^+ \stackrel{\text{def}}{=} \bigoplus_{k=1}^{\infty} A^{\otimes k}. \tag{2.2}$$

The element $[A^+]_{ij}$ yields the maximal weight of any path from j to i (the value $[A^+]_{ij} = +\infty$ is possible). Indeed, by definition

$$[A^+]_{ij} = \max \left\{ [A^{\otimes k}]_{ij} : k \geq 1 \right\},$$

where $[A^{\otimes k}]_{ij}$ is the maximal weight of a path from j to i of length k; see Theorem 2.1.

LEMMA 2.2 *Let $A \in \mathbb{R}_{\max}^{n \times n}$ be such that any circuit in $\mathcal{G}(A)$ has average circuit weight less than or equal to e. Then, it holds that*

$$A^+ = \bigoplus_{k=1}^{\infty} A^{\otimes k} = A \oplus A^{\otimes 2} \oplus A^{\otimes 3} \oplus \cdots \oplus A^{\otimes n} \in \mathbb{R}_{\max}^{n \times n}.$$

Proof. Since A is of dimension n, all paths in $\mathcal{G}(A)$ from i to j of length greater than n are necessarily made up of at least one circuit and a path from i to j of length at most n. Because circuits in $\mathcal{G}(A)$ have nonpositive weights, it follows that

$$[A^+]_{ji} \leq \max\{[A^{\otimes k}]_{ji} : k \in \underline{n}\},$$

which concludes the proof of the lemma. □

We conclude this section with a discussion of graph-theoretical concepts that we will need later on. Let $\mathcal{G} = (\mathcal{N}, \mathcal{D})$ denote a graph with node set $\mathcal{N}$ and arc set $\mathcal{D}$. For $i, j \in \mathcal{N}$, node j is said to be reachable from node i, denoted as $i\mathcal{R}j$, if there exists a path from i to j. A graph $\mathcal{G}$ is called *strongly connected* if for any two nodes $i, j \in \mathcal{N}$, node j is reachable from node i. A matrix $A \in \mathbb{R}_{\max}^{n \times n}$ is called *irreducible* if its communication graph $\mathcal{G}(A)$ is strongly connected. If a matrix is not irreducible, it is called *reducible*.

To better deal with graphs that are not strongly connected, we say for nodes $i, j \in \mathcal{N}$ that node j *communicates with* node i, denoted as $i\mathcal{C}j$, if either $i = j$ or there exists a path from i to j and a path from j to i. Hence, $i\mathcal{C}j \iff i = j$ or $[i\mathcal{R}j$

and $j\mathcal{R}i]$. Note that the relation "communicates with" is an equivalence relation. Indeed, its reflexivity and symmetry follow by definition, and its transitivity follows by the concatenation of paths.

If a graph $\mathcal{G} = (\mathcal{N}, \mathcal{D})$ is not strongly connected, then not all nodes of $\mathcal{N}$ communicate with each other. In this case, given a node, say, node i, it is possible to distinguish the subset of nodes that communicate with i from the subset of nodes that do not communicate with i. In the first subset all nodes communicate with each other, whereas in the second subset not all nodes necessarily communicate with each other. In the latter case a further subdivision of the nodes is possible. Repeated application of the previous idea therefore yields that the node set $\mathcal{N}$ can be partitioned as $\mathcal{N}_1 \cup \mathcal{N}_2 \cup \cdots \cup \mathcal{N}_q$, where $\mathcal{N}_r,\ r \in \underline{q}$, denotes a subset of nodes that communicate with each other but not with other nodes of $\mathcal{N}$. Recall that a partitioning of a set is a division into nonempty subsets such that the joint union is the whole set and the mutual intersections are all empty.

Given the above partitioning of $\mathcal{N}$, it is possible to focus on subgraphs of $\mathcal{G}$, denoted by $\mathcal{G}_r = (\mathcal{N}_r, \mathcal{D}_r),\ r \in \underline{q}$, where $\mathcal{D}_r$ denotes the subset of $\mathcal{D}$ of arcs that have both the begin node and the end node in $\mathcal{N}_r$. If $\mathcal{D}_r \neq \emptyset$, the subgraph $\mathcal{G}_r = (\mathcal{N}_r, \mathcal{D}_r)$ is known as a maximal strongly connected subgraph (m.s.c.s.) of $\mathcal{G} = (\mathcal{N}, \mathcal{D})$. By definition, nodes in $\mathcal{N}_r$ do not communicate with nodes outside $\mathcal{N}_r$. However, it can happen that $i\mathcal{R}j$ for some $i \in \mathcal{N}_r$ and $j \in \mathcal{N}_{r'}$ with $r \neq r'$, but then the converse (i.e., $j\mathcal{R}i$) does not hold. We denote by $[i] \stackrel{\text{def}}{=} \{j \in \mathcal{N} : i\mathcal{C}j\}$ the set of nodes containing node i that communicate with each other. These nodes together form the equivalence class in which i is contained. Hence, given node $i \in \mathcal{N}$, there exists an $r \in \underline{q}$ such that $i \in \mathcal{N}_r$ and $[i] = \mathcal{N}_r$.

Note that the above partitioning covers *all* nodes of $\mathcal{N}$. If a node of $\mathcal{G}$ is contained in one or more circuits, it communicates with certain other nodes or with itself in case one of the circuits actually is a self-loop. In any case, the arc set of the associated subgraph is not empty. However, if the graph $\mathcal{G}$ contains a node that is not contained in any circuit of $\mathcal{G}$, say, node i, then node i does not communicate with other nodes and it communicates only with itself. Then, by definition, node i forms an equivalence class on its own, so that $[i] = \{i\}$. Because there does not even exist an arc from i to itself, it follows that the associated subgraph is given by $([i], \emptyset)$; i.e., the node set consists of node i only and the arc set is empty. Further, although it is not strongly connected, $([i], \emptyset)$ will be referred to as an m.s.c.s. This is merely done for convenience. Hence, in the following all subgraphs $\mathcal{G}_r = (\mathcal{N}_r, \mathcal{D}_r),\ r \in \underline{q}$, introduced above are referred to as m.s.c.s.'s.

We define the *reduced graph*, denoted by $\widetilde{\mathcal{G}} = (\widetilde{\mathcal{N}}, \widetilde{\mathcal{D}})$, by $\widetilde{\mathcal{N}} = \{[i_1], \ldots, [i_q]\}$ and $([i_r], [i_s]) \in \widetilde{\mathcal{D}}$ if $r \neq s$ and there exists an arc $(k, l) \in \mathcal{D}$ for some $k \in [i_r]$ and $l \in [i_s]$. Hence, the number of nodes in the reduced graph is exactly the number of m.s.c.s.'s in the graph. The reduced graph models the interdependency of m.s.c.s.'s.

Note that the reduced graph does not contain circuits. Indeed, if the reduced graph would contain a circuit, then two or more m.s.c.s.'s would be connected to each other by means of a circuit, forming a new m.s.c.s. larger than the m.s.c.s.'s it contains. However, this would contradict the fact that these subgraphs already were maximal and strongly connected.

Let A_{rr} denote the matrix obtained by restricting A to the nodes in $[i_r]$, for all $r \in \underline{q}$, i.e., $[A_{rr}]_{kl} = a_{kl}$ for all $k, l \in [i_r]$. Notice that for all $r \in \underline{q}$ either A_{rr} is irreducible or $A_{rr} = \varepsilon$. It is easy to see that because the reduced graph does not contain any circuits, the original reducible matrix A, possibly after a relabeling of the nodes in $\mathcal{G}(A)$, can be written in the form

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & \cdots & A_{1q} \\ \mathcal{E} & A_{22} & \cdots & \cdots & A_{2q} \\ \mathcal{E} & \mathcal{E} & A_{33} & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathcal{E} & \mathcal{E} & \cdots & \mathcal{E} & A_{qq} \end{pmatrix},$$

with matrices $A_{sr}, 1 \leq s < r \leq q$, of appropriate size. Each finite entry in A_{sr} corresponds to an arc from a node in $[i_r]$ to a node in $[i_s]$. The block upper triangular form shown above is said to be a *normal form* of matrix A. Note that the normal form of a matrix is not unique.

The notion of the cyclicity of a graph plays an important role in this book.

DEFINITION 2.3 *The cyclicity of a graph $\mathcal{G}$, denoted by $\sigma_{\mathcal{G}}$, is defined as follows:*

- *If $\mathcal{G}$ is strongly connected, then its cyclicity equals the greatest common divisor of the lengths of all elementary circuits in $\mathcal{G}$. If $\mathcal{G}$ consists of just one node without a self-loop, then its cyclicity is defined to be one.*
- *If $\mathcal{G}$ is not strongly connected, then its cyclicity equals the least common multiple of the cyclicities of all maximal strongly connected subgraphs of $\mathcal{G}$.*

We continue with further definitions of some graph-theoretical notions. The set of *direct predecessors* of node i is denoted by $\pi(i)$; more formally,

$$\pi(i) \stackrel{\text{def}}{=} \{j \in \underline{n} : (j, i) \in \mathcal{D}\}.$$

Moreover, denote the set of all *predecessors* of node i by

$$\pi^+(i) \stackrel{\text{def}}{=} \{j \in \underline{n} : j\mathcal{R}i\},$$

and set $\pi^*(i) = \{i\} \cup \pi^+(i)$. In words, $\pi(i)$ is the set of nodes immediately upstream of i; $\pi^+(i)$ is the set of all nodes from which node i can be reached; and $\pi^*(i)$ is the set of all nodes from which node i can be reached, including node i itself. In the same vein, denote the set of *direct successors* of node i by $\sigma(i)$, more formally,

$$\sigma(i) \stackrel{\text{def}}{=} \{j \in \underline{n} : (i, j) \in \mathcal{D}\};$$

write

$$\sigma^+(i) \stackrel{\text{def}}{=} \{j \in \underline{n} : i\mathcal{R}j\}$$

for the set of all *successors* of node i; and set $\sigma^*(i) = \{i\} \cup \sigma^+(i)$. In words, $\sigma(i)$ is the set of nodes immediately downstream of i; $\sigma^+(i)$ is the set of all nodes that can be reached from node i; and $\sigma^*(i)$ is the set of all nodes that can be reached from i, including node i itself.

Example 2.1.3 *Let*

$$A = \begin{pmatrix} \varepsilon & 0 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & -3 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & 4 & \varepsilon & 0 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ 0 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & 16 & \varepsilon & \varepsilon & \varepsilon & -5 & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 0 & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & 9 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \frac{1}{2} & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 6 & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ 9 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \end{pmatrix}.$$

The communication graph of A is shown in Figure 2.2. The graph $\mathcal{G}(A)$ has node

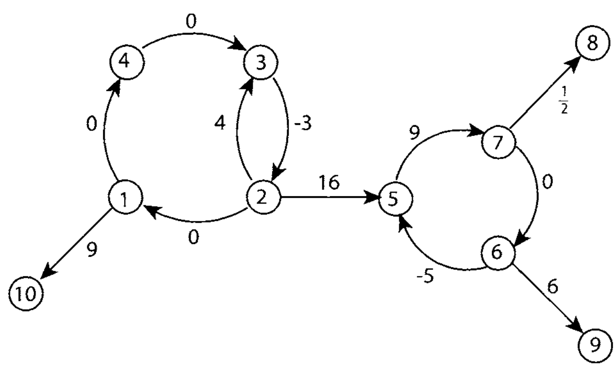

Figure 2.2: The communication graph of matrix A in Example 2.1.3.

set $\mathcal{N}(A) = \{1, 2, 3, \ldots, 10\}$ and arc set $\mathcal{D}(A) = \{$(2,1), (3,2), (2,3), (4,3), (1,4), $(2,5), (6,5), (7,6), (5,7), (7,8), (6,9), (1,10)\}$. Specifically, $\mathcal{G}(A)$ contains three elementary circuits $\rho = ((2,3),(3,2)), \theta = ((4,3),(3,2),(2,1),(1,4))$, and $\eta = ((6,5),(5,7),(7,6))$. The graph is not strongly connected, for example, $2\mathcal{R}5$, but it does not hold that $5\mathcal{R}2$. In words, node 5 is reachable from node 2, but the converse is not true.

The predecessor and successor sets are given for node 10, for example, by

$$\pi(10) = \{1\}\,,\ \pi^+(10) = \{1,2,3,4\}\,,\ \pi^*(10) = \{1,2,3,4,10\},$$

and

$$\sigma(10) = \sigma^+(10) = \emptyset\,,\ \sigma^*(10) = \{10\}.$$

For node 5 *these sets read*

$$\pi(5) = \{2,6\}\,,\ \pi^+(5) = \pi^*(5) = \{1,2,3,4,5,6,7\},$$

and

$$\sigma(5) = \{7\}\,,\ \sigma^+(5) = \sigma^*(5) = \{5,6,7,8,9\}.$$

There are five m.s.c.s.'s in $\mathcal{G}(A)$ *with the set of nodes* $[1] = [2] = [3] = [4] = \{1,2,3,4\}$, $[5] = [6] = [7] = \{5,6,7\}$, $[8] = \{8\}$, $[9] = \{9\}$, *and* $[10] = \{10\}$. *Because* $|\rho|_1 = 2$ *and* $|\theta|_1 = 4$, *the m.s.c.s. corresponding to, for instance,* $[2]$ *has cyclicity* 2, *being the greatest common divisor of all circuit lengths in* $[2]$. *Because* $|\eta|_1 = 3$, *the m.s.c.s. corresponding to, say* $[5]$, *has cyclicity* 3. *The other m.s.c.s.'s have cyclicity* 1 *by definition. Therefore, the graph* $\mathcal{G}(A)$ *has cyclicity* 6, *being the least common multiple of the cyclicity of all m.s.c.s.'s. Hence,* $\sigma_{\mathcal{G}(A)} = 6$.

The reduced graph is depicted in Figure 2.3.

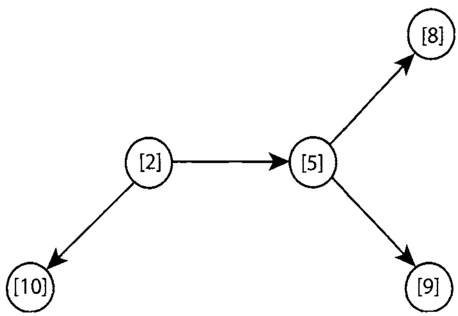

Figure 2.3: The reduced graph of $\mathcal{G}(A)$ in Example 2.1.3.

Based on the reduced graph, let $[i_1] = [8] = \{8\}$, $[i_2] = [9] = \{9\}$, $[i_3] = [5] = \{5,6,7\}$, $[i_4] = [10] = \{10\}$, *and* $[i_5] = [2] = \{1,2,3,4\}$. *The corresponding matrices are*

$$A_{11} = A_{22} = A_{44} = \varepsilon,\ A_{33} = \begin{pmatrix} \varepsilon & -5 & \varepsilon \\ \varepsilon & \varepsilon & 0 \\ 9 & \varepsilon & \varepsilon \end{pmatrix},\ A_{55} = \begin{pmatrix} \varepsilon & 0 & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & -3 & \varepsilon \\ \varepsilon & 4 & \varepsilon & 0 \\ 0 & \varepsilon & \varepsilon & \varepsilon \end{pmatrix}.$$

If both the rows and columns of A are placed into the order

$$8,9,5,6,7,10,1,2,3,4,$$

obtained from placing the elements of the sets $[i_1]$, $[i_2]$, $[i_3]$, $[i_4]$, *and* $[i_5]$ *one after* *another, then the following normal form of A is the result*

$$\begin{pmatrix} \varepsilon & \varepsilon & \varepsilon & \varepsilon & \frac{1}{2} & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & 6 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & -5 & \varepsilon & \varepsilon & \varepsilon & 16 & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & 0 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 9 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 9 & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 0 & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & -3 & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 4 & \varepsilon & 0 \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 0 & \varepsilon & \varepsilon & \varepsilon \end{pmatrix}.$$

In particular, the diagonal blocks of this normal form of A, starting in the upper left corner and going down to the lower right corner, are given by A_{11}, A_{22}, A_{33}, A_{44}, and A_{55}, respectively.

2.2 EIGENVALUES AND EIGENVECTORS

We start with the definition of one of the most important notions in this book.

DEFINITION 2.4 *Let $A \in \mathbb{R}_{\max}^{n \times n}$ be a square matrix. If $\mu \in \mathbb{R}_{\max}$ is a scalar and $v \in \mathbb{R}_{\max}^{n}$ is a vector that contains at least one finite element such that*

$$A \otimes v = \mu \otimes v,$$

then μ is called an eigenvalue *of A and v an* eigenvector *of A associated with eigenvalue μ.*

Note that Definition 2.4 allows an eigenvalue to be $\varepsilon = -\infty$. Also, eigenvectors are allowed to have elements equal to ε as long as they contain finite elements. Note further that a square matrix may have more than one eigenvalue; i.e., eigenvalues are not necessarily unique. Eigenvectors are certainly not unique. Indeed, if v is an eigenvector, then so is $\alpha \otimes v$, with α an arbitrary finite number. Furthermore, observe that the set of all eigenvectors associated with an eigenvalue is a vector space in the max-plus sense, called the eigenspace. Indeed, if v and w are eigenvectors of A associated with eigenvalue μ and α, β are numbers in $\mathbb{R}_{\max}$, then

$$\begin{aligned} A \otimes (\alpha \otimes v \oplus \beta \otimes w) &= \alpha \otimes A \otimes v \oplus \beta \otimes A \otimes w \\ &= \alpha \otimes \mu \otimes v \oplus \beta \otimes \mu \otimes w \\ &= \mu \otimes \big(\alpha \otimes v \oplus \beta \otimes w\big). \end{aligned}$$

Then, if at least one of the numbers α and β is finite, the vector $\alpha \otimes v \oplus \beta \otimes w$ is an eigenvector of A for the eigenvalue μ. The eigenspace of matrix A associated to an eigenvalue μ is denoted by $V(A, \mu)$. When it is clear from the context that the eigenvalue of matrix A is unique, then this eigenvalue often will be denoted by $\lambda(A)$ and the corresponding eigenspace will in that case be denoted by $V(A)$.

A first observation on eigenvalues and eigenvectors is that any finite eigenvalue μ of a square matrix A is the average weight of some circuit in $\mathcal{G}(A)$. To see this, notice that by definition an associated eigenvector v has at least one finite element; that is, there exists a node/index $\eta_1 \in \underline{n}$ such that $v_{\eta_1} \neq \varepsilon$. Then $[A \otimes v]_{\eta_1} = \mu \otimes v_{\eta_1} \neq \varepsilon$. Hence, there exists a node η_2 with

$$a_{\eta_1 \eta_2} \otimes v_{\eta_2} = \mu \otimes v_{\eta_1},$$

implying that $a_{\eta_1 \eta_2} \neq \varepsilon$, $v_{\eta_2} \neq \varepsilon$, and $(\eta_2, \eta_1) \in \mathcal{D}(A)$. Following the same reasoning, a node η_3 can be found such that

$$a_{\eta_2 \eta_3} \otimes v_{\eta_3} = \mu \otimes v_{\eta_2},$$

with $a_{\eta_2 \eta_3} \neq \varepsilon$, $v_{\eta_3} \neq \varepsilon$, and $(\eta_3, \eta_2) \in \mathcal{D}(A)$. Proceeding in this way, eventually some node, say, node η_h, must be encountered for a second time, because the number of nodes is finite. We have then found a circuit

$$\gamma = ((\eta_h, \eta_{h+l-1}), (\eta_{h+l-1}, \eta_{h+l-2}), \ldots, (\eta_{h+1}, \eta_h))$$

of length

$$|\gamma|_1 = l, \tag{2.3}$$

with weight

$$|\gamma|_w = \bigotimes_{k=0}^{l-1} a_{\eta_{h+k} \eta_{h+k+1}}, \tag{2.4}$$

where $\eta_h = \eta_{h+l}$. By construction,

$$\bigotimes_{k=0}^{l-1} \left(a_{\eta_{h+k} \eta_{h+k+1}} \otimes v_{\eta_{h+k+1}} \right) = \mu^{\otimes l} \otimes \bigotimes_{k=0}^{l-1} v_{\eta_{h+k}}.$$

Recall that $\otimes$ reads as $+$ in conventional algebra. Hence, the above equation reads

$$\sum_{k=0}^{l-1} \left(a_{\eta_{h+k} \eta_{h+k+1}} + v_{\eta_{h+k+1}} \right) = l \times \mu + \sum_{k=0}^{l-1} v_{\eta_{h+k}}. \tag{2.5}$$

Because $\eta_h = \eta_{h+l}$ it follows that

$$\sum_{k=0}^{l-1} v_{\eta_{h+k+1}} = \sum_{k=0}^{l-1} v_{\eta_{h+k}}.$$

Subtracting $\sum_{k=0}^{l-1} v_{\eta_{h+k}}$ on both sides of equation (2.5) yields

$$\sum_{k=0}^{l-1} a_{\eta_{h+k} \eta_{h+k+1}} = l \times \mu,$$

which by (2.4) means that

$$|\gamma|_w = l \times \mu = \mu^{\otimes l}.$$

By (2.3), the average weight of the circuit γ then equals

$$\frac{|\gamma|_w}{|\gamma|_1} = \frac{1}{l} \times \mu^{\otimes l} = \mu.$$

We have thus proved the following lemma.

LEMMA 2.5 *Let $A \in \mathbb{R}_{\max}^{n\times n}$ have a finite eigenvalue μ. Then, a circuit γ exists in $\mathcal{G}(A)$ such that*

$$\mu = \frac{|\gamma|_{\rm w}}{|\gamma|_1}.$$

According to the above lemma, average weights of circuits are candidates for eigenvalues. Unfortunately, the above lemma does not tell which circuits actually define an eigenvalue. So, why not try the maximal average circuit weight as a first candidate for an eigenvalue? This choice has the advantage that it is independent of any a priori knowledge about particular circuits. In the following this idea is pursued further.

Let $\mathcal{C}(A)$ denote the set of all elementary circuits in $\mathcal{G}(A)$ and write

$$\lambda = \max_{p\in\mathcal{C}(A)} \frac{|p|_{\rm w}}{|p|_1} \tag{2.6}$$

for the maximal average circuit weight. Notice that $\mathcal{C}(A)$ is a finite set, and if not empty, the maximum on the right-hand side in (2.6) is thus attained by (at least) one circuit in $\mathcal{G}(A)$. In the case where $\mathcal{C}(A) = \emptyset$, define $\lambda = -\infty$. Notice that if A is irreducible, then λ is finite (irreducibility of A implies that $\mathcal{G}(A)$ contains at least one circuit). Note that if $\mathcal{G}(A)$ contains no circuit (for example, if A is a strictly lower triangular matrix), then, according to this definition, $\lambda = -\infty$.

A circuit p in $\mathcal{G}(A)$ is called *critical* if its average weight is maximal, that is, if $\lambda = |p|_{\rm w}/|p|_1$. The *critical graph* of A, denoted by $\mathcal{G}^{\rm c}(A) = (\mathcal{N}^{\rm c}(A), \mathcal{D}^{\rm c}(A))$, is the graph consisting of those nodes and arcs that belong to critical circuits in $\mathcal{G}(A)$. A node $i \in \mathcal{N}^{\rm c}(A)$ will sometimes be referred to as a *critical node*. Similarly, a subpath of a critical circuit will be occasionally called a *critical path*. Note that the critical graph of an irreducible matrix does not have to be strongly connected.

Example 2.2.1 *Revisit the situation put forward in Example 2.1.3. The graph $\mathcal{G}(A)$ contains three circuits with average weight $1/2$, $-3/4$, and $4/3$, respectively. The maximal average circuit weight is therefore equal to $4/3$, and the critical graph consists of the circuit $\eta = ((6,5),(5,7),(7,6))$.*

Let A be a square matrix over $\mathbb{R}_{\max}$ whose communication graph contains at least one circuit. Let A' be the matrix obtained from A by subtracting $\tau \in \mathbb{R}$ from every finite element of A. Then the communication graphs of A and A' are the same, except for the arc weights. Furthermore, if a circuit in $\mathcal{G}(A)$ has average weight α, then the same circuit in $\mathcal{G}(A')$ has average weight $\alpha - \tau$. It follows that a circuit in $\mathcal{G}(A)$ is critical if and only if it is critical in $\mathcal{G}(A')$. Hence, the critical graphs $\mathcal{G}^{\rm c}(A)$ and $\mathcal{G}^{\rm c}(A')$ are the same, again except for the arc weights. Clearly, by taking τ equal to the maximal average circuit weight in $\mathcal{G}(A)$, the maximal average circuit weight in $\mathcal{G}(A')$ becomes zero.

LEMMA 2.6 *Assuming that $\mathcal{G}(A)$ contains at least one circuit, it follows that any circuit in $\mathcal{G}^{\rm c}(A)$ is critical.*

Proof. The proof will be given by contradiction. Take λ as defined in (2.6). As indicated above, it may, for ease of exposition, be assumed that $\lambda = 0$. Suppose that $\mathcal{G}^c(A)$ contains a circuit ρ with average weight different from zero. Note that ρ is also a circuit in $\mathcal{G}(A)$. If the average weight of ρ is larger than zero, then the maximal average circuit weight of A is larger than zero, which contradicts the starting point that $\lambda = 0$. Now suppose that the average weight of ρ is negative. Observe that ρ is the concatenation of paths ρ_i, i.e., $\rho = \rho_1 \circ \rho_2 \circ \cdots \circ \rho_\kappa$, where each ρ_i is a subpath of a critical circuit c_i, $i \in \underline{\kappa}$. Hence, there exist subpaths ξ_i such that $c_i = \xi_i \circ \rho_i$, $i \in \underline{\kappa}$. Since all circuits c_i, $i \in \underline{\kappa}$, have weight zero, the circuit composed of the concatenation of the ξ_i's (i.e., $\xi = \xi_1 \circ \xi_2 \circ \cdots \circ \xi_\kappa$) is thus a circuit with positive (average) weight, which contradicts again the starting point that $\lambda = 0$. □

If not stated otherwise, we adopt throughout the text the convention that for any $x \in \mathbb{R}$

$$\varepsilon = \varepsilon - x \qquad \text{and} \qquad \varepsilon - \varepsilon = e. \tag{2.7}$$

Let λ, defined in (2.6), be finite, and consider the matrix A_λ with elements

$$[A_\lambda]_{ij} \ = \ a_{ij} - \lambda. \tag{2.8}$$

Matrix A_λ is occasionally referred to as the *normalized* matrix. It is clear that the maximum average circuit weight of $\mathcal{G}(A_\lambda)$ is zero. Therefore, Lemma 2.2 implies that A_λ^+ is well defined, where A_λ^+ should be read as $(A_\lambda)^+$. As noticed before, the set of critical circuits of A and A_λ coincide, and consequently, $\mathcal{G}^c(A)$ and $\mathcal{G}^c(A_\lambda)$ coincide except for their weights. This gives

$$\forall \eta \in \mathcal{N}^c(A): \quad [A_\lambda^+]_{\eta\eta} = e = 0. \tag{2.9}$$

Indeed, every node of the critical graph is contained in a circuit and every circuit of the critical graph has weight zero. So, any path from a node in the critical graph to itself has weight zero. Next, define

$$A_\lambda^* \overset{\text{def}}{=} E \oplus A_\lambda^+ \ = \ \bigoplus_{k \geq 0} A_\lambda^{\otimes k}, \tag{2.10}$$

where A_λ^* stands for $(A_\lambda)^*$, and notice that

$$A_\lambda^+ \ = \ A_\lambda \otimes (E \oplus A_\lambda^+) \ = \ A_\lambda \otimes A_\lambda^*. \tag{2.11}$$

Let $[B]_{\cdot k}$ denote the kth column of a matrix B. The definition of A_λ^* implies that

$$[A_\lambda^*]_{\cdot\eta} \ = \ [E \oplus A_\lambda^+]_{\cdot\eta}\,. \tag{2.12}$$

It follows from (2.12) that the ith element of the vector $[A_\lambda^*]_{\cdot\eta}$ satisfies

$$[A_\lambda^*]_{i\eta} = [E \oplus A_\lambda^+]_{i\eta} = \begin{cases} \varepsilon \oplus [A_\lambda^+]_{i\eta} & \text{for } i \neq \eta, \\ e \oplus [A_\lambda^+]_{i\eta} & \text{for } i = \eta. \end{cases}$$

Then from (2.9) for $\eta \in \mathcal{N}^c(A)$ it follows that

$$[A_\lambda^+]_{\cdot\eta} \ = \ [A_\lambda^*]_{\cdot\eta}.$$

If we replace A_λ^+ by $A_\lambda \otimes A_\lambda^*$ (see (2.11)), then the above equality is equivalent to

$$[A_\lambda \otimes A_\lambda^*]_{\cdot\eta} \ = \ [A_\lambda^*]_{\cdot\eta},$$

which gives

$$A_\lambda \otimes [A_\lambda^*]_{\cdot\eta} = [A_\lambda^*]_{\cdot\eta}$$

or, equivalently,

$$A \otimes [A_\lambda^*]_{\cdot\eta} = \lambda \otimes [A_\lambda^*]_{\cdot\eta}.$$

Hence, it follows that λ is an eigenvalue of A and that the ηth column of A_λ^* is an associated eigenvector, for any $\eta \in \mathcal{N}^c(A)$. We summarize our analysis as follows.

LEMMA 2.7 *Let the communication graph $\mathcal{G}(A)$ of matrix $A \in \mathbb{R}_{\max}^{n\times n}$ have finite maximal average circuit weight λ. Then, the scalar λ is an eigenvalue of A, and the column $[A_\lambda^*]_{\cdot\eta}$ is an eigenvector of A associated with λ, for any node η in $\mathcal{G}^c(A)$.*

Lemma 2.7 establishes the existence of an eigenvalue and of associated eigenvectors provided that the maximal average circuit weight is indeed finite. Notice that the irreducibility of A already implies that the maximal average circuit weight is finite. As we will show next, the irreducibility of A moreover implies that any eigenvector associated with any finite eigenvalue of A has only finite elements.

Let v be an eigenvector of A associated with eigenvalue μ, and call the set of nodes of $\mathcal{G}(A)$ corresponding to finite entries of v the *support* of v. Suppose that the support of v does not cover the whole node set of $\mathcal{G}(A)$. If A is irreducible, then any node can be reached from any node and there have to be arcs from the nodes in the support of v going to nodes not belonging to the support of v. Hence, there exists a node j in the support of v and a node i not in the support of v with $a_{ij} \neq \varepsilon$. Then, $a_{ij} \neq \varepsilon$ implies that $[A \otimes v]_i \geq a_{ij} \otimes v_j > \varepsilon$, and the support of $A \otimes v$ is thus bigger than the support of v. Since $\mu \otimes v = A \otimes v$, this contradicts the fact that the support of v and $\mu \otimes v$ have to be equal for any finite μ. We summarize the above analysis as follows.

LEMMA 2.8 *Let $A \in \mathbb{R}_{\max}^{n\times n}$. If A is irreducible, then any vector $v \in \mathbb{R}_{\max}^{n}$, with at least one finite element, that solves*

$$\mu \otimes v = A \otimes v$$

for some finite μ has all elements different from ε.

Example 2.2.2 *Consider*

$$A = \begin{pmatrix} e & \varepsilon \\ 1 & e \end{pmatrix}.$$

The powers of A are as follows

$$A^{\otimes 2} = \begin{pmatrix} e & \varepsilon \\ 1 & e \end{pmatrix}, \qquad A^{\otimes 3} = \begin{pmatrix} e & \varepsilon \\ 1 & e \end{pmatrix}, \ldots,$$

and we conclude that

$$A^{\otimes(k+1)} = e \otimes A^{\otimes k}, \qquad k \geq 1.$$

Hence, we obtain $\lambda = e\,(=0)$ as an eigenvalue of A. Moreover, it follows that

$$A_\lambda = A = A_\lambda^* = A_\lambda^+.$$

Hence, in accordance with Lemma 2.7,

$$\begin{pmatrix} e \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \varepsilon \\ e \end{pmatrix}$$

are eigenvectors of A associated with λ. Notice that A is not irreducible and Lemma 2.8 does not apply, as is illustrated by the eigenvector $(\varepsilon, e)^\top$, which has a nonfinite element.

The existence of a finite eigenvalue has been shown in Lemma 2.7 for the case where A is irreducible. The next step is to show that irreducibility also implies that the eigenvalue of A is unique. Pick any circuit $\gamma = ((\eta_1, \eta_2), (\eta_2, \eta_3), \ldots, (\eta_l, \eta_{l+1}))$ in $\mathcal{G}(A)$ of length $l = |\gamma|_1$ with $\eta_{l+1} = \eta_1$. Then,

$$a_{\eta_{k+1}\eta_k} \neq \varepsilon\,, \qquad k \in \underline{l}. \tag{2.13}$$

Suppose that μ is a finite eigenvalue of A, and let v be an eigenvector associated with μ. Because it is assumed that $\mu \otimes v = A \otimes v$, it follows that

$$a_{\eta_{k+1}\eta_k} \otimes v_{\eta_k} \leq \mu \otimes v_{\eta_{k+1}}\,, \qquad k \in \underline{l}.$$

Now argue as in the proof of Lemma 2.5, except that the equalities are replaced by appropriate inequalities. See the text before Lemma 2.5. Proceeding in the above way, we find that the average weight of the circuit γ satisfies

$$\frac{|\gamma|_{\mathrm{w}}}{|\gamma|_1} \leq \frac{1}{l} \times \mu^{\otimes l} = \mu.$$

The above analysis holds for any circuit $\gamma \in \mathcal{C}(A)$. In other words, any finite eigenvalue μ has to be larger than or equal to the maximal average circuit weight λ. But, by Lemma 2.5, any finite eigenvalue μ can always be obtained as the average weight of a circuit in $\mathcal{G}(A)$. Hence, λ is a finite eigenvalue of A, and by (2.6) it is uniquely determined.

Suppose now that ε is an eigenvalue of A with corresponding eigenvector v. Then, v has at least one finite element, say, v_η. If A is irreducible, then there is a row γ of A such that $a_{\gamma\eta}$ is finite, which gives

$$\varepsilon = [\varepsilon \otimes v]_\gamma = [A \otimes v]_\gamma \geq a_{\gamma\eta} \otimes v_\eta.$$

But the expression on the above right-hand side is finite, and we conclude that ε cannot be an eigenvalue of an irreducible matrix. Consequently, λ is the unique eigenvalue of A.

For easy reference we summarize our analysis in the following theorem.

THEOREM 2.9 *Any irreducible matrix $A \in \mathbb{R}_{\max}^{n \times n}$ possesses one and only one eigenvalue. This eigenvalue, denoted by $\lambda(A)$, is a finite number and equal to the maximal average weight of circuits in $\mathcal{G}(A)$, i.e.,*

$$\lambda(A) = \max_{\gamma \in \mathcal{C}(A)} \frac{|\gamma|_{\mathrm{w}}}{|\gamma|_1}.$$

Example 2.2.3 *Revisit the situation put forward in Example 2.1.1. The communication graph of A, given in Figure 2.1, is strongly connected, and A is therefore irreducible. Elaborating on the notation already introduced in Example 2.1.1, we obtain* $\mathcal{C}(A) = \{\rho, \theta\}$. *The average circuit weights are* $|\rho|_{\mathrm{w}}/|\rho|_1 = 13$ *and* $|\theta|_{\mathrm{w}}/|\theta|_1 = 12$. *Theorem 2.9 applies, and we obtain* $\lambda(A) = \max(13, 12) = 13$ *for the eigenvalue of A.*

Theorem 2.9 characterizes the eigenvalue of an irreducible square matrix. Algorithms for computing the eigenvalue will be presented in Chapters 5 and 6.

2.3 SOLVING LINEAR EQUATIONS

Following (2.2) and (2.10), we formally define for any $A \in \mathbb{R}_{\max}^{n \times n}$

$$A^* \stackrel{\text{def}}{=} E \oplus A^+ = \bigoplus_{k \geq 0} A^{\otimes k}. \tag{2.14}$$

From Lemma2.2 it follows easily that A^* exists for any square matrix A with a communication graph $\mathcal{G}(A)$ having only nonpositive circuit weights. Note that $A^{\otimes n}$ refers to the maximal weight of paths of length n. Hence, these paths contain at least one circuit. If all circuits have nonpositive circuit weight, then

$$\left[A^{\otimes n}\right]_{ij} \leq \bigotimes_{k=0}^{n-1} \left[A^{\otimes k}\right]_{ij}, \qquad i, j \in \underline{n},$$

and under the conditions put forward in Lemma 2.2, A^* can be determined as

$$A^* = \bigoplus_{k=0}^{n-1} A^{\otimes k}. \tag{2.15}$$

As will be shown in the next theorem, the operator $(\cdot)^*$ provides the means for solving the equation $x = A \otimes x \oplus b$, with $A \in \mathbb{R}_{\max}^{n \times n}$ and $b \in \mathbb{R}_{\max}^{n}$. The precise statement is as follows.

THEOREM 2.10 *Let* $A \in \mathbb{R}_{\max}^{n \times n}$ *and* $b \in \mathbb{R}_{\max}^{n}$. *If the communication graph* $\mathcal{G}(A)$ *has maximal average circuit weight less than or equal to e, then the vector* $x = A^* \otimes b$ *solves the equation* $x = (A \otimes x) \oplus b$. *Moreover, if the circuit weights in* $\mathcal{G}(A)$ *are negative, then the solution is unique.*

Proof. It will be shown that

$$A^* \otimes b = A \otimes (A^* \otimes b) \oplus b.$$

By Lemma 2.2, A^* exists, implying that

$$\begin{aligned} A^* \otimes b &= \bigoplus_{k \geq 0} A^{\otimes k} \otimes b \\ &= \left(\bigoplus_{k \geq 1} A^{\otimes k} \otimes b \right) \oplus (E \otimes b) \\ &= A \otimes \left(\bigoplus_{k \geq 0} A^{\otimes k} \otimes b \right) \oplus (E \otimes b) \\ &= A \otimes (A^* \otimes b) \oplus b, \end{aligned}$$

which concludes the proof of the first part of the theorem.

In order to prove uniqueness under the condition that circuits have negative average weights, argue as follows. Suppose that x is a solution of $x = b \oplus (A \otimes x)$; then substituting the expression for x in $b \oplus (A \otimes x)$ it follows that

$$x = b \oplus (A \otimes b) \oplus (A^{\otimes 2} \otimes x),$$

and repeating the argument, one obtains

$$\begin{aligned} x &= b \oplus (A \otimes b) \oplus (A^{\otimes 2} \otimes b) \oplus (A^{\otimes 3} \otimes x) \\ &= b \oplus (A \otimes b) \oplus \cdots \oplus (A^{\otimes (k-1)} \otimes b) \oplus (A^{\otimes k} \otimes x) \\ &= \bigoplus_{l=0}^{k-1} (A^{\otimes l} \otimes b) \oplus (A^{\otimes k} \otimes x). \end{aligned} \tag{2.16}$$

The entries of $A^{\otimes k}$ are the maximal weights of paths of length k. For k large enough, any path necessarily contains one or more copies of certain elementary circuits as subpaths, and as k tends to ∞, the number of required elementary circuits tends to ∞. Since circuits have negative weight, the elements of $A^{\otimes k}$ tend to ε, as k tends to ∞, i.e.,

$$\lim_{k \to \infty} A^{\otimes k} \otimes x = \mathcal{E}.$$

Hence, letting k tend to ∞ in equation (2.16) yields that $x = A^* \otimes b$, where Lemma 2.2 is used to show that

$$\lim_{k \to \infty} \bigoplus_{l=0}^{k-1} (A^{\otimes l} \otimes b) = \left(\lim_{k \to \infty} \bigoplus_{l=0}^{k-1} A^{\otimes l} \right) \otimes b = A^* \otimes b.$$

□

To conclude this section, let us consider a different kind of linear equation, namely, $A \otimes x = b$, where A is not necessarily square! A study shows that a solution of this equation does not always exist. However, it will be shown below that one can always find a *greatest solution* to the max-plus inequality $A \otimes x \leq b$. This greatest solution is called the *principal solution* and will be denoted as $x^*(A, b)$.

THEOREM 2.11 *For $A \in \mathbb{R}_{\max}^{m \times n}$ and $b \in \mathbb{R}^m$ it holds that*

$$[x^*(A, b)]_j = \min\{b_i - a_{ij} \; : \; i \in \underline{m}\},$$

for $j \in \underline{n}$.

Proof. Notice that if $A \otimes x \leq b$, then the following equivalences hold:

$$\begin{aligned} \forall i \, \forall j : \quad & a_{ij} + x_j \leq b_i \\ \Leftrightarrow \forall i \, \forall j : \quad & x_j \leq b_i - a_{ij} \\ \Leftrightarrow \forall j : \quad & x_j \leq \min\{b_i - a_{ij} \ : \ i \in \underline{m}\}. \end{aligned}$$

□

Please note that in Theorem 2.11 a requirement is that $b \in \mathbb{R}^m$. This theorem may not hold for $b \in \mathbb{R}^m_{\max}$. A counterexample is the scalar equation $\varepsilon \otimes x = \varepsilon$.

From Theorem 2.11 it follows that $x^*(A, b)$, seen as greatest solution of the inequality $A \otimes x \leq b$, is uniquely determined. Indeed, any other solution x of the inequality is such that $x \leq x^*(A, b)$. This situation is different as far as the inequality $A \otimes x \geq b$ is concerned. For such an inequality a *smallest solution* need not be uniquely determined nor even exist.

By Theorem 2.11, any vector x with $x \leq x^*(A, b)$ satisfies $A \otimes x \leq b$. Notice that if A has a column, say, column j, with all elements equal to ε, then $[x^*(A, b)]_j = +\infty$, which means that $A \otimes x \leq b$ can be achieved by a vector whose jth component may take any value in $\mathbb{R}_{\max}$.

Let A be a matrix with the travel times of a train network. Suppose that the planned departure times are given by b; then $x^*(A, b)$ gives the latest departure times of trains from the previous stations such that b still can be met. This kind of *just-in-time* application will be dealt with in more detail in Section 9.1.

Theorem 2.11 is a special result coming from the so-called residuation theory by which optimal solutions can be obtained for inequalities that are generalizations of the above max-plus inequality $A \otimes x \leq b$. A recent and thorough reference on the subject of residuation theory is [25].

2.4 EXERCISES

1. Show that for any $k \geq 1$
$$\lambda^{\otimes k} \otimes A_\lambda^{\otimes k} = A^{\otimes k}.$$
2. Show that for any $k \geq 1$
$$A_\lambda^{\otimes k} = \left(A^{\otimes k}\right)_\nu,$$
with $\nu = k \times \lambda$ (conventional multiplication).
3. Show that if $A \in \mathbb{R}^{n \times n}_{\max}$ is irreducible, then A is regular.
4. Let $A \in \mathbb{R}^{n \times n}_{\max}$ be irreducible. Show that for any node i in $\mathcal{G}(A)$, it holds that $\pi^*(i) = \pi^+(i) = \sigma^+(i) = \sigma^*(i) = \underline{n}$.
5. Let λ denote the maximal average circuit weight in the communication graph of a square matrix A. Consider A_λ and show that
$$0 = \max_{p \in \mathcal{C}(A_\lambda)} \frac{|p|_{\mathrm{w}}}{|p|_1}.$$
6. Let $A \in \mathbb{R}^{n \times n}_{\max}$ be a strictly lower triangular matrix, and consider A^* defined in equation (2.14). Give a direct proof for the fact that A^* and A are related as in equation (2.15). Show that the above need not to be true if A is just lower triangular.

7. Consider the graph depicted in Figure 2.1. What is the maximum weight of a path from 1 to 3 of length 4?
8. Compute eigenvectors for matrix A given in Example 2.1.1.
9. Let
$$A = \begin{pmatrix} \varepsilon & e \\ \varepsilon & e \end{pmatrix}.$$
Show that A has eigenvalues ε and e, and give one corresponding eigenvector for each eigenvalue.
10. For the next matrices A, investigate the existence and uniqueness of a solution of the equation $x = A \otimes x \oplus b$ with $b = \mathbf{u}$. If a solution exists, give the complete solution set of the equation.
$$A = \begin{pmatrix} -1 & e \\ -2 & -1 \end{pmatrix}, \quad A = \begin{pmatrix} e & -2 \\ 2 & -1 \end{pmatrix}, \quad A = \begin{pmatrix} e & e \\ 2 & -1 \end{pmatrix}$$
11. Assume that
$$A = \begin{pmatrix} 1 & 2 \\ e & 1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} e \\ 1 \end{pmatrix}.$$
Compute $x^*(A, b)$, as defined in Theorem 2.11, and verify that this vector is the greatest solution of the inequality $A \otimes x \leq b$. Show that for the inequality $A \otimes x \geq b$ no (unique) smallest solution exists.

2.5 NOTES

In conventional algebra, a square matrix is called *irreducible* if no identical permutation of its rows and columns exists such that the matrix is transformed into a block upper triangular structure. As shown in [5], the definition in the current chapter of the irreducibility of a square max-plus matrix is equivalent to a max-plus version of the conventional definition.

Theorem 2.9 is the max-plus analogue of the Perron-Frobenius theorem in conventional linear algebra, which states that an irreducible square nonnegative matrix, say, B has a largest eigenvalue that is positive and real, where *largest* means largest in modulus. It is well known that this eigenvalue is given by $\limsup_{k\to\infty} \left(\mathrm{tr}(B^k)\right)^{1/k}$, where $\mathrm{tr}(B^k)$ denotes the trace of the kth power of the nonnegative matrix B in a conventional sense.

To see the parallel with max plus, notice that the maximal average circuit weight of circuits of length k crossing node i is given by $(1/k) \times [A^{\otimes k}]_{ii}$. Theorem 2.9 then yields

$$\lambda = \bigoplus_{k\geq 1} \left(\frac{1}{k} \times \max_{i\in\underline{n}} [A^{\otimes k}]_{ii} \right) = \bigoplus_{k\geq 1} \left(\frac{1}{k} \times \bigoplus_{i=1}^{n} [A^{\otimes k}]_{ii} \right) = \bigoplus_{k\geq 1} \left(\mathrm{tr}_{\oplus}(A^{\otimes k}) \right)^{\otimes(1/k)},$$

where $\mathrm{tr}_{\oplus}(A^{\otimes k}) \stackrel{\text{def}}{=} \bigoplus_{i=1}^{n} [A^{\otimes k}]_{ii}$ stands for the trace of $A^{\otimes k}$ in max-plus sense. Note that $\bigoplus_{k\geq 1}$ can be seen as $\sup_{k\geq 1}$. Circuits of length larger than n can be built up from circuits of length at most n. Therefore, the maximal average weight of these two types of circuits can be expressed in terms of each other. From this it easily follows that for all $h \geq n$

$$\lambda = \sup_{k\geq 1} \left(\mathrm{tr}_{\oplus}(A^{\otimes k}) \right)^{\otimes(1/k)} = \sup_{k\geq h} \left(\mathrm{tr}_{\oplus}(A^{\otimes k}) \right)^{\otimes(1/k)} = \limsup_{h\to\infty} \left(\mathrm{tr}_{\oplus}(A^{\otimes h}) \right)^{\otimes(1/h)},$$

yielding a similar expression for the max-plus eigenvalue as in the conventional case.

The operator $(\cdot)^*$ defined in (2.14) is in the literature referred to as the *Kleene star*. In the present chapter it is shown that the Kleene star of a square matrix over $\mathbb{R}_{\max}$ exists if any cycle weight in its communication graph is nonpositive. The Kleene star again is encountered in Section 4.2, where it is shown that eigenvectors can be characterized through the Kleene star of the normalized matrix, defined in (2.8). See also Lemma 2.7.

Almost twenty years ago, Professor Cuninghame-Green was invited to Delft University of Technology in order to give some lectures about "his" max-plus algebra. We came to talk about the well-known realization problem in mathematical systems theory. The simplest version is as follows. Given a series of scalars g_i, $i \in \mathbb{N}$, find $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$, and $C \in \mathbb{R}^{1\times n}$ such that $CA^iB = g_i$ for all $i \in \mathbb{N}$, everything in conventional algebra, and in such a way that n is as small as possible. The solution in conventional systems theory is well known, but what can one say about the same problem statement and set of equations in max-plus algebra? Professor Cuninghame-Green did not see the solution immediately. But when he left, thanking us for the hospitality, he seemed rather confident and said he would send us the solution the following week.... In the meantime, many papers with partial results have been published, but to the authors' current knowledge, the general solution is still unknown; see [76] and [38].