Technische Universität München

Department of Mathematics

An Algebraic Approach to Understanding Generalized Recursive Max-linear Model

Master Thesis

by

Zhongwei Zhang

Supervisor: Prof. Dr. Claudia Klüppelberg
Advisor: Dipl.-Math. Nadine Gissibl
Submission Date: April 22, 2018

I hereby declare that this thesis is my own work and that no other sources have been used except those clearly indicated and referenced.

Munich, April 22, 2018

Abstract

A generalized recursive max-linear model on a directed graph is a structural equation model where each variable is a max-linear function of their parental node variables and independent noise variables. It is a multidimensional fixed point equation in the max-times algebra $(\overline{\mathbb{R}}_+^{d\times d}, \vee, \odot)$ and its fixed point can be approached by iteration. When time is introduced into this model, it becomes a deterministic max-times system. The stability of this system is of high interest and is investigated when the maximum cycle mean of the coefficient matrix is less than or equal to 1. Inspired by the fact that in a Gaussian graphical model conditional independence given all remaining random variables corresponds to a zero entry in the inverse covariance matrix, the relationship between conditional dependence structure and tail dependence structure for a recursive max-linear model on a polytree is also investigated.

Contents

Chapter 1

Introduction

A recursive max-linear model on a directed acyclic graph (briefly DAG) was introduced in Gissibl and Klüppelberg (2018) and defined as

$$X_i := \bigvee_{k \in \mathrm{pa}(i)} c_{ki} X_k \vee c_{ii} Z_i \ , \ \ i = 1, 2, \ldots, d \tag{1.1}$$

where $Z_1, Z_2, \ldots, Z_d$ are independent noise variables on the support $\overline{\mathbb{R}}_+ := [0, \infty)$, $\mathrm{pa}(i)$ denotes the parents of i, and c_{ki} are non-negative weights, i.e., $c_{ki} > 0$ for $k \in \mathrm{pa}(i) \cup \{i\}$ and $c_{ki} = 0$ for $j \in V \backslash (\mathrm{pa}(i) \cup \{i\})$. It is a combination of graphical models (see Lauritzen (1996) and Koller and Friedman (2009)) and extreme value theory (see de Haan and Ferreira (2006) and Resnick (1987)). On the DAG each node i represents the random variable X_i, and the dependence structure among $X_i, i = \{1, 2, \ldots, d\}$ are encoded in the graph. This model is motivated by modeling extreme risks. Hence the noise variables are assumed to have extreme value distributions, for instance, $\alpha-$Fréchet.

One of the main results in Gissibl and Klüppelberg (2018) is that:

Let $\mathbf{X}$ *be a recursive max-linear model on a DAG* $\mathcal{D} = (V, E)$ *defined as Equation* (1.1)*, and let* $B = (b_{ij})_{d \times d}$ *be the max-linear coefficient matrix with entries defined by* $b_{ii} = c_{ii}, b_{ji} = 0$ *for* $j \in V \backslash (an(i) \cup \{i\})$*, and* $b_{ji} = \bigvee_{p \in P_{ji}} c_{k_0 k_0} \prod_{l=0}^{n-1} c_{k_l k_{l+1}}$ *for* $j \in an(i)$*, where* P_{ji} *is a set of* $j \to i$ *paths,* $p = [j = k_0 \to k_1 \to \cdots \to k_n = i]$*, and* $an(i)$ *denotes the ancestors of* i*, then* $\mathbf{X}$ *is a max-linear model and can be represented as*

$$X_i = \bigvee_{j \in an(i) \cup \{i\}} b_{ji} Z_j, \ i = 1, \ldots, d. \tag{1.2}$$

Furthermore, if we define an operation $\odot$ *as:* $F \odot G := (\bigvee_k f_{ik} g_{kj})$ *for matrices* $F = (f_{ij})$ *and* $G = (g_{ij})$ *with compatible sizes, then the max-linear coefficient matrix* B *can be*

calculated by

$$B = \tilde{C} \text{ for } d = 1 \text{ and } B = \tilde{C} \vee \bigvee_{k=0}^{d-2} (C_1 \odot C^{\odot k}) \text{ for } d \geq 2,$$

where $C = (c_{ij})$, $\tilde{C} := diag(c_{11}, \ldots, c_{dd})$, $C_1 := (c_{ii}c_{ij})_{d\times d}$.

Note that by using the operation $\odot$, Equation (1.1) can be written in matrix form as:

$$\mathbf{X} = (C - \tilde{C})^T \odot \mathbf{X} \vee \tilde{C} \odot \mathbf{Z} \ , \tag{1.3}$$

where $(C - \tilde{C})^T$ is the transpose of $(C - \tilde{C})$. Similarly Equation (1.2) can be written as $\mathbf{X} = B^T \odot \mathbf{Z}$. Hence the above result in fact implies that $\mathbf{X} = B^T \odot \mathbf{Z}$ is the unique solution of the fixed point Equation (1.3).

In Chapter 3 we introduce the max-times algebra $(\overline{\mathbb{R}}_+, \vee, \cdot)$, i.e., a semiring over non-negative real number, equipped with maximum and multiplication as the two binary operations. Similarly as in linear algebra, the operations $(\vee, \cdot)$ can be extended from $\overline{\mathbb{R}}_+$ to vectors in $\overline{\mathbb{R}}_+^d$ and matrices in $\overline{\mathbb{R}}_+^{d\times d}$. It can be easily checked that the algebra $(\overline{\mathbb{R}}_+^{d\times d}, \vee, \odot)$ is also a semiring, where the operation $\vee$ between matrices means that maximum is taken componentwise, and the operation $\odot$ is defined the same as above. Then we focus on the algebra $(\overline{\mathbb{R}}_+^{d\times d}, \vee, \odot)$, introduce the key concepts (maximum cycle mean, weak and strong transitive closures), present their properties and use that to investigate the asymptotic behavior of the $n-$th power of an arbitrary matrix A (here A^n denotes $A \odot A \odot \cdots \odot A$, where A appears n times). These concepts and results in max-times algebra are the fundamentals for later chapters.

It is natural in the usual algebra $(\mathbb{R}, +, \cdot)$ to approach a fix point by iteration (see Goldie and Maller (2000)). In Chapter 4 we use the same idea to solve the fixed point Equation (1.3). We provide an alternative proof, compared to theorem 2.2 in Gissibl and Klüppelberg (2018), for the statement that a recursive max-linear model on a DAG defined as Equation (1.1) is a max-linear model $\mathbf{X} = B^T \odot \mathbf{Z}$. Moreover, we show that the max-linear coefficient matrix B can be computed from matrix C in a much more simplified way: $B = \tilde{C} \odot (C')^{d-1}$, where $\tilde{C} = \text{diag}\{c_{11}, c_{22}, \ldots, c_{dd}\}$ and C' is defined by replacing all the diagonal entries in C by 1 and keeping non-diagonal entries unchanged. Then we investigate a general case, the generalized recursive max-linear model $\mathbf{X}$ on a directed graph $\mathcal{D}$. That is, $\mathbf{X}$ is still defined as in Equation (1.1), but the graph $\mathcal{D}$ is not necessarily acyclic. By using the results from Chapter 3, we show that if the maximum cycle mean of $C - \tilde{C}$ is less than 1, then as the iteration times goes to ∞, the model $\mathbf{X}$ tends to $\mathbf{X}_\infty = B^T \odot \mathbf{Z}$, where B is defined the same as above. Furthermore, it can be easily checked that $\mathbf{X} = B^T \odot \mathbf{Z}$ is a fixed point of Equation (1.3).

In Chapter 5 a deterministic max-times system $(\mathbf{X}_t)_{t\in\mathbb{N}}$, which is defined by $\mathbf{X}_t = A \odot \mathbf{X}_{t-1}$ with finite initial state $\mathbf{X}_0$ and coefficient matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$, is introduced. Its stability and periodicity is investigated, when the maximum cycle mean of A is less than or equal to 1. Then we introduce noise into this system and define $(\mathbf{X}_t)_{t\in\mathbb{N}}$ by $\mathbf{X}_t = A \odot \mathbf{X}_{t-1} \vee \mathbf{Z}_t$ with finite initial state $\mathbf{X}_0$, coefficient matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$ and i.i.d random variables $\mathbf{Z}_t, t \in \mathbb{N}$. We show that if the i.i.d. noise variables $Z_t^i, t \in \mathbb{N}, i = 1, \dots, d$ are regularly varying with index α and the maximum cycle mean of A is less than 1, then $\lim_{n\to\infty} \mathbf{X}_n \in MDA(F)$ with

$$F(\mathbf{x}) = \exp\big(- \sum_{i=1}^{d} (\frac{1}{x_i})^\alpha - \sum_{k=1}^{\infty} \sum_{i=1}^{d} \bigvee_{j=1}^{d} (\frac{a_{ji}^{(k)}}{x_j})^\alpha \big), \quad \mathbf{x} = (x_1, \dots, x_d)^T \in \overline{\mathbb{R}}_+^d \, .$$

Furthermore, if $Z_t^i, t \in \mathbb{N}, i = 1, \dots, d$ follow Fréchet-α distribution, then the distribution of $\mathbf{X}_n$ converges pointwise to F.

It is well known, that for a Gaussian graphical model, if a pair of variables is not contained in the edge set, then they are conditional independent given all remaining variables, and this also corresponds to a zero entry in the precision matrix (see Lauritzen (1996)). Inspired by that idea, in Chapter 6 we investigate the relationship between tail dependence structure (see Equation (2.2)) and conditional dependence structure for a recursive max-linear model on a polytree. We define the permanent of a square matrix in the max-times algebra (see Definition 6.4), and show that if the distribution $P(\mathbf{X})$ is faithful to the polytree $\mathcal{D}$, and for all $\mathbf{s} \subseteq V \backslash \{i, j\}$, $X^{(i)}$ and $X^{(j)}$ are conditionally dependent given $\{X^{(r)} : r \in \mathbf{s}\}$, then $\mathrm{maxper}(\chi_{\{i,k\}\times\{j,k\}}) = \chi_{ij}$ for any $k \in V$. Note that the assumption that $P(\mathbf{X})$ is faithful to the polytree $\mathcal{D}$ has already been proven by Claudia Klüppelberg and Steffen Lauritzen in their forthcoming paper.

Chapter 2

Prelimilaries

In this chapter we provide some background knowledge about graph terminology, graphical models and recursive max-linear models.

2.1 Basic Graph Theoretical Definitions

A directed graph (briefly *digraph*) is an ordered pair $\mathcal{D} = (V, E)$, where V is a non-empty finite set and $E \subseteq V \times V$. The set V is called the *node set* and E *edge set* of $\mathcal{D}$. For an edge $e = (u, v) \in E$, we say the edge e is *leaving* u and *entering* v. Every edge of the form (u, u) is called a *loop*. It is possible that in a digraph several edges exist between two nodes (called a multigraph). In this thesis, however, we will restrict to digraphs in which there is at most one edge between any two nodes.

A *subgraph* of $\mathcal{D}$ is a graph $\mathcal{D}_1$ such that $V(\mathcal{D}_1) \in V(\mathcal{D})$ and $E(\mathcal{D}_1) \in E(\mathcal{D})$. An induced *subgraph* by a set of nodes $A \in V$ is the subgraph that is constituted by all and only nodes in A and existing edges between two nodes in A. An induced *subgraph* by a set of edges $B \in E$ is the subgraph that is constituted by all and only edges in B and existing nodes in B.

Let $\mathcal{D} = (V, E)$ be a digraph. A sequence $\pi = (v_1, v_2, \ldots, v_p)$ with $p > 1$ and $v_1, v_2, \ldots, v_p \in V$ is called a *path* if $(v_i, v_{i+1}) \in E$ for all $i = 1, 2, \ldots, p-1$. The nodes v_1 and v_p are called the *starting node* and the *endnode* of π, respectively. $p - 1$ is called the *length* of π and denoted by $l(\pi)$. If π has the starting node u and endnode v, we say π is a $u \to v$ path. If π is a $u \to v$ path and π' is a $v \to w$ path, then the concatenation of these two paths is a $u \to w$ path and denoted by $\pi \circ \pi'$. A path $(v_1, v_2, \ldots, v_p)$ is called a *cycle* if $v_1 = v_p$ and $p > 1$. Moreover, it is called an *elementary cycle* if no node, except v_1, appears more than once. If there is no cycle in the path π, then π is called *acyclic*.

Furthermore, if there is no cycle in the digraph $\mathcal{D}$, then $\mathcal{D}$ will be called a *directed acyclic graph* (DAG).

Let $i, j \in V$, if $(i, j) \in E$ then i is called a *parent* of j. If there is an $i \to j$ path in $\mathcal{D}$, then i is called an *ancestor* of j, and j is called a *descendant* of i. We use $\mathrm{pa}(i), \mathrm{an}(i), \mathrm{de}(i)$ to denote the set of all parents, ancestors, and descendants of i in $\mathcal{D}$, respectively.

2.2 Recursive Max-linear Model on a DAG

A recursive max-linear model $\mathbf{X}$ on a DAG $\mathcal{D} = (V, E)$ was first defined in Gissibl and Klüppelberg (2018) as

$$X_i := \bigvee_{k \in \mathrm{pa}(i)} c_{ki} X_k \vee c_{ii} Z_i \ , \quad i = 1, 2, \ldots, d \tag{2.1}$$

with independent random variables $Z_1, Z_2, \ldots, Z_d$ on the support $\overline{\mathbb{R}}_+ := [0, \infty)$, and positive weights c_{ki} for all $i \in V$ and $k \in \mathrm{pa}(i) \cup \{i\}$. It is a combination of graphical models (see Lauritzen (1996) and Koller and Friedman (2009)) and extreme value theory (see de Haan and Ferreira (2006) and Resnick (1987)). In this thesis we focus on the directed graphical models, which is called Bayesian network in Koller and Friedman (2009).

On DAG $\mathcal{D}$ each node i represents its corresponding random variable X_i, and all the random variables $X_1, \ldots, X_d$ have a joint probability distribution P. A probability distribution P is called *faithful* with respect to a graph $\mathcal{D}$ if conditional independencies of the distribution can be inferred from d-separation on the graph $\mathcal{D}$ and vice versa. Formally, let $\mathbf{X} \in \overline{\mathbb{R}}_+^d$ have a joint probability distribution P, then P is said faithful with respect to $\mathcal{D}$ if for every $i, j \in V, i \neq j$ and every set $s \subseteq V$

$$\begin{aligned} & X_i \text{ and } X_j \text{ are conditionally independent given } \{X_r : r \in s\} \\ \Leftrightarrow \quad & \text{node } i \text{ and node } j \text{ are d-separated by the set } s. \end{aligned}$$

The definition and more details about d-separation are described in Chapter 3.2 of Lauritzen (1996). If the distribution P is faithful with respect to graph $\mathcal{D}$, then the conditional independence structure of $\mathbf{X}$ is encoded in the graph $\mathcal{D}$. That is, we can use Markov property (see Chapter 3.2 in Lauritzen (1996)) to learn the conditional independence structure of $\mathbf{X}$. The structural and graphical properties of a recursive max-linear model $\mathbf{X}$ on a DAG $\mathcal{D}$ are intensively investigated in Gissibl and Klüppelberg (2018).

Motivated by modeling extreme risks, in Gissibl, Klüppelberg, and Moritz (2017) the noise variables $Z_1, \ldots, Z_d$ are assumed to be independently and identically distributed

(briefly i.i.d.) with support $\overline{\mathbb{R}}_+ := (0, \infty)$ and *regularly varying* with index $\alpha \in \overline{\mathbb{R}}_+$ (denoted by $Z \in RV(\alpha)$). The latter means that

$$\lim_{t\to\infty} \frac{1 - F_Z(xt)}{1 - F_Z(t)} = x^{-\alpha},$$

where F_Z is the probability distribution function of Z. More details about regular variation are referred to Resnick (1987). Under this assumption it can be shown that the independent copies $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \ldots, \mathbf{X}^{(n)}$ of $\mathbf{X}$ satisfies

$$a_n^{-1} \bigvee_{r=1}^{n} \mathbf{X}^{(r)} \xrightarrow[n\to\infty]{d} \mathbf{M} = (M_1, \ldots, M_d),$$

where (a_n) is a positive normalizing sequence, $\mathbf{M}$ is a non-degenerate random vector with distribution function G and d stands for weak convergence. This is called that $\mathbf{X}$ is in the maximum domain of attraction of G and denoted by $\mathbf{X} \in \mathrm{MDA}(G)$. The proof and explicit formula of G are referred to Appendix A.2 in Gissibl, Klüppelberg, and Moritz (2017).

Another important result in Gissibl, Klüppelberg, and Moritz (2017) is about the *tail dependence matrix* $\chi = (\chi(i,j))_{d\times d}$ with entry $\chi(i,j)$ defined by

$$\chi(i,j) = \chi(j,i) = \lim_{u\uparrow 1} \mathrm{P}(X_i > F_i^{\leftarrow}(u) | X_j > F_j^{\leftarrow}(u)), \tag{2.2}$$

where X_i, X_j are components of $\mathbf{X}$ and F_i, F_j are their respective distribution functions. $\chi(i,j)$ is called the *tail dependence coefficient* between X_i and X_j. In that paper it is investigated how far the matrix χ can be used to identify the dependence structure of $\mathbf{X}$ and its associated DAGs. For instance, the equivalence between (a) X_i and X_j are independent and (b) $\chi(i,j) = 0$ is shown in Theorem 2.3.

Chapter 3

Max-times Algebra

In this chapter we introduce the key concepts in max-times algebras, provide their properties, and use these concepts to investigate the asymptotic behavior of the n-th power of a random matrix. The concepts and their basic properties are borrowed from chapter 1 of the book Butkovič (2010) (see also Saraf-Poor (2018)), where the max-plus algebra has been investigated. Since by taking logarithm we can change a multiplication to addition, the max-times algebra and max-plus algebra are essentially the same.

Our contributions to the max-times algebra are Theorem 3.16 and the whole Section 3.4. In Theorem 3.16 a fast method for calculating the strong transitive closure, if it exists, is provided. In Section 3.4 we investigate the asymptotic behavior of the n-th power of a random matrix, which turns out to mainly depend on the maximum cycle mean of the matrix. These contributions play a significant role in the following chapters, where they are the most important results we will use for proving the theorems there.

3.1 Notations, Definitions, and Basic Properties

We use the following notations

$$\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{0\}$$
$$a \vee b = \max(a, b)$$
$$a \cdot b = ab$$

and consider the triplet $(\overline{\mathbb{R}}_+, \vee, \cdot)$ throughout this thesis. It can be easily shown that $(\overline{\mathbb{R}}_+, \vee, \cdot)$ is a commutative idempotent semiring with 0 being the neutral element for $\vee$ and 1 for $\cdot$. Similarly as in linear algebra, we can extend the pair of binary operations $(\vee, \cdot)$ from $\overline{\mathbb{R}}_+$ to matrices in $\overline{\mathbb{R}}_+^{d\times d}$ and vectors in $\overline{\mathbb{R}}_+^{d}$. That is, for matrices $A = (a_{ij})$,

$B = (b_{ij})$ and $C = (c_{ij})$ with compatible sizes and non-negative entries, we have

$$\begin{aligned} C = A \vee B \text{ if } & c_{ij} = a_{ij} \vee b_{ij} \text{ for all } i, j \\ C = A \odot B \text{ if } & c_{ij} = \bigvee_k a_{ik} \cdot b_{kj} \text{ for all } i, j \end{aligned} \tag{3.1}$$

It can be checked that $(\overline{\mathbb{R}}_+^{d\times d}, \vee, \odot)$ is also a semiring, and the neutral elements for $\vee$ and $\odot$ are the matrix 0 and I with appropriate dimensions, where I is the unit matrix with diagonal entries 1 and non-diagonal entries 0.

If A is a square matrix in $\overline{\mathbb{R}}_+^{d\times d}$, then the product $A \odot A \odot A \odot \cdots \odot A$, where A appears k times, will be denoted as A^k. A^0 is simply equal to I for any square matrix A. For convenience we also denote the (i, j) entry in A^k as $a_{ij}^{(k)}$ and should not confuse it with a_{ij}^k, which is the k-th power of a_{ij} with respect to $\cdot$.

Note that although the operation $\vee$ is not invertible, it is idempotent, which provides us the following property.

Lemma 3.1. *[Lemma 1.1.4 in Butkovič (2010)] For every $A \in \overline{\mathbb{R}}_+^{d\times d}$ and non-negative integer $k \in \mathbb{N}$,*

$$(I \vee A)^k = I \vee A \vee A^2 \vee \cdots \vee A^k.$$

Proof. We prove the statement by induction. Clearly the equation $(I \vee A)^k = I \vee A \vee A^2 \vee \cdots \vee A^k$ holds for $k = 0$ and $k = 1$. Now we assume this equation holds for $k = n$. Then for $k = n + 1$, by using the fact $(\overline{\mathbb{R}}_+^{d\times d}, \vee, \odot)$ is a semiring we can get

$$\begin{aligned} (I \vee A)^{n+1} &= (I \vee A)^n \odot (I \vee A) \\ &= (I \vee A \vee A^2 \vee \cdots \vee A^n) \odot (I \vee A) \\ &= (I \vee A \vee A^2 \vee \cdots \vee A^n) \odot I \vee (I \vee A \vee A^2 \vee \cdots \vee A^n) \odot A \\ &= (I \vee A \vee A^2 \vee \cdots \vee A^n) \vee (A \vee A^2 \vee A^3 \vee \cdots \vee A^{n+1}) \\ &= I \vee A \vee A^2 \vee \cdots \vee A^n \vee A^{n+1}. \end{aligned}$$

Hence the statement holds. □

3.2 Weighted Digraphs and Matrices

A *weighted digraph* is $\mathcal{D} = (V, E, w)$, where (V, E) is a digraph and w is a positive real function on the edge set E. If $\pi = (v_1, v_2, \ldots, v_p)$ is a path in the weighted digraph $\mathcal{D}$, then the *weight* of π is the product of the weights of all the edges constituting the path, i.e. $w(\pi) = w(v_1, v_2)w(v_2, v_3) \cdots w(v_{p-1}, v_p)$. If the path π is a cycle, then $w(\pi)$ represents

the weight of the cycle. As w stands for the "weight" rather than the "length" of a path/cycle, the word "heaviest path/cycle" will be used throughout the thesis, instead of "longest path/cycle".

Lemma 3.2. *[Lemma 1.5.4 in Butkovič (2010)] If $\mathcal{D} = (V, E, w)$ is a weighted digraph and the weights of all the cycles in $\mathcal{D}$ are less than or equal to 1, then for every $u, v \in V$, a heaviest $u \to v$ path exists if at least one $u \to v$ path exists. In that case at least one heaviest $u \to v$ path has length $|V|$ or less.*

Proof. Assume π is a $u \to v$ path in $\mathcal{D}$ with $l(\pi) > |V|$, then π must contain a cycle as a subpath. Since the weights of all the cycles in $\mathcal{D}$ are less than or equal to 1, by successively deleting the cycles we get a $u \to v$ path π' with $l(\pi') \leq |V|$ and $w(\pi') \geq w(\pi)$. A heaviest $u \to v$ path of length $|V|$ or less exists since the set of such paths is finite. Hence the statement follows. □

For a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$, $\mathcal{D}_A$ will denote the weighted digraph (V, E, w), where $V = \{1, 2, \ldots, d\}$, $E = \{(i, j) : a_{ij} > 0\}$ and $w(i, j) = a_{ij}$ for all $(i, j) \in E$. If $\pi = (v_1, v_2, \ldots, v_p)$ is a path in $\mathcal{D}_A$, then $w(\pi, A) = w(\pi)$ will denote the weight of the path π and by definition we know $w(\pi) = a_{v_1v_2}a_{v_2v_3} \cdots a_{v_{p-1}v_p}$.

Proposition 3.3. *For a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$ and $k \in \mathbb{N}$, $a_{ij}^{(k)}$ is the maximum weight of all $i \to j$ paths of length k.*

Proof. Recall that $A^k = (a_{ij}^{(k)})$. By the definition of matrix multiplication in max-times algebra we have

$$a_{ij}^{(k)} = \bigvee_{v_1=1}^{d} \bigvee_{v_2=1}^{d} \cdots \bigvee_{v_{k-1}=1}^{d} a_{iv_1}a_{v_1v_2}a_{v_2v_3} \cdots a_{v_{k-2}v_{k-1}}a_{v_{k-1}j}.$$

Note that if the $i \to j$ path $(i, v_1, v_2, \ldots, v_{k-1}, j)$ exists, then $a_{iv_1}a_{v_1v_2}a_{v_2v_3} \cdots a_{v_{k-2}v_{k-1}}a_{v_{k-1}j}$ stands for the weight of this path. Otherwise $a_{iv_1}a_{v_1v_2}a_{v_2v_3} \cdots a_{v_{k-2}v_{k-1}}a_{v_{k-1}j}$ simply equals to 0. Hence $a_{ij}^{(k)} = \bigvee_{\pi\in P_{ij}^k} w(\pi)$, where P_{ij}^k denotes the set of all $i \to j$ paths of length k. Therefore the statement follows. □

3.3 Weak and Strong Transitive Closure

In this section we introduce the essential concepts, like maximum cycle mean and transitive closures, and provide some of their properties. All these concepts and properties will play a significant role in understanding the generalized recursive max-linear model in Chapter 4.

Definition 3.4. For a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$, the *maximum cycle mean* of A is defined as

$$\lambda(A) := \max_{\sigma} w(\sigma)^{\frac{1}{l(\sigma)}} , \tag{3.2}$$

where σ is an elementary cycle in $\mathcal{D}_A$.

Remark 3.5. Since the number of elementary cycles in a digraph with finite nodes is finite, we know the maximum cycle mean $\lambda(A)$ always exists. From this definition we also know that the corresponding digraph $\mathcal{D}_A$ of matrix A is acyclic, which means $\mathcal{D}_A$ is a DAG, if and only if $\lambda(A) = 0$.

Lemma 3.6. *[Lemma 1.6.2 in Butkovič (2010)] If we take the maximum in Definition 3.4 over all cycles in $\mathcal{D}_A$, then $\lambda(A)$ remains unchanged.*

Proof. Let σ be a cycle in $\mathcal{D}_A$. Then σ can be divided into elementary cycles $\sigma_1, \sigma_2, \ldots, \sigma_r$ with $l(\sigma_1) + l(\sigma_2) + \cdots + l(\sigma_r) = l(\sigma)$. By Definition 3.4 we know $\lambda(A) \geq w(\sigma_i)^{\frac{1}{l(\sigma_i}}$ for $i = 1, \ldots, r$. Hence we have

$$\begin{aligned} w(\sigma)^{\frac{1}{l(\sigma)}} &= \big(w(\sigma_1)w(\sigma_2)\cdots w(\sigma_r)\big)^{\frac{1}{l(\sigma)}} \\ &\leq \big(\lambda(A)^{l(\sigma_1)}\lambda(A)^{l(\sigma_2)}\cdots\lambda(A)^{l(\sigma_r)}\big)^{\frac{1}{l(\sigma)}} \\ &= \lambda(A)^{\frac{l(\sigma_1)+l(\sigma_2)+\cdots+l(\sigma_r)}{l(\sigma)}} \\ &= \lambda(A) \end{aligned}$$

Therefore the statement holds. □

From Proposition 3.3 that we know $a_{ij}^{(k)}$ represents the maximum weight of all $i \to j$ paths of length k in $\mathcal{D}_A$. This implies the following method for computing $\lambda(A)$.

Proposition 3.7. *For a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$,*

$$\lambda(A) = \bigvee_{k,i=1}^{d} (a_{ii}^{(k)})^{\frac{1}{k}}.$$

Proof. By Proposition 3.3 we know that $a_{ii}^{(k)}$ represents the maximum weight of all cycles of length k in $\mathcal{D}_A$ which pass through i. Hence by Lemma 3.6 we have $\bigvee_{k,i=1}^{d} (a_{ii}^{(k)})^{\frac{1}{k}} \leq \lambda(A)$. It is left to show the other direction. By Definition 3.4 we know in order to calculate $\lambda(A)$ it is enough to limit ourselves to elementary cycles. Assume σ is an elementary cycle of length k in $\mathcal{D}_A$. Since $a_{ii}^{(k)}$ represents the maximum weight of all cycles of length k in $\mathcal{D}_A$ which pass through i, we have $\bigvee_{i=1}^{d} (a_{ii}^{(k)})^{\frac{1}{k}} \geq w(\sigma)^{\frac{1}{k}}$. Moreover, as the length of an

elementary cycle in $\mathcal{D}_A$ cannot exceed d, we can get $\bigvee_{k,i=1}^{d} (a_{ii}^{(k)})^{\frac{1}{k}} \geq \lambda(A)$.

All in all, $\lambda(A) = \bigvee_{k,i=1}^{d} (a_{ii}^{(k)})^{\frac{1}{k}}$. □

Example 3.8. For a matrix

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

we have

$$A^2 = \begin{pmatrix} 0.5 & 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}, \qquad A^3 = \begin{pmatrix} 1 & 0.5 & 0.5 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0.25 & 0 & 1 & 0.5 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0.5 & 0 & 2 & 1 & 0 \end{pmatrix},$$

$$A^4 = \begin{pmatrix} 1 & 1 & 1 & 0.5 & 0 \\ 0.5 & 1 & 2 & 1 & 0 \\ 0.5 & 0.25 & 0.25 & 0 & 0.5 \\ 0.5 & 0 & 2 & 1 & 0 \\ 1 & 0.5 & 0.5 & 0 & 1 \end{pmatrix}, \qquad A^5 = \begin{pmatrix} 0.5 & 1 & 2 & 1 & 0.5 \\ 1 & 0.5 & 2 & 1 & 1 \\ 0.5 & 0.5 & 0.5 & 0.25 & 0 \\ 1 & 0.5 & 0.5 & 0 & 1 \\ 1 & 1 & 1 & 0.5 & 0 \end{pmatrix}.$$

Hence $\lambda(A) = \bigvee_{k,i=1}^{5} (a_{ii}^{(k)})^{\frac{1}{k}} = 1$.

Definition 3.9. For a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$, we define $S_n(A) := A \vee A^2 \vee A^3 \vee \cdots \vee A^n$. If $S_n(A)$ converges to a matrix without $+\infty$ as $n \to \infty$, we define

$$\Gamma(A) := \lim_{n\to\infty} S_n(A), \qquad \Delta(A) := I \vee \Gamma(A) \quad ,$$

$\Gamma(A)$ and $\Delta(A)$ are called the *weak* and *strong transitive closure* of A respectively.

Remark 3.10. Note that from Proposition 3.3 we know that $a_{ij}^{(k)}$ stands for the maximum weight of all $i \to j$ paths of length k in $\mathcal{D}_A$. Hence if $\Gamma(A) = (\gamma_{ij})$ exists as a matrix without $+\infty$, then γ_{ij} is the maximum weight of all $i \to j$ paths, whatever the path length is.

Proposition 3.11. *[Proposition 1.6.10 in Butkovič (2010)] For a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$, $\Gamma(A) = \lim_{n\to\infty} S_n(A)$ exists as a matrix without $+\infty$ if and only if $\lambda(A) \leq 1$. Furthermore, if $\lambda(A) \leq 1$, then $\Gamma(A) = S_k(A)$ for every integer $k \geq d$.*

Proof. If $\lambda(A) \leq 1$, by Lemma 3.2 we know that for every $i, j \in \{1, 2, \ldots, d\}$, in order to find the heaviest $i \to j$ path, we can limit ourselves to the $i \to j$ paths of length less than or equal to d. By Proposition 3.3 we also know that $a_{ij}^{(k)}$ represents the weight of the heaviest $i \to j$ path of length k. Hence we have $a_{ij}^{(n)} \leq a_{ij} \vee a_{ij}^{(2)} \vee \cdots \vee a_{ij}^{(d)}$ for every $n \in \mathbb{N}$, which implies $A^n \leq A \vee A^2 \vee \cdots \vee A^d$ for every $n \in \mathbb{N}$. Therefore $\Gamma(A)$ exists as a matrix without $+\infty$ and $\Gamma(A) = A \vee A^2 \vee \cdots \vee A^d = A \vee A^2 \vee \cdots \vee A^d \vee \cdots \vee A^k = S_k(A)$ for every $k \geq d$.
On the other hand, if $\Gamma(A)$ exists as a matrix without $+\infty$, then by Remark 3.10 the weight of every cycle in $\mathcal{D}_A$ must be less than or equal to 1, otherwise we can go over the cycle again and again, and the weight of the resulting path will tend to $+\infty$. Hence by definition, $\lambda(A) \leq 1$. □

Corollary 3.12. *For a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$, $\Delta(A) = I \vee \Gamma(A)$ exists as a matrix without $+\infty$ if and only if $\lambda(A) \leq 1$. Furthermore, if $\lambda(A) \leq 1$, then $\Delta(A) = I \vee S_k(A)$ for every integer $k \geq d$.*

Proof. By Proposition 3.11 we know that $\Gamma(A) = \lim_{n\to\infty} S_n(A)$ exists as a matrix without $+\infty$ if and only if $\lambda(A) \leq 1$ and if $\lambda(A) \leq 1$, then $\Gamma(A) = S_k(A)$ for every $k \geq d$. Since $\Delta(A) = I \vee \Gamma(A)$, we know $\Delta(A)$ exists as a matrix without $+\infty$ if and only if $\Gamma(A)$ exists as a matrix without $+\infty$. Hence the statement follows. □

If $\lambda(A) \leq 1$, the result $\Delta(A) = I \vee A \vee A^2 \vee \cdots \vee A^d$ can be further simplified by using the property of an increasing matrix.

Definition 3.13. A matrix $A = (a_{ij}) \in \overline{\mathbb{R}}_+^{d\times d}$ is called increasing if $a_{ii} \geq 1$ for every $i \in \{1, 2, \ldots, d\}$.

Lemma 3.14. *[Lemma 1.6.11 in Butkovič (2010)] If $A \in \overline{\mathbb{R}}_+^{d\times d}$ is increasing, then $x \leq A \odot x$ for every $x \in \overline{\mathbb{R}}_+^d$. Therefore,*

$$A \leq A^2 \leq A^3 \leq \cdots .$$

Proof. Note that $I \leq A$. Hence $x = I \odot x \leq A \odot x$. The second statement follows by taking every column of A as x and repeating the argument. □

Proposition 3.15. *[Proposition 1.6.12 in Butkovič (2010)] If $A \in \overline{\mathbb{R}}_+^{d\times d}$ is increasing and $\lambda(A) = 1$, then*

$$\Gamma(A) = A^{d-1} = A^d = A^k, \text{ for every integer } k \geq d.$$

Proof. By Proposition 3.11, $\Gamma(A) = A \vee A^2 \vee \cdots \vee A^k$ for every $k \geq d$, which is $A^k, k \geq d$ by Lemma 3.14. Hence it is left to show $A^{d-1} = A^d$. Now we consider the entries in these two

matrices, i.e., $a_{ij}^{(d-1)}$ and $a_{ij}^{(d)}$ for every $i, j \in \{1, 2, \ldots, d\}$. Assume that $a_{ij}^{(d-1)} < a_{ij}^{(d)}$, then by Proposition 3.3 the maximum weight of all $i \to j$ paths of length d, say $w(\pi)$, is greater than the maximum weight of all $i \to j$ paths of length $d-1$. Since $a_{ii}^{(d)} = a_{ii}^{(d-1)} = 1$, we know that $i \neq j$. From $l(\pi) = d$ we know π must contain a cycle σ. By deleting σ we get another $i \to j$ path π' with $l(\pi') < d$ and $w(\pi) = w(\sigma)w(\pi')$. Since $\lambda(A) \leq 1$, we know $w(\sigma) \leq 1$. Hence $w(\pi, A) \leq w(\pi', A)$ and $a_{ij}^{(d)} \leq a_{ij}^{(l(\pi'))} \leq a_{ij}^{(d-1)}$, which is a contradiction to the assumption $a_{ij}^{(d-1)} < a_{ij}^{(d)}$. Therefore $a_{ij}^{(d-1)} = a_{ij}^{(d)}$ for every $i, j \in \{1, 2, \ldots, d\}$. □

Theorem 3.16. *For a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$ with $\lambda(A) \leq 1$, if we replace all diagonal entries in A by 1 and denote the resulting matrix as A', then*

$$\Delta(A) = (A')^{d-1} \tag{3.3}$$

Proof. By Definition 3.13 we know A' is increasing. Since A and A' only differ in diagonal entries, we know the elementary cycles in $\mathcal{D}_A$ and $\mathcal{D}_{A'}$ only differ in the loops, i.e., cycles of length 1. Hence if σ is an elementary cycle of length $k, k > 2$ in $\mathcal{D}_{A'}$, then by Definition 3.4 and $\lambda(A) \leq 1$ we know $w(\sigma) \leq 1$. For the elementary cycles of length 1 in $\mathcal{D}_{A'}$, we know all their weights are 1, since all the diagonal entries in A' are 1. Hence by Definition 3.4, $\lambda(A') = 1$. By Proposition 3.15 we know $\Gamma(A') = (A')^{d-1}$. Hence all we need to show is $\Gamma(A') = \Delta(A) = I \vee \Gamma(A)$.
We denote $\Gamma(A') = (\gamma'_{ij})_{d\times d}$ and $\Gamma(A) = (\gamma_{ij})_{d\times d}$. Then by Remark 3.10 we know γ'_{ij} and γ_{ij} represent the maximum weight of all $i \to j$ paths in the digraphs $\mathcal{D}_{A'}$ and $\mathcal{D}_A$, respectively. Since $\lambda(A) \leq \lambda(A') = 1$, we know the weights of all loops, i.e., cycles of length 1, in both $\mathcal{D}_A$ and $\mathcal{D}_{A'}$ are less than or equal to 1. Hence for $i \neq j,\ i, j \in \{1, 2, \ldots, d\}$, in order to find the heaviest $i \to j$ path, we can limit ourselves to the $i \to j$ paths without loops. Since A' and A only differ in diagonal entries, we know $\mathcal{D}_A$ and $\mathcal{D}_{A'}$ only differ in the loops. Hence $\gamma'_{ij} = \gamma_{ij}$ for $i \neq j,\ i, j \in \{1, 2, \ldots, d\}$. For the diagonal entries, by Proposition 3.7 we know that $\lambda(A) \leq 1$ implies $\bigvee_{k,i=1}^{d} a_{ii}^{(k)} \leq 1$. Hence by Proposition 3.11, $\gamma_{ii} = \bigvee_{k=1}^{d} (a_{ii}^{(k)} \leq 1$ for every $i \in \{1, 2, \ldots, d\}$. Similarly by Proposition 3.7 and Proposition 3.11, we know that $\gamma'_{ii} \leq 1$ for every $i \in \{1, 2, \ldots, d\}$. On the other hand, since all diagonal entries in A' are 1, we have $\gamma'_{ii} = \bigvee_{k=1}^{d} a_{ii}^{\prime(k)} \geq a'_{ii} = 1$ for every $i \in \{1, 2, \ldots, d\}$. Hence $\gamma'_{ii} = 1$, and $\gamma'_{ii} = \gamma_{ii} \vee 1$ for every $i \in \{1, 2, \ldots, d\}$.
Therefore, $\Gamma(A') = I \vee \Gamma(A) = \Delta(A)$. □

Remark 3.17. From the proof in Theorem 3.16 we know that for a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$ with $\lambda(A) \leq 1$, if A' is defined as above, then $\lambda(A') = 1, \Gamma(A') = I \vee \Gamma(A) = \Delta(A)$, $\gamma'_{ii} = 1$ for every $i \in \{1, 2, \ldots, d\}$, and $\gamma'_{ij} = \gamma_{ij}$ for $i \neq j, i, j \in \{1, 2, \ldots, d\}$. By Proposition 3.15 we also have $\Gamma(A') = (A')^{d-1}$. Hence all the diagonal entries in matrices $\Delta(A)$ and $(A')^{d-1}$

must be 1. Moreover, if we denote $(A')^{d-1} = (a_{ij}'^{(d-1)})$, then $a_{ij}'^{(d-1)} = \gamma_{ij}$ for $i \neq j, i, j \in \{1, 2, \ldots, d\}$.

3.4 Asymptotic Behavior of A^n

In this section we investigate the asymptotic behavior of A^n for $A \in \overline{\mathbb{R}}_+^{d\times d}$ as $n \to \infty$. We show that it is determined by the maximum cycle mean $\lambda(A)$ as defined in Definition 3.4. We start with the simplest case $\lambda(A) > 1$, then discuss $\lambda(A) < 1$, and finally investigate the most complicated case $\lambda(A) = 1$.

Proposition 3.18. *For every matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$ with $\lambda(A) > 1$, $\lim\limits_{n\to\infty} A^n$ does not exist finitely, i.e. there are entries in $\lim\limits_{n\to\infty} A^n$ which are $+\infty$.*

Proof. Since $\lambda(A) > 1$, by Proposition 3.7 there must exist some $k, i \in \{1, 2, \ldots, d\}$ such that $a_{ii}^{(k)} > 1$. Recall from Remark 3.10 that $a_{ii}^{(k)}$ represents the maximum weight of the cycle $i \to i$ of length k, going n times over this cycle gives $a_{ii}^{(nk)} \geq (a_{ii}^{(k)})^n$. Hence $a_{ii}^{(nk)}$ converges to $+\infty$ as $n \to \infty$. □

From Proposition 3.11 we know that for a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$, if $\lambda(A) \leq 1$, then $\Gamma(A)$ exists and $A^n \leq \Gamma(A)$ for every positive integer n. In the following we denote $\Gamma(A) = (\gamma_{ij})$ and give a much more precise upper bound for every entry in A^n.

Lemma 3.19. *For every matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$ with $\lambda(A) \leq 1$, if $n \leq d$, then $a_{ij}^{(n)} \leq \gamma_{ij}$ for every $i, j \in \{1, 2, \ldots, d\}$. If $n > d$, then $a_{ij}^{(n)} \leq \gamma_{ij}(\lambda(A))^{n-d}$ for every $i, j \in \{1, 2, \ldots, d\}$.*

Proof. The first statement follows directly from Definition 3.9. For the case $n > d$, if $a_{ij}^{(n)} = 0$, then $a_{ij}^{(n)} \leq \gamma_{ij}(\lambda(A))^{n-d}$ clearly holds. If $a_{ij}^{(n)} > 0$, by Proposition 3.3 there must exist at least one $i \to j$ path of length n. Without loss of generality we assume π is such a path. Since $l(\pi) = n > d$, π must contain cycles and can be divided into cycles $\sigma_1, \sigma_2, \ldots, \sigma_k$ and a $i \to j$ subpath π' such that π' does not go through any node twice except i. Such a division gives us $l(\sigma_1) + l(\sigma_2) + \cdots + l(\sigma_k) + l(\pi') = n$, $l(\pi') \leq d$, $l(\sigma_1)+l(\sigma_2)+\cdots+l(\sigma_k) \geq n-d$ and $w(\pi) = w(\sigma_1)w(\sigma_2)\cdots w(\sigma_k)w(\pi')$. By Definition 3.4 we know $w(\sigma_1)w(\sigma_2)\cdots w(\sigma_k) \leq \lambda(A)^{l(\sigma_1)+l(\sigma_2)+\cdots+l(\sigma_k)} \leq (\lambda(A))^{n-d}$. Furthermore, by Remark 3.10 we know $w(\pi') \leq \gamma_{ij}$. Hence $w(\pi) \leq w(\pi')(\lambda(A))^{n-d} \leq \gamma_{ij}(\lambda(A))^{n-d}$. Since π can be any arbitrary $i \to j$ path of length n, we get $a_{ij}^{(n)} = \max\limits_{\pi\in P_{ij}^n} w(\pi) \leq \gamma_{ij} \odot (\lambda(A))^{n-d}$, where P_{ij}^n denotes the set of all $i \to j$ paths of length n. □

Example 3.20. For the matrix A defined in Example 3.8, we know $\lambda(A) = 1$. Hence by Proposition 3.11 we know $\Gamma(A)$ exists and

$$\Gamma(A) = A \vee A^2 \vee A^3 \vee A^4 \vee A^5 = \begin{pmatrix} 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \end{pmatrix}.$$

With the help of software R or Matlab we can easily calculate that

$$A^{1000} = \begin{pmatrix} 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \end{pmatrix}.$$

Hence it is seen that $a_{ij}^{1000} \leq \gamma_{ij}(\lambda(A))^{1000-5} = \gamma_{ij}$ for every $i, j \in \{1, 2, 3, 4, 5\}$.

Proposition 3.21. *For every matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$ with $\lambda(A) < 1$, $\lim_{n\to\infty} A^n = 0$.*

Proof. By Lemma 3.19, $a_{ij}^{(n)} \leq \gamma_{ij} \odot (\lambda(A))^{n-d}$ for every $i, j \in \{1, 2, \ldots, d\}$ and $n > d$. Hence $\lim_{n\to\infty} a_{ij}^{(n)} = 0$ if $\lambda(A) < 1$. □

Definition 3.22. For a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$, if a cycle $\sigma = (v_1, v_2, \ldots, v_p = v_1)$ in $\mathcal{D}_A$ satisfies $w(\sigma)^{\frac{1}{l(\sigma)}} = \lambda(A)$ and $v_i \neq v_j$ for $i, j = 1, 2, \ldots, p-1, i \neq j$, then the cycle σ is called a critical cycle in $\mathcal{D}_A$ and all the nodes in σ are called critical nodes.

Note that given a matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$, the critical cycle in $\mathcal{D}_A$ may not be unique, i.e., there can be several critical cycles in $\mathcal{D}_A$. But the number of critical cycles in $\mathcal{D}_A$ is finite due to the fact that the number of nodes in $\mathcal{D}_A$ is finite. In the following we discuss the asymptotic behavior of A^n when $\lambda(A) = 1$.

Proposition 3.23. *Let $A \in \overline{\mathbb{R}}_+^{d\times d}$ be a matrix with $\lambda(A) = 1$. For $i, j \in \{1, 2, \ldots, d\}$, if there is no $i \to j$ path in $\mathcal{D}_A$ containing a critical cycle, then $\lim_{n\to\infty} a_{ij}^{(n)} = 0$. If there exists one $i \to j$ path π in $\mathcal{D}_A$ containing a critical cycle σ and $w(\pi) = \gamma_{ij}$, then $a_{ij}^{(l(\pi)+kl(\sigma))} = w(\pi) = \gamma_{ij}$ for every integer k such that $l(\pi) + kl(\sigma) > 0$.*

Proof. If there is no $i \to j$ path in $\mathcal{D}_A$ containing a critical cycle, then $\lim_{n\to\infty} a_{ij}^{(n)} = 0$ follows from the proof of Lemma 3.19.
If there exists one $i \to j$ path π in $\mathcal{D}_A$ containing a critical cycle σ and $w(\pi) = \gamma_{ij}$, then

by Remark 3.10 we know $w(\pi) = \gamma_{ij} = a_{ij}^{(l(\pi))}$. Furthermore, since σ is a critical cycle and $\lambda(A) = 1$, we know $w(\sigma) = 1$. Hence going over σ for whatever many times or deleting σ does not change the weight of the resulting path π'. Therefore $a_{ij}^{(l(\pi)+kl(\sigma))} \geq w(\pi') = w(\pi)$. By Remark 3.10 we know $a_{ij}^{(n)} \leq \gamma_{ij}$ for every $i, j \in \{1, 2, \ldots, d\}$ and every positive integer n. Hence $a_{ij}^{(l(\pi)+kl(\sigma))} = w(\pi) = \gamma_{ij}$. □

However, if $w(\pi) \neq \gamma_{ij}$, then the equation $a_{ij}^{(l(\pi)+kl(\sigma))} = w(\pi)$ does not necessarily hold as we can see from the following example.

Example 3.24. We use the matrix A defined in Example 3.8 again. From Example 3.8 we know $\lambda(A) = 1$. The corresponding digraph of matrix A is as follows.

$\mathcal{D}_A$

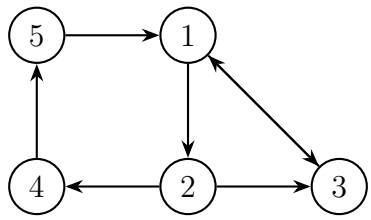


Figure 3.1: The corresponding digraph $\mathcal{D}_A$ of matrix A.

It can be seen that $\mathcal{D}_A$ contains two critical cycles, $\sigma_1 = (1, 2, 3, 1)$ and $\sigma_2 = (1, 2, 4, 5, 1)$. It can also be seen that $\pi = (1, 2, 3, 1, 2, 3, 1, 3)$ is a $1 \to 3$ path of length 7 which goes through the critical cycle σ_1. By using software R we can calculate A^7 and get $a_{13}^{(7)} = w(\pi) = 1 < \gamma_{13}$, since from Example 3.20 we know $\gamma_{13} = 2$. However, by using R we can also get $a_{13}^{(7+l(\sigma_1))} = a_{13}^{(10)} = 2 = \gamma_{13} > w(\pi)$ This is because the $i \to j$ path $\pi' = (1, 2, 4, 5, 1, 2, 4, 5, 1, 2, 3)$ satisfies $l(\pi') = 10$ and $w(\pi') = 2 = \gamma_{13}$.

Therefore the condition $w(\pi) = \gamma_{ij}$ in Proposition 3.23 can not be dropped. Otherwise the conclusion $a_{ij}^{(l(\pi)+kl(\sigma))} = w(\pi)$ does not necessarily hold.

In the following we investigate the special case, where all the cycles in $\mathcal{D}_A$ are either critical cycles or combinations of the critical cycles, and for every $i, j \in \{1, 2, \ldots, d\}$ there exists at most one $i \to j$ path without cycles. The notation $a \equiv b (\text{mod } m)$ means a and b are congruent modulo m. That is, their difference $a - b$ is an integer multiple of m.

Corollary 3.25. *let $A \in \overline{\mathbb{R}}_+^{d\times d}$ be a matrix with $\lambda(A) = 1$. If all cycles in $\mathcal{D}_A$ are either critical cycles or combinations of critical cycles, and for every $i, j \in \{1, 2, \ldots, d\}$ there exists at most one $i \to j$ path without cycles, then*

$$A^{n_1} = A^{n_2} \text{ for } n_1 \equiv n_2 (mod\ m), \qquad n_1, n_2 \geq d,$$

where m is the least common multiple of the lengths of all critical cycles in $\mathcal{D}_A$.

Proof. We first show that for $i, j \in \{1, 2, \ldots, d\}$, every $i \to j$ path π satisfies $w(\pi) = \gamma_{ij}$. By Remark 3.10 we know that γ_{ij} stands for the maximum weight of all the $i \to j$ paths. Since for every $i, j \in \{1, 2, \ldots, d\}$ there exists at most one $i \to j$ path without cycles and all the cycles in $\mathcal{D}_A$ are either critical cycles or combinations of the critical cycles, we know all the $i \to j$ paths actually have the same weight. Hence for every $i \to j$ path π, we have $w(\pi) = \gamma_{ij}$.
For $i, j \in \{1, 2, \ldots, d\}$, if there is no $i \to j$ path containing a cycle, then clearly $a_{ij}^{(n)} = 0$ for every integer $n \geq d$. For $i, j \in \{1, 2, \ldots, d\}$, if there exists one $i \to j$ path π containing a cycle σ, then $w(\pi) = \gamma_{ij} > 0$ and σ is either a critical cycle or a combination of critical cycles. If σ is a combination of k critical cycles, deleting $k - 1$ critical cycles will give us a $i \to j$ path π' which contains only one critical cycle and $w(\pi') = w(\pi) = \gamma_{ij}$. Since m is the least common multiple of the lengths of all the critical cycles, by Proposition 3.23 we get $a_{ij}^{(n_1)} = a_{ij}^{(n_2)} = \gamma_{ij}$ for $n_1 \equiv n_2 (\text{mod } m),\ n_1, n_2 \geq d$.
All in all, $A^{n_1} = A^{n_2}$ for $n_1 \equiv n_2 (\text{mod } m),\ n_1, n_2 \geq d$. □

$$B = \begin{pmatrix} 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

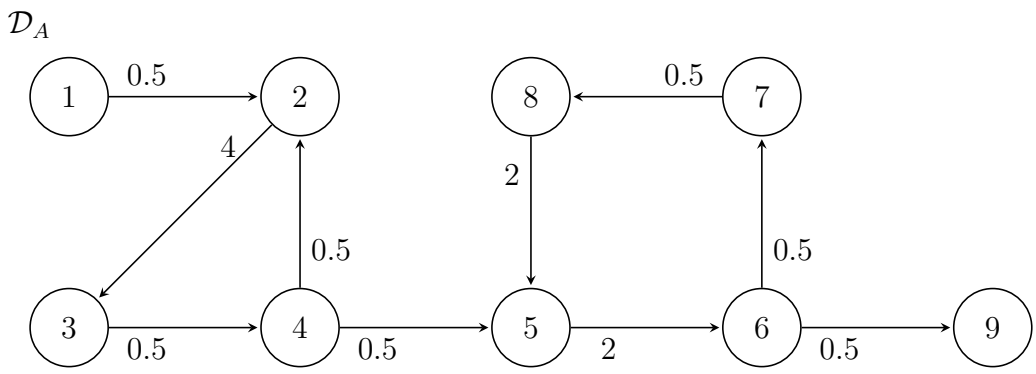


Figure 3.2: Matrix B and its corresponding digraph $\mathcal{D}_B$.

Example 3.26. The matrix B is defined as in Figure 3.2 and $\mathcal{D}_B$ is its corresponding digraph. It is easily seen that $\mathcal{D}_B$ only contains two elementary cycles $\sigma_1 = (2, 3, 4, 2)$ and $\sigma_2 = (5, 6, 7, 8, 5)$ with $w(\sigma_1) = w(\sigma_2) = 1$, $l(\sigma_1) = 3$, $l(\sigma_2) = 4$. By Definition 3.4 we

know $\lambda(B) = w(\sigma_1) = w(\sigma_2) = 1$, which implies that both σ_1 and σ_2 are critical cycles. Hence all the assumptions in Corollary 3.25 are satisfied. By using software R or Matlab we can validate that $A^{n_1} = A^{n_2}$ for $n_1 \equiv n_2 (\text{mod } 12),\ n_1, n_2 > 9$, where 12 is the least common multiple of $l(\sigma_1)$ and $l(\sigma_2)$.

Chapter 4

Generalized Recursive Max-linear Models on Digraphs

Recall from Equation (1.1) that a recursive max-linear model on a directed acyclic graph (DAG) $\mathcal{D} = (V, E)$ is defined by

$$X_i := \bigvee_{k \in \mathrm{pa}(i)} c_{ki} X_k \vee c_{ii} Z_i \ , \quad i = 1, 2, \ldots, d$$

where $Z_1, Z_2, \ldots, Z_d$ are independent non-negative random variables with support $\overline{\mathbb{R}}_+ = [0, \infty)$, and c_{ki} are non-negative weights for all $i \in V$ and $k \in \mathrm{pa}(i) \cup \{i\}$. Denote $c_{ki} = 0$ for $k \notin \mathrm{pa}(i) \cup \{i\}$, then by using the max-times algebra $(\overline{\mathbb{R}}_+, \vee, \odot)$ the model $\mathbf{X}$ can be written as

$$\mathbf{X} = (C - \tilde{C})^T \odot \mathbf{X} \vee \tilde{C} \odot \mathbf{Z} \ ,$$

where $\tilde{C} = \mathrm{diag}\{c_{11}, c_{22}, \ldots, c_{dd}\}$ and $(C - \tilde{C})^T$ means the transpose of $(C - \tilde{C})$.

Observe that in the above model $C \in \overline{\mathbb{R}}_+^{d \times d}$ is upper triangular if the nodes in the DAG $\mathcal{D}$ are well-ordered. In this chapter we discuss the general case, i.e. the *generalized recursive max-linear model* on a digraph $\mathcal{D}$. It is defined the same as in Equation (1.3), but the weighted adjacency matrix C is not necessarily upper triangular, even if the nodes in $\mathcal{D}$ are well-ordered. That is, the digraph $\mathcal{D}$ corresponding to the matrix $(C - \tilde{C})$ can contain cycles. We call the model generalized recursive because it is recursive in the sense that X_i is still defined as in Equation (1.1) and is determined by its parental variables, but differs from classical recursive model in the way that the digraph $\mathcal{D}$ can be cyclic. In fact, the generalized recursive max-linear model should be more suitable for the real-world data than the classical recursive max-linear model, because many real-world influences among factors are basically cyclic. An interesting question is to find the fixed point of Equation (1.3).

Also note that for Equation (1.1), substituting the representations of $X_k, k \in \mathrm{pa}(i)$ into its right-hand side gives a new representation of X_i, in terms of its further ancestral variables and noise variables. Here we say the node i_1 is a further ancestor of j than the node i_2 if there is a $i_1 \to i_2 \to j$ path in $\mathcal{D}$. Hence by substitution we try to find the representation of X_i in terms of its further ancestral variables and noise variables. Now we do the substitution for all $X_i, i = 1, 2, \ldots, d$, i.e., we iterate Equation (1.3). After one iteration we get $\mathbf{X} = \big((C - \tilde{C})^2\big)^T \mathbf{X} \vee \big((C - \tilde{C}) \vee I\big)^T \odot \tilde{C} \odot \mathbf{Z}$, where I is the identity matrix in $\overline{\mathbb{R}}_+^{d\times d}$. After n-times' iteration, the fixed point equation becomes

$$\mathbf{X} = ((C - \tilde{C})^{n+1})^T X \vee \big((C - \tilde{C})^n \vee (C - \tilde{C})^{n-1} \vee \cdots (C - \tilde{C}) \vee I\big)^T \odot \tilde{C} \odot \mathbf{Z} \quad . \tag{4.1}$$

We shall show that $(C - \tilde{C})^{n+1}$ and $(C - \tilde{C})^n \vee (C - \tilde{C})^{n-1} \vee \cdots (C - \tilde{C}) \vee I$ play an essential role in the iteration and finding the fixed point of Equation (1.3). By using the results from max-times algebra in Chapter 3, we are able to provide a different approach for that, compared to the one in Gissibl and Klüppelberg (2018).

Recall maximum cycle mean from Definition 3.4. By Proposition 3.11 and Proposition 3.18 we know that if the maximum cycle mean of the matrix $(C - \tilde{C})$ satisfies $\lambda(C - \tilde{C}) > 1$, then both $\lim_{n\to\infty} (C - \tilde{C})^n$ and $\lim_{n\to\infty} I \vee (C - \tilde{C}) \vee (C - \tilde{C})^2 \vee \cdots \vee (C - \tilde{C})^n$ do not exist finitely, that is, some of their entries are $+\infty$. Hence in the following we focus on the case $\lambda(C - \tilde{C}) \leq 1$. We start with the simplest case $\lambda(C - \tilde{C}) = 0$, where the corresponding digraph $\mathcal{D}$ of the matrix $(C - \tilde{C})$ is acyclic. Then we investigate the cases where $0 < \lambda(C - \tilde{C}) < 1$ and $\lambda(C - \tilde{C}) = 1$.

Lemma 4.1. *For a matrix $C \in \overline{\mathbb{R}}_+^{d\times d}$ with $\lambda(C) \leq 1$, we define $B = (b_{ij}) := \tilde{C} \odot (C')^{d-1}$, where $\tilde{C} = diag\{c_{11}, c_{22}, \ldots, c_{dd}\}$ and C' is defined by replacing all diagonal entries in C by 1 and keeping non-diagonal entries unchanged. Then on the digraph $\mathcal{D}_{C-\tilde{C}}$,*

$$b_{ij} = c_{ii} \bigvee_{\pi \in P_{ij}} w(\pi) \text{ for } j \in an(i), b_{ii} = c_{ii} \text{ and } b_{ij} = 0 \text{ for } j \in V \backslash (an(i) \cup \{i\}),$$

where P_{ij} is the set of all $i \to j$ paths.

Proof. By Remark 3.17 we know that if we denote $(C')^{d-1} = (c_{ij}^{\prime(d-1)})$ and $\Gamma(C) = (\gamma_{ij})$, then $c_{ii}^{\prime(d-1)} = 1$ and $c_{ij}^{\prime(d-1)} = \gamma_{ij}$ for every $i, j \in \{1, 2, \ldots, d\}, i \neq j$. Since the matrix B is defined as $B =:= \tilde{C} \odot (C')^{d-1}$, by the multiplication between matrices in max-times algebra we get $b_{ii} = c_{ii}$ for every $i \in \{1, 2, \ldots, d\}$, $b_{ij} = c_{ii}\gamma_{ij} = c_{ii} \bigvee_{\pi \in P_{ij}} w(\pi)$ for $j \in \mathrm{an}(i)$ and $b_{ij} = c_{ii}\gamma_{ij} = 0$ for $j \in V \backslash (an(i) \cup \{i\})$. □

Theorem 4.2. *If $\mathbf{X}$ is a recursive max-linear model on a DAG $\mathcal{D}$ with weighted adjacency matrix $C \in \overline{\mathbb{R}}_+^{d\times d}$, then $\mathbf{X}$ is a generalized max-linear model and can be represented as*

$$\mathbf{X} = B^T \odot \mathbf{Z} \ , \tag{4.2}$$

where $B := \tilde{C} \odot (C')^{d-1}$ *is the max-linear coefficient matrix,* $\tilde{C} = diag\{c_{11}, c_{22}, \ldots, c_{dd}\}$ *and* C' *is defined by replacing all the diagonal entries in* C *by 1 and keeping non-diagonal entries unchanged.*

Proof. By Equation (4.1) we know that the random vector X satisfies

$$\mathbf{X} = ((C - \tilde{C})^{n+1})^T X \vee \left((C - \tilde{C})^n \vee (C - \tilde{C})^{n-1} \vee \cdots (C - \tilde{C}) \vee I\right)^T \odot \tilde{C} \odot \mathbf{Z} \quad (4.3)$$

for every $n \in \mathbb{N}$. Since $\mathcal{D}$ is acyclic, we know from Remark 3.5 that $\lambda(C - \tilde{C}) = 0$ and the lengths of all the paths in $\mathcal{D}$ do not exceed $d - 1$. Hence by Remark 3.10 $(C - \tilde{C})^n = 0$ for every $n \geq d$. From Corollary 3.12 and Theorem 3.16 we also know that

$$\begin{aligned}\Delta(C - \tilde{C}) &= \lim_{m \to \infty} \left(I \vee (C - \tilde{C}) \vee (C - \tilde{C})^2 \vee \cdots \vee (C - \tilde{C})^m\right) \\ &= I \vee (C - \tilde{C}) \vee (C - \tilde{C})^2 \vee \cdots \vee (C - \tilde{C})^n \text{ for every } n \geq d \\ &= (C')^{d-1}.\end{aligned}$$

Therefore, if we take $n \geq d$, then Equation (4.3) gives $\mathbf{X} = B^T \odot \mathbf{Z}$.

By Lemma 4.1 we know that all entries in B satisfy $b_{ij} = c_{ii} \bigvee_{\pi \in P_{ij}} w(\pi)$ for $j \in \text{an}(i), b_{ii} = c_{ii}$ and $b_{ij} = 0$ for $j \in V \backslash (an(i) \cup \{i\})$, which is exactly the definition (1.6) of max-linear coefficient in Gissibl and Klüppelberg (2018), hence the matrix $B := \tilde{C} \odot (C')^{d-1}$ is the max-linear coefficient matrix. □

Remark 4.3. Note that by theorem 2.2 in Gissibl and Klüppelberg (2018), a recursive max-linear model is a generalized max-linear model. Here we have provided an alternative algebraic approach for the proof. Furthermore, we have also provided a simpler representation of the max-linear coefficient matrix B in terms of C, compared to Theorem 2.4 in Gissibl and Klüppelberg (2018).

Theorem 4.4. *If* $\mathbf{X}$ *is a generalized recursive max-linear model on a digraph* $\mathcal{D}$ *defined as in Equation* (1.3) *with* $0 < \lambda(C - \tilde{C}) < 1$*, then as the iteration times* n *in Equation* (4.1) *goes to* ∞*, the model* $\mathbf{X}$ *tends to the generalized max-linear model*

$$\mathbf{X}_\infty = B^T \odot \mathbf{Z} \ ,$$

where $B = \tilde{C} \odot (C')^{d-1}$ *is the max-linear coefficient matrix and* C' *is defined the same as in Theorem 4.2.*

Proof. Equation (4.1) tells us the model $\mathbf{X}$ becomes $\mathbf{X} = ((C - \tilde{C})^{n+1})^T \mathbf{X} \vee \left((C - \tilde{C})^n \vee (C - \tilde{C})^{n-1} \vee \cdots (C - \tilde{C}) \vee I\right)^T \odot \tilde{C} \odot \mathbf{Z}$ after n-times' substitution. Hence in order to get the limit model as $n \to \infty$, we only need to understand the limit behavior of $(C - \tilde{C})^n$

and $I \vee (C-\tilde{C}) \vee (C-\tilde{C})^2 \vee \cdots \vee (C-\tilde{C})^n$. Since $0 < \lambda(C-\tilde{C}) < 1$, from Corollary 3.12 and Theorem 3.16 we know that

$$\begin{aligned}\Delta(C-\tilde{C}) &= \lim_{m\to\infty} \left(I \vee (C-\tilde{C}) \vee (C-\tilde{C})^2 \vee \cdots \vee (C-\tilde{C})^m\right) \\ &= I \vee (C-\tilde{C}) \vee (C-\tilde{C})^2 \vee \cdots \vee (C-\tilde{C})^n \text{ for every } n \geq d \\ &= (C')^{d-1}.\end{aligned}$$

By Proposition 3.21 we also know $\lim_{n\to\infty}(C-\tilde{C})^n = 0$. Therefore, as $n \to \infty$, the model $\mathbf{X}$ tends to the generalized max-linear model $\mathbf{X}_\infty = B^T \odot \mathbf{Z}$. □

Remark 4.5. In fact, it can be checked that the limit model $\mathbf{X}_\infty = B^T \odot \mathbf{Z}$ satisfies Equation (1.3), which implies that it is a solution of Equation (1.3). Hence, by trying to represent $\mathbf{X}$ in terms of its most initial ancestors and noise variables, we find the fixed point of Equation (1.3).

Note that by Remark 3.5, $\lambda(C-\tilde{C}) > 0$ if and only if there are cycles in the corresponding digraph $\mathcal{D}$ of matrix $(C-\tilde{C})$. Hence in the above case $\mathcal{D}$ is not a DAG any more.

Theorem 4.6. *Let $\mathbf{X}$ be a generalized recursive max-linear model on a digraph $\mathcal{D}$ defined as in Equation* (1.3) *with $\lambda(C-\tilde{C}) = 1$ and $\mathcal{D}$ satisfies the following conditions:*

(a) all cycles in $\mathcal{D}$ are either critical cycles or combinations of critical cycles,

(b) for every $i, j \in \{1, 2, \ldots, d\}$ there exists at most one $i \to j$ path without cycles.

Then the model $\mathbf{X}$ has at most $(m+d-1)$ different representations, where m is the least common multiple of the lengths of all the critical cycles.

Proof. By Equation (4.1) we know the model $\mathbf{X}$ can be represented as $\mathbf{X} = ((C-\tilde{C})^{n+1})^T X \vee \left((C-\tilde{C})^n \vee (C-\tilde{C})^{n-1} \vee \cdots (C-\tilde{C}) \vee I\right)^T \odot \tilde{C} \odot \mathbf{Z}$ for every positive integer n. For convenience we denote this representation as $\mathbf{X}_n$. Similarly as the proof of Theorem 4.4, from Corollary 3.12 and Theorem 3.16 we obtain $I \vee (C-\tilde{C}) \vee (C-\tilde{C})^2 \vee \cdots \vee (C-\tilde{C})^k = (C')^{d-1}$ for every $k \geq d$, where the matrix C' is defined by replacing all the diagonal entries in C by 1 and keeping non-diagonal entries unchanged. Hence for $n \geq d$, the representation $\mathbf{X}_n$ can be simplified as $\mathbf{X}_n = ((C-\tilde{C})^{n+1})^T X \vee \left((C')^{d-1}\right)^T \odot \tilde{C} \odot \mathbf{Z}$. By Corollary 3.25 we also know $(C-\tilde{C})^{n_1} = (C-\tilde{C})^{n_2}$ for $n_1 \equiv n_2 (\text{mod } m)$, $n_1, n_2 \geq d$. Hence $\mathbf{X}_{n_1} = \mathbf{X}_{n_2}$ for $n_1 \equiv n_2 (\text{mod } m)$, $n_1, n_2 \geq d$. Therefore $\mathbf{X}$ has at most $(m+d-1)$ different representations. □

Chapter 5

Deterministic Max-times System

5.1 Deterministic Max-times System

In this section we introduce the deterministic max-times system and apply our results about max-times algebra to investigate the stability of this system. A deterministic max-times system $(\mathbf{X}_t)_{t\in\mathbb{N}}$ is defined by

$$\mathbf{X}_t = A \odot \mathbf{X}_{t-1} \tag{5.1}$$

with finite initial state $\mathbf{X}_0$ and coefficient matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$. Note that here $\odot$ is the matrix multiplication in the max-times algebra (see Equation (3.1)).

From a theoretic system point of view, the most interesting question is how the system evolves, i.e. the limiting behavior of $\mathbf{X}_t$. Note that the max-times algebra $(\overline{\mathbb{R}}_+, \vee, \cdot)$ and max-plus algebra $(\overline{\mathbb{R}}, \vee, +)$ are essentially the same, since a multiplication can be changed to an addition by taking logarithm and the operation logarithm commutes with maximum. In the max-plus algebra, when entries of the coefficient matrix A are non-negative, the ergodic theory about both deterministic and stochastic max-plus linear system is well studied in Heidergott (2006). The assumption that entries of A are non-negative is a natural condition within the book Heidergott (2006), because the author focuses on using max-plus linear system to model queuing systems. In that case entries of A represent service times at different stations in a queuing system, and they are assumed to be non-negative. However, if we consider other applications, entries of A might be negative, which corresponds to a matrix with entries in the interval $[0, 1)$ in the max-times algebra setting. Hence, the maximum cycle mean $\lambda(A)$ (see Definition 3.4) might be less than or equal to 1. In the following we use our results from Chapter 3 to investigate the limiting behavior of $\mathbf{X}_t$ when $\lambda(A) \leq 1$.

Proposition 5.1. *For a deterministic max-times system* $(\mathbf{X}_t)_{t\in\mathbb{N}}$ *defined by* $\mathbf{X}_t = A\odot\mathbf{X}_{t-1}$

with finite initial state $\mathbf{X}_0$ *and coefficient matrix* $A \in \overline{\mathbb{R}}_+^{d\times d}$*, if* $\lambda(A) < 1$*, then*

$$\lim_{t\to\infty} \mathbf{X}_t = 0.$$

Proof. Since the system is defined by $\mathbf{X}_t = A \odot \mathbf{X}_{t-1}$ with finite initial state $\mathbf{X}_0$, by induction we get $X_t = A^t \odot \mathbf{X}_0$. By Proposition 3.21, if $\lambda(A) < 1$, then $\lim_{t\to\infty} A^n = 0$. Hence $\lim_{t\to\infty} \mathbf{X}_t = \lim_{t\to\infty} A^t \odot \mathbf{X}_0 = 0$. □

The proposition implies that if the maximum cycle mean of the coefficient matrix $\lambda(A)$ is less than 1, then the deterministic max-times system converges to 0, whatever the finite initial state is. That is, if the coefficient matrix A satisfies $\lambda(A) < 1$, then the system is globally asymptotically stable.

Example 5.2. Consider the deterministic max-times system $\mathbf{X}_t = A \odot \mathbf{X}_{t-1}$ with finite initial state X_0 and coefficient matrix

$$A = \begin{pmatrix} 0 & 1.5 \\ 0.6 & 0.8 \end{pmatrix}.$$

By Equation (3.1) we can calculate that

$$A^2 = \begin{pmatrix} 0.9 & 1.2 \\ 0.48 & 0.9 \end{pmatrix}.$$

Hence by Proposition 3.7,

$$\lambda(A) = \bigvee_{k,i=1}^{2} (a_{ii}^{(k)})^{\frac{1}{k}} = \sqrt{0.9} < 1.$$

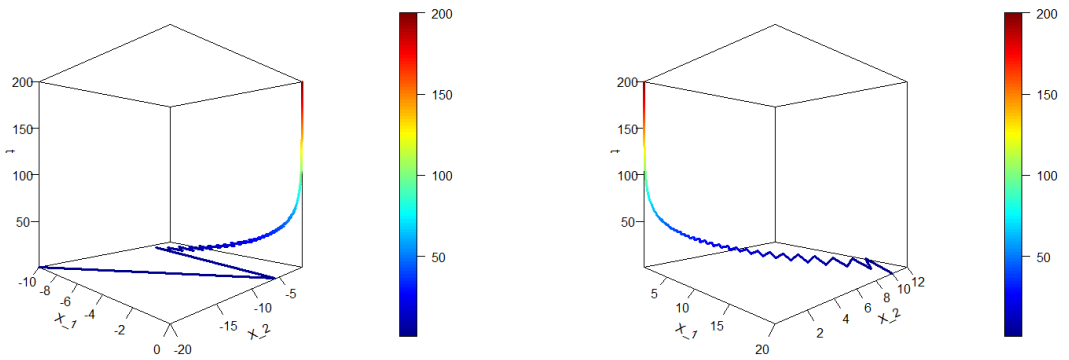
From Figure 5.1 we can see that the system $\mathbf{X}_t$ converges to 0 for different initial states $\mathbf{X}_0 = (-10, -20)$ and $\mathbf{X}_0 = (20, 10)$. In fact it can be easily checked that $\mathbf{X}_t$ converges to 0 for any initial state.

Proposition 5.3. *For a deterministic max-times system* $(\mathbf{X}_t)_{t\in\mathbb{N}}$ *defined by* $\mathbf{X}_t = A \odot \mathbf{X}_{t-1}$ *with finite initial state* $\mathbf{X}_0$ *and coefficient matrix* $A \in \overline{\mathbb{R}}_+^{d\times d}$*, if* $\lambda(A) = 1$*, all cycles in* $\mathcal{D}_A$ *are either critical cycles or combinations of critical cycles, and for every* $i, j \in \{1, 2, \ldots, d\}$ *there exists at most one* $i \to j$ *path without cycles, then*

$$\mathbf{X}_{t_1} = \mathbf{X}_{t_2} \text{ for } t_1 \equiv t_2 (mod\ m), \qquad t_1, t_2 \geq d,$$

where m *is the least common multiple of the lengths of all critical cycles in* $\mathcal{D}_A$*.*

Proof. By induction we have $X_t = A^t \odot \mathbf{X}_0$ for $t \in \mathbb{N}$. By Corollary 3.25 we know that $A^{t_1} = A^{t_2}$ for $t_1 \equiv t_2 (\text{mod } m), t_1, t_2 \geq d$. Hence $\mathbf{X}^{t_1} = \mathbf{X}^{t_2}$ for $t_1 \equiv t_2 (\text{mod } m), t_1, t_2 \geq d$. □



(a) $\mathbf{X}_0 = (-10, -20)$ (b) $\mathbf{X}_0 = (20, 10)$

Figure 5.1: Evolution of $\mathbf{X}_t$ for different initial state $\mathbf{X}_0$

Remark 5.4. This proposition implies that if $\lambda(A) = 1$, and the conditions that all cycles in $\mathcal{D}_A$ are either critical cycles or combinations of critical cycles, and for every $i, j \in \{1, 2, \ldots, d\}$ there exists at most one $i \to j$ path without cycles, are satisfied, then the system $\mathbf{X}_t$ will have periodicity m when $t \geq d$.

5.2 Deterministic Max-linear System with Noise

In this section we introduce noise into the deterministic max-times system, and then investigate the stability of this resulting system. A deterministic max-linear system with noise is a stochastic process $(\mathbf{X}_t)_{t\in\mathbb{N}}$ defined by

$$\mathbf{X}_t = A \odot \mathbf{X}_{t-1} \vee \mathbf{Z}_t \tag{5.2}$$

with finite initial state $\mathbf{X}_0$, coefficient matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$ and i.i.d random variables $\mathbf{Z}_t, t \in \mathbb{N}$.

The deterministic max-linear system with noise corresponds to the linear state space model in the usual $(\mathbb{R}, +, \cdot)$ algebra. A linear state space model is a stochastic process $(\mathbf{X}_t)_{t\in\mathbb{N}}$ satisfying $\mathbf{X}_t = F\mathbf{X}_{t-1} + G\mathbf{Z}_t$ with arbitrary initial state $\mathbf{X}_0$ and i.i.d. random variables $\mathbf{Z}_t$. The stability about this model has been intensively studied in Meyn and Tweedie (1993).

In the following we assume all noise variables $Z_t^i, i = 1, \ldots, d, t \in \mathbb{N}$ are i.i.d. and regularly varying with index $\alpha \in \overline{\mathbb{R}}_+$. We focus on the case $\lambda(A) < 1$. By induction we know that the deterministic max-linear system with noise can be written as

$$\mathbf{X}_n = A \odot \mathbf{X}_{n-1} \vee \mathbf{Z}_n$$

$$
\begin{aligned}
&= A^2 \odot \mathbf{X}_{n-2} \vee \mathbf{Z}_n \vee A \odot \mathbf{Z}_{n-1} \\
&= A^3 \odot \mathbf{X}_{n-3} \vee \mathbf{Z}_n \vee A \odot \mathbf{Z}_{n-1} \vee A^2 \odot \mathbf{Z}_{n-2} \\
&= \cdots \\
&= A^n \odot \mathbf{X}_0 \vee \mathbf{Z}_n \vee A \odot \mathbf{Z}_{n-1} \vee \cdots \vee A^{n-1} \odot \mathbf{Z}_1 \text{ for } n \in \mathbb{N}. \qquad (5.3)
\end{aligned}
$$

Proposition 3.21 implies that $\lim_{n\to\infty} A^n = 0$. Hence in order to investigate the limit behavior of $\mathbf{X}_n$, it is enough to only investigate the limit behavior of the term $\mathbf{Z}_n \vee A \odot \mathbf{Z}_{n-1} \vee \cdots \vee A^{n-1} \odot \mathbf{Z}_1$.

Lemma 5.5. *Let $\mathbf{Z}_n = (Z_n^1, \ldots, Z_n^d)^T$ and $\mathbf{M}_n = (m_1^n, \ldots, m_d^n)^T := \mathbf{Z}_n \vee A \odot \mathbf{Z}_{n-1} \vee \cdots \vee A^{n-1} \odot \mathbf{Z}_1$. If $Z_t^i, i = 1, \ldots, d, t \in \mathbb{N}$ are i.i.d. and regularly varying with index $\alpha \in \overline{\mathbb{R}}_+$ (denoted as $Z \in RV(\alpha)$), then $\mathbf{M}_n \in MDA(G)$ with*

$$
G(\mathbf{x}) = \exp\big(- \sum_{i=1}^{d} (\frac{1}{x_i})^\alpha - \sum_{k=1}^{n-1} \sum_{i=1}^{d} \bigvee_{j=1}^{d} (\frac{a_{ji}^{(k)}}{x_j})^\alpha \big), \quad \mathbf{x} = (x_1, \ldots, x_d)^T \in \overline{\mathbb{R}}_+^d .
$$

Proof. We denote $A^n = (a_{ij}^{(n)})$, and $\mathrm{An}^{(n)}(i)$ as the set of ancestors of i and i itself, in the corresponding digraph of matrix A^n. From $\mathbf{M}_n = \mathbf{Z}_n \vee A \odot \mathbf{Z}_{n-1} \vee A^{n-1} \odot \mathbf{Z}_1$ we have

$$
\begin{cases}
m_1^n &= Z_n^1 \vee (a_{11} Z_{n-1}^1 \vee \cdots \vee a_{1d} Z_{n-1}^d) \vee \cdots \vee (a_{11}^{(n-1)} Z_1^1 \vee \cdots \vee a_{1d}^{(n-1)} Z_1^d) \\
&\vdots \\
m_d^n &= Z_n^d \vee (a_{d1} Z_{n-1}^1 \vee \cdots \vee a_{dd} Z_{n-1}^d) \vee \cdots \vee (a_{d1}^{(n-1)} Z_1^1 \vee \cdots \vee a_{dd}^{(n-1)} Z_1^d)
\end{cases}
$$

Since $Z_t^i, i = 1, \ldots, d, t \in \mathbb{N}$ are i.i.d., their common distribution can be denoted as F_Z, then the distribution function of $\mathbf{M}_n$ is

$$
\begin{aligned}
F_n(\mathbf{x}) &= \mathrm{P}(\mathbf{M}_n \le \mathbf{x}) \\
&= \mathrm{P}(m_1^n \le x_1, \ldots, m_d^n \le x_d) \\
&= \mathrm{P}(Z_n^i \vee (a_{i1} Z_{n-1}^1 \vee \cdots \vee a_{id} Z_{n-1}^d) \vee \cdots \vee (a_{i1}^{(n-1)} Z_1^1 \vee \cdots \vee a_{id}^{(n-1)} Z_1^d) \le x_i, i = 1, \ldots, d) \\
&= \mathrm{P}(Z_n^i \le x_i, Z_{n-1}^i \le \bigwedge_{j \in \mathrm{An}(i)} \frac{x_j}{a_{ji}}, \ldots, Z_1^i \le \bigwedge_{j \in \mathrm{An}^{(n-1)}(i)} \frac{x_j}{a_{ji}^{(n-1)}}, i = 1, \ldots, d) \\
&= \mathrm{P}(Z_n^i \le x_i, Z_k^i \le \bigwedge_{j \in \mathrm{An}^{(n-k)}(i)} \frac{x_j}{a_{ji}^{(n-k)}}, i = 1, \ldots, d, k = 1, \ldots, n-1) \\
&= \prod_{i=1}^{d} \mathrm{P}(Z_n^i \le x_i) \cdot \prod_{k=1}^{n-1} \prod_{i=1}^{d} \mathrm{P}(Z_k^i \le \bigwedge_{j \in \mathrm{An}^{(n-k)}(i)} \frac{x_j}{a_{ji}^{(n-k)}}) \\
&= \prod_{i=1}^{d} F_Z(x_i) \cdot \prod_{k=1}^{n-1} \prod_{i=1}^{d} F_Z(\bigwedge_{j \in \mathrm{An}^{(k)}(i)} \frac{x_j}{a_{ji}^{(k)}})
\end{aligned}
$$

By Proposition 1.11 in Resnick (1987), $Z \in RV(\alpha)$ if and only if the distribution of Z is in the maximum domain of attraction of $\Phi_\alpha(x) = \exp(-x^{-\alpha})$, i.e., there exists a positive sequence $(a_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} F_Z^n(a_n x) = \Phi_\alpha(x), \text{ for every } x \in \overline{\mathbb{R}}_+.$$

Hence we have

$$\begin{aligned}
\lim_{m\to\infty} \mathrm{P}^m(\mathbf{M}_n \le a_m \mathbf{x}) &= \lim_{m\to\infty} \prod_{i=1}^d F_Z^m(a_m x_i) \cdot \prod_{k=1}^{n-1}\prod_{i=1}^d F_Z^m(a_m \bigwedge_{j\in \mathrm{An}^{(k)}(i)} \frac{x_j}{a_{ji}^{(k)}}) \\
&= \prod_{i=1}^d \Phi_\alpha(x_i) \cdot \prod_{k=1}^{n-1}\prod_{i=1}^d \Phi_\alpha(\bigwedge_{j\in \mathrm{An}^{(k)}(i)} \frac{x_j}{a_{ji}^{(k)}}) \\
&= \exp\big(-\sum_{i=1}^d (\frac{1}{x_i})^\alpha - \sum_{k=1}^{n-1}\sum_{i=1}^d \bigvee_{j=1}^d (\frac{a_{ji}^{(k)}}{x_j})^\alpha\big)
\end{aligned}$$

Therefore the statement $\mathbf{M}_n \in MDA(G)$ follows. □

Theorem 5.6. *For a deterministic max-linear system with noise $(\mathbf{X}_t)_{t\in\mathbb{N}}$ defined by $\mathbf{X}_t = A \odot \mathbf{X}_{t-1} \vee \mathbf{Z}_t$ with finite initial state $\mathbf{X}_0$, coefficient matrix $A \in \overline{\mathbb{R}}_+^{d\times d}$ and i.i.d regularly varying noise variables $Z_t^i \in RV(\alpha), t \in \mathbb{N}, i = 1, \dots, d$, if $\lambda(A) < 1$, then $\lim_{n\to\infty} \mathbf{X}_n \in MDA(F)$ with*

$$F(\mathbf{x}) = \exp\big(-\sum_{i=1}^d (\frac{1}{x_i})^\alpha - \sum_{k=1}^{\infty}\sum_{i=1}^d \bigvee_{j=1}^d (\frac{a_{ji}^{(k)}}{x_j})^\alpha\big), \quad \mathbf{x} = (x_1, \dots, x_d)^T \in \overline{\mathbb{R}}_+^d .$$

Proof. By Equation (5.3) and Proposition 3.21, we have $\lim_{n\to\infty} \mathbf{X}_n = \lim_{n\to\infty} \mathbf{M}_n$. If we denote $\lim_{n\to\infty} \mathbf{M}_n$ as $\mathbf{M}_\infty$ and its distribution function as F_∞, then by the proof of Lemma 5.5 we have

$$\begin{aligned}
F_\infty(\mathbf{x}) &= \mathrm{P}(\mathbf{M}_\infty \le \mathbf{x}) \\
&= \prod_{i=1}^d F_Z(x_i) \cdot \prod_{k=1}^{\infty}\prod_{i=1}^d F_Z(\bigwedge_{j\in \mathrm{An}^{(k)}(i)} \frac{x_j}{a_{ji}^{(k)}})
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\lim_{m\to\infty} \mathrm{P}^m(\mathbf{M}_\infty \le a_m \mathbf{x}) &= \lim_{m\to\infty} \prod_{i=1}^d F_Z^m(a_m x_i) \cdot \prod_{k=1}^{\infty}\prod_{i=1}^d F_Z^m(a_m \bigwedge_{j\in \mathrm{An}^{(k)}(i)} \frac{x_j}{a_{ji}^{(k)}}) \\
&= \prod_{i=1}^d \Phi_\alpha(x_i) \cdot \prod_{k=1}^{\infty}\prod_{i=1}^d \Phi_\alpha(\bigwedge_{j\in \mathrm{An}^{(k)}(i)} \frac{x_j}{a_{ji}^{(k)}}) \\
&= \exp\big(-\sum_{i=1}^d (\frac{1}{x_i})^\alpha - \sum_{k=1}^{\infty}\sum_{i=1}^d \bigvee_{j=1}^d (\frac{a_{ji}^{(k)}}{x_j})^\alpha\big).
\end{aligned}$$

Note that for any given $\mathbf{x} = (x_1, \ldots, x_d)^T$ with $x_i > 0, i = 1, \ldots, d$, the series $\sum_{k=1}^{\infty}\sum_{i=1}^{d}\bigvee_{j=1}^{d}(\frac{a_{ji}^{(k)}}{x_j})^\alpha$ always converges. The reason is as follows. By Proposition 3.11 we know that if $\lambda(A) < 1$, then the weak transitive closure $\Gamma(A) = (\gamma_{ij})$ of A exists and every $\gamma_{ij}, i, j \in \{1, 2, \ldots, d\}$ is finite. Moreover, by Lemma 3.19, $a_{ij}^{(n)} \le \gamma_{ij}(\lambda(A))^{n-d}$ for every $i, j \in \{1, 2, \ldots, d\}$ and $n > d$. If we denote $\gamma = \bigvee_{i,j=1}^{d}\gamma_{ij}$, then we have $a_{ij}^{(n)} \le \gamma(\lambda(A))^{n-d}$ for every $i, j \in \{1, 2, \ldots, d\}$ and $n > d$. Therefore

$$\begin{aligned}
&\sum_{k=1}^{\infty}\sum_{i=1}^{d}\bigvee_{j=1}^{d}(\frac{a_{ji}^{(k)}}{x_j})^\alpha \\
&= \sum_{k=1}^{d}\sum_{i=1}^{d}\bigvee_{j=1}^{d}(\frac{a_{ji}^{(k)}}{x_j})^\alpha + \sum_{k=d+1}^{\infty}\sum_{i=1}^{d}\bigvee_{j=1}^{d}(\frac{a_{ji}^{(k)}}{x_j})^\alpha \\
&\le \sum_{k=1}^{d}\sum_{i=1}^{d}\bigvee_{j=1}^{d}(\frac{a_{ji}^{(k)}}{x_j})^\alpha + \sum_{k=d+1}^{\infty}\sum_{i=1}^{d}\bigvee_{j=1}^{d}(\frac{\gamma(\lambda(A))^{k-d}}{x_j})^\alpha \\
&= \sum_{k=1}^{d}\sum_{i=1}^{d}\bigvee_{j=1}^{d}(\frac{a_{ji}^{(k)}}{x_j})^\alpha + \sum_{k=d+1}^{\infty}(\lambda(A))^{\alpha(k-d)}\bigvee_{j=1}^{d}(\frac{\gamma}{x_j})^\alpha
\end{aligned}$$

Since $\lambda(A) < 1$, series $\sum_{k=d+1}^{\infty}(\lambda(A))^{\alpha(k-d)}\bigvee_{j=1}^{d}(\frac{\gamma}{x_j})^\alpha$ converges. Hence by dominated convergence theorem the series $\sum_{k=1}^{\infty}\sum_{i=1}^{d}\bigvee_{j=1}^{d}(\frac{a_{ji}^{(k)}}{x_j})^\alpha$ also converges. Therefore the statement $\lim_{n\to\infty}\mathbf{X}_n \in MDA(F)$ follows. □

Corollary 5.7. *If the noise variables $Z_t^i, t \in \mathbb{N}, i = 1, \ldots, d$ follow Fréchet-α distribution, then the distribution of $\mathbf{X}_n$ converges pointwise to F with*

$$F(\mathbf{x}) = \exp\big(-\sum_{i=1}^{d}(\frac{1}{x_i})^\alpha - \sum_{k=1}^{\infty}\sum_{i=1}^{d}\bigvee_{j=1}^{d}(\frac{a_{ji}^{(k)}}{x_j})^\alpha\big), \quad \mathbf{x} = (x_1, \ldots, x_d)^T \in \overline{\mathbb{R}}_+^d\ .$$

Proof. From the proof of Lemma 5.5 we know that the distribution function of $\mathbf{M}_n$ is

$$\begin{aligned}
F_n(\mathbf{x}) &= \mathrm{P}(\mathbf{M}_n \le \mathbf{x}) \\
&= \prod_{i=1}^{d}F_Z(x_i)\cdot\prod_{k=1}^{n-1}\prod_{i=1}^{d}F_Z(\bigwedge_{j\in\mathrm{An}^{(k)}(i)}\frac{x_j}{a_{ji}^{(k)}})
\end{aligned}$$

By Equation (5.3), $\mathbf{X}_n = A^n \odot \mathbf{X}_0 \vee \mathbf{M}_n$. For simplicity we denote $\mathbf{A}_n := A^n \odot \mathbf{X}_0 \in \overline{\mathbb{R}}_+^d$, then the distribution of $\mathbf{X}_n$ is

$$\mathrm{P}(\mathbf{X}_n \le \mathbf{x}) = \mathrm{P}(\mathbf{A}_n \vee \mathbf{M}_n \le \mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} < \mathbf{A}_n \\ \mathrm{P}(\mathbf{M}_n \le \mathbf{x}), & \text{if } \mathbf{x} \ge \mathbf{A}_n \end{cases}$$

Using Equation (5.3) and Proposition 3.21, we know that $\lim_{n\to\infty} \mathbf{A}_n = \lim_{n\to\infty} A^n \odot \mathbf{X}_0 = 0$. Hence for every $\mathbf{x} \in \overline{\mathbb{R}}_+^d$,

$$\begin{aligned}
\lim_{n\to\infty} \mathrm{P}(\mathbf{X}_n \le \mathbf{x}) &= \lim_{n\to\infty} \mathrm{P}(\mathbf{A}_n \vee \mathbf{M}_n \le \mathbf{x}) \\
&= \lim_{n\to\infty} \mathrm{P}(\mathbf{M}_n \le \mathbf{x}) \\
&= \lim_{n\to\infty} \prod_{i=1}^{d} F_Z(x_i) \cdot \prod_{k=1}^{n-1} \prod_{i=1}^{d} F_Z\Big(\bigwedge_{j\in \mathrm{An}^{(k)}(i)} \frac{x_j}{a_{ji}^{(k)}}\Big) \\
&= \prod_{i=1}^{d} F_Z(x_i) \cdot \prod_{k=1}^{\infty} \prod_{i=1}^{d} F_Z\Big(\bigwedge_{j\in \mathrm{An}^{(k)}(i)} \frac{x_j}{a_{ji}^{(k)}}\Big) \\
&= \prod_{i=1}^{d} \Phi_\alpha(x_i) \cdot \prod_{k=1}^{\infty} \prod_{i=1}^{d} \Phi_\alpha\Big(\bigwedge_{j\in \mathrm{An}^{(k)}(i)} \frac{x_j}{a_{ji}^{(k)}}\Big) \\
&= \exp\Big(-\sum_{i=1}^{d} (\frac{1}{x_i})^\alpha - \sum_{k=1}^{\infty} \sum_{i=1}^{d} \bigvee_{j=1}^{d} (\frac{a_{ji}^{(k)}}{x_j})^\alpha \Big).
\end{aligned}$$

□

Chapter 6

Recursive Max-linear Models on Polytrees

Recall from Equation (2.2) the definition of tail dependence matrix χ for a recursive max-linear model (RMLM) $\mathbf{X}$ on a DAG $\mathcal{D} = (V, E)$. For simplicity we denote $\chi = (\chi_{ij})$ and $\chi_{\{i_1,i_2,\ldots,i_d\}\times\{j_1,j_2,\ldots,j_d\}}$ as a square submatrix of χ with its rows $i_1, \ldots, i_d$ and columns $j_1, \ldots, j_d$. In this chapter we investigate the relationship between conditional dependence structure and the tail dependence structure among the random variables.

We first recall some results from Gissibl, Klüppelberg, and Moritz (2017), which will be used throughout this chapter. The proofs for these results can be found in the paper Gissibl, Klüppelberg, and Moritz (2017).

By equation (7) and (8) in Gissibl, Klüppelberg, and Moritz (2017), we know the relation between the standardized MLCM and the tail dependence coefficient (TDC) matrix. That is, for a RMLM $\mathbf{X}$, if we denote the standardized MLCM by

$$\overline{B} = (\overline{b}_{ij})_{d\times d} = \Big(\frac{b^{\alpha}_{ij}}{\sum_{k\in \mathrm{An}(j)} b^{\alpha}_{kj}}\Big)_{d\times d} \ ,$$

then the TDC can be computed as

$$\chi_{ij} = \chi_{ji} = \sum_{k\in \mathrm{An}(i)\cap \mathrm{An}(j)} \overline{b}_{ki} \wedge \overline{b}_{kj} \ . \tag{6.1}$$

Lemma 6.1. *[Lemma 2.1(b) in Gissibl, Klüppelberg, and Moritz (2017)] Let $i \in V$ and $j \in an(i)$, then $\chi(i,j) \leq \sum_{k\in An(j)} \overline{b}_{ki} < 1$.*

Lemma 6.2. *[Lemma 3.4 in Gissibl, Klüppelberg, and Moritz (2017)] $\mathbf{X}$ is a recursive max-weighted model (RMWM), i.e. a recursive max-linear model with all paths being max-weighted. Let $\mathcal{D} = (V, E)$ be the associated DAG and $i \in V$, then*

(a) for $j \in An(i)$, $\chi(i,j) = \frac{\bar{b}_{ji}}{\bar{b}_{jj}} = \sum_{k \in An(j)} \bar{b}_{ki} = \sum_{k \in An(j)} \bar{b}_{kk}\chi(k,i)$.

(b) for $k \in an(i)$ *and* $j \in an(k)$, $\chi(j,i) = \chi(j,k)\chi(k,i) < \chi(j,k) \wedge \chi(k,i)$.

Lemma 6.3. *If the nodes* i, $j \in V$ *have a common ancestor, then there must be a lowest common ancestor. That is, the other common ancestors must be ancestors of the lowest common ancestor. Moreover, if we denote the lowest common ancestor as* h*, then* $\chi_{ij} = \chi_{hi} \wedge \chi_{hj}$.

Proof. The first statement follows from the definition of polytree. For the second statement, by Equation (6.1) we know $\chi_{ij} = \sum_{k \in \mathrm{An}(i) \cap \mathrm{An}(j)} \bar{b}_{ki} \wedge \bar{b}_{kj} = \sum_{k \in \mathrm{An}(h)} \bar{b}_{ki} \wedge \bar{b}_{kj}$. Then using Lemma 6.2 (a) and (b) we can get

$$\begin{aligned}\chi_{ij} &= \sum_{k \in \mathrm{An}(h)} \bar{b}_{ki} \wedge \bar{b}_{kj} = \sum_{k \in \mathrm{An}(h)} \bar{b}_{kk}(\chi_{ki} \wedge \chi_{kj}) = \sum_{k \in \mathrm{An}(h)} \bar{b}_{kk}(\chi_{kh}\chi_{hi} \wedge \chi_{kh}\chi_{hj}) \\ &= (\chi_{hi} \wedge \chi_{hj}) \sum_{k \in \mathrm{An}(h)} \bar{b}_{kk}\chi_{kh} = (\chi_{hi} \wedge \chi_{hj})\chi_{hh} = \chi_{hi} \wedge \chi_{hj}.\end{aligned}$$

□

Definition 6.4. Let $A = (a_{ij}) \in \overline{\mathbb{R}}_+^{d \times d}$, we define the *max-algebraic permanent* (or briefly permanent) of A as follows:

$$maxper(A) = \bigvee_{\pi \in P_n} \prod_{i \in N} a_{i,\pi(i)} \quad ,$$

where P_n is the set of all permutations of the set $N = \{1, 2, \ldots, d\}$.

Theorem 6.5. **X** *is a recursive max-linear model on a polytree* $\mathcal{D} = (V, E)$. χ *is its tail dependence coefficient matrix. Then for* i, $j \in V$*, if there is an edge between* i *and* j*, we have*

$$maxper(\chi_{\{i,k\} \times \{j,k\}}) = \chi_{ij} \text{ for every } k \in V.$$

Proof. The submatrix $\chi_{\{i,k\} \times \{j,k\}}$ is:

$$\chi_{\{i,k\} \times \{j,k\}} = \begin{pmatrix} \chi_{ij} & \chi_{ik} \\ \chi_{kj} & \chi_{kk} \end{pmatrix} = \begin{pmatrix} \chi_{ij} & \chi_{ik} \\ \chi_{kj} & 1 \end{pmatrix} ,$$

hence

$$\mathrm{maper}(\chi_{\{i,k\} \times \{j,k\}}) = \chi_{ij} \cdot 1 \vee \chi_{ik} \cdot \chi_{kj} = \chi_{ij} \vee (\chi_{ik}\chi_{kj}).$$

Without loss of generality we assume the edge is from i to j in $\mathcal{D}$. Then it is left to show $\chi_{ij} \vee \chi_{ik}\chi_{kj} = \chi_{ij}$, or equivalently $\chi_{ij} \geq \chi_{ik}\chi_{kj}$ for every $k \in V$. We prove this statement by covering all cases, i.e. all possible positions of k in the polytree $\mathcal{D}$. For the trivial case $k = i$ or $k = j$, we directly get $\chi_{ik}\chi_{kj} = \chi_{ij}$, hence in the following we assume $k \neq i, k \neq j$.

(i) If there is a path from i to k, but no path from j to k, then $i \in \text{an}(k), j \notin \text{an}(k)$. Since $i \in \text{pa}(j), i \in \text{an}(k)$, and $\mathcal{D}$ is a polytree, then i must be the lowest common ancestor of j and k. By Lemma 6.3, $\chi_{kj} = \chi_{ik} \wedge \chi_{ij} \leq \chi_{ij}$. From Lemma 6.1 we know $\chi_{ik} < 1$, so $\chi_{ik}\chi_{kj} < \chi_{ij}$.

(ii) If there is a path from k to i, then $k \in \text{an}(i), k \in \text{an}(j)$. Using Lemma 6.2(b) and Lemma 6.1, $\chi_{ik}\chi_{kj} < \chi_{kj} = \chi_{ki}\chi_{ij} < \chi_{ij}$.

(iii) If there is a path from j to k, then $i \in \text{an}(k), j \in \text{an}(k)$. By Lemma 6.2(b) and Lemma 6.1, $\chi_{ik}\chi_{kj} < \chi_{ik} = \chi_{ij}\chi_{jk} < \chi_{ij}$.

(iv) If there is path from k to j, then k and i cannot be adjacent or have a common ancestor. So $\chi_{ki} = 0$ and $\chi_{ik}\chi_{kj} = 0 < \chi_{ij}$.

(v) If neither k and i nor k and j are connected, but k and i have common ancestors. By Lemma 6.3, k and i must have a lowest common ancestor, say h. Since $i \in \text{pa}(j)$, h must also be the lowest common ancestor for k and j. From Lemma 6.1, Lemma 6.2(b) and Lemma 6.3 we have $\chi_{ki} = \chi_{hk} \wedge \chi_{hi} < 1$ and $\chi_{kj} = \chi_{hj} \wedge \chi_{hk} \leq \chi_{hj} = \chi_{hi}\chi_{ij} < \chi_{ij}$. So $\chi_{ik}\chi_{kj} < \chi_{ij}$.

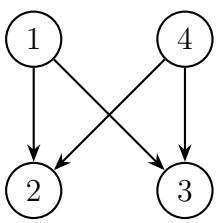
(vi) If neither k and i nor k and j are connected, k and j have a common ancestor but k and i have no common ancestors, then $\chi_{ik} = 0$, so $\chi_{ik}\chi_{kj} = 0 < \chi_{ij}$.

(vii) If neither k and i nor k and j are connected, neither k and j nor k and i have a common ancestor, then $\chi_{ik} = \chi_{kj} = 0$ and $\chi_{ik}\chi_{kj} = 0 < \chi_{ij}$.

□

Note that the condition $\mathcal{D}$ is a polytree can not be dropped, otherwise the theorem does not hold. Consider the following counter-example.

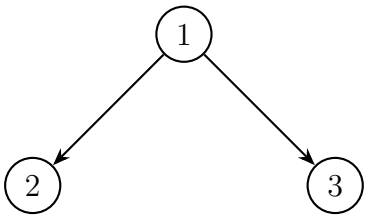
Example 6.6. Assume the recursive max-weighted model $\mathbf{X}$ has the following associated DAG $\mathcal{D}$ with standardized max-linear coefficients $\bar{b}_{11} = \bar{b}_{44} = 1, \bar{b}_{12} = 0.1, \bar{b}_{13} = 0.5, \bar{b}_{43} = \bar{b}_{42} = 0.4$.



Then $\chi_{12} = \frac{b_{12}}{b_{11}} = 0.1$, $\chi_{13} = \frac{b_{13}}{b_{11}} = 0.5$, $\chi_{23} = b_{12} \wedge b_{13} + b_{42} \wedge b_{43} = 0.5$. In this model although we have the edge $1 \to 2$, the inequality $\chi_{12} \geq \chi_{13}\chi_{23}$ does not hold.

Theorem 6.5 shows that there is an edge between i and j is a sufficient condition for $\text{maxper}(\chi_{\{i,k\}\times\{j,k\}}) = \chi_{ij}$. Unfortunately it is not a necessary condition, as the following example shows.

Example 6.7. Assume the recursive max-weighted model $\mathbf{X}$ has the following associated DAG $\mathcal{D}$.



Then by Lemma 6.3 and Lemma 6.1, $\chi_{23} = \chi_{12} \wedge \chi_{13} > \chi_{12}\chi_{13}$. Since $\chi_{ii} = 1$ for any $i \in \{1, 2, 3\}$, we also know $\chi_{23} \geq \chi_{k2}\chi_{k3}$ for $k = 2, 3$. Hence we have $\chi_{23} \geq \chi_{k2}\chi_{k3}$ for any $k \in V$. However, there is no edge between the node 2 and 3. Therefore there is an edge between i and j is not a necessary condition for $\text{maxper}(\chi_{\{i,k\}\times\{j,k\}}) = \chi_{ij}$.

From theorem 3.4 in Spirtes, Glymour, and Scheines (2000) we know, if the distribution $P(\mathbf{X})$ is faithful to the DAG $\mathcal{D}$, then the following statements are equivalent:

(a) there is an edge between nodes i and j in DAG $\mathcal{D}$

(b) for all $\mathbf{s} \subseteq V\backslash\{i, j\}$, $X^{(i)}$ and $X^{(j)}$ are conditionally dependent given $\{X^{(r)} : r \in \mathbf{s}\}$

Combining this property and Theorem 6.5, we know the relationship between the conditional dependence structure and the max-algebraic permanent of χ.

Corollary 6.8. *If the distribution $P(\mathbf{X})$ is faithful to the polytree $\mathcal{D}$, and for all $\mathbf{s} \subseteq V\backslash\{i, j\}$, $X^{(i)}$ and $X^{(j)}$ are conditionally dependent given $\{X^{(r)} : r \in \mathbf{s}\}$, then maxper$(\chi_{\{i,k\}\times\{j,k\}}) = \chi_{ij}$ for any $k \in V$.*

Acknowledgements

First of all, I want to thank my supervisor, Claudia Klüppelberg, for her remarkable guidance in the past six months. Her continuous presence, unceasing passion for math, and unlimited inspiration and creativity has always motivated me to think deeper and produce better work. Her great kindness and optimistic attitude towards research and life has shown me how to become a better person. There are no words to appropriately express how grateful and fortunate I am to have the opportunity to learn from and work with her.

I would like to thank Steffen Lauritzen for many interesting and fruitful discussions. And also many thanks to the Mathematical Statistics Chair at TUM, especially Mario Krali, Nadine Gissbl, Viet Son Pham, Carlos Enrique Améndola Cerón, Thiago do Rego Sousa, and Thomas Delerue, for the pleasant research atmosphere and many fun talks.

Last but not least, I would like to thank my parents for their consistent support and strong belief in me. Without them, I could not have even taken my first step to pursue a study in Germany.

Bibliography

Butkovič, Peter (2010). *Max-linear Systems: Theory and Algorithms*. Springer.

de Haan, Laurens and Ana Ferreira (2006). *Extreme Value Theory: An Introduction*. Springer.

Gissibl, Nadine and Claudia Klüppelberg (2018). "Max-linear models on directed acyclic graphs". *Bernoulli* 24.4A, pp. 2693–2720.

Gissibl, Nadine, Claudia Klüppelberg, and Otto Moritz (2017). "Tail dependence of recursive max-linear models with regularly varying noise variables".

Goldie, Charles M. and Ross A. Maller (2000). "Stability of Perpetuities". *The Annals of Probability* 28.3, pp. 1195–1218.

Heidergott, Bernd (2006). *Max-Plus Linear Stochastic Systems and Pertubation Analysis*. Springer.

Koller, Daphne and Nir Friedman (2009). *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press.

Lauritzen, Steffen L. (1996). *Graphical Models*. Oxford University Press.

Meyn, Sean P. and Richard L. Tweedie (1993). *Markov Chains and Stochastic Stability*. Springer.

Resnick, Sidney I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer.

Saraf-Poor, Kian (2018). "Die Untersuchung des Maximum-Operators als algebraische Struktur". *Projektarbeit Technische Universität München*.

Spirtes, Peter, Clark Glymour, and Richard Scheines (2000). *Causation, Prediction, and Search*. The MIT Press, 2nd edition.