Notes for a course

Opinion Dynamics on Social Networks

Claudio Altafini

This version: January 10, 2023

Contents

Chapter 1

Introduction

These notes are inspired (when not copied) from various sources. In order of importance the following references are useful for the course:

A. V. Proskurnikov and R. Tempo. A Tutorial on Modeling and Analysis of Dynamic Social Networks. Part I and II, Annual Reviews in Control, 43:65-79, 2017, and 45:166-190, 2018.

F. Bullo: Lectures on Network Systems, 2020 (available online)

M. Mesbahi and M. Egerstedt. Graph theoretic methods in multiagent networks. Princeton Univ Press. 2010.

F. Fagnani and P. Frasca. Introduction to averaging dynamics over networks. Springer. 2017.

Version: this notes are very much work in progress. There are many unclear points, inconsistencies and missing links. The current version is from January 10, 2023.

Chapter 2

A preliminary overview

2.1 Problems we want to investigate

A "social network" in this course is a network of "agents" (e.g., individuals), exchanging opinions and influencing each other. For a social network, the main question we ask is what is the "collective emerging behavior", that is, what can we deduce on the opinions that are being formed by these interactions?

The ingredients that enter into our models are:

1. A graph having the agents as nodes and the interactions between pairs of individuals as edges, also called an *influence graph*, see Fig. 2.1.
2. *State variables* associated to the nodes, representing the *opinions*.
3. A method to describe how the agents influence each other, and how the opinions get modified by these influences. In practice, a *dynamical system* living on the graph and evolving according to its topology (i.e., to the presence or less of edges on the graph).

More formally:

1. A graph $\mathcal{G}$ is specified as a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of nodes $\mathcal{V} = \{1, 2, \ldots, n\}$ and $\mathcal{E}$ is a set of edges, $\mathcal{E} = \{(i, j) \in \mathcal{V} \times \mathcal{V}\}$ (if the graph is directed, "(i, j)" means the edge $i \to j$).
2. The state variables are denoted x_i, $i = 1, \ldots, n$, with either $x_i \in \mathbb{R}$ or $x_i \in \mathbb{R}^m$, meaning that we represent the opinions are real-valued quantities. The state vector is then

   $$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

   which lives in $\mathbb{R}^n$ if $x_i \in \mathbb{R}$, or in $\mathbb{R}^{nm}$ if $x_i \in \mathbb{R}^m$.
3. The dynamical system can be represented in Continuous Time (CT) or in Discrete Time (DT). For instance assuming $x \in \mathbb{R}^n$, in CT it has the form

   $$\dot{x} = f(x) \tag{2.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a (smooth) vector field of components

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix}$$

meaning that in components, (2.1) reads

$$\begin{cases} \dot{x}_1 & = f_1(x) \\ & \vdots \\ \dot{x}_n & = f_n(x) \end{cases}$$

and in each component of f_i we assume that the argument contains only the state of node i and of the nodes which are incoming neighbors of i according to $\mathcal{G}$, as in the following example.

Example 2.1 Consider the graph in Fig. 2.1. The pattern of influences is the following:

- agent 1 is influenced by agents 2, 3, and 4;
- agent 2 is influenced by agent 1;
- agent 3 is influenced by agents 2 and 4;
- agent 4 is influenced by agent 1.

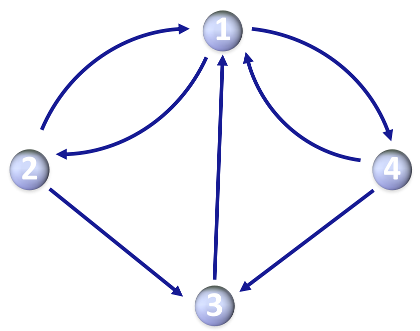


Figure 2.1: An influence graph: the agents are the nodes and the edges represent how they influence each other.

This means that we can rewrite (2.1) in components as

$$\begin{cases} \dot{x}_1 & = f_1(x_1, x_2, x_3, x_4) \\ \dot{x}_2 & = f_2(x_1, x_2) \\ \dot{x}_3 & = f_3(x_2, x_3, x_4) \\ \dot{x}_4 & = f_4(x_1, x_4) \end{cases} \tag{2.2}$$

□

Example 2.2 If the dynamics in (2.1) is linear, then for the graph in Fig. 2.1 we get

$$\dot{x} = Ax \quad \text{where} \quad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} \\ a_{41} & 0 & 0 & a_{44} \end{bmatrix}$$

Notice that the structure of A reflects that of $\mathcal{G}$: $a_{ij} \neq 0 \iff (j, i) \in \mathcal{E}$, i.e., A an be used to describe the graph of the social network (up to the diagonal elements, which we have added in A but which were not given in the graph $\mathcal{G}$). In fact we will call the matrix A the adjacency matrix associated to the graph $\mathcal{G}$ (and write $\mathcal{G}(A)$ to specify the graph whose adjacency matrix is A). □

Some basic questions that we want to answer for our opinions dynamics are the following.

- do the opinions converge somewhere?
- do they get closer to each other or do they get further apart?
- do they stabilize or keep changing (e.g. periodically, or unpredictably, etc.)?

Answering these questions means understanding the asymptotic behavior of the system (2.1). In turn, the asymptotic properties of the opinions are determined by the type of dynamics we insert in the model, and all are possible as we change $f(x)$.

2.2 Types of behavior of interest in opinion dynamics

At a more complex level, we are interested to investigate if there is any collective behavior that "emerges" from the dynamics that can have a meaningful sociological interpretation, i.e., that describes reasonably well some phenomena known to occur in social networks of individuals. Some examples of such behaviors are given in Table 2.1.

In this course we will be interested only in some of these types of behavior.

1. The agents discuss and reach an agreement on their opinions: $x(t)$ converge to a common value (i.e., a consensus).
2. The opinions at the end of a discussion are closer to each other but not identical.
3. The discussion renders the opinions more polarized.
4. Clusters of similar opinions appear out of the discussion.

We shall now give a brief preview of each of these four situations. Since each will be investigated in detail in a subsequent chapter, we do not provide here any definition nor rigorous argument.

Phenomenon	Description	Model
Consensus	The agents achieve an agreement	DeGroot
Stubborness	The agents tend to defend their opinions in a discussion	Friedkin-Johnsen
Polarization	The agents achieve two different, opposite consensus values	Bipartite consensus
Bounded confidence	The agents "trust" only other agents having similar opinions. They cluster according to the initial opinions	Hegselmann-Krause
Biased assimilation	Describes how the agents give more importance to their own opinions	Pravan-Goed-Lee
Reflected self-appraisal	Describes the evolution of the "social power" of the agents on a sequence of discussions	DeGroot-Friedkin
Multi-stage concatenated negotiations	Stubborn negotiations on a sequence of concatenated discussions, leading to consensus	Concatenated Friedkin-Johnsen
Eco chambers	Similar opinions tend to reinforce each other	?
Herding	Leader-follower or "bandwagon" behavior	?
Wisdom of crowds	Crowds tend to reject extremisms	?
⋮		

Table 2.1: Common types of opinion dynamics phenomena.

2.2.1 The agents achieve a common value: consensus

Consider the adjacency matrix $A \in \mathbb{R}^{n\times n}$ and assume that $A \geq 0$ (elementwise, i.e., $a_{ij} \geq 0$). Consider the following linear system

$$\dot{x} = -Lx \tag{2.3}$$

where L is the Laplacian matrix associated to A

$$L = \mathrm{diag}(A\mathbb{1}) - A, \quad \text{with} \quad \mathbb{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

which in components reads

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij}(x_j - x_i), \qquad i = 1, \ldots, n \tag{2.4}$$

As we will see in later on, one of the properties of the Laplacian is that its eigenvalues have all nonnegative real part, and in particular by construction

$$0 = \lambda_1(L) \leq \mathrm{Re}[\lambda_2(L)] \leq \ldots \leq \mathrm{Re}[\lambda_n(L)]$$

which implies that

$$\mathrm{Re}[\lambda_n(-L)] \leq \ldots \leq \mathrm{Re}[\lambda_2(-L)] \leq \lambda_1(-L) = 0$$

Notice that the eigenvalue $\lambda_1(L) = 0$ has eigenvector $\mathbb{1}$: $L\mathbb{1} = 0$ by construction. Under certain assumptions (connectivity of the graph $\mathcal{G}$), we have that $-L$ is "marginally stable" (i.e., stable but not asymptotically stable) which implies that in (2.3) we have

$$x(t) \overset{t\to\infty}{\longrightarrow} x^* \in \ker(-L) = \operatorname{span}(\mathbb{1})$$

or, in words, the opinions in a system like (2.3) must converge to a vector of the form $\alpha\mathbb{1}$ with α a scalar. But this asymptotic value corresponds in components to $x_i^* = x_j^*$, i.e., all opinions become equal. This is called a consensus state: all agents eventually agree on the same value.

The assumption $A \geq 0$ means that the agents influence each other in a positive way, and we will see that it implies that the agents collaborate with each other. It is a necessary condition for achieving a common goal, such as consensus.

2.2.2 The opinions get closer but do not become identical: Friedkin-Johnsen model

This problem formulation is similar to a consensus, but in addition the agents have a tendency to remain attached to their opinions, i.e., they have a certain degree of stubbornness. The Friedkin-Johnsen (FJ) model encodes this attachment by taking a convex combination between the update rule in (2.3) and the initial opinion of the agents, i.e., their initial conditions.

$$\dot{x} = -((I - \Theta)L + \Theta)x + \Theta x(0), \tag{2.5}$$

where

$$\Theta = \begin{bmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{bmatrix}, \quad \theta_i \in [0, 1]$$

is the diagonal matrix of stubbornness coefficients. The higher is the stubbornness coefficient $\theta_i \in [0,\, 1]$, the more stubborn the i-th agent is. The system (2.5) is a linear inhomogeneous system (i.e., a system which is affine in the state x). We will show that its state update matrix $-((I - \Theta)L$ is Hurwitz, i.e., all its eigenvalues have a negative real part. Therefore the matrix can be inverted, and the equilibrium point

$$x^* = ((I - \Theta)L + \Theta)^{-1}\Theta x(0)$$

is globally asymptotically stable. It is also $x^* \in \operatorname{co}(x(0))$, the convex hull of the initial conditions, meaning that in the model (2.5) indeed the opinions get closer to each other, although in x^* it is in general $x_i^* \neq x_j^*$, i.e., the asymptotic value achieved by the opinion is not a consensus.

2.2.3 The opinions become polarized: bipartite consensus

So far, we have assumed that all agents collaborate to achieve a common goal, and collaboration has been encoded as nonnegativity of the adjacency matrix A. In social networks, it is not very realistic to assume that all agents are always friendly and collaborative with each other. Depending on the context, competition and antagonism may be facts of life one has to live with. If positive edges on our graphs correspond to collaboration (or any other form of "friendly" interaction, like trust, or cooperation) among the agents, then negative edges should correspond to antagonism (or any other form of "unfriendly" interaction, like mistrust, rivalry).

Our aim is to build a model on a signed graph and to study its behavior. Given a signed adjacency matrix $A_s \in \mathbb{R}^{n\times n}$, construct the signed Laplacian as

$$L_s = \text{diag}(|A_s|\mathbb{1}) - A_s \tag{2.6}$$

where $|A_s|$ is the elementwise absolute value. Consider the equivalent of the dynamical system (2.3) on a signed graph:

$$\dot{x} = -L_s x \tag{2.7}$$

In components, (2.7) reads:

$$\dot{x}_i = \sum_{j=1}^{n} |a_{ij}|(\text{sgn}(a_{ij})x_j - x_i)$$

where $\text{sgn}(\cdot)$ is the sign function.

Let us consider only a special case, when $\mathcal{G}(A_s)$ is structurally balanced, i.e., when the graph can be partitioned into two parts, such that on each part there are only positive edges and through the cut set that splits the graph there are only negative edges (see Fig. 4.8(a)). In this case, the system (2.7) is marginally stable, and a convergence result similar to the consensus problem holds, but only for the absolute values

$$|x(t)| \overset{t\to\infty}{\longrightarrow} |x^*| = \alpha\mathbb{1} \in \text{span}(\mathbb{1})$$

while

$$x_i^* = \begin{cases} +\alpha & \text{on one side the of the partition} \\ -\alpha & \text{on the other side of the partition} \end{cases}$$

This phenomenon is called bipartite consensus, and corresponds to opinions that become completely polarized.

2.2.4 The opinions form clusters: bounded confidence models

The models we have seen so far are linear, i.e., the influences act identically on an agent regardless of how far the opinions are. However, it could be more realistic to assume that agents are sensitive to (and influenced by) opinions that are similar to their own opinions and not influenced by opinions that are far away from their own opinions. There are many possible ways to encode this principle. The one we show here is called bounded confidence, or Hegselmann-Krause model. It assumes that only opinions that are below a certain distance threshold from the opinion of an agent are used in the update law for that agent. If we assume for simplicity that this distance is 1, then the bounded confidence model has the following structure

$$\dot{x}_i = \sum_{j \text{ s.t. } |x_i - x_j|\le 1}^{n} a_{ij}(x_j - x_i) \tag{2.8}$$

Notice that the difference w.r.t. (2.4) is in the summation index: only agents j that are neighbours of i according to the graph $\mathcal{G}(A)$ and that at the same time have similar opinions (as measured by $|x_i - x_j| \le 1$) exert an influence on node i.

The model (2.8) tends to form clusters of opinions. Once opinions start to group into clusters, then their distance typically remains below the threshold 1, hence within each cluster the consensus rule of (2.4) applies and a consensus is achieved. Different clusters however will have different consensus values. The state-depended summation makes the model (2.8) nonlinear, and in fact this type of clustering behavior cannot be obtained with a linear model.

Chapter 3

Matrices

In this notes we consider only real-valued matrices $A \in \mathbb{R}^{n\times n}$, of elements a_{ij}, $i, j = 1, \dots, n$. A is symmetric if it coincides with its transpose: $A = A^\top$.

Positivity/Nonnegativity. $A > 0$ means elementwise positive, i.e., $a_{ij} > 0$ for all $i, j = 1, \dots, n$, while $A \geq 0$ means element-wise nonnegative, i.e., $a_{ij} \geq 0$ for any $i, j = 1, \dots, n$. In principle, $A \geq 0$ is valid also when $A = 0$. When we wan to specify $A \geq 0$ *and* $A \neq 0$, we will use the symbol $A \gneq 0$. Such A is sometimes called semipositive. This notation is used also for vectors.

Positive (semi)definite. A matrix A is called positive semidefinite (psd) if $x^\top Ax = x^\top \frac{A+A^\top}{2} x \geq 0 \;\forall x \in \mathbb{R}^n$ and it is called positive definite (pd) if $x^\top Ax = x^\top \frac{A+A^\top}{2} x > 0 \;\forall x \in \mathbb{R}^n \setminus \{0\}$. A compact notation for pd (psd) is $A \succ 0$ (resp. $A \succeq 0$). Analogous definitions hold for negative (semi)definitness.

Spectral radius. The spectrum of A is denoted $\text{spec}(A) = \{\lambda_1(A), \dots, \lambda_n(A)\}$, where $\lambda_i(A)$, $i = 1, \dots, n$, are the eigenvalues of A, and the vector space generated by its columns is $\text{span}(A)$ (or $\text{range}(A)$). The *spectral radius* of A is the maximum norm of its eigenvalues

$$\rho(A) = \max_i \{|\lambda_i(A)| \text{ s. .t. } \lambda_i(A) \in \text{spec}(A)\},$$

and corresponds to the radius of the least disk containing all eigenvalues of A. The *spectral abscissa* of A is the largest real part of its eigenvalues

$$\mu(A) = \max_i \{\text{Re}[\lambda_i(A)] \text{ s. .t. } \lambda_i(A) \in \text{spec}(A)\},$$

and delimits to the right (in the vertical direction) the half plane in $\mathbb{C}$ containing all eigenvalues of A.

Irreducibility. Recall that a permutation matrix Π is just reordering of the rows (or of the columns) of the identity (i.e., each row and column of Π contains one and only one entry equal to 1). A matrix A is said *reducible* if there exists a permutation matrix Π such that

$$\Pi A \Pi^\top = \tilde{A} = \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}$$

is block triangular, with $\tilde{A}_{11}$ and $\tilde{A}_{22}$ nontrivial square matrices (i.e., of dimension ≥ 1). A matrix is *irreducible* if it is not reducible. An equivalent characterization of irreducibility is that $(I + |A|)^{n-1} > 0$, where $|A|$ is the matrix whose entries are the absolute values of the entries of A (see [12], Thm. 6.2.24). Consequently, another equivalent condition is that

$$I + |A| + \ldots + |A|^{n-1} > 0. \tag{3.1}$$

A reducible matrix can always be put into triagular form by a suitable permutation:

$$\Pi A \Pi^\top = \tilde{A} = \begin{bmatrix} \tilde{A}_{11} & 0 & \ldots & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} & & \\ \vdots & & \ddots & \\ \tilde{A}_{\ell 1} & \ldots & & \tilde{A}_{\ell\ell} \end{bmatrix} \tag{3.2}$$

In (3.2), each block $\tilde{A}_{ii}$ is irreducible or has dimension 1 (and possibly is equal to 0).

Example 3.1 The matrix

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ a_{21} & 0 & 0 & 0 \\ 0 & a_{32} & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.3}$$

is reducible, since the permutation matrix Π that reorders the nodes as $1 \to 2$, $2 \to 3$, $3 \to 4$ and $4 \to 1$, i.e.,

$$\Pi = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

leads to

$$\tilde{A} = \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline \tilde{a}_{21} & 0 & \tilde{a}_{23} & \tilde{a}_{24} \\ 0 & \tilde{a}_{32} & 0 & 0 \\ \tilde{a}_{41} & 0 & \tilde{a}_{43} & 0 \end{array}\right] = \left[\begin{array}{c|c} \tilde{A}_{11} & 0 \\ \hline \tilde{A}_{21} & \tilde{A}_{22} \end{array}\right] = \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline a_{14} & 0 & a_{12} & a_{13} \\ 0 & a_{21} & 0 & 0 \\ a_{34} & 0 & a_{32} & 0 \end{array}\right]$$

which has the upper right corner block of zeros, hence it is reducible. $\tilde{A}$ is in Frobenius normal form, with $\tilde{A}_{11} = 0$ and $\tilde{A}_{22}$ block which is 3×3 irreducible. □

Neumann series. Given $A \in \mathbb{R}^{n \times n}$, the *Neumann series* of A is the infinite sum $I + A + A^2 + \ldots = \sum_{t=0}^{\infty} A^t$.

Proposition 3.2 *Given $A \in \mathbb{R}^{n \times n}$, the following are equivalent.*

1. *The Neumann series of A converges.*

2. *$\rho(A) < 1$.*

3. *$\lim_{t \to \infty} A^t = 0$.*

When these conditions hold then $(I - A)^{-1}$ *exists and it is*

$$(I - A)^{-1} = \sum_{t=0}^{\infty} A^t$$

Proof. See [20], page 618.

A generalization of Proposition 3.3 is the following, (see [20], page 630).

Proposition 3.3 *Given* $A \in \mathbb{R}^{n\times n}$*, the following are equivalent.*

1. $\lim_{t\to\infty} A^t$ *exists.*
2. $\rho(A) < 1$*, or else* $\rho(A) = 1$ *is a strictly dominating eigenvalue of multiplicity 1 (i.e., any other eigenvalue of* A*,* $\lambda \neq \rho(A)$*, is such that* $|\lambda| < 1$*).*

Stochastic matrix. A nonnegative matrix $A \geq 0$ is said *row-stochastic* if $A\mathbb{1} = \mathbb{1}$. In words, the rows of A sum to one $\sum_{j=1}^{n} a_{ij} = 1$. This means that for all entries of A we have $a_{ij} \in [0, 1]$. A nonnegative matrix is *column stochastic* if $\mathbb{1}^\top A = \mathbb{1}^\top$ and doubly stochastic when it is both row and column stochastic.

Z-matrices, M-matrices, H-matrices, and comparison matrices. When A is such that $a_{ij} \lneq 0 \ \forall \ i \neq j$, it is said a *Z-matrix*. It is said a (nonsingular) *M-matrix* if it is a Z-matrix and the real part of its eigenvalues is positive. When the real part of its eigenvalues is instead nonnegative, then A is said a *singular M-matrix*. A Z-matrix A is often written $A = \alpha I - B$, with $B \geq 0$ and $\alpha > 0$. The Z-matrix A is then an M-matrix if in addition $\alpha \geq \rho(B)$. It is a nonsinglular M-matrix if $\alpha > \rho(B)$.

For any $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, its *comparison matrix* is given by $A^{\rm cmp} = [a_{ij}^{\rm cmp}] \in \mathbb{R}^{n\times n}$ where

$$a_{ij}^{\rm cmp} = \begin{cases} |a_{ij}| & \text{if } i = j \\ -|a_{ij}| & \text{if } i \neq j. \end{cases}$$

Matrices A_1 and A_2 such that $A_1^{\rm cmp} = A_2^{\rm cmp}$ are sometimes called *equimodular*. A matrix A whose comparison matrix is a (singular or nonsingular) M-matrix is said an *H-matrix*. It is said an H_+-matrix if in addition $a_{ii} \geq 0$, for all $i = 1, \ldots, n$. As we will see below, for singular comparison matrices, the H-matrices can be singular or nonsingular.

Mezler matrices. A matrix $A \in \mathbb{R}^{n\times n}$ is said *Metzler* if $a_{ij} \geq 0$ for all $i \neq j$. The diagonal elements a_{ii} can be anything, but obviously we are interested in the case in which at least one oft hem is negative (otherwise A is a nonnegative matrix). In other words, a Metzler matrix is a negated Z-matrix.

3.1 Perron-Frobenius theory for nonnegative matrices

The theorems given in this and next section can be found in standard textbooks such as [12, 20]. No proof is provided here.

For positive matrices we have the following theorem (dating back to 1907).

Theorem 3.4 (Perron theorem for positive matrices) *Consider $A > 0$. Then:*

1. $\rho(A) > 0$;
2. $\rho(A)$ *is an eigenvalue of* A;
3. $\rho(A)$ *is algebraically simple as an eigenvalue of* A *(i.e., it has multiplicity 1);*
4. *the right (resp. left) eigenvector* v *(resp.* w*) corresponding to* $\rho(A)$ *is positive:* $v > 0$ *(resp.* $w > 0$*);*
5. *for every other eigenvalue* $\lambda(A) \in \text{spec}(A)$, $\lambda(A) \neq \rho(A)$, *it is* $|\lambda(A)| < \rho(A)$;
6. $\lim_{k\to\infty} \frac{A^k}{\rho^k(A)} = \frac{vw^\top}{w^\top v}$;
7. *no other nonnegative eigenvector can exist.*

The proof can be found in [12], Thm 8.2.8. Meaning: $A > 0$ has a simple, real, positive, strictly dominating eigenvalue equal to $\rho(A)$. In other words, we can localize the position of the dominant eigenvalue, see Fig. 3.1, and, since $v > 0$ $w > 0$, also of the dominant eigenvector. The powers of A tend to fall into the "spectral projector space" corresponding to this dominant

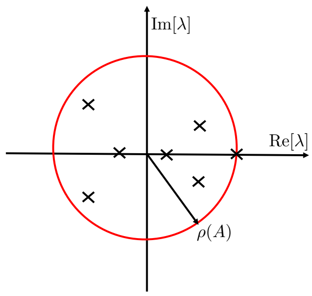


Figure 3.1: Spectrum of a matrix obeying Perron-Frobenius theorem. The matrix is either positive, or nonnegative and primitive.

eigenvalue. In condition 6, the scalar at the denominator, $w^\top v$, is a normalization factor, while the rank-1 matrix in numerator, $vw^\top$, corresponds to the spectral projector onto the subspace determined by $\rho(A)$. Obviously uniqueness of a positive eigenvector v (or w) is up to a scalar multiplication: cv (or cw) is also an eigenvector for all $c \in \mathbb{R}$.

The assumption $A > 0$ is quite restrictive for large matrices (it corresponds to full connectivity of the corresponding graph). It is more realistic to assume that $A \geq 0$. For nonnegative matrices we can have several versions of the Perron-Frobenius theorem, depending on the assumption we make on A. We see three of them, with progressively stronger assumptions.

Theorem 3.5 (Perron-Frobenius for nonnegative matrices, weakest version) *Consider* $A \geq 0$. *Then:*

1. $\rho(A)$ *is an eigenvalue of* A *(not necessarily* > 0*);*
2. *the right (resp. left) eigenvector* v *(resp.* w*) corresponding to* $\rho(A)$ *is nonnegative:* $v \geq 0$ *(resp.* $w > 0$*).*

What is not guaranteed by this theorem is that $\rho(A) > 0$ (the matrix could be nilpotent, hence all eigenvalues could be 0), and neither are multiplicity, strict dominance and positivity of the eigenvector.

Example 3.6 The nonnegative matrix

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

has eigenvalues $\lambda_1(A) = \lambda_2(A) = 0$, i.e., $\rho(A) = 0$ of multiplicity 2. □

Adding the assumption of irreducibility we have:

Theorem 3.7 (Perron-Frobenius for nonegative matrices, intermediate version) *Consider $A \geq 0$ irreducible. Then:*

1. $\rho(A) > 0$*;*
2. $\rho(A)$ *is an eigenvalue of* A*;*
3. $\rho(A)$ *is algebraically simple as an eigenvalue of* A *(i.e., it has multiplicity 1);*
4. *the right (resp. left) eigenvector* v *(resp.* w*) corresponding to* $\rho(A)$ *is positive:* $v > 0$ *(resp.* $w > 0$*);*
5. *no other nonnegative eigenvector can exist.*

Proof: [12], Thm 8.4.4. What is not guaranteed by this version is strict dominance of $\rho(A)$: there could be other eigenvalues of A having the same modulus, i.e., lying on the circle of radius $\rho(A)$. Hence it cannot be guaranteed that the powers of A converge to the subspace of the spectral projector corresponding to $\rho(A)$, i.e., to the rank-1 matrix $vw^\top$. Actually, convergence is not guaranteed at all.

Example 3.8 The irreducible matrix

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

has eigenvalues $\lambda_{1,2}(A) = \pm 1$ of right eigenvectors $v_1 = [1\ 1]^\top$ and $v_2 = [1\ -1]^\top$. Since it is $A^2 = I$, $A^3 = A$ and so on, $\lim_{k\to\infty} \frac{A^k}{\rho^k(A)}$ is not converging. □

In order to add convergence to the Perron-Frobenius theorem we need to restrict to A which is primitive.

Definition 3.9 *An irreducible matrix $A \geq 0$ is said* primitive *if there exists a $k \in \mathbb{N}$ such that A^k is positive.*

In words, after a certain power a nonnegative matrix which is primitive becomes positive, hence all properties related to power series must be similar to those of a positive matrix.

Theorem 3.10 (Perron-Frobenius for nonnegative matrices, strongest version) *Consider $A \geq 0$ primitive. Then:*

1. $\rho(A) > 0$;
2. $\rho(A) \in \text{spec}(A)$;
3. *algebraic multiplicity of* $\rho(A)$ *is 1;*
4. *the right (resp. left) eigenvector* v *(resp.* w*) corresponding to* $\rho(A)$ *is positive:* $v > 0$ *(resp.* $x > 0$*);*
5. *for every other* $\lambda(A) \in \text{spec}(A)$*,* $\lambda(A) \neq \rho(A)$*, it is* $|\lambda(A)| < \rho(A)$*;*
6. $\lim_{k\to\infty} \frac{A^k}{\rho^k(A)} = \frac{vw^\top}{w^\top v}$;
7. *no other nonnegative eigenvector can exist.*

Nonnegative primitive matrices have the same properties as positive matrices: they have only one dominant eigenvalue, and it is real, positive, simple and strictly dominating.

A sufficient condition for primitivity of $A \geq 0$ irreducible is that A has a positive diagonal element $a_{ii} > 0$. Another is the existence of a triplet of arcs $(i, j), (i, k), (j, k)$. The reason is that a matrix which is irreducible but not primitive (called *imprimitive*) can be seen as the adjacency matrix of a graph having cycles of length which is multiple of an integer $r > 1$, and the paths given by the powers of A "preserve" this cyclicity.

Proposition 3.11 (Primitivity) *Consider* $A \geq 0$ *irreducible.* A *is primitive if and only if the greatest common divisor of the length of all cycles of the graph* $\mathcal{G}(A)$ *is* $r = 1$.

Such greatest common divisor is called the cyclicity index of A. An irreducible A with cyclicity index 1 is then primitive (it is sometimes called aperiodic).

Proposition 3.12 (Imprimitivity and dominant eigenvalues) *Consider* $A \geq 0$ *irreducible, imprimitive of cyclicity index* r*. Then the* r *eigenvalues of* A *on the spectral circle are the* r*-th roots of unity, and must have multiplicity 1.*

In other words, $\lambda_1, \ldots, \lambda_r$ such that $|\lambda_i| = \rho(A)$ are the solutions of the equation $\lambda^r = \rho(A)$, and can be expressed as $\{\rho(A), \rho(A)\omega, \ldots, \rho(A)\omega^{r-1}\}$ where $\omega = e^{\frac{2\pi i}{r}}$.

Example 3.13 The matrix A of Example 3.8 is imprimitive with cyclicity index 2. □

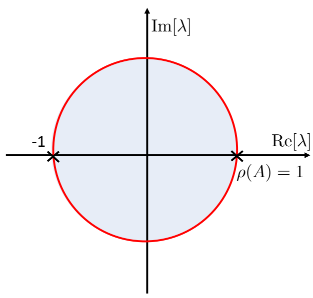
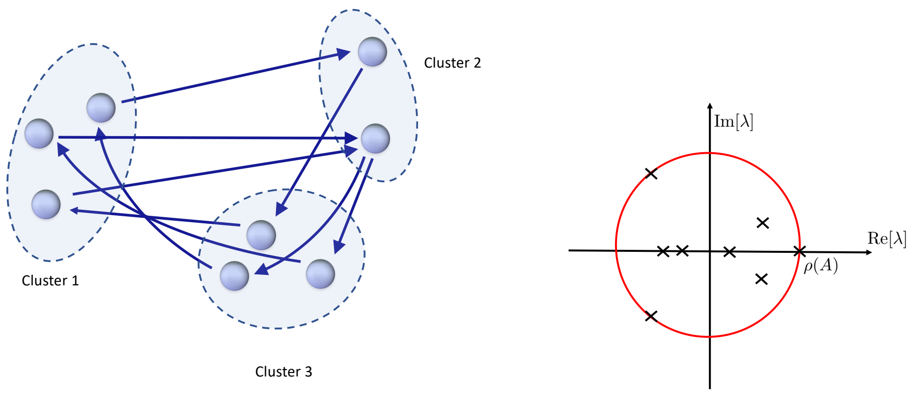


Figure 3.2: Spetrum of the imprimitive matrix of Example 3.8.

Example 3.14 The matrix

$$A = \begin{bmatrix} 0 & 0 & A_{13} \\ A_{21} & 0 & 0 \\ 0 & A_{32} & 0 \end{bmatrix} \geq 0$$

where A_{ij} are block matrices, is irreducible but has all cycles of length 3 or multiples of 3, hence it is not primitive and has cyclicity index 3, see Fig. 3.3. In this case A has 3 eigenvalues on the spectral circle of radius $\rho(A)$. □



(a) (b)

Figure 3.3: Graph and spectrum of the imprimitive matrix of Example 3.14.

3.2 Perron-Frobenius theorem for Metzler matrices

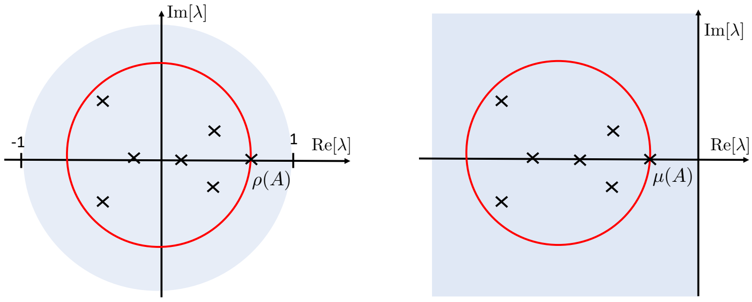
The class of Metzler matrices enjoys properties very similar to those of nonnegative matrices. We have already seem that a matrix A is Metzler if $a_{ij} \geq 0$ for all $i \neq j$. Its diagonal elements a_{ii} can be anything, but if they are all nonnegative, then A is nonnegative and the problem become not particularly interesting, so we typically consider the case in which some a_{ii} are negative, often all of them. In fact, a Metzler matrix A can be written as $A = B - \alpha I$ with $B \geq 0$ and $\alpha \in \mathbb{R}_+$. In this case we typically consider α sufficiently big ($\alpha \geq -\min_i(a_{ii})$). The eigenvalues of A and B differ only by $-\alpha$

$$\lambda_i(B) \quad \longrightarrow \quad \lambda_i(A) = \lambda_i(B) - \alpha$$

see Fig. 3.4. Hence we can use properties of nonnegative matrices to study the eigenvalues also for Metzler matrices. In particular, we know from Perron-Frobenius that if $B \geq 0$ then B has an eigenvalue equal to $\rho(B)$, therefore A must have an eigenvalue equal to the spectral abscissa, see Fig. 3.4(b). We have then proven the following proposition:

Proposition 3.15 *A Metzler matrix always has an eigenvalue which is real and equal to the spectral abscissa* $\mu(A) = \rho(B) - \alpha$.

When we will investigate continuous-time (CT) linear systems with state update matrix equal to A, then the dominant eigenvalue that matters is the one with biggest real part.



(a) eigenvalues of B (b) eigenvalues of $A = B - \alpha I$

Figure 3.4: Eigenvalues of Metzler matrices.

Proposition 3.16 *A Metzler matrix A has always a single dominant eigenvalue equal to $\mu(A)$.*

If we do not specify anything more on the matrix A, then the multiplicity of $\mu(A)$ can be larger than 1. However, it is equal to 1 when A is irreducible. Note further that for CT linear systems there is no need of primitivity assumptions, only irreducibility matters.

Theorem 3.17 (Perron-Frobenius for Metzler matrices) *Consider A Metzler and irreducible. Then*

1. *$\mu(A)$ is a real eigenvalue of A;*
2. *the algebraic multiplicity of $\mu(A)$ is equal to 1;*
3. *right (resp. left) eigenvector of $\mu(A)$ is positive: $v > 0$ (resp. $w > 0$);*
4. *$\lambda \in \mathrm{spec}(A)$, $\lambda \neq \mu(A)$ implies $\mathrm{Re}(\lambda) < \mu(A)$;*
5. *$\lim_{t\to\infty} \frac{e^{At}}{e^{\mu(A)t}} = \frac{vw^\top}{w^\top v}$;*
6. *A has no other nonnegative eigenvector.*

Chapter 4

Graphs

A graph $\mathcal{G}$ is a pair $(\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, 2, \ldots, n\}$ are the nodes, and $\mathcal{E} = \{(i,j) \in \mathcal{V} \times \mathcal{V}\}$ is a set of edges connecting the nodes. Associated to $\mathcal{G}$ is an adjacency matrix $A \in \mathbb{R}^{n \times n}$ having nonzero entries in correspondence of the edges: $a_{ij} \neq 0 \iff (j,i) \in \mathcal{E}$. a_{ij} is called the weight associated to the edge (i,j). The graph is then denoted $\mathcal{G}(A)$. Note that $\mathcal{G}(A)$ is undirected when A is symmetric, and directed (i.e., a digraph) otherwise. Note also that we use the "dynamical systems" convention for the adjacency matrix: a_{ij} corresponds to the edge $(j,i) \in \mathcal{E}$, or, more explicitly, $j \to i$. The convention used in the graph theory literature is obtained by taking the transpose of A, see Fig. 4.1. For undirected graphs, of course, $a_{ij} = a_{ji}$

(a) our convention (b) graph theory convention

Figure 4.1: Edges and adjacency matrix.

i.e., $(i,j) \in \mathcal{E} \iff (j,i) \in \mathcal{E}$. Diagonal elements a_{ii} correspond to *self-loops* i.e., to edges $(i,i) \in \mathcal{E}$.

Degree

For an undirected graph, the degree of a node i is the number of its neighbors. Denoting $\mathcal{N}_i = \{j \in \mathcal{V} \text{ s. t. } (j,i) \in \mathcal{E}\}$ the neighborhood of i, then the degree is $\deg(i) = \text{card}(\mathcal{N}_i)$. For digraphs, one distinguishes between in-degree and out-degree:

$$\deg^{\text{in}}(i) = \text{card}(\mathcal{N}_i^{\text{in}}) \quad \text{where} \quad \mathcal{N}_i^{\text{in}} = \{j \in \mathcal{V} \text{ s. t. } (j,i) \in \mathcal{E}\} = \text{ in-neighborhood.}$$
$$\deg^{\text{out}}(i) = \text{card}(\mathcal{N}_i^{\text{out}}) \quad \text{where} \quad \mathcal{N}_i^{\text{out}} = \{j \in \mathcal{V} \text{ s. t. } (i,j) \in \mathcal{E}\} = \text{ out-neighborhood.}$$

A node with zero in-degree is called a *root* while one with zero out-degree is called a *leaf* of $\mathcal{G}(A)$. Note that for any graph

$$\text{card}(\mathcal{E}) = \sum_i \deg^{\text{in}}(i) = \sum_i \deg^{\text{out}}(i)$$

(a) $\deg^{\text{in}}(i) = 3$ (b) $\deg^{\text{out}}(i) = 4$

Figure 4.2: Node in- and out-degree.

One can also consider the *weighted in/out degree*, easily computed in terms of the adjacency matrix A:

$$\text{wdeg}^{\text{in}}(i) = \sum_{j=1}^{n} a_{ij}$$
$$\text{wdeg}^{\text{out}}(i) = \sum_{j=1}^{n} a_{ji}$$

Again, in most of the graph theory literature, the roles of rows and columns are exchanged (i.e., A should be replaced with $A^\top$). Obviously the weighted in-degree of a root is zero, and so is the weighted out-degree of a leaf.

The graph $\mathcal{G}(A^\top)$ obtained by flipping the direction of all edges is called the *reverse graph* of $\mathcal{G}(A)$. Roots of $\mathcal{G}(A)$ becomes leaves of $\mathcal{G}(A^\top)$ and viceversa.

Connectedness

For an undirected graph $\mathcal{G}$, a *path* between nodes i and j ($i,\ j \in \mathcal{V}$) is a sequence of distinct nodes $k_1, \ldots, k_\ell \in \mathcal{V}$ such that all consecutive pairs are edges of $\mathcal{G}$: $(j, k_1), (k_1, k_2), \ldots, (k_\ell, i) \in \mathcal{E}$.

Definition 4.1 (Connected) *An undirected graph is* connected *if there exists an undirected path between every pair of nodes.*

For a digraph $\mathcal{G}$ a *(directed) path* between i and j is a sequence of distinct nodes $k_1, \ldots, k_\ell \in \mathcal{V}$ such that all consecutive pairs are directed edges of $\mathcal{G}$: $(j, k_1), (k_1, k_2), \ldots, (k_\ell, i) \in \mathcal{E}$ (the directed path is $j \to k_1 \to \ldots \to k_\ell \to i$). The weight of the path is defined as the product of the edge weights:

$$\mathcal{P}_{ji} = a_{ik_\ell} \cdot \ldots \cdot a_{k_2 k_1} \cdot a_{k_1 j} \tag{4.1}$$

Definition 4.2 (Strongly connected) *A digraph is* strongly connected *if there exists a directed path between every pair of nodes. It is said* weakly connected *if it is connected when we disregard the direction of its edges.*

A (directed or undirected) path that begins and ends at the same node is called a *cycle*. The cycles we consider in these notes are always simple, i.e., they traverse a node at most once. Notice that for digraphs also the cycles are intended as directed, i.e., the paths follow the direction of the arrows. A graph is said *acyclic* if it contains no cycle. The weight of a cycle can be denifed analogously to (4.1).

The meaning of (strong) connectivity is that there exists a (directed) path between any two nodes of the graph. Hence the graph must have cycles. Connectivity can be fully characterized by looking at the adjacency matrix A.

Example 4.3 For the adjacency matrix (3.3) of Example 3.1, the graph $\mathcal{G}(A)$ is shown in Fig. 4.3. Notice that in the digraph $\mathcal{G}(A)$ it is impossible to reach node 4 from any other node,

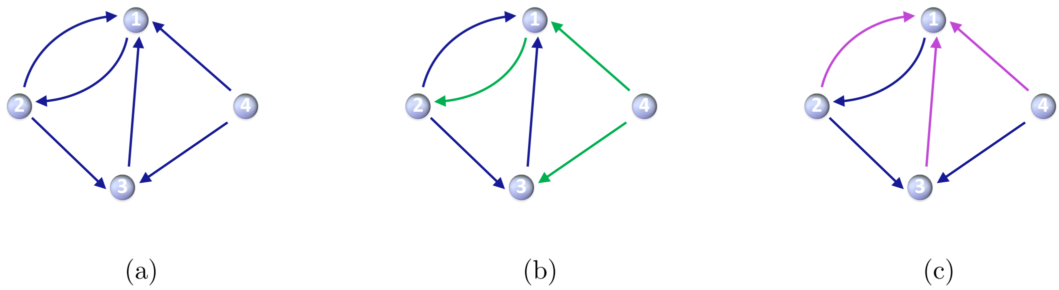


Figure 4.3: Example 4.3: reducible digraph with a directed spanning tree. A possible directed spanning tree is shown in green in (b), while a possible terminal spanning tree is shown in violet in (c).

meaning that $\mathcal{G}(A)$ cannot be strongly connected. It has however a directed spanning tree rooted at node 4, see Fig. 4.3(b). □

This is not a coincidence, as shown in the next proposition.

Proposition 4.4 (Strong connectivity and irreducibility) *A digraph $\mathcal{G}(A)$ is strongly connected if and only if A is irreducible.*

An analogous result holds also for undirected graphs.

Spanning trees and forests

For undirected graphs, a *tree* is a connected graph without cycles. A tree is minimally connected: between each pair of nodes there is a single path, but if you remove a single edge the graph becomes disconnected. It is also maximally acyclic: if you add an edge then the graph has at least a cycle. Given a connected undirected graph $\mathcal{G}(A)$, a *spanning tree* of $\mathcal{G}(A)$ is a subgraph of $\mathcal{G}(A)$ which is a tree and includes all vertices of $\mathcal{V}$.

For a digraph $\mathcal{G}(A)$, one can define a *directed tree* as a subgraph of $\mathcal{G}(A)$ which is a tree and in which the direction of the edges follows that of $\mathcal{G}(A)$, i.e., a minimally connected directed acyclic subgraph of $\mathcal{G}(A)$. In particular, a *directed spanning tree* is a directed tree of $\mathcal{G}(A)$ which is spanning (i.e., it touches all vertices of $\mathcal{V}$). If in the directed spanning tree there exists a distinguished node (a root) which has a directed path to any other node, then it is called a *rooted directed spanning tree*. If instead there is a node (a leaf) which is globally reachable through directed paths from any other node, then it is called a *terminal directed spanning tree*. Since roots become leaves of the reverse graph $\mathcal{G}(A^\top)$ and viceversa, each concept we define on a rooted (or terminal) subgraph of $\mathcal{G}(A)$ admits an equivalent concept defined for a terminal (or rooted) subgraph of $\mathcal{G}(A^\top)$. The rooted/terminal directed spanning tree just introduced is one of them: an acyclic subgraph of $\mathcal{G}(A)$ is a terminal directed tree if and only if its reverse subgraph on $\mathcal{G}(A^\top)$ is a rooted directed tree. A directed spanning tree can also be neither rooted nor terminal, see Fig. 4.4, but such objects are not considered in these notes. In

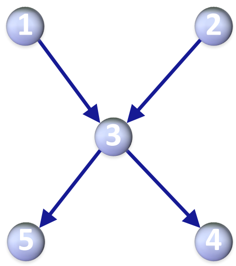
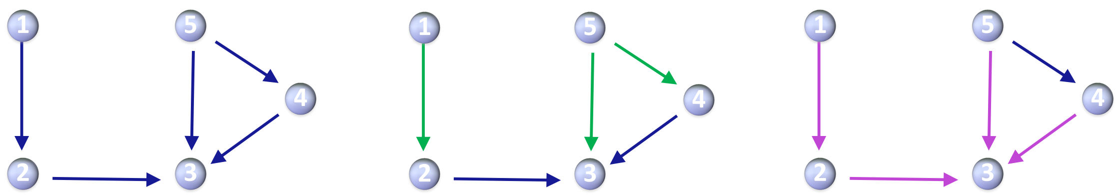


Figure 4.4: Reducible digraph with a directed spanning tree which is neither rooted nor terminal, but with a directed spanning forest rooted at nodes 1 and 2, and terminating at nodes 4 and 5.

the following, when dealing with directed spanning trees, we always specify which of the two categories, rooted or terminal, we are referring to.

While strongly connected graphs (obviously) always have directed spanning trees (both rooted and terminal), graphs that are not strongly connected (i.e., digraphs $\mathcal{G}(A)$ for which A is reducible or, for short, reducible digraphs) may or may not have a directed spanning tree. In particular, a weakly connected digraph always has a directed spanning tree, but this may be neither rooted not terminal, see Fig. 4.4. Obviously, disconnected digraphs cannot have any directed spanning tree. However, reducible digraphs always have a *directed spanning forest*, i.e., a collection of directed spanning trees rooted at different nodes and terminating at different nodes, see examples in Figs. 4.4 and 4.5.



(a) (b) (c)

Figure 4.5: Reducible digraph without a rooted directed spanning tree, but with a directed spanning forest (shown in green in (b)) rooted at nodes 1 and 5 (and terminating at nodes 2, 3, 4). The graph has a terminal directed spanning tree terminating at node 3 (shown in violet in (c)).

Notice that sometimes non strongly connected digraphs may have strongly connected components that serve as roots/leaves. We call these *rooted/terminal strongly connected components*. For instance, in Fig. 4.6(a) the strongly connected component formed by nodes 1-2-3 is rooted, i.e., none of the nodes has incoming edges originating outside the strongly connected component itself. Hence this graph has a directed spanning tree rooted at any node of the strongly connected component. Similarly, nodes 5-6-7 form a terminal strongly connected component. In Fig. 4.6(b) there are instead two rooted strongly connected components, and the graph admits a directed spanning forest rooted at any node of the corresponding strongly connected component (one for each strongly connected component, of course). Notice that the strongly connected component formed by the nodes 5-6-7 in Fig. 4.6(b) is neither rooted nor terminal. A terminal (resp. rooted) directed spanning tree of $\mathcal{G}(A)$ is a rooted (resp. terminal)

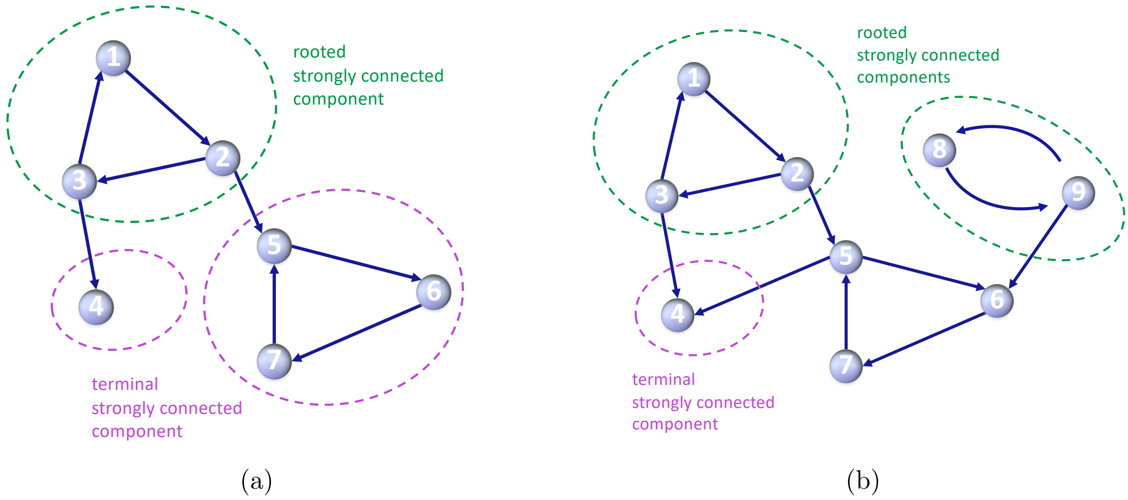


Figure 4.6: (a): A graph with a rooted strongly connected component and two terminal strongly connected components; (b): A graph with two rooted strongly connected components and one terminal strongly connected component.

directed spanning tree of $\mathcal{G}(A^\top)$.

Graphs with inputs and outputs

Assume that some of the nodes of $\mathcal{G}(A) = (\mathcal{V}, \mathcal{E})$ are interfacing the graph with the external environment. For digraphs, such interface can be two-way, meaning that we can consider input nodes $\mathcal{V}^{\text{in}} \subseteq \mathcal{V}$ and output nodes $\mathcal{V}^{\text{out}} \subseteq \mathcal{V}$. A digraph with input nodes $\mathcal{V}^{\text{in}} \subseteq \mathcal{V}$ and output nodes $\mathcal{V}^{\text{out}} \subseteq \mathcal{V}$ is said *input connected* if each node in $\mathcal{V}$ can be reached through a directed path from a node in $\mathcal{V}^{\text{in}}$. It is said *output connected* if there is a directed path from any node in $\mathcal{V}$ to a node in $\mathcal{V}^{\text{out}}$.

Proposition 4.5 (Graph input and output connectivity) *Consider a digraph $\mathcal{G}(A) = (\mathcal{V}, \mathcal{E})$ with input nodes $\mathcal{V}^{\text{in}} \subseteq \mathcal{V}$ and output nodes $\mathcal{V}^{\text{out}} \subseteq \mathcal{V}$. The following conditions are equivalent*

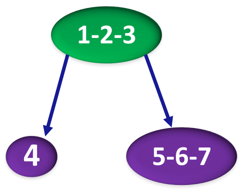
1. *$\mathcal{G}(A)$ is input connected from $\mathcal{V}^{\text{in}}$.*
2. *$\mathcal{G}(A)$ admits a directed spanning forest rooted at (a subset of) $\mathcal{V}^{\text{in}}$.*

Also the following conditions are equivalent.

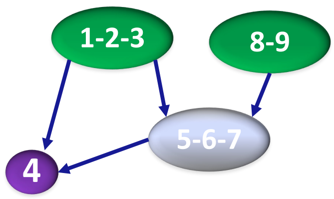
1. *$\mathcal{G}(A)$ is output connected to $\mathcal{V}^{\text{out}}$.*
2. *$\mathcal{G}(A)$ admits a terminal directed spanning forest terminating at (a subset of) $\mathcal{V}^{\text{out}}$.*
3. *The reverse graph $\mathcal{G}(A^\top)$ is input connected from $\mathcal{V}^{\text{out}}$.*
4. *The reverse graph $\mathcal{G}(A^\top)$ admits a directed spanning forest rooted at (a subset of) $\mathcal{V}^{\text{out}}$.*

Condensation graph

The *condensation digraph* of $\mathcal{G}(A)$, call it $\mathcal{G}^{\text{cond}}(A) = \mathcal{G}(A^{\text{cond}})$, is obtained by replacing each strongly connected component of $\mathcal{G}(A)$ with a single node, and by placing a directed edge from node i to node j of $\mathcal{G}^{\text{cond}}(A)$ if there is at least one directed edge from any node of the ith strongly connected component to any node of the jth strongly connected component of $\mathcal{G}(A)$. By construction, $\mathcal{G}^{\text{cond}}(A)$ is a directed acyclic graph. On Fig. 4.7 the condensation graphs of the examples in Fig. 4.6 are shown. In the condensation graph associated to $\mathcal{G}(A)$, rooted (resp. terminal) strongly connected components become roots (resp. leaves) of $\mathcal{G}^{\text{cond}}(A)$. Different rules can be used to compute the adjacency matrix A^{cond} of $\mathcal{G}^{\text{cond}}(A)$ from A.



(a)



(b)

Figure 4.7: Condensation graphs for the two graphs of Fig. 4.6. Rooted and terminal strongly connected components are in green and violet respectively.

4.1 Graphical properties of nonnegative/Metzler reducible matrices

So far the entries of the adjacency matrix A have been left unspecified with respect to their sign: $a_{ij} \neq 0 \iff (j, i) \in \mathcal{E}$. The case by far more common is however when edges have nonnegative weights: $a_{ij} > 0 \iff (j, i) \in \mathcal{E}$. In this case, in fact, one can express paths as powers of A. For $A \geq 0$, the kth power of A corresponds to the sum of all possible length-k paths between pairs of nodes, weighted by the edge weights along the path. For instance,

$$\left[A^2\right]_{ij} = a_{i1}a_{1j} + \ldots + a_{in}a_{nj}$$

and $[A^2]_{ij} > 0$ if and only if there is at least a path of length 2 connecting j to i, at least as long as $a_{ii} = 0$ for all i. In later chapters we are interested in graphs that are associated to positive systems, which correspond indeed to adjacency matrices $A \geq 0$ but also to adjacency matrices that are Metzler, for which $a_{ij} \geq 0$ when $i \neq j$, but a_{ii} may also be negative. It is therefore convenient to consider the graph obtained from $\mathcal{G}(A)$ disregarding the self-loops, or, equivalently, the graph obtained from A after dropping the diagonal terms ("nd" = no diagonal):

$$\mathcal{G}^{\text{nd}}(A) = \mathcal{G}(A^{\text{nd}}) \quad \text{where} \quad A^{\text{nd}} = A - \text{diag}(a_{11}, \ldots, a_{nn})$$

where $\text{diag}(a_{11}, \ldots, a_{nn})$ is the diagonal matrix having $a_{11}, \ldots, a_{nn}$ on the diagonal. Once we have eliminated potentially negative self-loops, the properties presented below are valid also for Metzler matrices.

The first property is that the existence of a directed spanning tree in $\mathcal{G}(A)$ can be formulated in terms of powers of A.

Proposition 4.6 (Spanning tree and positive columns) *Consider the digraph $\mathcal{G}(A)$, $A^{\rm nd} \geq 0$.*

1. *$\mathcal{G}(A)$ has a directed spanning tree rooted at $j \in \mathcal{V}$ if and only if the jth column of $I + A^{\rm nd} + \ldots + (A^{\rm nd})^{n-1}$ is positive.*
2. *$\mathcal{G}(A)$ has a directed spanning tree terminating at $j \in \mathcal{V}$ if and only if the jth row of $I + A^{\rm nd} + \ldots + (A^{\rm nd})^{n-1}$ is positive.*

Proof. Existence of a directed spanning tree rooted at j means existence of a path from j to any $i \in \mathcal{V}$, of length varying between 1 and $n-1$. If the length is k, from (4.1), this corresponds to a nonzero entry $\left[(A^{\rm nd})^k\right]_{ij}$. This must be true for each i. Since $A^{\rm nd} \geq 0$, all such paths are present also in $\sum_{k=0}^{n-1}(A^{\rm nd})^k$, hence this matrix must have all nonzero entries in the jth column. The opposite implication follows from similar considerations. As terminal directed spanning trees are rooted directed spanning trees for $\mathcal{G}(A^\top)$ the second condition follows from the first. □

Notice that the self-loops of $\mathcal{G}(A)$ (not considered in $A^{\rm nd}$) do not play any role in forming the paths (they only add extra paths to those produced by $A^{\rm nd}$). The presence of the identity matrix in the summation $\sum_{k=0}^{n-1}(A^{\rm nd})^k$ takes care of placing a nonzero entry on the diagonal. Notice further that if A (and hence $A^{\rm nd}$) is irreducible, it follows from (3.1) that $\sum_{k=0}^{n-1}(A^{\rm nd})^k > 0$. i.e., all columns and rows are positive.

An argument similar to Proposition 4.6 is valid also for spanning forests, even though now rather than a positive column we have positive columns in different submatrices of A, one for each tree of the forest.

Proposition 4.7 (Spanning forest and positive columns) *Consider the digraph $\mathcal{G}(A)$, $A^{\rm nd} \geq 0$.*

1. *$\mathcal{G}(A)$ has a directed spanning forest rooted at $\{v_1^{\rm in}, \ldots, v_{k_{\rm in}}^{\rm in}\} \subseteq \mathcal{V}^{\rm in}$ if and only if there exists a partition of $\mathcal{V}$ into $\{v_1^{\rm in}, v_{1,1}, \ldots, v_{1,j_1}\}, \ldots, \{v_{k_{\rm in}}^{\rm in}, v_{k_{\rm in},1}, \ldots, v_{k_{\rm in},j_{k_{\rm in}}}\}$ such that if $A_1^{\rm in} = A^{\rm nd}\left[\{v_1^{\rm in}, v_{1,1}, \ldots v_{1,j_1}\}\right]$, …, $A_{k_{\rm in}}^{\rm in} = A^{\rm nd}\left[\{v_{k_{\rm in}}^{\rm in}, v_{k_{\rm in},1}, \ldots, v_{k_{\rm in},j_{k_{\rm in}}}\}\right]$ are the square submatrices of $A^{\rm nd}$ obtained selecting the rows/columns $\{v_1^{\rm in}, v_{1,1}, \ldots v_{1,j_1}\}$, …, $\{v_{k_{\rm in}}^{\rm in}, v_{k_{\rm in},1}, \ldots, v_{k_{\rm in},j_{k_{\rm in}}}\}$, then $\sum_{i=0}^{j_r-1}\left(A_r^{\rm in}\right)^i$, $r = 1, \ldots, k_{\rm in}$, all have at least a positive column (the first).*
2. *$\mathcal{G}(A)$ has a directed spanning forest terminating at $\{v_1^{\rm out}, \ldots, v_{k_{\rm out}}^{\rm out}\} \subseteq \mathcal{V}^{\rm out}$ if and only if $\mathcal{G}(A^\top)$ has a directed spanning forest rooted at $\{v_1^{\rm out}, \ldots, v_{k_{\rm out}}^{\rm out}\} \subseteq \mathcal{V}^{\rm out}$.*

Proposition 4.7 gives us an algebraic way to check input and output connectivity of a graph.

Example 4.8 For the two reducible graphs of Fig. 4.3(a) and Fig. 4.5(a), consider the following adjacency matrices:

$$A_1 = \begin{bmatrix} 0 & 0.7 & 0.5 & 1.2 \\ 0.3 & 0 & 0 & 0 \\ 0 & 1.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0.6 & 1.1 \\ 0 & 0 & 0 & 0 & 0.8 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\mathcal{G}(A_1)$ has a directed spanning tree rooted at node 4, and $I + A_1 + A_1^2$ already has the fourth column which is positive (there are spanning trees in which all nodes have a maximal distance to the root equal to 2, like the one shown in green in Fig. 4.3(b), hence it is not necessary to go to the power $n-1 = 3$ in this case). Instead, $\mathcal{G}(A_2)$ has no rooted directed spanning tree, but it has a directed spanning forest rooted at nodes 1 and 5. In this case, splitting rows/columns of A_2 into $\{1, 2\}$ and $\{3, 4, 5\}$, the two submatrices $A[\{1,2\}]$ and $A[\{3,4,5\}]$ each get a positive column already at power 1 (i.e, in $I + A_2$), in correspondence of the spanning forest shown in green in Fig. 4.5(b). Both graphs have terminal spanning trees, terminating at nodes 1 and 3 respectively, see Fig. 4.3(c) and Fig. 4.5(c). □

As Fig. 4.6 shows, graphs that are not strongly connected may have strongly connected subcomponents. To investigate their structure in fine grade, it is useful to resort to a canonical form called the *Frobenius normal form for reducible matrices.* We are going to use two different versions of the Frobenius normal form, one emphasizing the rooted strongly connected components, and the other the terminal strongly connected components. The first, which we denote A^{root}, is based on reordering the diagonal blocks of the triangular form (3.2). If $\mathcal{G}(A)$ has $\ell > 1$ strongly connected components (possibly also of dimension 1), then A in the triangular form (3.2) has ℓ irreducible square blocks on the diagonal (possibly equal to 0 when they have dimension 1). Some of the strongly connected components will be rooted, say $k < \ell$, while $\ell - k$ others will be not. Let us apply a permutation matrix to bring these k rooted strongly connected components to the first k blocks. Then the "rooted" version of the Frobenius normal form has the following pattern:

$$A^{\text{root}} = \left[\begin{array}{cccc|cccc} A_{11} & 0 & \dots & 0 & 0 & & \dots & 0 \\ 0 & A_{22} & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & & \ddots & & & & & \\ 0 & \dots & 0 & A_{kk} & 0 & & \dots & 0 \\ \hline A_{k+1,1} & \dots & & A_{k+1,k} & A_{k+1,k+1} & 0 & \dots & 0 \\ \vdots & & & \vdots & A_{k+2,k+1} & A_{k+2,k+2} & \ddots & \vdots \\ & & & & \vdots & \ddots & \ddots & 0 \\ A_{\ell 1} & \dots & & A_{\ell k} & A_{\ell k+1} & \dots & A_{\ell,\ell-1} & A_{\ell\ell} \end{array}\right]. \tag{4.2}$$

The empty rows in the first k blocks (diagonal blocks excluded) capture the fact that rooted strongly connected components have no incoming edge (other than from nodes in the same strongly connected component which end up in the diagonal block). Notice that the diagonal blocks A_{ii}^{root} can also be equal to 0 when they have dimension 1. In the associated condensation graph $\mathcal{G}^{\text{cond}}(A) = \mathcal{G}(A^{\text{cond}})$, each block A_{ij}^{root} becomes a single entry of A^{cond}.

When instead of emphasizing the rooted strongly connected components it is more useful to highlight the terminal strongly connected components, then using a different permutation matrix it is convenient to consider an alternative ("terminal") Frobenius normal form:

$$A^{\text{term}} = \left[\begin{array}{cccc|cccc} A_{11} & 0 & \dots & 0 & A_{1,k+1} & & \dots & A_{1,\ell} \\ 0 & A_{22} & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & \ddots & \ddots & 0 & & & & \\ 0 & \dots & 0 & A_{kk} & A_{k+1,k} & & \dots & A_{k+1,\ell} \\ \hline 0 & \dots & & 0 & A_{k+1,k+1} & A_{k+1,k+2} & \dots & A_{k+1,\ell} \\ \vdots & & & & 0 & A_{k+2,k+2} & \ddots & \vdots \\ \vdots & & & \vdots & \vdots & \ddots & \ddots & A_{\ell-1,\ell} \\ 0 & \dots & & 0 & 0 & \dots & 0 & A_{\ell\ell} \end{array}\right] \tag{4.3}$$

which highlights that, of the ℓ strongly connected components of $\mathcal{G}(A^{\text{term}})$, the first k are terminal (there are no edges leaving the first k strongly connected components) while the remaining $\ell - k$ are not. As the examples show, the index k in (4.2) and (4.3) need not be the same.

$\mathcal{G}(A)$ reducible		
Frobenius normal form	A^{root}, Eq. (4.2)	A^{term}, Eq. (4.3)
spanning tree/forest	rooted	terminal
CT: diagonal dominance/equipotency	by row	by columns
condition for CT stability	$A\mathbb{1} \leq 0$	$\mathbb{1}^\top A \leq 0$
DT: (sub)stochasticity	by row	by columns
condition for DT stability	$A\mathbb{1} \leq \mathbb{1}$	$\mathbb{1}^\top A \leq \mathbb{1}^\top$
Nodes with affine term	$\mathcal{V}^{\text{out}}$	$\mathcal{V}^{\text{in}}$
Nodes with strict diagonal dominance / substochasticity	$\mathcal{V}^{\text{in}}$	$\mathcal{V}^{\text{out}}$
Applications	consensus FJ model	Markov chains compartmental systems

Table 4.1: Properties of reducible matrices.

Example 4.9 For Fig. 4.6(a), grouping the nodes according to the strongly connected components $1 = \{1, 2, 3\}$, $2 = \{4\}$ and $3 = \{5, 6, 7\}$, we have the "rooted" Frobenius normal form

$$A_1^{\text{root}} = \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & 0 & 0 \\ A_{31} & 0 & A_{33} \end{bmatrix}$$

where $A_{22} = 0$ corresponds to the node $\{4\}$. Since $A_{21} \neq 0$, only the first component is rooted. If instead we reverse the order of the components we obtain (using identity blocks of suitable dimension) the "terminal" Frobenius normal form

$$A_1^{\text{term}} = \begin{bmatrix} 0 & 0 & I_3 \\ 0 & 1 & 0 \\ I_3 & 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & 0 & 0 \\ A_{31} & 0 & A_{33} \end{bmatrix} \begin{bmatrix} 0 & 0 & I_3 \\ 0 & 1 & 0 \\ I_3 & 0 & 0 \end{bmatrix} = \begin{bmatrix} A_{33} & 0 & A_{31} \\ 0 & 0 & A_{21} \\ 0 & 0 & A_{11} \end{bmatrix} \tag{4.4}$$

meaning that A_{33} and $A_{22} = 0$ are the terminal strongly connected components of the terminal spanning forest for this graph. For the graph in Fig. 4.6(b), to get the structure (4.2), we must order the strongly connected components in the order $1 = \{1, 2, 3\}$, $2 = \{8, 9\}$, $3 = \{5, 6, 7\}$ and $4 = \{4\}$, getting

$$A_2^{\text{root}} = \begin{bmatrix} A_{11} & 0 & 0 & 0 \\ 0 & A_{22} & 0 & 0 \\ A_{31} & A_{32} & A_{33} & 0 \\ A_{41} & 0 & A_{43} & A_{44} \end{bmatrix}$$

The first two components are rooted, as can be seen by the fact that $A_{21} = 0$. Also in this case, reversing the order of the strongly connected components gives

$$A_2^{\text{term}} = \begin{bmatrix} A_{44} & A_{43} & 0 & A_{41} \\ 0 & A_{33} & A_{32} & A_{31} \\ 0 & 0 & A_{22} & 0 \\ 0 & 0 & 0 & A_{11} \end{bmatrix}$$

from which we see that only A_{44} (i.e., node 4 in Fig. 4.6(b)) is a terminal strongly connected component for this graph.

□

4.2 Signed graphs, Structural balance.

So far we have considered only graphs with nonnegative edge weights, i.e., associated to nonnegative adjacency matrices: $A \geq 0$. It is however also possible to consider graphs with negative edge weights, which we call *signed graphs*. For signed graphs the adjacency matrix A_s has one or more negative entries $a_{s,ij} < 0$ for some $(i, j) \in \mathcal{E}$. Consequently, also the weight of the paths and cycles of $\mathcal{G}(A_s)$ can be positive or negative. In particular we say that a cycle $i \to k_1 \to \ldots \to k_\ell \to i$ is positive if its weight $\mathcal{P}_{ii} = a_{s,ik_\ell} \cdot \ldots \cdot a_{s,k_1 i} > 0$, i.e., if it contains an even number of negative edges (including the case of zero negative edges). It is said negative if instead it contains an odd number of negative edges.

The following property of signed graph will be used. Recall that $\mathcal{G}^{\text{nd}}(A_s)$ is the graph $\mathcal{G}(A_s)$ without self-loops.

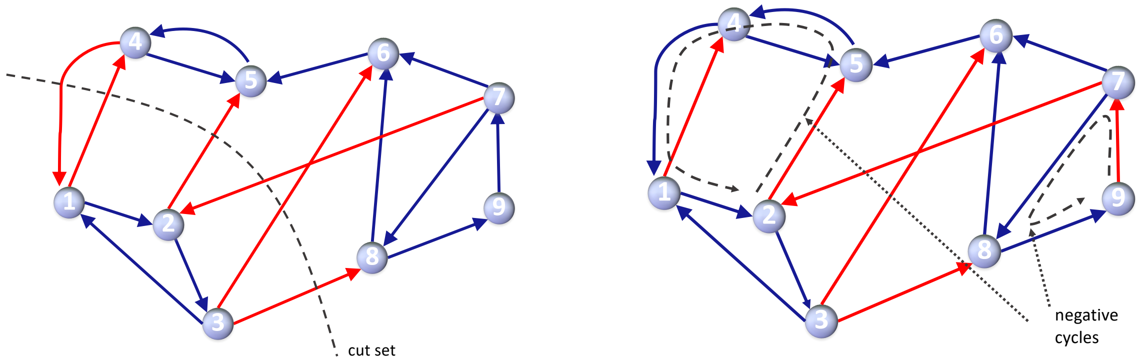
Definition 4.10 (Structural balance) *A signed graph $\mathcal{G}^{\text{nd}}(A_s)$ is said* structurally balanced *if $\exists$ partition of the nodes $\mathcal{V}_1$, $\mathcal{V}_2$, with $\mathcal{V}_1 \cup \mathcal{V}_2 = \mathcal{V}$, $\mathcal{V}_1 \cap \mathcal{V}_2 = 0$ such that*

- $a_{s,ij} \geq 0 \ \forall \, i, \, j \in \mathcal{V}_q$, $q = 1, 2$;
- $a_{s,ij} \leq 0 \ \forall \, i \in \mathcal{V}_q$, $j \in \mathcal{V}_r$, $q \neq r$.

It is said structurally unbalanced *otherwise.*

The meaning of this definition is shown in Fig. 4.8 (a): $\mathcal{G}^{\text{nd}}(A_s)$ structurally balanced means that the graph can be partitioned into two groups of nodes such that on each side of the partition you only have positive edges, while through the partition itself you have all and only negative edges. The following Proposition gives conditions that are equivalent to structural balance.

Proposition 4.11 (Characterization of structural balance) *A signed graph $\mathcal{G}^{\text{nd}}(A_s)$ is structurally balanced if and only if any of the following equivalent conditions holds:*



(a) structurally balanced (b) structurally unbalanced

Figure 4.8: Signed graphs (blue edge: positive; red edge: negative).

1. *all cycles of* $\mathcal{G}^{\rm nd}(A_s)$ *are positive;*
2. $\exists$ *a diagonal signature matrix* $S = \mathrm{diag}(s_1 \ldots s_n)$, $s_i = \pm 1$, *such that* $SA_s^{\rm nd}S$ *is nonnegative;*

Example 4.12 Consider the signed graphs of Fig. 4.8, whose associated signed adjacency matrices have sign pattern

$$A_1 = \begin{bmatrix} 0 & 0 & + & - & 0 & 0 & 0 & 0 & 0 \\ + & 0 & 0 & 0 & 0 & 0 & - & 0 & 0 \\ 0 & + & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ - & 0 & 0 & 0 & + & 0 & 0 & 0 & 0 \\ 0 & - & 0 & + & 0 & + & 0 & 0 & 0 \\ 0 & 0 & - & 0 & 0 & 0 & + & + & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & + \\ 0 & 0 & - & 0 & 0 & 0 & + & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & + & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0 & 0 & + & + & 0 & 0 & 0 & 0 & 0 \\ + & 0 & 0 & 0 & 0 & 0 & - & 0 & 0 \\ 0 & + & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ - & 0 & 0 & 0 & + & 0 & 0 & 0 & 0 \\ 0 & - & 0 & + & 0 & + & 0 & 0 & 0 \\ 0 & 0 & - & 0 & 0 & 0 & + & + & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & - \\ 0 & 0 & - & 0 & 0 & 0 & + & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & + & 0 \end{bmatrix}$$

The signed graph $\mathcal{G}(A_1)$ ($= \mathcal{G}^{\rm nd}(A_1)$ in this example) of Fig. 4.8(a) is structurally balanced and all its cycles are positive. Notice that the diagonal matrix S mentioned in Proposition 4.11 is given by $S = \mathrm{diag}(-1, -1, -1, 1, 1, 1, 1, 1, 1)$, i.e., $SA_1S \geq 0$. The signs of S also identify the partition into the two factions $\mathcal{V}_1$ and $\mathcal{V}_2$. Graphically, such transformation "flips" the sign of all edges through the cut set determining the partition. In terms of adjacency matrices, A_1 can be decomposed according to the $\mathcal{V}_1$ and $\mathcal{V}_2$ splitting as

$$A_1 = \begin{bmatrix} A_{1,11} & A_{1,12} \\ A_{1,21} & A_{1,22} \end{bmatrix}$$

with $A_{1,11} \geq 0$, $A_{1,22} \geq 0$ and $A_{1,12} \leq 0$, $A_{1,21} \leq 0$. The transformation SA_1S renders these off-diagonal blocks nonnegative. The graph $\mathcal{G}(A_2)$ of Fig. 4.8(b) is unbalanced: it has some negative cycles.

Chapter 5

Stability of linear and nonlinear systems: a primer

5.1 Basic stability definitions

A general nonlinear system of continuous-time (CT) ordinary differential equations (ODEs) is written as

$$\dot{x} = f(x) \tag{5.1}$$

where $x \in \mathbb{R}^n$ is the state vector and $f : \mathbb{R}^n \to \mathbb{R}^n$ is a vector field, which for simplicity we assume smooth enough so that the solution exists unique for each initial condition $x(0)$.

An analogous system can be defined in discrete-time (DT). For simplicity, we assume that the time interval between to consecutive steps is fixed and equal to 1, so that the time axis is described simply by an index $k = 1, 2, \ldots$, and the discrete difference equation (DDE) can be recursively written in terms of a vector field $\phi : \mathbb{R}^n \to \mathbb{R}^n$:

$$x(k+1) = \phi(x(k)) \tag{5.2}$$

Example 5.1 (Euler discretization) If we approximate the differentiation operation with an Euler difference: $\dot{x} \simeq \frac{x((k+1)\Delta t) - x(k\Delta t)}{\Delta t}$, where Δt is the step size of the discretization, then

$$x((k+1)\Delta t) = x(k\Delta t) + \Delta t f(x(k\Delta t))$$

Omitting the Δt in the argument of $x(\cdot)$, in (5.2) we get $\phi(x(k)) = x(k) + \Delta t f(x(k))$. □

Consider an equilibrium point of (5.1) or (5.2): x^* such that $f(x^*) = 0$ for (5.1) or $x^* = \phi(x^*)$ for (5.2). Let us give first standard definitions of stability for a CT/DT system like (5.1) or (5.2).

Definition 5.2 (Stability) *An equilibrium point x^* of a system like (5.1) or (5.2) is said*

1. marginally stable *(or stable in the sense of Lyapunov) if for all finite initial conditions $x(0)$ the solution $x(t)$ stays bounded for all $t > 0$.*
2. asymptotically stable *if for all finite $x(0)$ the solution $x(t)$ stays bounded for all $t > 0$ and $\lim_{k\to\infty} x(k) = x^*$.*
3. unstable *otherwise.*

5.2 Stability of linear systems

When the vector field in (5.1) or (5.2) is linear (i.e., $f(x) = Ax$ or $\phi(x) = Ax$), then we obtain a linear system, represented in continuous time as

$$\dot{x} = Ax \tag{5.3}$$

and in discrete time as

$$x(k+1) = Ax(k) \tag{5.4}$$

5.2.1 Basic stability characterizations based on eigenvalues

For both the systems (5.3) and (5.4), necessary and sufficient conditions for marginal and asymptotic stability can be obtained by looking at $\mathrm{spec}(A)$. In both cases, it is enough to look at $x^* = 0$ (which is always an equilibrium).

CT linear system

Using the Definitions 5.2 for stability we have the following characterization for (5.3):

Theorem 5.3 (Eigenvalue characterization of stability; CT system) *Consider the equilibrium point $x^* = 0$ of the system (5.3).*

1. *x^* is marginally stable if and only if all $\lambda_i(A) \in \mathrm{spec}(A)$ are s.t. $\mathrm{Re}[\lambda_i(A)] \leq 0$ and $\lambda_i(A)$ s.t. $\mathrm{Re}[\lambda_i(A)] = 0$ have algebraic multiplicity 1 (i.e., the corresponding Jordan blocks have all dimension 1).*

2. *x^* is asymptotically stable if and only if all $\lambda_i(A) \in \mathrm{spec}(A)$ are s.t. $\mathrm{Re}[\lambda_i(A)] < 0$.*

3. *x^* is unstable otherwise.*

A matrix A with all eigenvalues $\lambda_i(A)$ such that $\mathrm{Re}[\lambda_i(A)] < 0$ is called *Hurwitz stable*.

DT linear system

For the DT linear system (5.4) we have the following characterization, based on the eigenvalues of the state matrix.

Theorem 5.4 (Eigenvalue characterization of stability; DT system) *Consider the equilibrium point $x^* = 0$ of the system (5.4).*

1. *x^* is marginally stable if and only if all $\lambda_i(A) \in \mathrm{spec}(A)$ are s.t. $|\lambda_i(A)| \leq 1$ and $\lambda_i(A)$ s.t. $|\lambda_i(A)| = 1$ have algebraic multiplicity 1 (i.e., the corresponding Jordan blocks have all dimension 1).*

2. *x^* is asymptotically stable if and only if all $\lambda_i(A) \in \mathrm{spec}(A)$ are s.t. $|\lambda_i(A)| < 1$.*

3. *x^* is unstable otherwise.*

A matrix A with all eigenvalues $\lambda_i(A)$ s.t. $|\lambda_i(A)| < 1$ is called *Schur stable*.

5.2.2 Other stability characterizations

CT linear system

An equivalent characterization to A Hurwitz can be obtained via the Lyapunov equation.

Theorem 5.5 (Stability via Lyapunov equation; CT system) *A is Hurwitz if and only if for each Q symmetric positive definite there exists P symmetric positive definite that solves the Lyapunov equation:*

$$A^\top P + PA = -Q \tag{5.5}$$

The meaning of Theorem 5.5 is that we can define a so-called Lyapunov function, i.e., a positive definite function whose derivative computed along the trajectories of the system is negative definite. In particular, choosing the quadratic function $V(x) = x^\top Px$ with P symmetric and positive definite, then $V(x) > 0$ for all $x \neq 0$, $V(0) = 0$. Differentiating V along the trajectories of (5.3) and using (5.5), $\dot{V}(x) = x^\top(A^\top P+PA)x = -x^\top Qx < 0$ for all $x \neq 0$ and $\dot{V}(0) = 0$. The meaning is that the trajectories of (5.3) move towards progressively decreasing level surfaces of $V(x)$, and asymptotically must reach $x^* = 0$. This guarantees the asymptotic stability of the system.

A stronger form of stability is obtained in correspondence of P which is a positive definite *diagonal* matrix.

Definition 5.6 (Diagonal stability) *A matrix $A \in \mathbb{R}^{n\times n}$ is said* diagonally stable *if there exists a positive definite diagonal matrix $D = \mathrm{diag}(d)$, $d \in \mathbb{R}^n$, $d > 0$ such that $A^\top D + DA$ is negative definite.*

As a special case, when $D = I$ then we have that A itself is negative definite (i.e., the eigenvalues of $A + A^\top$ have negative real part).

When instead a Hurwitz matrix remains Hurwitz for all possible *rescalings*, i.e., multiplication of its row (or columns) by positive factors then the matrix is said D-stable, which is another form of stability stronger than Hurwitz.

Definition 5.7 (D-stability) *A matrix $A \in \mathbb{R}^{n\times n}$ is said D-stable if the matrix DA is Hurwitz for all positive definite diagonal matrices $D = \mathrm{diag}(d)$, $d \in \mathbb{R}^n$, $d > 0$.*

The property of D-stability is a form of robustness: when you perturb A to DA (which corresponds to rescaling $\dot{x}_i$ to $\dot{x}_i/d_i$) the stability character is unchanged.

These different stability conditions are related to each other in the following way:

$$\begin{array}{r} \text{negative definitness of } A \\ \Downarrow \\ \text{diagonal stability of } A \\ \Downarrow \\ \text{D-stability of } A \\ \Downarrow \\ A \text{ Hurwitz} \end{array} \tag{5.6}$$

Example 5.8 The matrix

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

is Hurwitz and D-stable, but cannot be diagonally stable because of the 0 entry in the (1, 1) slot, see Example 2.1.6 of [14] for the details. □

A special case: normal matrices

A matrix A is said *normal* if it commutes with its transpose: $AA^\top = A^\top A$. For a normal matrix all stability characterizations in (5.6) are equivalent, see [5, 14] for more details.

Lemma 5.9 (Stability characterization for normal matrices) *For a normal matrix A, negative definitness, diagonal stability, and D-stability are all equivalent to Hurwitz stability.*

Proof. If A is normal, then there exists an orthonormal matrix U such that $A = UHU^\top$, where, if $\mu_1, \ldots, \mu_\ell$ are the real eigenvalues of A and $\nu_1 \pm i\omega_1, \ldots, \nu_{\frac{n-\ell}{2}} \pm i\omega_{\frac{n-\ell}{2}}$ are its complex conjugate eigenvalues:

$$H = \begin{bmatrix} \mu_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \mu_\ell & & & & & \\ & & & \nu_1 & \omega_1 & & & \\ & & & -\omega_1 & \nu_1 & & & \\ & & & & & \ddots & & \\ & & & & & & \nu_{\frac{n-\ell}{2}} & \omega_{\frac{n-\ell}{2}} \\ & & & & & & -\omega_{\frac{n-\ell}{2}} & \nu_{\frac{n-\ell}{2}} \end{bmatrix}$$

If follows that $A^{\rm sym} = \frac{1}{2}(A + A^\top) = \frac{1}{2}U(H + H^\top)U^\top$ and therefore that $\mathrm{Re}[\lambda_i(A)] = \lambda_i(A^{\rm sym})$. Hurwitz stability then corresponds to negative definitness of $A^{\rm sym}$, meaning that all implications in (5.6) can be reversed. □

Geršgorin theorem and diagonal dominance

Geršgorin theorem gives a way to localize the eigenvalues of a matrix A inside the so-called Geršgorin disks, i.e., disks in the complex plane centered at the diagonal elements of A and of radius given by the reduced row (or columns) absolute sums of A.

Theorem 5.10 (Geršgorin theorem) *The eigenvalues of A are located in the union of the n disks*

$$\left\{ z \in \mathbb{C} \ s.t. \ |z - a_{ii}| \leqslant \sum_{j \neq i} |a_{ij}| \right\}. \tag{5.7}$$

Proof: [12], Thm 6.1.1. Since A is real, all disks are centered in the real axis.

When a diagonal element a_{ii} is in absolute value larger than the associated row sum $\sum_{j\neq i} |a_{ij}|$, then the corresponding Geršgorin disk does not touch the imaginary axis. Such a situation is captured by the notion of diagonal dominance.

Definition 5.11 (Diagonal dominance) *A matrix A is* diagonally dominant by rows *if*

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}| \quad i = 1, \dots, n \tag{5.8}$$

It is said strictly diagonally dominant by rows *if all inequalities in* (5.8) *are strict*

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}| \quad i = 1, \dots, n$$

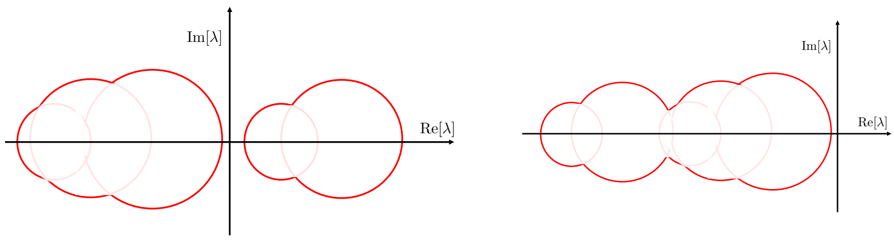
and weakly diagonally dominant by rows *if* (5.8) *holds and at least one inequality is strict. If instead none of the inequalities in* (5.8) *is strict, i.e.,*

$$|a_{ii}| = \sum_{j \neq i} |a_{ji}| \quad i = 1, \dots, n \tag{5.9}$$

then A is called diagonally equipotent by rows.

Analogous definitions hold for diagonal dominance by columns, if for instance we replace (5.8) with

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ji}| \quad i = 1, \dots, n \tag{5.10}$$



(a) A invertible (b) A Hurwitz

Figure 5.1: Examples of Geršgorin disks for a matrix A.

If we have strict diagonal dominance, i.e., if no Geršgorin disk is intersecting (not even at the boundary) the imaginary axis, then the matrix A is nonsingular, see Fig. 5.1(a). If, in addition, all disks lie to the left of the imaginary axis, then the matrix A is Hurwitz, see Fig. 5.1(b). The case of weak diagonal dominance, in which some of the disks can touch at the boundary (but not cross through) the imaginary axis, can also be treated if in addition we have irreducibility of A (see Theorem 5.12).

The case of diagonal equipotency (5.9) corresponds to all disks touching the imaginary axis. This is the only case in which a boundary point on the Geršgorin region (5.7) may be an eigenvalue. However, while the comparison matrix of A, A^{cmp}, is singular by construction, the same may or may not be true for A. When in addition the disks are on the left of the imaginary axis, then if A is irreducible it may be Hurwitz or marginally stable, depending on the sign pattern of A. When irreducibility is missing, then A diagonally equipotent may even be unstable (think of a nilpotent matrix, for instance).

Given A, consider $\mathcal{G}^{\text{nd}}(A)$, the subgraph of $\mathcal{G}(A)$ obtained disregarding the self-loops.

Theorem 5.12 (Stability based on diagonal dominance) *Consider $A \in \mathbb{R}^{n\times n}$.*

1. *If A is strictly diagonally dominant, or irreducible and weakly diagonally dominant, then A is nonsingular. If in addition $a_{ii} < 0$ for all $i = 1, \dots, n$, then A is also Hurwitz.*
2. *Consider A irreducible and diagonally equipotent, with $a_{ii} < 0$ for all $i = 1, \dots, n$.*
   (a) A is Hurwitz if and only if $\mathcal{G}^{\mathrm{nd}}(A)$ is not structurally balanced.
   (b) A is marginally stable if and only if $\mathcal{G}^{\mathrm{nd}}(A)$ is structurally balanced.

Proof. Part 1 combines results known as Levy-Desplanques theorem and Taussky theorem, and can be found in standard books, like [12] Thm 6.1.10, Cor. 6.2.9 or Cor. 6.2.27. Part 2 is from [1], Theorem 5. Here we treat the case of diagonal dominance by rows, diagonal dominance by columns being completely equivalent. From Geršgorin theorem, we know that A is at least marginally stable. Since both conditions of Part 2 are necessary and sufficient, they are mutually exclusive, hence showing one of them is enough. Here we choose Hurwitz stability. Consider $\mathcal{G}^{\mathrm{nd}}(A)$ which is not structurally balanced and assume, by contradiction, that A is not Hurwitz stable, i.e., that 0 is an eigenvalue of A of eigenvector $x \neq 0$. Lemma 6.2.3 of [12] implies that for the components of x, $|x_i| = |x_j| = \xi \neq 0 \; \forall i, j, = 1, \dots, n$, and that all Geršgorin disks pass through 0. We can then write $Ax = 0$, $x^\top A^\top = 0$ and

$$x^\top A^{\mathrm{sym}} x = 0, \tag{5.11}$$

where A^{sym} denotes the symmetric part of A. Consider the reduced row sums $R_i = \sum_{j\neq i} |a_{ij}|$, $i = 1, \dots, n$ and the reduced column sums $C_i = \sum_{j\neq i} |a_{ji}|$, $i = 1, \dots, n$. Since $a_{ii} < 0$, A diagonally equipotent implies $-a_{ii} = R_i$, $i = 1, \dots, n$ (while in general $|a_{ii}| \neq C_i$). Denote $R = \mathrm{diag}(R_1, \dots, R_n)$ and $C = \mathrm{diag}(C_1, \dots, C_n)$ and let A_d and A_o be respectively the diagonal and off-diagonal parts of A. Then $A = A_d + A_o = -R + A_o$ and $A^{\mathrm{sym}} = -R + \frac{A_o + A_o^\top}{2}$. If we rewrite A^{sym} in the following form $A^{\mathrm{sym}} = -\frac{R-C}{2} - \frac{R+C}{2} + \frac{A_o + A_o^\top}{2}$, then (5.11) becomes

$$0 = -x^\top \left(\frac{R-C}{2}\right) x - x^\top \left(\frac{R+C}{2}\right) x + x^\top \left(\frac{A_o + A_o^\top}{2}\right) x. \tag{5.12}$$

Expanding the first term of (5.12) and using the fact that $x_i^2 = \xi^2 \; \forall i = 1, \dots, n$

$$\sum_{i=1}^{n} x_i^2 (R_i - C_i) = \sum_{i=1}^{n} x_i^2 \left(\sum_{j\neq i} (|a_{ij}| - |a_{ji}|) \right) = \xi^2 \sum_{\substack{i,j=1\\ j\neq i}}^{n} (|a_{ij}| - |a_{ji}|) = 0.$$

The rest of (5.12) can be written as

$$0 = \frac{1}{2} \sum_{i=1}^{n} \sum_{j>i} \left(-x_i^2 (|a_{ij}| + |a_{ji}|) - x_j^2 (|a_{ij}| + |a_{ji}|) + 2 x_i x_j (a_{ij} + a_{ji}) \right).$$

In the case $a_{ij} a_{ji} \geq 0 \; \forall i, j = 1, \dots, n$ (i.e., all length-2 cycles of $\mathcal{G}(A)$ are positive), then

$|a_{ij}| + |a_{ji}| = |a_{ij} + a_{ji}|$, which implies

$$\begin{aligned} 0 &= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j>i}\left(x_i^2(|a_{ij}+a_{ji}|) + x_j^2(|a_{ij}+a_{ji}|) - 2x_i x_j(a_{ij}+a_{ji})\right) \\ &= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j>i}|a_{ij}+a_{ji}|\left(x_i^2 + x_j^2 - 2x_i x_j \operatorname{sgn}(a_{ij}+a_{ji})\right) \\ &= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j>i}|a_{ij}+a_{ji}|\left(x_i - \operatorname{sgn}(a_{ij}+a_{ji})x_j\right)^2. \end{aligned} \tag{5.13}$$

This has the form of a sum of squares, implying that each term of the summation must be equal to 0 in correspondence of the eigenvector x. On the graph $\mathcal{G}(A)$, consider a negative cycle $\mathcal{C} = \{(i_1, i_2), (i_2, i_3), \ldots, (i_p, i_1)\} \subset \mathcal{E}$. Denote $\mathcal{S}_+ = \{(j,i) \in \mathcal{C} \text{ s. t. } a_{ij} > 0\}$ its subset of positive edges and $\mathcal{S}_- = \{(j,i) \in \mathcal{C} \text{ s. t. } a_{ij} < 0\}$ that of negative edges. Since $\operatorname{sgn}(a_{ij} + a_{ji}) = \operatorname{sgn}(a_{ij})$ in (5.13) we have for the subset of edges in $\mathcal{C}$

$$0 = -\frac{1}{2}\sum_{(j,i)\in\mathcal{S}_+}|a_{ij}+a_{ji}|\left(x_i - \operatorname{sgn}(a_{ij}+a_{ji})x_j\right)^2 - \frac{1}{2}\sum_{(j,i)\in\mathcal{S}_-}|a_{ij}+a_{ji}|\left(x_i - \operatorname{sgn}(a_{ij}+a_{ji})x_j\right)^2$$

which implies $\begin{cases} x_i = x_j & (j,i) \in \mathcal{S}_+ \\ x_i = -x_j & (j,i) \in \mathcal{S}_-. \end{cases}$

We claim that this system has no nonzero solution of the type $|x_i| = \xi \neq 0$, $i = 1, \ldots, n$, when $\mathcal{S}_-$ contains an odd number of edges. Since, in correspondence of $(j,i) \in \mathcal{S}_+$, $x_i = x_j$, it is possible to reduce the cycle dropping all edges in $\mathcal{S}_+$ and merging the corresponding node variables. If k is the number of edges in $\mathcal{S}_-$, then we are left with the subset of equations

$$\begin{aligned} x_{i_1} &= -x_{i_2} \\ x_{i_2} &= -x_{i_3} \\ &\vdots \\ x_{i_k} &= -x_{i_1}, \end{aligned} \tag{5.14}$$

from which we have, since k is odd,

$$\begin{aligned} x_{i_1} = x_{i_3} = \ \ldots \ &= x_{i_k} \\ x_{i_2} = x_{i_4} = \ \ldots \ &= x_{i_{k-1}}. \end{aligned}$$

However, from the last eq. of (5.14), we also have $x_{i_1} = -x_{i_1}$, $|x_{i_1}| = \xi \neq 0$, which is a contradiction. It implies that the terms of (5.13) belonging to a negative cycle cannot be all equal to 0 and hence that 0 cannot be an eigenvalue of A. In the case $a_{ij}a_{ji} < 0$ for at least a pair of edges in $\mathcal{E}$, then $|a_{ij} + a_{ji}| < |a_{ij}| + |a_{ji}|$, i.e., in place of (5.13) we have

$$0 = -\sum_{i_k, j_k \in \mathcal{K}} \epsilon_k (x_{i_k}^2 + x_{j_k}^2) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j>i}|a_{ij}+a_{ji}|\left(x_i - \operatorname{sgn}(a_{ij}+a_{ji})x_j\right)^2 \tag{5.15}$$

where $\mathcal{K}$ is the set of index pairs i_k, j_k for which $a_{i_k,j_k}a_{j_k,i_k} < 0$ and $\epsilon_k > 0$. Clearly (5.15) can never be true for x such that $|x_i| = \xi \neq 0$, $i = 1, \dots, n$ unless $\mathcal{K}$ is empty. Therefore even in presence of a single length-2 negative cycle 0 cannot be an eigenvalue of A.
To show the reverse implication, assume A Hurwitz stable and by contradiction that $\mathcal{G}^{\text{nd}}(A)$ has no negative cycle (i.e., that $\mathcal{G}(A)$ has no negative cycle of length > 1). Then, from Proposition 4.11, $\exists\ S = \text{diag}(s_1, \dots, s_n)$, $s_i \in \{\pm 1\}$, such that $\hat{A} = SAS$ has all non-negative off-diagonal entries and $\hat{a}_{ij} = |a_{ij}|\ \forall i, j = 1, \dots, n$, $i \neq j$. Since $\hat{a}_{ii} = a_{ii} < 0$, it is $\hat{A} = -A^{\text{cmp}}$. For diagonally equipotent matrices however, A^{cmp} must be singular, hence we have a contradiction.
□

DT linear system

In DT, the analogous of Theorem 5.5 is the following: Schur stability is equivalent to existence of a P symmetric pd that solves the DT Lyapunov equation

$$A^\top PA - A = -Q \tag{5.16}$$

for each given Q symmetric pd. Also some of the other properties can be formulated analogously to the CT case. For instance, diagonal Schur stability corresponds to the special case in which (5.16) is solved by a diagonal pd matrix P.

5.3 Stability of nonlinear systems

Note: this Section will be added later on. A standard reference is [15].

Chapter 6

Basic network dynamics: positive systems and their stability

6.1 Network-adapted dynamics

We aim at determining ways to define reasonable dynamics on a given graph $\mathcal{G}(A)$. To each node i we assign one or more state variables, for instance $x_i \in \mathbb{R}$ or $x_i \in \mathbb{R}^n$.

In some applications, such variable(s) can be constrained, typically $x_i \geq 0$. This happens for instance when x_i represents a physical quantity, like a mass or a concentration, or a frequency, or a probability.

We assume that the dynamics occurs in a distributed way, i.e., as interactions between first neighbors on the graph $\mathcal{G}(A)$. So an update rule for the dynamics at node i must be based only on the own state of node i, x_i, and on the states x_j of the incoming neighbours of node i: $\mathcal{N}_i^{\text{in}} = \{j \in \mathcal{V} \text{ s. t. } (j,i) \in \mathcal{E}\}$.

As general form for the dynamics we can write, in continuous time,

$$\dot{x}_i = f(x_i, x_j, j \in \mathcal{N}_i^{\text{in}}), \quad i = 1, \dots, n \tag{6.1}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$ is a (smooth) functional. In vector form, then

$$\dot{x} = f(x) = \begin{bmatrix} f_1(x_1,\, x_j \in \mathcal{N}_1^{\text{in}}) \\ \vdots \\ f_n(x_n,\, x_j \in \mathcal{N}_n^{\text{in}}) \end{bmatrix} \tag{6.2}$$

Notice that with this formulation, if we denote

$$J(x) = \frac{\partial f(x)}{\partial x}$$

the Jacobian of (6.2) computed in some point x (often, but not always, an equilibrium point), then it is $J_{ij}(x) \neq 0 \Longrightarrow a_{ij} \neq 0$ for some $j \neq i$. (The opposite implication will also be true in most of the examples we study). The meaning of this formulation is that the topology of the interaction graph is reflected on the dynamics (the Jacobian captures the dynamics locally): with perhaps the exception of the self-loops, $\mathcal{G}(A)$ corresponds to $\mathcal{G}(J)$, i.e., to the graph having the Jacobian J as its adjacency matrix. The diagonal terms of J (i.e., the self-loops of $\mathcal{G}(J)$)

may violate this correspondence, in the sense that J may contain elements on the diagonal, even if A does not. Recall that we treat self-loops differently from the rest of the graph. To highlight this fact, we have already introduced $\mathcal{G}^{\text{nd}}(J) = \mathcal{G}(J^{\text{nd}})$, the graph associated to J in which we omit self-loops, where J^{nd} is the Jacobian matrix with 0 diagonal.

When $x^* = 0$ is an equilibrium point, these diagonal terms are normally negative, as they help to render $x^* = 0$ stable.

Example 6.1 (Ex. 2.1 cont'd). For the graph $\mathcal{G}(A)$ of Example 2.1 (see Fig. 2.1) the vector field $f(x)$ is given in (2.2). Computing $J(x) = \frac{\partial f(x)}{\partial x}$ shows that indeed $J_{ij} \neq 0 \Longrightarrow a_{ij} \neq 0$ for all $i \neq j$. □

The simplest possible case is when the vector field $f(x)$ is linear : $f(x) = Ax$. In this case, in continuous time,

$$\dot{x} = Ax \tag{6.3}$$

and in discrete time

$$x(k+1) = Ax(k) \tag{6.4}$$

In both cases the state update matrix A plays the role of adjacency matrix of the graph $\mathcal{G}(A)$.

Example 6.2 Consider the following interconnected system with saturated nonlinearities

$$\dot{x} = -x + \kappa A \psi(x) \tag{6.5}$$

where the irreducible off-diagonal matrix $A \geq 0$ is a weighted adjacency matrix of spectral radius $\rho(A) > 0$ describing the interactions among the agents, $\kappa > 0$ is a scalar coefficient, and $\psi(x) = \begin{bmatrix} \psi_1(x_1) & \dots & \psi_n(x_n) \end{bmatrix}^\top$, $\psi_i(x_i) = \tanh(x_i)$, is a vector of saturated sigmoidal functions depending only on the state of the sending node x_i. The system (6.5) is used e.g. in [9] to describe collective distributed decision-making systems. If we consider the linearization at $x^* = 0$ and $\kappa = 1$, then $J = -(I - A)$ is a Metzler matrix. If we impose the condition $\kappa < \frac{1}{\rho(A)}$, then $x^* = 0$ is a globally asymptotically stable equilibrium point for (6.5). In fact, in this case a simple quadratic Lyapunov function $V = \frac{1}{2}\|x\|^2$ leads to

$$\dot{V} = -x^\top x + \kappa x^\top A \psi(x) \leq x^\top(-I + \kappa A)x < 0$$

because $\psi_i(x_i)$ obeys to the sector inequality $0 \leq \psi_i(x_i) x_i \leq 1$. □

6.2 A special case: positive systems

If we have a dynamical system in which the state $x \in \mathbb{R}^n$ represent a nonnegative quantity (e.g., again, mass, concentration, etc.), and the dynamics obey to $x(t) \geq 0$ for all time t, then we have a positive system. Denote $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \;\text{ s. t. }\; x \geq 0\}$ the positive orthant of $\mathbb{R}^n$.

Definition 6.3 (Positive system) *A system is* positive *if the positive orthant* $\mathbb{R}^n_+$ *is* forward invariant *for its solution:* $x(0) \geq 0 \implies x(t) \geq 0$ *for all times* $t \geq 0$.

The simplest case is that of a *positive linear system*. Positivity in (6.3) and (6.4) is obtained by imposing constraints on the signs of the entries of A. In DT, for instance, we have that if $A \geq 0$ and $x(0) \geq 0$ then the system (6.4) is positive, i.e., it has solution $x(t) \geq 0$ for all $t \geq 0$.

The following is a necessary and sufficient condition for positivity of a DT linear system.

Proposition 6.4 (Positivity of DT linear systems) *The system (6.4) with $x(0) \geq 0$ is positive if and only if $A \geq 0$.*

Proof. "$\Longrightarrow$" By contradiction, if $a_{ij} < 0$ for some (i, j), then it is enough to choose the initial condition $x(0) = e_j^\top$ (e_j = unit vector having 1 on the j-th slot and 0 elsewhere) to get $x(1)$ which is not nonnegative ($x_i(1) < 0$).
"$\Longleftarrow$" Simply notice that we can write recursively

$$x_i(k+1) = \sum_{j=1}^{n} \underbrace{a_{ij}}_{\geq 0} \underbrace{x_j(k)}_{\geq 0} \geq 0$$

□

To discuss positive linear systems in continuous-time, we need to make of use of Metzler matrices.

Proposition 6.5 (Positivity of CT linear systems) *The system (6.3) with $x(0) \geq 0$ is positive if and only if A is Metzler.*

Proof. "$\Longrightarrow$" Same proof by contradiction as in Prop. 6.4.
"$\Longleftarrow$" Let us assume that A has at least a diagonal term $a_{ii} < 0$ (if not, the proof is straightforward). Consider a point $x \geq 0$ and let us look at the i-th component of (6.3)

$$\dot{x}_i = \underbrace{\sum_{j \neq i} a_{ij} x_j}_{\geq 0} + \underbrace{a_{ii} x_i}_{\leq 0} \tag{6.6}$$

The second term in the right hand side is non-positive, so assume that the entire derivative decreases. This means that x_i gets closer to the boundary of $\mathbb{R}^n_+$. However, when it approaches this boundary x_i decreases, and it reaches 0 on the boundary itself, meaning that the negative part of (6.6) vanishes, or $\dot{x}_i|_{x_i=0} \geq 0$. This shows that $\mathbb{R}^n_+$ is a forward invariant set for (6.3), hence (6.3) is a positive system. □

By the same argument one can prove positivity also for nonlinear systems that are homogeneous in the state variables, i.e., of the form $f_i(x) = x_i \phi_i(x)$ for all i, and also the following:

Proposition 6.6 (Positivity of CT nonlinear systems) *If a system $\dot{x} = f(x)$ is such that $f_i(x) \geq 0$ for all $x \in \mathbb{R}^n_+$ such that $x_i = 0$, $i = 1, \ldots, n$, then $\mathbb{R}^n_+$ is a forward invariant set for it.*

The meaning is that if on the entire boundary of $\mathbb{R}^n_+$ the vector field $f(x)$ does not point outside of $\mathbb{R}^n_+$, then the solution of the ODE cannot escape $\mathbb{R}^n_+$, see Fig. 6.1.

Example 6.7 (Example 6.2 cont'd) The system (6.5) obeys to Proposition 6.6, hence it is a positive system if $x(0) \in \mathbb{R}^n_+$. □

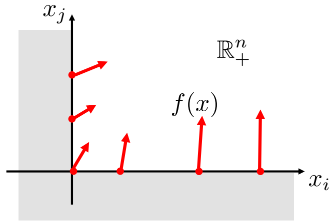


Figure 6.1: Positive nonlinear system.

6.3 Another special case: cooperative systems

If instead of focusing on positive states we are interested in a system having the properties of a positive system but for which the state is unconstrained, $x \in \mathbb{R}^n$, then we have a *cooperative system.*

Definition 6.8 (Cooperative linear system) *The CT linear system* (6.3) *is cooperative if A is Metzler. The DT linear system* (6.4) *is cooperative if $A \geq 0$.*

So linear cooperative systems are the most general class of systems corresponding to non-negative (or Metzler) matrices. The behavior of a cooperative linear system is identical to that of a positive linear system when $x(0) \geq 0$. However, for it, also the negative orthant $\mathbb{R}^n_- = \{x \in \mathbb{R}^n \text{ s. t. } x \leq 0\}$ is forward invariant: $x(0) \in \mathbb{R}^n_- \Longrightarrow x(t) \in \mathbb{R}^n_-$ for all $t \geq 0$, see Fig. 6.2. What happens for $x(0) \in \mathbb{R}^n \setminus \{\mathbb{R}^n_+ \cup \mathbb{R}^n_-\}$ will be investigated when we study the asymptotic behavior of our systems.

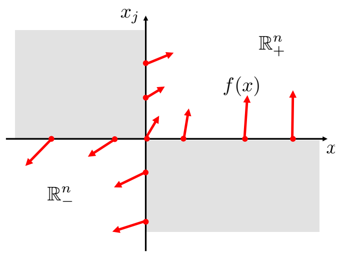


Figure 6.2: Cooperative system.

Notice that also for the general form of the network dynamics (6.2) we can check if the system is cooperative (and hence also if its restriction to $x(0) \in \mathbb{R}^N_+$ is positive).

Proposition 6.9 (Cooperativity of CT nonlinear systems) *Consider the system* (6.2)*. If the Jacobian $J(x) = \frac{\partial f(x)}{\partial x}$ is such that $J_{ij}(x) \geq 0 \ \forall \ i \neq j$ and for all $x \in \mathbb{R}^n$, then the system* (6.2) *is cooperative.*

The proof of this proposition can be found in [22] and is beyond the scope of this course. Its interpretation is however simple: if the linearization behaves as a cooperative system at all points in $\mathbb{R}^n$, then so must the original nonlinear system.

A consequence is that a cooperative system has a somewhat simple dynamical behavior. For instance its trajectories cannot be periodic or have limit cycles, and cannot exhibit chaotic behavior.

Example 6.10 (Example 6.2 cont'd) For the system (6.5), $J(x) = \frac{\partial f(x)}{\partial x} = -I + \kappa A \frac{\partial \psi(x)}{\partial x}$ has the same sign pattern as A in the off-diagonal part, since $\frac{\partial \psi_i(x_i)}{\partial x_i} > 0 \, \forall \, x_i$ and $\kappa > 0$. When A is off-diagonal and nonnegative, it means that the system (6.5) obeys to Proposition 6.9, hence it is a cooperative system. □

6.4 A third special case: monotone systems

Monotone systems are the equivalent of cooperative systems but for a pair of orthants other than $\{\mathbb{R}^n_+, \mathbb{R}^n_-\}$. Consider a signature vector $s = \begin{bmatrix} s_1 & \dots & s_n \end{bmatrix}$, $s_i = \pm 1$, and let $S = \text{diag}(s)$ be the associated diagonal signature matrix, as already seen in Section 4.2. Denote $\mathbb{R}^n_s$ the orthant identified by s: $\mathbb{R}^n_s = \{x \in \mathbb{R}^n \text{ s. t. } Sx \geq 0\}$ and "$\leq_s$" the partial order generated by s: $x_1 \leq_s x_2$ if and only if $x_2 - x_1 \in \mathbb{R}^n_s$.

Definition 6.11 (Monotone system) *The nonlinear system* (6.2) *is monotone with respect to the partial order s if for all initial conditions $x_1(0)$, $x_2(0)$ in some domain $\mathcal{D} \subseteq \mathbb{R}^n$ such that $x_1(0) \leq_s x_2(0)$ it is $x_1(t) \leq_s x_2(t) \; \forall \, t > 0$.*

See Fig. 6.3 for a graphical interpretation. The partial order is strict when in addition $x_1(t) \neq x_2(t)$, i.e., the inequality is strict in at least one of the coordinates.

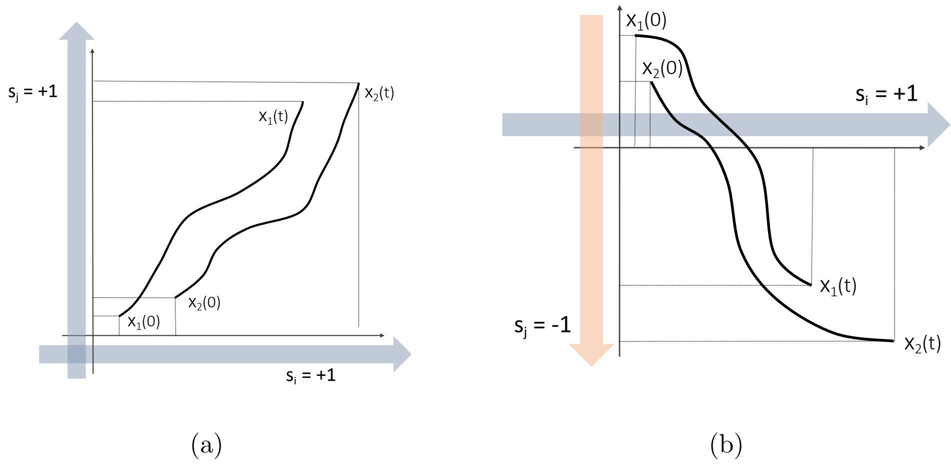


Figure 6.3: 2D slice of the phase portrait of a system (6.2). (a): Cooperative system. (b): Monotone system.

Proposition 6.12 (Jacobian of monotone system) *The system* (6.2) *is monotone with respect to the partial order s if and only if $SJ(x)S$ is Metzler $\forall \, x \in \mathcal{D} \subseteq \mathbb{R}^n$, where $J(x) = \frac{\partial f(x)}{\partial x}$ and $S = \text{diag}(s)$.*

Since $S^{-1} = S$, the meaning of Proposition 6.12 is that after a change of basis with S (which is simply a change of orthant order in the coordinates for which $s_i = -1$) the system (6.2) becomes cooperative. For instance, in Fig.6.3(b) the Jacobian has $\text{sgn}(J_{ij}(x)) = -1$, and choosing S such that $s_j = -1$, $s_i = +1$ leads to $\text{sgn}(s_i s_j J_{ij}) = +1$. Fig.6.3(a) shows the associated cooperative system after this transformation. In the linear case, as $\frac{\partial f(x)}{\partial x} = A$, the situation is even more straightforward. Consequently, from Definition 6.11, a nonlinear cooperative system could be defined as follows.

Definition 6.13 (Nonlinear cooperative system) *The system $\dot{x} = f(x)$ is cooperative if for all $x_1(0)$, $x_2(0)$ in some domain $\mathcal{D} \subseteq \mathbb{R}^n$ such that $x_1(0) \leq x_2(0)$ it is $x_1(t) \leq x_2(t)$ $\forall\, t > 0$.*

If we reason as in Section 6.1, and consider the graph $\mathcal{G}(J(x))$, where $J(x) = \frac{\partial f(x)}{\partial x}$, or rather $\mathcal{G}^{\mathrm{nd}}(J(x))$ (obtained after disregarding the diagonal entries of $J(x)$) then the condition of Proposition 6.12 can be rephrased in graphical terms in terms of the structural balance property we saw in Chapter 4 (Definition 4.10).

Proposition 6.14 (Monotone system and structural balance) *The system (6.2) is monotone with respect to the partial order s if and only if the graph $\mathcal{G}^{\mathrm{nd}}(J(x))$ is structurally balanced $\forall\, x \in \mathcal{D} \subseteq \mathbb{R}^n$ and with the same signature vector s for all $x \in \mathbb{R}^n$, where $J(x) = \frac{\partial f(x)}{\partial x}$.*

Example 6.15 (Example 6.2 cont'd) For the system (6.5), since $\psi_i(x_i)$ is an odd function, A is off-diagonal and $\kappa > 0$, it is $\mathrm{sgn}(J_{ij}) = \mathrm{sgn}(a_{ij})$. When A is signed, the system is monotone if and only if $SAS \geq 0$, i.e., if and only if $\mathcal{G}(A)$ is structurally balanced. □

It then follows that also for monotone systems the dynamical behavior is somewhat simple. In particular, if all trajectories stay bounded then they generically converge to equilibria and cannot admit limit cycles.

6.5 Stability of DT linear positive/cooperative systems

If we have a DT positive or cooperative system, then the matrix A is assumed nonnegative. As stability of a linear system depends only on the properties of A, the two cases of positive and cooperative system can be treated identically. Since $A \geq 0$, the Perron-Frobenius theorem holds. We would like now to combine the stability properties of Chapter 5 with the Perron-Frobenius theorem of Section 3.1, whose main message is that we can describe the dominant eigenvalue $\rho(A)$. The associated dominant eigenspace is in fact the subspace in which trajectories converge asymptotically, at least as long as domination is strict.

Stability of A primitive

Proposition 6.16 (Convergence to dominant eigenspace) *Consider the system (6.4) and assume that $A \geq 0$ is primitive. Then*

$$x(t) \overset{t\to\infty}{\longrightarrow} \mathrm{span}(v)$$

where $v > 0$ is the right eigenvector relative to the Perron-Frobenius eigenvalue $\rho(A)$.

Proof. Assuming that $A \geq 0$ is primitive, then $\rho(A) > 0$ is strictly dominant, hence from Theorem 3.10,

$$\lim_{t\to\infty} \frac{A^t}{\rho^t(A)} = \frac{vw^\top}{w^\top v}$$

where $v > 0$ and $w > 0$ are right and left eigenvectors relative to $\rho(A)$. Therefore

$$\lim_{t\to\infty} x(t) = \lim_{t\to\infty} A^t x(0) = \lim_{t\to\infty} \rho^t(A)\frac{vw^\top}{w^\top v}x(0) = \underbrace{\lim_{t\to\infty} \rho^t(A)}_{\text{scalar}} \frac{\overbrace{w^\top x(0)}^{\text{scalar}}}{\underbrace{w^\top v}_{\text{scalar}}} v = \alpha v \tag{6.7}$$

for some scalar α, i.e., $x(t)$ converges to the eigenspace corresponding to v, $\text{span}(v)$. □

The "magnitude" of $\rho(A)$ decides the stability character of (6.4).

Proposition 6.17 (Stability of DT positive/cooperative system) *Consider the system (6.4) and assume that $A \geq 0$ is primitive. Then*

1. *If $\rho(A) < 1$ then $x^* = \lim_{t\to\infty} x(t) = 0$ is the only asymptotically stable equilibrium point of the system;*
2. *If $\rho(A) = 1$, then $x^* = \lim_{t\to\infty} x(t) = \frac{w^\top x(0)}{w^\top v} v$ is a marginally stable equilbrium point;*
3. *If $\rho(A) > 1$, then $x^* = \lim_{t\to\infty} x(t) = \pm\infty$ and the system is unstable.*

Proof. From (6.7),

1. If $\rho(A) < 1 \implies \lim_{t\to\infty} \rho^t(A) = 0 \implies x(t) \overset{t\to\infty}{\longrightarrow} 0$.
2. If $\rho(A) = 1 \implies \lim_{t\to\infty} \rho^t(A) = 1 \implies x(t) \overset{t\to\infty}{\longrightarrow} \frac{w^\top x(0)}{w^\top v} v$.
3. If $\rho(A) > 1 \implies \lim_{t\to\infty} \rho^t(A) = \infty \implies x(t) \overset{t\to\infty}{\longrightarrow} \infty$.

□

Example 6.18 Consider a positive DT system of the form (6.4) of size $n = 10$. The three cases described in Proposition 6.17 are shown in Fig. 6.4, whose first row shows the trajectories of an entire $x(t)$ vector, while the second row shows a phase portrait of 2 coordinates x_i and x_j for a set of trajectories from different initial conditions. The matrix A is nonnegative and irreducible, and $x_i(0) \geq 0$ in all trajectories. Hence the system is positive.

□

Stability of A irreducible and imprimitive

When $A \geq 0$ is irreducible but imprimitive, then the dominance of $\rho(A)$ is not strict. Asymptotically, the situation is identical if $\rho(A) < 1$ or if $\rho(A) > 1$, but it is more complicated in the $\rho(A) = 1$ case, as the number of dominant eigenvalues is equal to the cyclicity index of A and they are all simple roots of the characteristic of A. They are distributed on the unit circle as the roots of unit: if r is the cyclicity index of A, then the dominant eigenvalues are the complex numbers $\lambda_j = e^{\frac{i2\pi j}{r}}$, $j = 0, 1, \ldots, r-1$ (i = imaginary unit), see Proposition 3.12. From Theorem 5.4, the system (6.4) is marginally stable, but its trajectories do not converge to a specific eigenspace of the dominant eigevectors (unless $x(0)$ is already belonging to one of the dominant eigenspaces). Rather, the trajectories keep oscillating in a linear combination of the dominant eigenvectors, see Example 6.19.

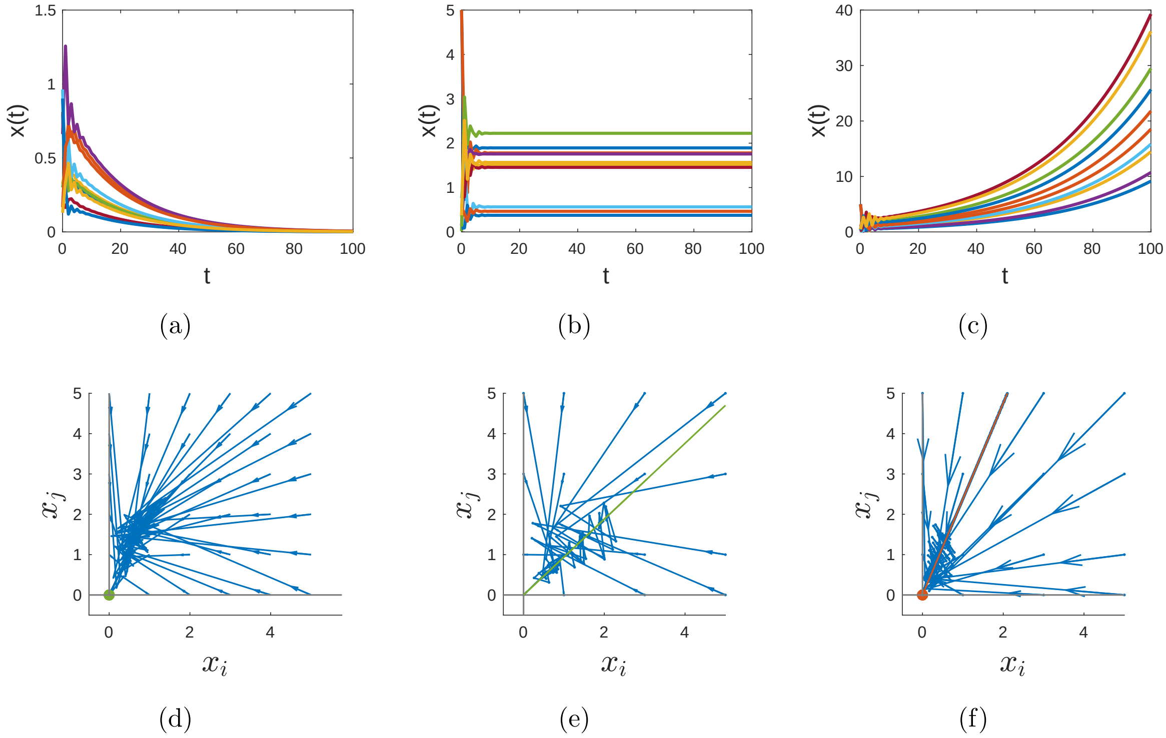


Figure 6.4: Example 6.18: asymptotic behavior of a positive DT linear system for different choices of $\rho(A)$. The top row shows a single trajectory for the entire state vector $x(t)$, while the bottom row shows the phase portrait of a pair of states x_i and x_j for different initial conditions. Left column: $\rho(A) < 1$ (the origin is an attractor for all trajectories); Middle column: $\rho(A) = 1$ (all trajectories converge to the dominant eigenspace, shown in green). Right column: $\rho(A) > 1$ (the trajectories diverge to $+\infty$ and they do so by aligning themselves along the dominant eigenvector (shown in red).

Example 6.19 (Example 3.8 cont'd). For the imprimitive matrix of Example 3.8, $\lambda(A) = \{-1, 1\}$, hence $\rho(A) = 1$ and the cyclicity index is 2. the solutions of the DT system (6.4) keep oscillating with period 2, see Fig. 6.5. □

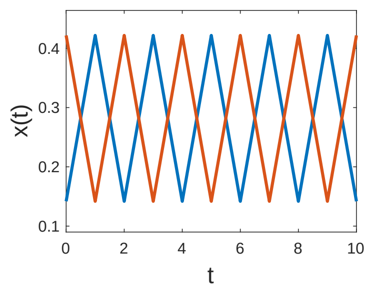


Figure 6.5: Example 6.19. Trajectory of a positive DT linear system with imprimitive A of cyclicity index 2.

Stability of reducible A

When $A \geq 0$ is reducible, then only the weaker form of the Perron-Frobenius theorem (Theorem 3.5) can be applied. Again the cases $\rho(A) < 1$ or if $\rho(A) > 1$ behave analogously, while the case $\rho(A) = 1$ is more difficult, and to study it it is convenient to use the "tterminal" Frobenius normal form (4.3). Each of the blocks $\hat{A}_{11}, \ldots, \hat{A}_{\ell,\ell}$ is irreducible, and may have a spectral radius equal to 1 or less than 1, since the spectrum of $\hat{A}$ is the union of the eigenvalues of $\hat{A}_{11}, \ldots, \hat{A}_{\ell,\ell}$, and $\rho(\hat{A}) = 1$. Each irreducible block $\hat{A}_{ii}$ may be primitive or imprimitive, hence the location and multiplicity of the dominant eigenvalues depends on the numerical values of A. Even though in each diagonal block the dominant eigenvalues are all simple, that does not necessarily lead to a marginally stable system, since the overall algebraic multiplicity may be larger than 1 (i.e., Jordan blocks of dimension larger than 1 can be produced for some of the dominant eigenvalues. As the following example shows, stability may be lost in some cases.

Example 6.20 The matrix

$$A = \left[\begin{array}{cc|cc} 0 & 1 & 0.4 & 0.7 \\ 1 & 0 & 0.3 & 0.1 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

is in the Frobenius normal form (4.3), with 2×2 block diagonal matrices that are imprimitive of cyclicity index 2, see Example 3.8. The eigenvalues of A are $\lambda(A) = \pm 1$ each of multiplicity 2, but each is associated to a Jordan block of dimension 2, hence A is unstable, see Fig. 6.6. □

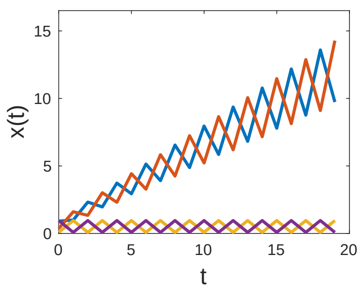


Figure 6.6: Example 6.20. The trajectory of a positive DT linear system with A reducible and $\rho(A) = 1$ can become unstable.

We will see e.g. when studying a special type of linear DT systems, called compartmental systems (Section 7.3.2), that in order to exclude unstable cases an extra assumption is needed: $\mathbb{1}^\top A \leq \mathbb{1}^\top$. This condition is called substochasticity. In CT, it corresponds to diagonal dominance, and its effect on positive systems is investigated in detail in Section 6.6.1. Equivalent results for substochastic matrices are then presented in Section 6.7, and allow to complete the picture for what concerns our stability analysis of the reducible case.

Equivalent characterizations of Schur stability

Let us investigate more in detail the asymptotically stable case. We have already defined positive definitness for a real-valued function $V(x)$: $V(x) > 0$ for all $x \neq 0$, $V(0) = 0$. A real-valued function $V(x)$ is said *co-positive* if $V(x) > 0$ for all $x \geq 0$, $x \neq 0$, $V(0) = 0$.

Theorem 6.21 (Equivalent stability characterizations for nonnegative matrices) *Consider the system (6.4) with $A \geq 0$. The following are equivalent*

1. $\rho(A) < 1$ *(i.e., A is Schur stable);*
2. $A - I$ *is Metzler Hurwitz;*
3. *the characteristic polynomial* $\det(sI - (A - I))$ *has all positive coefficients;*
4. *there exists a linear co-positive Lyapunov function $V(x) = w^\top x$, $w > 0$, such that $\Delta V(x) = V(Ax) - V(x) < 0$ for all $x \geq 0$;*
5. *A is diagonally Schur stable.*

Any of the conditions above corresponds to $x^* = 0$ being asymptotically stable for the system (6.4).

6.6 Stability of CT linear positive/cooperative systems

In CT, linear positive and cooperative systems are both characterized by A which is Metzler. Just like Metzler matrices are negated Z-matrices, so stable Metzler matrices are negated M-matrices. In fact, a Metzler matrix can always be written as $A = B - \alpha I$, with $B \geq 0$ and $\alpha \in \mathbb{R}_+$. When $\alpha < \rho(B)$ then the system is unstable. Stability holds if $\alpha \geq \rho(B)$. This implies that $\mathrm{Re}[\lambda_i(A)] \leq 0$, which can be subdivided in two cases:

1. nonsingular A: when $\alpha < \rho(B) \implies \mathrm{Re}[\lambda_i(A)] < 0 \implies A$ Metzler Hurwitz (i.e., A is both Metzler and Hurwitz);
2. singular A: when $\alpha = \rho(B) \implies \mathrm{Re}[\lambda_i(A)] \leq 0$ and $\mu(A) = 0$ is an eigenvalue. If in addition A is also irreducible, then $\mu(A) = 0$ is a simple eigenvalue and A is marginally stable.

Summing up, the equivalent of Proposition 6.17 for Metzler matrices is the following.

Proposition 6.22 (Stability of CT positive/cooperative systems) *Consider the system (6.3) and assume that A is Metzler and irreducible. Then*

1. *If $\mu(A) < 0$ then $x^* = \lim_{t\to\infty} x(t) = 0$ is the only asymptotically stable equilibrium point of the system;*
2. *If $\mu(A) = 0$, then $x^* = \lim_{t\to\infty} x(t) = \frac{w^\top x(0)}{w^\top v} v$ is a marginally stable equilbrium point;*
3. *If $\mu(A) > 0$, then $x^* = \lim_{t\to\infty} x(t) = \infty$ and the system is unstable.*

Example 6.23 Consider a cooperative system of the form (6.3) of size $n = 10$. The three cases described in Proposition 6.22 are shown in Fig. 6.7, whose first row shows the trajectories of the $x(t)$ vector, while the second row shows a phase portrait of 2 coordinates. The state matrix A is Metzler and irreducible, but the initial conditions are now in $\mathbb{R}^n$, hence the system is a cooperative system.

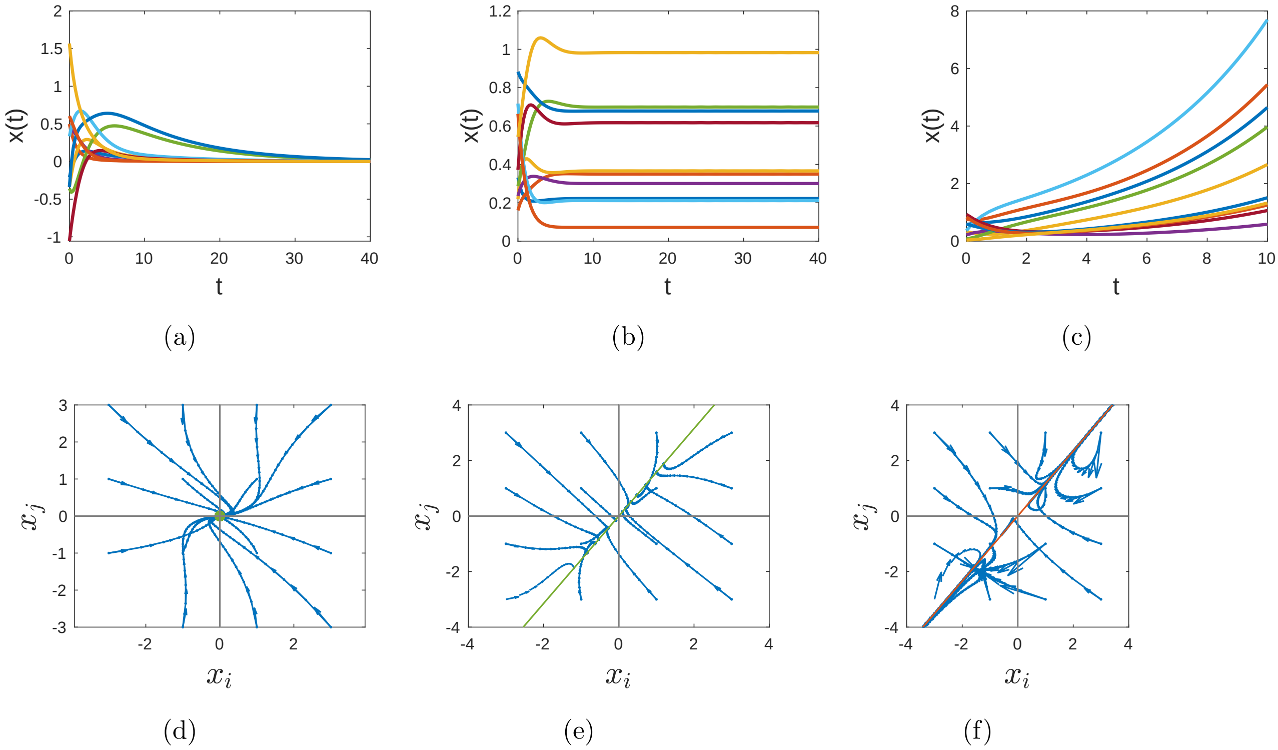


Figure 6.7: Cooperative CT linear system for different choices of $\mu(A)$. The top row shows a single trajectory for the entire state vector $x(t)$, while the bottom row shows the phase portrait of a pair of states x_i and x_j for different initial conditions. Left column: $\mu(A) < 0$ (the origin is asymptotically stable for all trajectories); Middle column: $\mu(A) = 0$ (marginal stability: all trajectories converge to the dominant eigenspace, shown in green). Right column: $\mu(A) > 0$ (unstable system: the trajectories diverge to $\pm\infty$ and they do so by aligning themselves along the dominant eigenvector (shown in red).

We can zoom more in detail at the stable cases. The following asymptotic stability characterizations are valid for Metzler systems.

Theorem 6.24 (Equivalent stability characterizations for Metzler matrices) *Consider the system (6.3) with A Metzler. The following are equivalent*

1. *A is Hurwitz;*
2. *there exists $v > 0$ s.t. $Av < 0$;*
3. *there exists $w > 0$ s.t. $w^\top A < 0$;*
4. *A is diagonally Hurwitz stable;*
5. *A is D-stable;*
6. *the coefficients of the characteristic polynomial $\det(sI - A)$ are all positive;*
7. *A can be written as $A = B - \alpha I$ with $B \geq 0$ and $\alpha > \rho(B)$.*

Any of the conditions above corresponds to $x^* = 0$ being asymptotically stable for the system (6.3).

From Theorem 6.24, when comparing stability characterizations of a general $A \in \mathbb{R}^{n\times n}$ with those of A which is Metzler, then the hierarchy (5.6) is replaced by the following one.

$$\begin{array}{r} \text{negative definitness of } A \text{ Metzler} \\ \Downarrow \\ \text{diagonal stability of } A \text{ Metzler} \\ \Updownarrow \\ \text{D-stability of } A \text{ Metzler} \\ \Updownarrow \\ A \text{ Metzler Hurwitz} \end{array} \tag{6.8}$$

Example 6.25 The matrix

$$A = \begin{bmatrix} -0.7 & 0.9 & 0 & 0.4 \\ 0.1 & -0.8 & 0.1 & 0 \\ 0 & 0.5 & -0.9 & 0 \\ 0 & 0.6 & 0.8 & -0.8 \end{bmatrix}$$

is Metzler Hurwitz, but $\lambda(A^{\text{sym}}) = \{-1.388, -1.194, -0.693, 0.076\}$, i.e., A^{sym} is not pd. The left and right eigenvectors associated to $\mu(A) = -0.279$ are

$$w = \begin{bmatrix} 0.213 & 0.896 & 0.355 & 0.1633 \end{bmatrix}^\top$$
$$v = \begin{bmatrix} 0.852 & 0.193 & 0.156 & 0.462 \end{bmatrix}^\top.$$

For them of course $A^\top w < 0$ and $Av < 0$. $P = \text{diag}(w)(\text{diag}(v))^{-1}$ is a diagonal stability matrix for A. □

6.6.1 Stability of Metzler matrices from diagonal dominance

Diagonal dominance.

The idea that diagonal dominance can help in establishing stability, described in Theorem 5.12, can be specialized to Metzler matrices. In particular, when for a Metzler matrix A it is $a_{ii} \leq 0$ for all i, then diagonal dominance can be written without absolute values. Instead of diagonal dominance by rows used in (5.8), let us consider diagonal dominance by columns:

$$-a_{ii} \geq \sum_{j\neq i} a_{ji} \quad i = 1, \dots, n \tag{6.9}$$

Theorem 6.26 (Diagonal dominance Theorem 5.12 adapted for Metzler matrices) *Consider A Metzler, with $a_{ii} \leq 0$ for all i.*

1. *If A is strictly diagonally dominant, or irreducible and weakly diagonally dominant, then A is nonsingular and Hurwitz.*

2. *If A is diagonally equipotent, i.e.,*

$$- a_{ii} = \sum_{j \neq i} a_{ji} \quad i = 1, \ldots, n \tag{6.10}$$

then A is singular.

3. *If A is diagonally equipotent and irreducible, then the multiplicity of the eigenvalue $\lambda = 0$ is 1 and A is marginally stable.*

Proof. Part 1 is straightforward from Theorem 5.12.
Part 2: If A is irreducible, then for each column i there exist a $j \neq i$ such that $a_{ji} > 0$. Hence, since A Metzler, (6.10) implies that $a_{ii} < 0$ for each i. Writing $-a_{ii} = \sum_{j \neq i} a_{ji}$ for all i is equivalent to writing $\mathbb{1}^\top A = 0$, i.e., $0 \in \text{spec}(A)$, hence A is singular.
Part 3: From Theorem 3.17, since the eigenvector associated to 0 is $\mathbb{1} > 0$, $\mu(A) = 0$ is the spectral abscissa of A, and since A is irreducible, $\mu(A) = 0$ must be a simple eigenvalue of A. Since $\lambda(A) \in \text{spec}(A)$, $\lambda(A) \neq \mu(A)$ implies $\text{Re}[\lambda(A)] < 0$ then A is marginally stable. □

The Metzler structure of A implies that more can be said. In particular, when A is weakly diagonally dominant and reducible, a necessary and sufficient condition for Hurwitz stability can be found. In this case, it is useful to specify that diagonal dominance is by columns, because that is associated to terminal spanning forest. Later on we formulate also a version of the same result for row diagonal dominance (and rooted spanning forests).

Theorem 6.27 (Stability for reducible matrices; diagonal dominance by columns) *Consider A Metzler and weakly diagonally dominant by columns, with $a_{ii} \leq 0$ for all i. Denote $\mathcal{V}^{\text{out}} \subseteq \mathcal{V}$ the set of indices for which the strict diagonal dominance inequality (6.9) is verified. The following conditions are equivalent.*

1. *A is Hurwitz.*
2. *$\mathcal{G}(A)$ has a spanning forest terminating at a subset of $\mathcal{V}^{\text{out}}$.*
3. *$\mathcal{G}(A)$ is output connected to $\mathcal{V}^{\text{out}}$.*

Proof.
2 $\implies$ 1. Recall that $\mathcal{G}(A)$ having a spanning forest terminating at nodes $v_1, \ldots, v_k$ means that each node of $\mathcal{G}(A)$ is connected via a directed path to one of $v_1, \ldots, v_k$. Furthermore, there exists a permutation of rows/columns that brings A into the transpose of the Frobenius form (4.3). Since this a block triangular matrix, the eigenvalues of A are given by the union of the eigenvalues of the diagonal blocks. Assume for now that each v_i $i = 1, \ldots, k$ belongs to the strongly connected component of index i in the Frobenius normal form, i.e., are part of the diagonal blocks $A_{11}, \ldots, A_{kk}$ of (4.3). Recall that these diagonal blocks correspond each to an irreducible matrix (possibly equal to 0 when the block is of dimension 1). Furthermore, on $\mathcal{G}(A)$ these blocks have no outgoing edge (apart from edges within the block itself), meaning that (6.9) reduces to a (possibly nonstrict) inequality within the nodes forming a block for all columns forming the block. To each of them, Theorem 6.26 can be applied. Consider for instance the block A_{jj}. Since A_{jj} is irreducible, if it is (at least) weakly diagonally dominant, then A_{jj} is Hurwitz, while if it is diagonally equipotent it is singular and marginally stable.

Having a spanning forest terminating at v_1 (part of A_{11}), ..., v_k (part of A_{kk}), where each v_i is associated to strict diagonal dominance on the corresponding column, means that each A_{jj} is weakly diagonally dominant, hence that each block A_{jj} is Hurwitz. Concerning the remaining $\ell - k$ diagonal blocks, these are non-terminal, meaning that at least one node in each of them has an outgoing edge (going outside the diagonal block itself). Consider for instance one such node i belonging to the non-terminal block A_{jj}, $j \in \{k+1, \ldots, \ell\}$. Denote $j_1, \ldots, j_\nu$ the indices of the nodes in this block (i is one of them). The node i is also connected to some other node η outside of the A_{jj} block (i.e., $a_{\eta i} > 0$). Then (6.9) can be written as

$$-a_{ii} \geq \sum_{\mu \in \{j_1, \ldots, j_\nu\}} a_{\mu i} + \sum_{\mu \in \mathcal{V} \setminus \{j_1, \ldots, j_\nu\}} a_{\mu i}$$

Since $\eta \in \mathcal{V} \setminus \{j_1, \ldots, j_\nu\}$, the second summand is certainly positive. When computing the eigenvalues of A_{jj}, because of the block triangular structure, the idea of diagonal dominance can be restricted to the block itself. In this case, it leads to strict inequality

$$-a_{ii} > \sum_{\mu \in \{j_1, \ldots, j_\nu\}} a_{\mu i}$$

meaning that each block A_{jj}, $j = k+1, \ldots, \ell$ is irreducible and weakly diagonally dominant. Hence all diagonal blocks A_{jj}, $j = 1, \ldots, \ell$, are Hurwitz and therefore A is Hurwitz.

Now let us consider the case in which there is no one-to-one match between the terminal nodes of the terminal spanning forest, denoted $\{v_1, \ldots, v_\rho\}$, and the terminal strongly connected components of the graph, and/or not all the terminal nodes are part of the set $\mathcal{V}^{\text{out}}$ of nodes associated to strict diagonal dominance. If $\rho \geq k$ and for each terminal strongly connected component there is at least one node in $\{v_1, \ldots, v_\rho\}$ (but some may have more than one), then the same conclusion can be easily drawn (possibly by modifying the terminal spanning forest). If instead some terminal strongly connected components are not represented in the set $\{v_1, \ldots, v_\rho\}$, then we are in contradiction, as a terminal spanning forest must touch all terminal strongly connected components.

1 $\Longrightarrow$ 2. To show the converse, observe that when $\mathcal{V}^{\text{out}}$ does not contain any node from some of the terminal strongly connected components, then these terminal strongly connected component have no outgoing edge (outside of the block itself) and the associated A_{jj} is diagonally equipotent. From condition 3 of Theorem 6.26, the block A_{jj} is marginally stable, but not Hurwitz. Hence A cannot be Hurwitz.

2 $\Longleftrightarrow$ 3. If we interpret $\mathcal{V}^{\text{out}}$ as outputs of $\mathcal{G}(A)$ then, from Proposition 4.5, item 2 corresponds to requiring that $\mathcal{G}(A)$ is output-connected. □

We will see in Section 7.3 when dealing with compartmental systems that the condition of Theorem 6.27 corresponds to having a "compartmental matrix without traps".

There is also a "dual" version of Theorem 6.27, valid when diagonal dominance is by rows, which will be needed when investigating e.g. consensus in Chapter 8 and Friedkin-Johnsen models in Chapter 11. In this case we deal with spanning forests which are rooted.

Theorem 6.28 (Stability for reducible matrices; diagonal dominance by rows) *Consider A Metzler and weakly diagonally dominant by rows, with $a_{ii} \leq 0$ for all i. Denote $\mathcal{V}^{\text{in}} \subseteq \mathcal{V}$ the set of indices in which diagonal dominance is strict. The following conditions are equivalent.*

1. *A is Hurwitz.*
2. *$\mathcal{G}(A)$ has a spanning forest rooted at a subset $\mathcal{V}^{\text{in}}$.*
3. *$\mathcal{G}(A)$ is input connected from $\mathcal{V}^{\text{in}}$.*

Diagonal equipotency and conservation laws

In Theorem 6.26, the diagonal equipotency by columns condition (6.10) is equivalent to $\mathbb{1}^\top A = 0$. In a dynamical system like (6.3), this condition can be written also as $\mathbb{1}^\top Ax = 0$, or $\mathbb{1}^\top \dot{x} = 0$. This is a conservation law on the dynamics, as, after integration, $\mathbb{1}^\top x(t) = \mathbb{1}^\top x(0) = c$ for all t, where c is a constant depending on the initial conditions. In general, conservation laws in the state are incompatible with asymptotic stability: changing initial conditions from $x(0)$ to $\tilde{x}(0) = x(0) + \epsilon$ for some small vector $\epsilon \in \mathbb{R}^n$ changes the equilibrium point achieved asymptotically, as $\mathbb{1}^\top \tilde{x}(t) = \mathbb{1}^\top x(t) + \mathbb{1}^\top \epsilon$ for all t. Hence for Metzler matrices, equipotence leads at most to marginal stability (when $a_{ii} \leq 0$ and A irreducible; Part 3 of Theorem 6.26). One exception is when the conservation laws is used to impose a constraint on the state, for instance the fact that fractions or probabilities must sum to 1: $\mathbb{1}^T x = 1$. In that case, the value of the conservation law is fixed, and the system evolves on the "slice" determined by the conservation law for all initial conditions. A Markov chain is an application which exploits this fact, see Section 7.1.

Diagonal dominance/equipotency and multiplicity of the 0 eigenvalue for reducible matrices

When A is reducible, the situation with the 0 eigenvalue of A is more complex. We have in fact multiple terminal strongly connected components, corresponding to the diagonal blocks A_{jj} in the Frobenius normal form (4.3). Since the blocks A_{jj} are irreducible, it follows from Theorem 6.26 that each of these blocks may contribute at most one $\lambda = 0$ eigenvalue to A, depending on whether A_{jj} is associated to diagonal equipotence or contains some strictly dominant column. The following is a straightforward consequence of Theorem 6.26.

Theorem 6.29 (Multiplicity of $\lambda = 0$ eigenvalue; columns dominance) *Consider A Metzler and diagonally dominant (possibly diagonally equipotent) by columns, with $a_{ii} \leq 0$ for all i. Then the multiplicity of the $\lambda = 0$ eigenvalue is equal to the number of terminal strongly connected components of $\mathcal{G}(A)$ for which, in the Frobenius normal form (4.3), the associated block diagonal matrix A_{jj} is diagonally equipotent by columns. If the multiplicity of $\lambda = 0$ is larger than 0, then A is marginally stable.*

In the theorem, only the strongly connected components which are terminal matter. In the Frobenius normal form (4.3), these are the blocks with indices $1, \ldots, k$, i.e., correspond to the blocks having all columns of zero entries, apart from the diagonal block itself. All other blocks (those indexed by $k+1, \ldots, \ell$ in (4.3)), do not contribute a 0 eigenvalue.

The final statement of Theorem 6.29 follows from the fact that in Frobenius normal form all $\lambda = 0$ eigenvalues belong to different terminal diagonal blocks (and end up in different blocks in the Jordan form).

Example 6.30 Let us consider again the graph $\mathcal{G}(A_1)$ of Fig. 4.6(a), discussed in Example 4.9. In particular, recall that this graph has two terminal strongly connected components, consisting of the nodes {5, 6, 7} and {4}. Consider the Frobenius normal form (4.3) A^{term} associated to $\mathcal{G}(A)$, obtained reordering the strongly connected components as mentioned in (4.4), i.e., {5, 6, 7} followed by {4} and then {1, 2, 3}, and associate to it the Metzler matrix

$$A_1^{\text{term}} = \left[\begin{array}{ccc|c|ccc} -\alpha & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -\beta & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -\gamma \end{array}\right]$$

where α, β, $\gamma \geq 0$ are parameters. We have the following cases for the spectrum of A_1^{term}.

- If $\alpha = 1$, $\beta = 0$ and $\gamma = 2$, then $\mathbb{1}^\top A_1^{\text{term}} = 0$ i.e., we have diagonal equipotence. Since A_1^{term} is reducible and has two terminal strongly connected components, the multiplicity of $\lambda = 0$ is 2.
- If $\alpha = 1$ and $\beta > 0$ (or $\alpha > 1$ and $\beta = 0$) and $\gamma = 2$, then A_1^{term} is diagonally dominant in exactly one column associated to a terminal strongly connected component. Since A_1 has two terminal strongly connected components, the multiplicity of $\lambda = 0$ is 1.
- If $\alpha > 1$, $\beta > 0$ and $\gamma = 2$, then the matrix is Hurwitz (it has two strictly diagonally dominant columns, and each is contained inside a terminal strongly connected components).

In other words, in order to have a Hurwitz matrix, each terminal strongly connected component has to contain a node which corresponds to strict diagonal dominance. Notice that strict diagonal dominance on nodes that do not belong to the terminal strongly connected components is irrelevant. For instance choosing $\alpha = 1$, $\beta > 0$ and $\gamma = 3$ does not lead to Hurwitz.

In the language of compartmental system we will see in Section 7.3 the system has to have no (simple) trap (intended as terminal strongly connected components not associated to a strict diagonal dominance). Finally, notice that weak diagonal dominance is not necessary for Hurwitz stability of Metzler matrices: choosing $\alpha > 1$, $\beta > 0$ and $\gamma = 1$ we loose dominance in the last column, but the matrix is nevertheless Metzler Hurwitz. □

Just like Theorem 6.28 is a "dual" version of Theorem 6.27 corresponding to replacing dominance by columns with dominance by rows, so the next theorem is the "dual" of Theorem 6.29.

Theorem 6.31 (Multiplicity of $\lambda = 0$ eigenvalue; row dominance) *Consider A Metzler and diagonally dominant (possibly diagonally equipotent) by rows, with $a_{ii} \leq 0$ for all i. Then the multiplicity of the $\lambda = 0$ eigenvalue is equal to the number of rooted strongly connected components of $\mathcal{G}(A)$ for which the associated block diagonal matrix A_{jj} in the Frobenius normal form (4.2) is diagonally equipotent by rows. If the multiplicity of $\lambda = 0$ is larger than 0, then A is marginally stable.*

Regular splitting and stability

We have seen that Metzler matrices and nonnegative matrices are related by a "splitting" like $A = B - \alpha I$, where A Metzler, $B \geq 0$ and a scalar $\alpha > 0$ corresponding how much we shift the eigenvalues of B to the left.

There are other ways to related nonnegative and Metzler matrices through a splitting. In particular we are interested in a so-called *regular splitting*. For a Metzler matrix A, the splitting $A = \Lambda + C$ is said regular if Λ is Metzler and Hurwitz and $C \geq 0$.

Proposition 6.32 (Regular splittings and stability) *Let $A = \Lambda + C$ be a regular splitting of a Metzler matrix A. Then A is Hurwitz if and only if $\rho(-\Lambda^{-1}C) < 1$.*

Proof in [23], Thm. 3.13. See also Berman-Plemmon Ch 6, Th. 2.3, N_{45} (where it is formulated in terms of nonsingular M-matrices). In other words:

- $\mu(A) < 0 \iff \rho(-\Lambda^{-1}C) < 1$;
- $\mu(A) = 0 \iff \rho(-\Lambda^{-1}C) = 1$;
- $\mu(A) < 0 \iff \rho(-\Lambda^{-1}C) > 1$.

6.6.2 Other properties of Metzler matrices

Exponential nonnegativity of Metzler matrices.

Metzler matrices are sometimes called exponentially positive matrices in the literature. The reason is the following property.

Proposition 6.33 *A matrix A is Metzler if and only if $e^{At} \geq 0$ for all $t \geq 0$. If, in addition, A is irreducible, the $e^{At} > 0$ for all $t > 0$.*

Inverse positivity of Metzler matrices.

Another useful property of a Metzler matrix is its inverse positivity.

Theorem 6.34 (Inverse positivity of Metzler matrices) *Consider a Metzler matrix A. The following are equivalent:*

1. *A is Hurwitz;*
2. *A is* inverse positive, *i.e., A is invertible and $-A^{-1} \gneq 0$.*

If, in addition, A is irreducible, then $-A^{-1} > 0$.

Proposition 6.35 *Consider $M \gneq 0$ irreducible. If $\exists x \gneq 0$ and $\lambda > 0$ such that $Mx \gneq \lambda x$, then $\rho(M) > \lambda$. If instead $Mx \lneq \lambda x$ then $\rho(M) < \lambda$. Analogous conditions hold for Metzler matrices (replacing $\rho(M)$ with $\mu(M)$).*

6.7 Back to DT: stability from substochasticity

For the DT positive linear system (6.4), the condition that A is Metzler is replaced by $A \geq 0$. The equivalence of diagonal dominance by columns is A *column substochastic*: $\mathbb{1}^\top A \leq \mathbb{1}^\top$, while for diagonal dominance by rows it is A *row substochastic*: $A\mathbb{1} \leq \mathbb{1}$. Diagonal equipotence by columns corresponds to A *columns stochastic*: $\mathbb{1}^\top A = \mathbb{1}^\top$, while diagonal equipotence by rows corresponds to A *row stochastic*: $A\mathbb{1} = \mathbb{1}$. For DT positive systems we can then obtain the equivalent of the conditions investigated in the CT case in the previous section. For instance Theorem 6.26 becomes the following.

Theorem 6.36 (Stability from substochasticity) *Consider $A \geq 0$ column (resp. row) substochastic.*

1. *If A is strictly substochastic, i.e., $\mathbb{1}^\top A < \mathbb{1}^\top$ (resp. $A\mathbb{1} < \mathbb{1}$) or A irreducible and substochastic with at least one strict inequality, then A is Schur stable.*
2. *If A is column (resp. row) stochastic, then $\rho(A) = 1$ is an eigenvalue.*
3. *If A is column (resp. row) stochastic and A is irreducible, then the multiplicity of the eigenvalue $\lambda = 1$ is 1 and A is marginally stable. In addition*
   - (a) *If A is primitive, then the eigenvalue $\rho(A) = 1$ is strictly dominant.*
   - (b) *If A is imprimitive of cyclicity index r, then A has r simple eigenvalues on the unit circle: $\lambda_j = e^{\frac{i2\pi j}{r}}$, $j = 0, \ldots, r-1$.*

The proof is a straightforward adaptation of that of Theorem 6.26, using the equivalent characterizations of Theorem 6.21. The properties of the primitive/imprimitive cases follow from the considerations in Section 6.5, see also Proposition 3.12.

Since A is reducible if and only if the associated Metzler matrix $A - I$ is reducible, also the conditions of Theorems 6.27-6.28 can be rephrased for DT systems.

Theorem 6.37 (Stability for reducible matrices; substochasticity) *Consider $A \geq 0$ column (resp. row) substochastic. Denote $\mathcal{V}^{\text{out}} \subseteq \mathcal{V}$ (resp. $\mathcal{V}^{\text{in}} \subseteq \mathcal{V}$) the set of column (resp. row) indices for which substochasticity is strict. The following conditions are equivalent.*

1. *A is Schur stable.*
2. *$\mathcal{G}(A)$ has a spanning forest terminating at a subset of $\mathcal{V}^{\text{out}}$ (resp. rooted at a subset of $\mathcal{V}^{\text{in}}$).*
3. *$\mathcal{G}(A)$ is output connected to $\mathcal{V}^{\text{out}}$ (resp. input connected to $\mathcal{V}^{\text{in}}$).*

Similarly, also the multiplicity of the $\lambda = 1$ eigenvalue of A can be investigated in an analogous way to the CT case.

Theorem 6.38 (Multiplicity of $\lambda = 1$ eigenvalue) *Consider $A \geq 0$ column (resp. row) substochastic (possibly stochastic). Then the multiplicity of the $\lambda = 1$ eigenvalue is equal to the number of terminal (resp. rooted) strongly connected components of $\mathcal{G}(A)$ for which the associated block diagonal matrix A_{jj} in the Frobenius normal form (4.3) is stochastic by columns (resp. by rows). The presence of other eigenvalues in the unit circle depends on the cyclicity index of the A_{ii} blocks. Whenever $\lambda = 1$ is an eigenvalue, the matrix A is marginally stable.*

The proof follows from that of Theorem 6.29, with again the extra argument that if a terminal (resp. rooted) diagonal block A_{ii} is primitive, then $\lambda = 1$ is a strictly dominating eigenvalue for that block, while if A_{ii} is imprimitive of cyclicity index r_i, then A_{ii} has r_i eigenvalues on the unit circle, radially equispaced, see Section 6.5. Because of the structure of the Frobenius normal form, the algebraic multiplicity of such dominant eigenvalue is one no matter the number k of terminal (resp. rooted) blocks, i.e., the associated Jordan blocks are all of dimension 1.

6.8 Positive affine systems and positive equilibria

In our analysis of equilibria of positive linear systems, so far we have only considered the case of an equilibrium in the origin: $x^* = 0$. Now, for most applications of positive systems one does not expect $x^* = 0$ to be an interesting equilibrium point: if the vector x represents e.g. masses, then having as equilibrium $x^* = 0$ means that all masses are vanishing asymptotically, which is not what happens in practice. We would like to have $x^* > 0$, (or at least $x^* \gneq 0$, i.e., $x_i^* > 0$ in some coordinates). Notice that one cannot simply translate the equilibrium with a change of state: for instance, for the CT linear system (6.3), if $z = x - x^*$, with $x^* \gneq 0$ then the new dynamics in z, $\dot{z} = Az$, indeed has $z^* = 0$ as equilibrium. However, $x \geq 0$ does not imply $z \geq 0$, but rather $z \geq -x^*$, which is a constraint difficult to impose in a model.

Another possibility is to consider linear systems with an affine term, i.e., add an extra term (independent from the state x) to the right hand side of (6.3) or (6.4), obtaining respectively,

$$\dot{x} = Ax + u \tag{6.11}$$

where A is Metzler, and

$$x(t+1) = Ax(t) + u \tag{6.12}$$

where $A \geq 0$. In both (6.11) and (6.12), $u \in \mathbb{R}^n$, $u \geq 0$ can be intended as a constant additive term to the dynamics. In principle, it can also be intended as an external input, even a time-varying one. However, if we are interested in studying the equilibria of (6.11), then we must restrict ourself to constant u. Constant $u \geq 0$ (and $u \gneq 0$ to avoid trivial cases) is a standing assumption in this Section.

6.8.1 CT positive affine system

Consider the system (6.11) with A is Metzler. According to Definition 6.3, a system like (6.11) is positive if $x(0) \geq 0$ implies $x(t) \geq 0 \ \forall t \geq 0$. A necessary and sufficient condition for positivity of the system (6.11) is that A is Metzler and $u \geq 0$ see [8].

Given $u \geq 0$, the system (6.11) admits a nonnegative (resp. positive) equilibrium x^* when there exists $x^* \geq 0$ (resp. $x^* > 0$) such that $Ax^* + u = 0$. We have the following cases:

- If A nonsingular, then x^* solving $Ax = -u$ exists unique.
- If A is singular, then $Ax = -u$ has infinitely-many solutions if $u \in \text{range}(A)$ and no solution otherwise.

In the following we focus on the first case.

Proposition 6.39 (Nonnegative stable equilibrium; CT case) *Consider the system* (6.11) *with A Metzler. The two conditions are equivalent.*

1. *A is Hurwitz.*
2. *For each $u \gneq 0$ there exists a unique $x^* \geq 0$ that solves $Ax^* + u = 0$, i.e., $x^* \geq 0$ is an equilibrium of* (6.11) *and it is asymptotically stable with domain of attraction $\mathbb{R}^n_+$.*

Proof. 1 $\implies$ 2. If A is Metzler and Hurwitz, A is invertible and, from Theorem 6.34, $-A^{-1} \gneq 0$. Hence for any $u \gneq 0$, $x^* = -A^{-1}u \geq 0$.
2 $\implies$ 1. By contradiction, if there exists $\lambda_o \in \text{spec}(A)$ such that $\text{Re}(\lambda_o) > 0$, then also $\mu(A) > 0$, and no matter what the multiplicity of $\mu(A)$ is, there is always a nonnegative left eigenvector $w \gneq 0$. This can be deduced adapting to Metzler matrices the weakest version of the Perron-Frobenius theorem (Theorem 3.5). But then it is enough to observe that for $u = w$

$$w^\top (Ax^* + u) = \underbrace{w^\top A}_{\geq 0} \underbrace{x^*}_{\geq 0} + \underbrace{w^\top u}_{>0} = 0$$

cannot have solutions. The argument hold also when $\mu(A) = 0$, no matter what the multiplicity of $\mu(A) = 0$ is. To show asymptotic stability, it is enough to notice that the translated system $z = x - x^*$ has ODE $\dot{z} = Az$, and its equilibrium point $z^* = 0$ is asymptotically stable. Since the system is linear, stability is global. □

In general, in the previous proposition $x^* \geq 0$, i.e., some of the components x_i^* can be 0. To guarantee strict positivity of x^*, i.e., $x_i^* > 0$ in all components, we need an extra condition, which we can call *u-connectivity*. Loosely speaking, this property characterizes the possibility of u to influence, directly or indirectly, the entire state vector x. In the positive systems literature [8], it is referred to as *excitability from u*: the system (6.11) is excitable from u if and only if each state variable can be made positive by applying u to the system initially at rest (i.e., in $x(0) = 0$).

As long as we deal with an irreducible A, it is enough to think of u-connectivity as input connectivity of $\mathcal{G}(A)$, a notion introduced in Chapter 4.

A irreducible

When A is irreducible, any $u \gneq 0$ leads to positivity of x^*.

Proposition 6.40 (Positive stable equilibria: CT irreducible case) *If A is Metzler Hurwitz and irreducible, then we have the following.*

- *For any $u \gneq 0$ the system* (6.11) *is u-connected.*
- *For any $u \gneq 0$ the unique equilibrium point is $x^* > 0$ and it is asymptotically stable with domain of attraction $\mathbb{R}^n_+$.*

Proof. From Proposition 6.39, an equilibrium $x^* \geq 0$ that solves $Ax^* + u = 0$ exists unique. From Theorem 6.34, irreducibility of A implies $-A^{-1} > 0$. Hence in $x^* = -A^{-1}u$ it is enough to have one of the components $u_i > 0$ for some i, to obtain $x^* > 0$. To show asymptotic stability, just apply the change of state $z = x - x^*$, whose dynamics is $\dot{z} = Az$ and which has $z^* = 0$ as a globally asymptotically stable equilibrium point. This implies that in the original basis $x = x^*$ is asymptotically stable. Combined with forward invariance of $\mathbb{R}^n_+$, this gives the result. □

Any of the equivalent conditions that we have seen earlier for A which is Hurwitz can be used also to investigate the strictly positive equilibrium point x^*, for instance the following one which makes use of diagonal dominance (from Theorem 6.26).

Proposition 6.41 *If A is Metzler, irreducible and weakly diagonally dominant with $a_{ii} \leq 0$, then for any $u \gneq 0$ the equilibrium point $x^* > 0$ is asymptotically stable for the system (6.11) with domain of attraction $\mathbb{R}^n_+$.*

A reducible

When A is reducible and diagonally dominant, then we need to distinguish between diagonal dominance by columns or by rows. In particular, there are two ways to formulate u-connectivity, linked to the two ways we have introduced to represent reducible systems, through the two Frobenius normal forms A^{term} and A^{root}. The most "natural" one is the one for A^{term}, because this is just the notion of input connectivity which we have used so far: $\mathcal{V}^{\text{in}} = \{i \in \mathcal{V} \text{ s. t. } u_i > 0\}$ is the input set associated to a given u. This is used for diagonal dominance by columns. From Proposition 4.5, u-connectivity then also corresponds to $\mathcal{G}(A)$ having a spanning forest rooted at a subset of $\mathcal{V}^{\text{in}}$.

The second representation of the u-connectivity notion is instead associated to the Frobenius normal form A^{root}, i.e., to diagonal dominance by rows. As this is associated to $\mathcal{G}(A^\top)$, the role of "inputs" and "outputs" are exchanged: the set $\{i \in \mathcal{V} \text{ s. t. } u_i > 0\}$ becomes now the set of "outputs" $\mathcal{V}^{\text{out}}$, and u-connectivity becomes our notion of output connectivity from Chapter 4, i.e., existence of a spanning forest terminating at a subset of $\mathcal{V}^{\text{out}}$. This second notion is somewhat counterintuitive, especially if we think of u as an external input axting on the system. It is however coherent with the representation we gave of strictly dominated columns and rows in Theorems 6.27, and 6.28. In both cases of column and row dominance, a necessary and sufficient condition for positivity of the equilibrium point can be stated in terms of input and output connectivity.

Theorem 6.42 (Positive stable equilibrium; CT reducible case) *Consider the system (6.11), and assume that A is Metzler and weakly diagonally dominant by columns (resp. by rows), with $a_{ii} \leq 0$. Consider $u \gneq 0$, and denote $\mathcal{V}^{\text{in}}$ (resp. $\mathcal{V}^{\text{out}}$) the set $\{i \in \mathcal{V} \text{ s. t. } u_i > 0\}$. Denote further $\mathcal{V}^{\text{out}}$ (resp. $\mathcal{V}^{\text{in}}$) the set of indices for which the diagonal dominance by column (resp. by row) is strict. The following are equivalent.*

1. *$\mathcal{G}(A)$ is input connected from $\mathcal{V}^{\text{in}}$ and output connected to $\mathcal{V}^{\text{out}}$.*

2. *$\mathcal{G}(A)$ has a spanning forest rooted at a subset of $\mathcal{V}^{\text{in}}$ and another spanning forest terminating at a subset of $\mathcal{V}^{\text{out}}$.*

3. *The system has a unique equilibrium point $x^* > 0$, and it is asymptotically stable with domain of attraction $\mathbb{R}^n_+$.*

Proof. We only prove the theorem for column dominance (the other case being identical).
1 $\iff$ 2. The equivalence is a straightforward application of Proposition 4.5.

1 $\Longrightarrow$ 3. From Theorem 6.27, A is Hurwitz, and, from Theorem 6.34, $-A^{-1} \geq 0$, hence, from Proposition 6.39, $x^* \geq 0$ asymptotically stable exists unique. From Theorem 6.24, A Mezler Hurwitz can be written as $A = B - \alpha I$ with $B \geq 0$ and $\alpha > \rho(B)$. Equivalently, $A = \alpha(\tilde{B} - I)$ where $\tilde{B} = B/\alpha$ is such that $\rho(\tilde{B}) < 1$. Hence for $\tilde{B}$ the Neumann series is converging, and we get

$$-A^{-1} = \frac{1}{\alpha}(I - \tilde{B})^{-1} = \frac{1}{\alpha}\sum_{j=0}^{\infty} \tilde{B}^j$$

Since $\mathcal{G}^{\rm nd}(A) = \mathcal{G}^{\rm nd}(B)$, while $\mathcal{G}^{\rm nd}(\tilde{B})$ differs only for a rescaling of the edges weights, we can apply to $\tilde{B}$ Proposition 4.7, which affirms that after a certain power ν, $\sum_{j=0}^{\nu} \tilde{B}^j$ has a number of positive subcolumns equal to the number of roots at most equal to card$(\mathcal{V}^{\rm in})$, and that these positive subcolumns touch all nodes of $\mathcal{V}$. Since $\tilde{B} \geq 0$, these positive subcolumns are present also in the Neumann series $\sum_{j=0}^{\infty} \tilde{B}^j$, hence they are present also in the expression for $-A^{-1}$.

By construction, the column indices of these subcolumns correspond to indices of the nonzero entries in the vector u, hence when taking the product $-A^{-1}u$ at least one nonzero (and therefore positive, since $-A^{-1} \geq 0$ and $u \geq 0$) term appears on each row of x^*. Therefore $x^* > 0$.

3 $\Longrightarrow$ 1. If $x^* > 0$ is asymptotically stable, then A must be Hurwitz (just apply the translation $z = x - x^*$). Therefore, from Theorem 6.27, $\mathcal{G}(A)$ has to be output connected. To show that $\mathcal{G}(A)$ is also input connected, let us reason by contradiction. Assume there exists a node, say $v_1 = 1$, that is not reachable by any of the nodes in $\mathcal{V}^{\rm in}$, and such that $u_1 = 0$. Assume for simplicity that this is the only node of $\mathcal{V}$ not reachable from $\mathcal{V}^{\rm in}$ (the situation is analogous in the more general case). Then we can deduce that A has to have an empty first row (diagonal element excluded). Consequently so do B and $\tilde{B}$, and all powers of $\tilde{B}$. Therefore, using again the Neumann series, also $-A^{-1}$ has a vanishing first row, diagonal element excluded. Since $u_1 = 0$, we have reached a contradiction, because in this way it cannot be $x_1^* > 0$. □

Example 6.43 For the system shown in the graph $\mathcal{G}(A_1)$ of Fig. 4.6(a), discussed in Examples 4.9 and 6.43, let us consider the terminal Frobenius normal form already introduced in Example 6.43. In particular, the matrix

$$A_1^{\rm term} = \left[\begin{array}{ccc|c|ccc} -1.8 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 \end{array}\right]$$

is Metzler Hurwitz and diagonally dominant by columns, with $\mathcal{V}^{\rm out} = \{1, 4\}$ (corresponding in the original $\mathcal{G}(A_1)$ of Fig. 4.6(a) to nodes 5 and 4). $\mathcal{V}^{\rm out}$ touches both the terminal strongly connected components, so the output connectivity condition of Theorem 6.42 is satisfied. Obtaining $x^* > 0$ requires choosing u such that we have also input connectivity, which means requiring that one of nodes of the last block of $A_1^{\rm term}$ has to have an input (in the original node enumeration of Fig. 4.6(a) this strongly connected component corresponds to nodes 1, 2, 3). In the current basis used for $A_1^{\rm term}$, for instance this corresponds to taking $\mathcal{V}^{\rm in} = \{7\}$, i.e.,

$$u = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^\top$$

and leads indeed to $x^* > 0$, see left panel in Fig. 6.8. In instead we choose as input set e.g., $\mathcal{V}^{\text{in}} = \{1, 4\}$ then we obtain only $x^* \geq 0$, and indeed the three components of the strongly connected component which is not terminal vanish, see right panel in Fig. 6.8. □

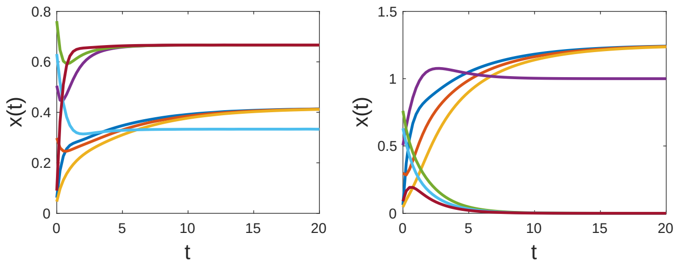


Figure 6.8: Example 6.43: positive (left) and nonnegative (right) equilibrium x^* for the CT affine positive system.

6.8.2 DT positive affine systems

In DT, consider the positive affine system (6.12) where $A \geq 0$ and $u \geq 0$. These conditions form a well-known necessary and sufficient condition for positivity of (6.12), see [8]. Given $u \geq 0$, an equilibrium of (6.12) is x^* such that $x^* = Ax^* + u$. Also in DT there are two possible cases:

- If $I - A$ nonsingular, then $(I - A)x^* = u$ admits a unique solution.
- If $I - A$ singular, then $(I - A)x^* = u$ admits an infinite number of solutions when $u \in \text{range}(I - A)$, no solution otherwise.

From Theorem 6.21, A is Schur stable if and only if $A - I$ is Metzler Hurwitz, hence we can directly map the main results obtained for the CT case into equivalent results for the DT case, given here without proofs.

Proposition 6.44 (Nonnegative/positive stable equilibria; DT case) *Consider the system (6.12) with $A \geq 0$. The two conditions are equivalent.*

1. *A is Schur stable.*
2. *For each $u \gneq 0$ there exists a unique $x^* \geq 0$ that solves $(I - A)x^* = u$, i.e., $x^* \geq 0$ is an equilibrium of (6.12) and it is asymptotically stable with domain of attraction $\mathbb{R}^n_+$.*

If in addition A is also irreducible, then for each $u \gneq 0$ the equilibrium point x^* is positive, $x^* > 0$.

As we have already seen in Section 6.7, the equivalent of diagonal dominance in DT is substochasticity. We can therefore formulate the conditions of Theorem 6.42 for reducible matrices A also in the DT case.

Theorem 6.45 (Positive stable equilibrium: DT reducible case) *Consider the system (6.12), and assume that $A \geq 0$ is substochastic by columns (resp. by rows). Consider $u \gneq 0$, and denote $\mathcal{V}^{\rm in}$ (resp. $\mathcal{V}^{\rm out}$) the set $\{i \in \mathcal{V}$ s. t. $u_i > 0\}$. Denote further $\mathcal{V}^{\rm out}$ (resp. $\mathcal{V}^{\rm in}$) the set of indices for which substochasticity by column (resp. by row) holds strictly. The following are equivalent.*

1. *$\mathcal{G}(A)$ is input connected from $\mathcal{V}^{\rm in}$ and output connected to $\mathcal{V}^{\rm out}$.*
2. *$\mathcal{G}(A)$ has a spanning forest rooted at a subset of $\mathcal{V}^{\rm in}$ and another spanning forest terminating at a subset of $\mathcal{V}^{\rm out}$.*
3. *The system has a unique equilibrium point $x^* > 0$, and it is asymptotically stable with domain of attraction $\mathbb{R}^n_+$.*

6.9 Stability of CT linear monotone systems

According to Propositions 6.12 and 6.14 a monotone linear system is characterized by a state matrix A whose graph $\mathcal{G}^{\rm nd}(A)$ is structurally balanced. The matrix A itself becomes Metzler after a change of orthant order. All conditions for stability of Metzler matrices discussed above can be extended in a somewhat trivial way to A monotone.

Proposition 6.46 *If $\exists$ a signature matrix $S = {\rm diag}(s_1, \ldots, s_n)$, $s_i = \pm 1$, such that $A_{\rm M} = SAS$ is Metzler and Hurwitz, then A is Hurwitz and the dominant eigenvector associated to $\mu(A)$ is Sv, where $v > 0$ solves $A_{\rm M} v = \mu(A_{\rm M}) v$.*

Proof. Since $S^{-1} = S$, SAS is a change of basis that does not change the eigenvalues, hence $\mu(A) = \mu(A_{\rm M}) < 0$ if $A_{\rm M}$ is Metzler and Hurwitz. From the usual eigenvalue-eigenvector equation for $A_{\rm M}$, $SASv = \mu(A)v$, one gets $ASv = \mu(A)Sv$. □

The orthant order s decides the signature of the dominant eigenvalue. Notice that this is also a fully-fledged Perron-Frobenius result, in the sense that all properties of PF are still valid. In order to compute the orthant order s the dominant eigenpair of A must be computed. This is computationally much less expensive than computing the entire spectrum of A.

Theorem 6.27 can be extended straightforwardly to A such that $\mathcal{G}^{\rm nd}(A)$ is structurally balanced.

Proposition 6.47 (Stability from diagonal dominance, structurally balanced case) *Consider A such that $\mathcal{G}^{\rm nd}(A)$ is structurally balanced, with $a_{ii} \leq 0$ for all i. A is Hurwitz if and only if A is weakly diagonally dominant and $\mathcal{G}^{\rm nd}(A)$ has a terminal spanning forest terminating at $\{v_1, \ldots, v_k\} \subseteq \mathcal{V}^{\rm out}$, where $\mathcal{V}^{\rm out}$ is the set of indices for which the diagonal dominance inequality (5.8) is strict.*

Chapter 7

Applications of network dynamics: a gallery

7.1 An application of DT positive systems: Markov chains

A *finite Markov chain* is a memoryless stochastic process, i.e., a sequence of random variables $\{X_t\}_{t=0,1,2,\dots}$ taking values in some alphabet $S_1, \dots S_n$, with the Markov property

$$\mathbb{P}\left(X_{t+1} = S_i | X_t = S_j, X_{t-1} = S_{j_{t-1}}, \dots, X_0 = S_{j_0}\right) = \mathbb{P}\left(X_{t+1} = S_i | X_t = S_j\right) = p_{ij}$$

where p_{ij} is the probability of transition from state S_j to state S_i. In words, the probability of the transition $S_j \to S_i$ does not depend on the history of the process. Denote $P = [p_{ij}]$ the transition matrix. Then $P \geq 0$, and P is a column stochastic matrix: $\mathbb{1}^\top P = \mathbb{1}^\top$. In fact the total probability of going from S_j to any other state must be one, which in formula reads $\sum_{i=1}^n p_{ij} = 1$. This implies that the columns of P are probability vectors

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \mathbb{P}(X = S_1) \\ \vdots \\ \mathbb{P}(X = S_n) \end{bmatrix} \quad \text{s. t.} \quad x_i \geq 0 \quad \text{and} \quad \mathbb{1}^\top x = 1$$

A *DT finite-state Markov chain* corresponds then to the positive system

$$x(t+1) = Px(t) \tag{7.1}$$

where $x(0)$ is a probability vector and P is a transition matrix. The Markov chain represents the chain of events that occur at the time points $t = 0, 1, 2, \dots$ in which $x(t)$ represents the state of the event that occurs at time t. The state $x(t)$ is a probability vector, and the equilibria of (7.1), x^* such that $x^* = Px^*$, are the stationary probability distributions of the Markov chain.

Notice that in the literature, instead of (7.1), you will normally see it represented with row vectors

$$x(t+1)^\top = x(t)^\top P^\top$$

with $P^\top$ a row-stochastic matrix. Our representation is of course equivalent, but more coherent with a system theory notation used in this book. If the probability vector $x(0)$ is the starting point of the chain, then $x(1) = Px(0)$ is also a probability vector:

$$\mathbb{1}^\top x(1) = \underbrace{\mathbb{1}^\top P}_{\mathbb{1}^\top} x(0) = \mathbb{1}^\top x(0) = 1$$

and so on:

$$\mathbb{1}^\top x(t) = 1 \quad \forall t \geq 0 \tag{7.2}$$

Eq. (7.2) is a conservation law for the dynamics of (7.1), which in this case corresponds to conservation of probability. We have already shown that for positive systems the positive orthant $\mathbb{R}_+^N$ is forward-invariant. Combining this fact with (7.2) corresponds geometrically to say that the evolution of $x(t)$ can only occur on the simplex $\{\mathbb{1}^\top x(t) = 1\} \cap \mathbb{R}_+^n$, see fig. 7.1.

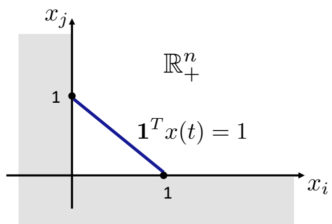


Figure 7.1: Conservation of probability in $\mathbb{R}_+^n$. In blue a 2D projection of the simplex which corresponds to the state space for (7.1).

The state after t steps is $x(t) = Px(t-1) = P^t x(0)$, i.e., P^t is the transition matrix of t steps. Since $x(t) \geq 0$ and $P \geq 0$, a Markov chain is a positive DT system, and we can apply Perron-Frobenius theorem to it. In particular, from $\mathbb{1}^\top P = \mathbb{1}^\top$, $\mathbb{1}$ is a left eigenvector of P of eigenvalue 1. If P is primitive, then $\rho(P) = 1$ is the spectral radius, since it corresponds to a positive eigenvector and there cannot be any other positive eigenvector (Theorem 3.10). When P is not irreducible, then only the weakest form of the Perron-Frobenius theorem (Theorem 3.5) holds, and from this theorem we cannot even conclude directly that $\rho(P) = 1$, let alone compute the multiplicity of the eigenvalue equal to $\rho(P)$.

Example 7.1 If $P = I$ then all eigenvalues of P are equal to 1 and all eigenvectors are ≥ 0. □

We can however conclude that $\rho(P) = 1$ indirectly, if we look at the conservation law (7.2), because

- if $\rho(P) > 1$ then the system (7.1) becomes unstable and $x(t)$ cannot be a probability vector as $x(t) \overset{t\to\infty}{\longrightarrow} \infty$;
- if $\rho(P) < 1$ then the system (7.1) becomes Shur stable and also in this case $\mathbb{1}^\top x(t) = 1$ is violated, as $x(t) \overset{t\to\infty}{\longrightarrow} 0$.

Hence the only possible case (compatible with the system (7.1) representing a probability vector) is $\rho(P) = 1$. Notice that we could have reached the same conclusion by simply applying Theorem 6.36 (condition 2).

What happen to the stability of the Markov chain (7.1)? From the column stochastic conditions, we expect to have marginal stability for P, see Section 6.7, but this need not imply that $\lim_{t\to\infty} x(t)$ exists. To understand when this leads to convergence of the Markov chain (7.1) (and why) it is convenient to analyze all possible cases in detail.

- P is irreducible

  1. P primitive. From Theorem 3.10, $\rho(P) = 1$ is a strictly dominant eigenvalue, and

$$\lim_{t\to\infty} P^t = \frac{v\mathbb{1}^\top}{\mathbb{1}^\top v} = v\mathbb{1}^\top = \begin{bmatrix} v & \dots & v \end{bmatrix} \tag{7.3}$$

where v is the right Perron-Frobenius eigenvector and we have chosen it normalized to 1: $\sum_{i=1}^n v_i = \mathbb{1}^\top v = 1$. Notice that $v\mathbb{1}^\top = \begin{bmatrix} v & \dots & v \end{bmatrix}$ is a rank-1 matrix having all columns equal to v. Consequently

$$\lim_{t\to\infty} x(t) = v \underbrace{\mathbb{1}^\top x(0)}_{=1} = v \tag{7.4}$$

i.e., all initial conditions converge to v, which is the unique stationary probability distribution of the Markov chain.

Notice that for the case $\rho(P) = 1$, Proposition 6.17 and Theorem 6.36 are predicting only marginal stability, and convergence is to $\text{span}(v)$ (or, more precisely, to the positive ray in the vector space $\text{span}(v)$), while (7.4) means convergence to a single point in $\mathbb{R}^n_+$, i.e., asymptotic stability. The reconciliation of these two results comes from the conservation law (7.2). In fact, since $v > 0$, $\text{span}(v)$ and the simplex $\mathbb{1}^\top x(t) = 1$ are always non-parallel, hence their intersection is given by a single point, see Fig. 7.2 for an illustration. In summary, even though we could infer only

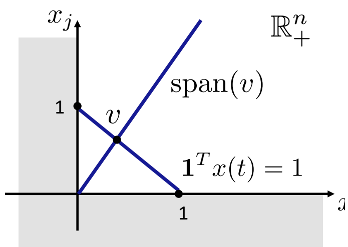


Figure 7.2: Stationary probability: intersection between $\text{span}(v)$ and the conservation law $\mathbb{1}^\top x(t) = 1$.

marginal stability, the extra constraints imposed by the structure of the problem imply that in reality the system (7.1) is asymptotic stable.

Another way to see it is to make use of a spectral projection on the orthogonal complement of the conservation law. Computing the Jordan form associated to P: $P = TJT^{-1}$, it is

$$J = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & J_1 \end{array}\right], \quad T = \left[\begin{array}{c|c} v & T_1 \end{array}\right], \quad T^{-1} = \begin{bmatrix} \mathbb{1}^\top \\ T_2 \end{bmatrix} \tag{7.5}$$

with $\rho(J_1) < 1$, meaning that

$$J^t \xrightarrow{t\to\infty} \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \end{array}\right] \quad \Longrightarrow \quad P^t \xrightarrow{t\to\infty} v \cdot 1 \cdot \mathbb{1}^\top = v\mathbb{1}^\top \tag{7.6}$$

In this spectral projection, the probability vector $x(t)$ is decomposed as

$$x(t) = \underbrace{v}_{\in \text{span}(v)} + \underbrace{y}_{\in \text{span}(v)^{\perp}}$$

and $y \overset{t\to\infty}{\longrightarrow} 0$.

2. P imprimitive, with a cyclicity index $r > 1$. In this case Theorem 3.10 and Proposition 6.17 do not hold. According to Section 6.5, the number of eigenvalues on the unit circle is equal to r, and these are radially equispaced (r roots of unity). In spite of marginal stability, $\lim_{t\to\infty} P^t$ does not exists generically, meaning that $\lim_{t\to\infty} x(t)$ is not defined. The probability vector $x(t)$ has a periodic pattern, of periodicity r. As mentioned in Section 3.1, a sufficient condition for primitivity is that at least one of the diagonal elements of P is positive, hence a necessary (but not sufficient) condition for imprimitivity is that $P_{ii} = 0$ for all i.

A word on terminology: in the literature, a Markov chain in which P is irreducible is often called an *ergodic* chain. It is called a *regular* chain if P is primitive. Regular chains are ergodic, but not viceversa, as the following example shows.

Example 7.2 If

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

P is irreducible, hence the associated Markov chain is ergodic. P is however imprimitive, i.e., the chain is not regular. In fact, $P^2 = I$, $P^3 = P$ and so on, meaning that $x(1) = Px(0)$, $x(2) = x(0)$, and so on. In particular, in $t = 1, 3, \ldots$ the values of the components of $x(t)$ are swapped w.r.t. $x(0)$, i.e., $x(t)$ oscillates with period $r = 2$. □

- P reducible. Let us assume (without loss of generality) that P is already in the Frobenius normal form (4.3), repeated here for convenience:

$$P = \hat{P} = \left[\begin{array}{c|c} \bar{P}_{11} & \bar{P}_{12} \\ \hline 0 & \bar{P}_{22} \end{array}\right] = \left[\begin{array}{cccc|cccc} \hat{P}_{11} & 0 & \ldots & 0 & \hat{P}_{1,k+1} & & \ldots & \hat{P}_{1,\ell} \\ 0 & \hat{P}_{22} & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & \ddots & & 0 & & & & \\ 0 & \ldots & 0 & \hat{P}_{kk} & \hat{P}_{k,k+1} & & \ldots & \hat{P}_{k,\ell} \\ \hline 0 & \ldots & & 0 & \hat{P}_{k+1,k+1} & \hat{P}_{k+1,k+2} & \ldots & \hat{P}_{k+1,\ell} \\ \vdots & & & & 0 & \hat{P}_{k+2,k+2} & \ddots & \vdots \\ \vdots & & & \vdots & \vdots & \ddots & \ddots & \hat{P}_{\ell-1,\ell} \\ 0 & \ldots & & 0 & 0 & \ldots & 0 & \hat{P}_{\ell\ell} \end{array}\right]. \tag{7.7}$$

In the context of Markov chains, its k terminal strongly connected components $\hat{P}_{11}, \ldots, \hat{P}_{kk}$ are normally called *ergodic* (or *closed*) *classes*, while the $\ell - k$ non-terminal strongly connected components are called *transient classes*. Each ergodic class represents itself a Markov chain with some external contribution coming from the transient classes that are connected to it in $\mathcal{G}(P)$. The probability of finding the chain in the states associated to the transient classes vanishes as $t \to \infty$, and asymptotically it gets completely absorbed by the states associated to the ergodic classes (also called *absorbing states*). Each ergodic class corresponds to a diagonal block $\hat{P}_{ii}$, $i = 1, \dots, k$, which is irreducible and column stochastic, as no edge is outgoing from these diagonal blocks. Hence $\rho(P) = 1$ is always an eigenvalue, and from Theorem 6.38, its multiplicity is equal to the number of ergodic classes (i.e., of terminal strongly connected components in $\mathcal{G}(P)$). The presence of other eigenvalues of modulus 1 depends on the primitivity of the $\hat{P}_{ii}$, $i = 1, \dots, k$, blocks. All the eigenvalues on the unit circle have nevertheless algebraic multiplicity equal to 1, meaning that P is always marginally stable, although $\lim_{t\to\infty} x(t)$ may or may not exist.

The following theorem summarizes the results on existence and uniqueness of stationary distributions in Markov chains (including the irreducible case as a special case).

Theorem 7.3 (Convergence in Markov chains) *Consider the Markov chain* (7.1) *in the Frobenius normal form* (7.7).

1. $\lim_{t\to\infty} P^t$ *exists if and only if all terminal strongly connected components of* $\mathcal{G}(P)$ *are primitive (i.e., all ergodic classes are regular). When it exists,*

$$\lim_{t\to\infty} P^t = \left[\begin{array}{c|c} R & S \\ \hline 0 & 0 \end{array}\right] \tag{7.8}$$

   where

$$R = \begin{bmatrix} v_1 \mathbb{1}_{n_1}^\top & & \\ & \ddots & \\ & & v_k \mathbb{1}_{n_k}^\top \end{bmatrix} \quad \text{and} \quad S = R\bar{P}_{12}(I - \bar{P}_{22})^{-1} \tag{7.9}$$

   with v_i *the right eigenvector of* $\hat{P}_{ii}$ *relative to* $\rho(\hat{P}_{ii}) = 1$*, normalized so that* $v_i^\top \mathbb{1}_{n_i} = 1$*, and* $\mathbb{1}_{n_i}$ *is the vector of 1 of size* $n_i = \dim(\hat{P}_{ii})$*.*

2. $x^* = \lim_{t\to\infty} x(t)$ *exists unique and is asymptotically stable for all* $x(0)$ *if and only if the graph* $\mathcal{G}(P)$ *has a terminal spanning tree (i.e., it has a single ergodic class) with associated terminal strongly connected component which is primitive (i.e., the single ergodic class is regular). When* x^* *exists, it is equal to* $x^* = v$*, with* $v = \begin{bmatrix} v_1 \\ 0 \end{bmatrix}$.

Proof. Condition 1. In the Frobenius normal form (7.7), each diagonal block $\hat{P}_{11}, \dots, \hat{P}_{kk}$ is column stochastic, while each diagonal block $\hat{P}_{k+1,k+1}, \dots, \hat{P}_{\ell\ell}$ is column substochastic, with strict inequality in at least one column in each block. From condition 1 of Theorem 6.36, the blocks $\hat{P}_{k+1,k+1}, \dots, \hat{P}_{\ell\ell}$ are all Schur stable, meaning that $\bar{P}_{22}$ is Schur stable. Since $\hat{P}$ is block triangular, so is $\hat{P}^t$, for any $t \in \mathbb{N}$, with diagonal blocks $\hat{P}^t_{jj}$. In particular, $\rho(\bar{P}_{11}) = 1$ and $\rho(\bar{P}_{22}) < 1$, and the k eigenvalues equal to 1 in $\bar{P}_{11}$ belong to k different Jordan blocks by construction. Hence the powers of $\bar{P}_{11}$ are bounded and each Jordan block admits a decomposition such as (7.5), meaning that the limit (7.6) exists. Therefore

$$\hat{P}^\infty = \lim_{t\to\infty} \hat{P}^t = \lim_{t\to\infty} \left[\begin{array}{c|c} \bar{P}_{11}^t & \star \\ \hline 0 & \bar{P}_{22}^t \end{array}\right] = \left[\begin{array}{c|c} R & S \\ \hline 0 & 0 \end{array}\right]$$

where R has the expression in (7.9). Concerning the expression for S in (7.9), this is obtained e.g. in [20] (p. 698) by exploiting the fact that, since all k dominant eigenvalues equal to 1 belong to different Jordan blocks, $\hat{P}^\infty$ is the projector onto $\ker(I-\hat{P})$ along $\mathrm{range}(I-P)$, hence $\mathrm{range}(\hat{P}^\infty) = \ker(I-\hat{P})$ and therefore also $\mathrm{range}(I-\hat{P}) = \ker(\hat{P}^\infty)$, which can be written as $\hat{P}^\infty(I-\hat{P}) = 0$. i.e.,

$$\left[\begin{array}{c|c} R & S \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c|c} I-\bar{P}_{11} & -\bar{P}_{12} \\ \hline 0 & I-\bar{P}_{22} \end{array}\right] = 0$$

and leads to (7.9).

Condition 2 can be proven combining the results of Theorems 6.36 and 6.38 with the existence of the conservation law (7.2). If $\mathcal{G}(P)$ has a terminal spanning tree with associated strongly connected component which is primitive, then in the Frobenius normal form (7.7), $k=1$, and the diagonal block $\hat{P}_{11}$ is column stochastic. From Theorems 6.38, then, the eigenvalue $\rho(P)=1$ is simple and strictly dominating all other eigenvalues. Hence the expressions (7.3) and (7.4) are both valid. The converse can be easily shown by contradiction, relying still on Theorems 6.38. Both cases of $\mathcal{G}(P)$ lacking a terminal spanning tree (i.e., having two or more terminal strongly connected components) and of imprimitive terminal strongly connected component lead to a Jordan form for P which has more than one eigenvalue of modulus 1 (either $\rho(P)=1$ appearing in $k\geq 2$ Jordan blocks each of algebraic multiplicity 1, or eigenvalues of the form $\lambda_j = e^{\frac{i2\pi j}{r}}$, $j=0,\ldots,r-1$ for some $r>1$, or both). In any case, P^t does not converge to a rank-1 matrix, hence $\lim_{t\to\infty} x(t)$ does not exist for a generic initial condition. Concerning the value of x^*, the expression for the stationary probability distribution is

$$x^* = \hat{P}^\infty x(0) = \left[\begin{array}{c|c} R & S \\ \hline 0 & 0 \end{array}\right] x(0).$$

When condition 2 holds (i.e., when $k=1$ in (7.7) and (7.9)), the expression for R simplifies to $R = v_1 \mathbb{1}_{n_1}^\top$ and that for S to $S = v_1 \mathbb{1}_{n_2}^\top$, where $n_2 = \dim(\bar{P}_{22})$. In fact, from column stochasticity of P,

$$\mathbb{1}^\top \begin{bmatrix} \bar{P}_{12} \\ \bar{P}_{22} \end{bmatrix} = \mathbb{1}_{n_1}^\top \bar{P}_{12} + \mathbb{1}_{n_2}^\top \bar{P}_{22} = \mathbb{1}_{n_2}^\top$$

from which one gets

$$\mathbb{1}_{n_1}^\top \bar{P}_{12}(I-\bar{P}_{22})^{-1} = \mathbb{1}_{n_2}^\top. \tag{7.10}$$

Splitting x into x_1 (of dimension n_1) and x_2 (of dimension n_2) and using (7.10)

$$\begin{aligned} x_1^* &= v_1 \mathbb{1}_{n_1}^\top x_1(0) + v_1 \mathbb{1}_{n_1}^\top \bar{P}_{12}(I-\bar{P}_{22})^{-1} x_2(0) \\ &= v_1 \mathbb{1}_{n_1}^\top x_1(0) + v_1 \mathbb{1}_{n_2}^\top x_2(0) \\ &= v_1 \underbrace{\left(\mathbb{1}_{n_1}^\top x_1(0) + \mathbb{1}_{n_2}^\top x_2(0)\right)}_{=1} \\ x_2^* &= 0. \end{aligned}$$

Asymptotic stability follows combining $\rho(\bar{W}_{22}) < 1$ with the same considerations on the "transversality" of the conservation law (7.2) we discussed earlier for the P irreducible case. □

It is worth emphasizing that under Condition 2 of the theorem, $\hat{P}^\infty$ is a rank-1 matrix: $\hat{P}^\infty = v\mathbb{1}^\top$, but based on the partially vanishing eigenvector v.

Example 7.4 The column stochastic matrix

$$P = \left[\begin{array}{cccc|cc} 0.2 & 0.4 & 0 & 0 & 0.15 & 0.17 \\ 0.8 & 0.6 & 0 & 0 & 0.15 & 0.2 \\ 0 & 0 & 0.7 & 0.5 & 0.11 & 0.1 \\ 0 & 0 & 0.3 & 0.5 & 0.22 & 0.23 \\ \hline 0 & 0 & 0 & 0 & 0 & 0.3 \\ 0 & 0 & 0 & 0 & 0.37 & 0 \end{array}\right]$$

is already in the Frobenius normal form (7.7): it has two ergodic classes both of dimension 2 and a transient class also of dimension 2. Both the ergodic classes are primitive. Hence condition 1 of Theorem 7.3 holds, but not condition 2. Using (7.9), for $\hat{P}^\infty$ it is

$$R = \begin{bmatrix} 0.333 & 0.333 & 0 & 0 \\ 0.667 & 0.667 & 0 & 0 \\ 0 & 0 & 0.625 & 0.625 \\ 0 & 0 & 0.375 & 0.375 \end{bmatrix}, \quad \text{and} \quad S = \begin{bmatrix} 0.157 & 0.149 \\ 0.313 & 0.299 \\ 0.331 & 0.345 \\ 0.199 & 0.207 \end{bmatrix}.$$

Notice how each diagonal block in R is a rank-1 matrix as expected. When evolving the Markov chain (7.1), the stationary probability distribution depends on the initial condition $x(0)$. See examples in Fig. 7.3.

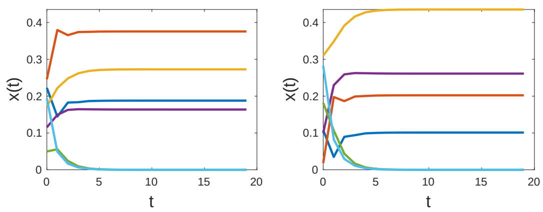


Figure 7.3: Example 7.4. Different initial conditions gives different absorbing states x^*.

7.2 Another application of DT positive systems: PageRank algorithm

PageRank is the algorithm at the base of the Google search engine. It was the feature that enabled this search engine to make a difference with its competitors, with the results we all know. (Nowadays it is said that there are more than 200 methods that make the page ranking in Google, and PageRank is just one of them).

The philosophy of PageRank is that a web page is important if it is pointed to by other important web pages. The web is represented as a graph of hyperlinks (i.e. of snippets of html code pointing to other pages). The outgoing links from a page are represented as columns of a sub-stochastic matrix. A matrix $A \geq 0$ is column sub-stochastic if $\mathbb{1}^\top A \leq \mathbb{1}^\top$.

Example 7.5 A toy internet of 4 nodes is shown in Figure 7.4. Node 1 has 3 outgoing hyperlinks, node 2 has 1, node 3 has 2 and node 4 has 0. The link matrix is the sub-stochastic matrix

(a) (b)

Figure 7.4: Pagerank graph of Example 7.5. (a): original graph with dangling nodes. (b): strongly connected graph with back edges.

$$A = \begin{bmatrix} 0 & 0 & 1/2 & 0 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1 & 0 & 0 \\ 1/3 & 0 & 1/2 & 0 \end{bmatrix}$$

Since the 4th columns is identically 0, A is not column stochastic. □

The basic search method consists of a random surfer model (or crawler), which works by browsing web pages and clicking randomly on hyperlinks to get to new web pages, and so on. Statistics and importance measures are then computed for the web pages visited. A crawler can run into two types of problems:

1. Dangling nodes: nodes without outlinks, meaning that the crawler stops there forever, unless e.g. one adds a back edge to the dangling nodes. In this case, the matrix becomes column stochastic.

   Example 7.6 (Example 7.5 cont'd). Node 4 is a dangling node. Assuming we get there from node 3, and add a back edge to node 3 as in Figure 7.4(b), the new link matrix is

   $$A = \begin{bmatrix} 0 & 0 & 1/2 & 0 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1 & 0 & 1 \\ 1/3 & 0 & 1/2 & 0 \end{bmatrix}$$

   which is now column stochastic. □

2. Lack of strong connectivity. If a group of nodes is strongly connected but isolated from the rest of the network, then the crawler stays there forever.

The solution in the latter problem is to restart the crawler somewhere else choosing a node at random (this is called "teleportation" in the language of Google search engine). On these jumps all nodes can for instance be equiprobable. In the PageRank algorithm, this is implemented by making a convex combination between the link matrix A and a matrix of all 1:

$$B = (1-m)A + \frac{m}{n}\mathbb{1}\mathbb{1}^\top, \qquad m \in (0,1)$$

The convex combinator m is chosen equal 0.15 in PageRank. B is positive and still column stochastic:

$$\mathbb{1}^\top B = (1-m)\underbrace{\mathbb{1}^\top A}_{\mathbb{1}^\top} + \frac{m}{n}\underbrace{\mathbb{1}^\top\mathbb{1}}_{n}\mathbb{1}^\top = (1-m)\mathbb{1}^\top + m\mathbb{1}^\top = \mathbb{1}^\top$$

From this expression, applying Perron-Frobenius theorem for positive matrices (Theorem 3.4) we have that $\rho(B) = 1$ is an eigenvalue of left eigenvector $\mathbb{1}$, and that the right eigenvector v is positive. v is the PageRank, i.e., the ranking of web-pages suggested by Google. It is an eigenvalue centrality measure for internet. By construction it satisfies $v = Bv$. To compute v in practice, one can set up an iteration scheme:

$$x(t+1) = Bx(t) \tag{7.11}$$

where $x(0)$ is a probability vector $x_i(0) \geq 0$, $\mathbb{1}^\top x(0) = 1$. In numerical linear algebra, the system (7.11) is called the *power method* for computing the dominating eigenvalue/eigenvector pair of a positive matrix.

In the terminology we saw in Section 7.1, B is a transition matrix, it is irreducible and even primitive by construction. We are in the case 1 of the Markov chains described in Section 7.1: (7.11) is marginally stable, but since we have the conservation law $\mathbb{1}^\top x(t) = 1$, asymptotic stability to v follows for all $x(0)$ which are probability vectors. If the eigenvalues of B, $\lambda_i(B)$, are sorted according to their modulus, then $\lambda_1(B) = \rho(B) = 1$ and $\lambda_2(B)$ is such that $|\lambda_2(B)| < 1$. $\lambda_2(B)$ is called the algebraic connectivity, and decides the speed of convergence of the power method. In fact, for (7.11) convergence is exponential, with rate

$$\left|\frac{\lambda_1(B)}{\lambda_2(B)}\right| = \frac{1}{|\lambda_2(B)|}$$

For PageRank it is known that $|\lambda_2(B)| < 1 - m = 0.85$, i.e., the error decays as $(0.85)^t$ for each calculation step. For instance, in 50 steps, $(0.85)^{50} \sim 10^{-4}$. If m is bigger than 0.15 then the convergence is faster, but the calculation gives a lower importance to the link matrix A. For internet, the size of A is $\sim 10^{10}$ (there are billions of web-pages). The computation of v is said to take around 1 week, and to occur once every month (but these could be very outdated estimates...).

7.3 An application of CT positive affine systems: Compartmental systems

Compartmental systems (also called dynamical flow networks) are distributed dynamical systems that describe flows of mass/substance among units called compartments, see the survey

[13]. They are used in many applications (biology, bioengineering, hydrology, economy, industrial and transport systems, etc). The assumption that is normally done is that a compartment is homogeneous i.e., the substance that enters a compartment is instantaneously mixed with the substance already present. A typical model of a compartmental system is shown in Fig. 7.5. The terms u_i represent inflow from the environment outside the system, the edges F_{0i} outflows

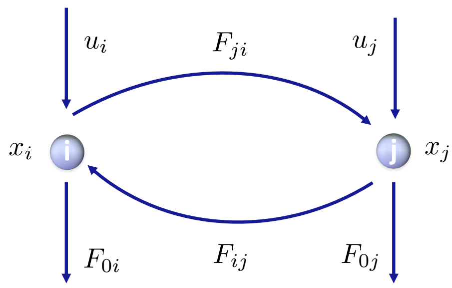


Figure 7.5: Compartmental system.

to outside the system, and F_{ij} transfer flow from compartment j to compartment i. To each compartment i is associated a state variable x_i representing concentration of mass or substance. Since mass or substance is a nonnegative quantity, it is natural to assume that $x_i \geq 0$, hence the natural setting in which to model compartmental systems is that of positive systems. We also put constraints on the flows: $u_i \geq 0$, $F_{0i} \geq 0$ and $F_{ij} \geq 0$. Notice that in general the flows are function of the state $F_{ij} = F_{ij}(x)$. Mass-balance at compartment i gives

$$\dot{x}_i = \underbrace{-F_{0i}(x) - F_{1i}(x) - \ldots - F_{ni}(x)}_{\text{outflows}} + \underbrace{F_{i1}(x) + \ldots + F_{in}(x) + u_i}_{\text{inflows}} \tag{7.12}$$

In oder to guarantee that the state stays positive it has to be that when $x_i = 0$ then $F_{0i}(x) = F_{ji}(x) = 0$. This can be imposed by assuming that the $F_{ji}(x)$ (including $F_{0i}(x)$) are all homogeneous function of degree at least 1 in the "donor" state x_i (i.e. the state of the node upstream to the edge): $F_{ji}(x) = h_{ji}(x)x_i$. In this way, forward invariance in $\mathbb{R}^n_+$ is obtained, see Prop. 6.6.

7.3.1 Linear compartmental systems

Assuming a linear model means considering the functions $F_{ji}(x) = h_{ji}x_i$, i.e., the flow functions are all linear in the donor node. In particular, the term $-h_{0i} \leq 0$ is called the outflow rate coefficient from node i because it represents an outflow of mass from the i-th node. Rewriting (7.12) in the linear case:

$$\dot{x}_i = -(h_{0i} + \sum_{j \neq i} h_{ji})x_i + \sum_{j \neq i} h_{ij}x_j + u_i \tag{7.13}$$

Expressing the i-th diagonal element as $a_{ii} = -(h_{0i} + \sum_{j\neq i} h_{ji})$ and the off-diagonal elements as $a_{ij} = h_{ij}$, we have a state update matrix

$$A = \begin{bmatrix} -h_{01} - \sum_{j\neq 1} h_{j1} & h_{12} & \dots & h_{1n} \\ h_{21} & -h_{02} - \sum_{j\neq 2} h_{j2} & & \\ \vdots & & & \\ h_{n1} & & & -h_{0n} - \sum_{j\neq n} h_{jn} \end{bmatrix} \tag{7.14}$$

and the system (7.13) in vector form is $\dot{x} = Ax + u$, where $u = \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix}^\top$, i.e., it is in the form (6.11). In particular, we have

- The matrix A in (7.14) is Metzler by construction, and it obeys to the condition $\mathbb{1}^\top A \leq 0$. Metzler matrices such that $\mathbb{1}^\top A \leq 0$ are called *compartmental matrices*.
- The condition $\mathbb{1}^\top A \leq 0$ corresponds to diagonal dominance. In particular, $\mathbb{1}^\top A = -h_0^\top = -\begin{bmatrix} h_{01} & \dots & h_{0n} \end{bmatrix} \leq 0$, i.e., the vector of outflow rate coefficients h_0 determines what in Section 6.8 we denoted output set: $\mathcal{V}^{\text{out}} = \{i \in \mathcal{V} \text{ s. t. } h_{0i} > 0\}$.
- the inflow vector u determines what in Section 6.8 we denoted input set: $\mathcal{V}^{\text{in}} = \{i \in \mathcal{V} \text{ s. t. } u_i > 0\}$.

In the compartmental systems literature [3, 2, 13], one uses frequently notions such as inflow connectivity, outflow connectivity, and traps. In particular, a trap is defined as a set of compartments without outflows to the rest of the system or to the environment. Traps may have subtraps inside. Simple traps are defined as traps not containing any subtrap. These can be rephrased in our terminology as follows:

- *Simple trap*

  = terminal strongly connected component of $\mathcal{G}(A)$ having empty intersection with $\mathcal{V}^{\text{out}}$.

  = terminal strongly connected component of $\mathcal{G}(A)$ whose associated block diagonal submatrix in the Frobenius normal form (4.2) of A is diagonally equipotent.

- *Inflow connectivity*

  = input connectivity of $\mathcal{G}(A)$.

  = existence of a spanning forest for $\mathcal{G}(A)$ rooted at (a subset of) $\mathcal{V}^{\text{in}}$.

- *Outflow connectivity*

  = output connectivity of $\mathcal{G}(A)$.

  = existence of a spanning forest for $\mathcal{G}(A)$ terminating at (a subset of) $\mathcal{V}^{\text{out}}$.

  = existence of a spanning forest for $\mathcal{G}(A^\top)$ rooted at (a subset of) $\mathcal{V}^{\text{out}}$.

  = absence of traps.

The results of Section 6.8, in particular Theorem 6.27 and Theorem 6.29, can be used to investigate linear compartmental systems. Consider a compartmental matrix A. A is always either Hurwitz or marginally stable. In particular A nonsingular means Hurwitz. We distinguish the following cases.

- A irreducible
  - A is Hurwitz if and only if $\mathbb{1}^\top A \lneq 0$ (Theorem 6.26). In this case it is $a_{ii} < 0$ and $-A^{-1} > 0$ (Theorem 6.34).
  - Consequently, A is singular if and only if $\mathbb{1}^\top A = 0$. In this case, it is marginally stable, with multiplicity of the $\lambda = 0$ eigenvalue equal to 1 (the system has a single simple trap) (Theorem 6.26).
  - Any $u \gneq 0$ is such that the system (6.11) is inflow connected (Proposition 6.40).
- A reducible
  - A is Hurwitz if and only if A is outflow connected (i.e., it has no traps) (Theorem 6.27). In this case it is $a_{ii} \leq 0$ and $-A^{-1} \geq 0$ (Theorem 6.34).
  - Consequently, A is singular if and only if it is not outflow connected (i.e., it has traps). In this case, A is marginally stable, with multiplicity of the $\lambda = 0$ eigenvalue equal to the number of terminal strongly connected components of $\mathcal{G}(A)$ for which the associated block diagonal matrix in (4.3) is diagonally equipotent (each such block gives a simple trap) (Theorem 6.29).
  - How about $x^* > 0$? Does it exist (perhaps non-unique)?

Analogous of the theorems of Section 6.8 can be set up. For instance for Theorem 6.42 we have the following.

Theorem 7.7 (Stability and positive equilibria in linear compartmental systems) *Consider the positive affine system (6.11) with A compartmental. The following conditions are equivalent.*

1. *$\mathcal{G}(A)$ is input connected from $\mathcal{V}^{\text{in}}$ and output connected to $\mathcal{V}^{\text{out}}$.*
2. *The system has a unique strictly positive equilibrium point $x^* = -A^{-1}u > 0$, and it is asymptotically stable with domain of attraction $\mathbb{R}^n_+$.*

In summary, in linear compartmental systems we can distinguish the following cases.

Case 1: $h_{0i} = 0$ for all i (i.e., the compartmental system has no outflow), and $u_i = 0$ for all i (i.e., the system has no inflow). The system (6.11) has a global conservation law:

$$\mathbb{1}^\top A = 0 \quad \Longrightarrow \quad \mathbb{1}^\top A x = \mathbb{1}^\top \dot{x} = 0 \quad \Longrightarrow \quad \mathbb{1}^\top x(t) = \text{const} \ \forall\, t$$

i.e., the total mass is conserved during the evolution, and in particular it is equal to $\mathbb{1}^\top x(0)$ = initial mass. In this case we have ∞-many equilibrium points, one for each conservation law. Since these form a continuum, none of them can be asymptotically stable.

Case 2: $h_{0i} = 0$ for all i and $u_i > 0$ for some i. The system accumulates mass:

$$1^\top \dot{x} = \underbrace{\mathbb{1}^\top A x}_{=0} + \mathbb{1}^\top u > 0$$

Consequently, the system cannot have an equilibrium point x^*, since the sum of the derivatives of the concentrations is strictly increasing. If we have inflow connectivity, then $x_i(t) \to +\infty$ for all i. If A irreducible, then any $u \gneq 0$ leads to $x_i(t) \to +\infty$ for all i.

Case 3: $h_{0i} > 0$ for some i, and $u_i = 0$ for all i. It is

$$1^\top \dot{x} = \underbrace{\mathbb{1}^\top A}_{\leq 0} x + 0 \leq 0$$

and the system dissipates. If $x_i > 0$ for all i (i.e., we start from nonzero concentration in all variables), the dissipation is certainly strict: $1^\top \dot{x} < 0$. In particular, if A irreducible, we have asymptotic stability, but since $u_i = 0 \; \forall \, i$, the only equilibrium point is $x^* = 0$. The same hold if instead of irreducibility we have outflow connectivity.

Case 4: $h_{0i} > 0$ for some i, and $u_j > 0$ for some j. Irreducibility of A leads to existence of a unique, asymptotically stable positive equilibrium point $x^* > 0$. The same holds if, in place of irreducibility, we have both inflow and outflow connectivity.

7.3.2 DT linear compartmental systems

A DT linear compartmental system can be obtained from (7.13) through analogous mass transfer and mass conservation reasoning. We assume that mass transfer between compartments or with the external environment happens at the discrete time instants $t = 0, 1, \ldots$, in a synchronous way for all compartments. Using the notation of Section 7.3.1, we can express the update law via an Euler discretization as

$$x_i(t+1) = \Big(1 - h_{0,i} - \sum_{j \neq i} h_{ji}\Big) x_i(t) + \sum_{j \neq i} h_{ij} x_j(t) + u_i(t). \tag{7.15}$$

Since the outflow from compartment i at time instant t can be at most equal to $x_i(t)$, it must be

$$\Big(h_{0,i} + \sum_{j \neq i} h_{ji}\Big) x_i(t) \leq x_i(t)$$

meaning that

$$1 - h_{0,i} - \sum_{j \neq i} h_{ji} \geq 0$$

i.e., (7.15) has positive diagonal elements. In matrix form, this can be written as

$$x(t+1) = Bx(t) + u(t) \tag{7.16}$$

where $B = I + A$, with A given in (7.14). Obviously $\mathcal{G}^{\text{nd}}(B) = \mathcal{G}^{\text{nd}}(A)$. i.e., the flow routing is the same, only the representation (in particular the self-loops) change when passing from CT to DT. The system (7.16) has the following properties.

- The matrix B is nonnegative and it obeys to the condition $\mathbb{1}^\top B \leq \mathbb{1}^\top$. By analogy to the CT case, such matrices are called *DT compartmental matrices*.
- The system (7.16) is a positive system.
- The conditions $\mathbb{1}^\top B \leq \mathbb{1}^\top$ and $B \geq 0$ imply that B is column *substochastic*, i.e., $\rho(B) \leq 1$.
- Inflows are determined by $\mathcal{V}^{\text{in}} = \{i \in \mathcal{V} \text{ s.t. } u_i > 0\}$, and outflows by $\mathcal{V}^{\text{out}} = \{i \in \mathcal{V} \text{ s.t. } h_{0,i} > 0\}$.

- Inflow connectivity (resp. outflow connectivity) can be computed in terms of spanning forests of $\mathcal{G}(B)$ rooted at $\mathcal{V}^{\text{in}}$ (resp. terminating at $\mathcal{V}^{\text{out}}$).
- Similarly to the CT model, by construction B can be either Schur stable or marginally stable. From Theorem 6.21, Schur stability of $B = I + A$ (i.e., $\rho(B) < 1$) is equivalent to Hurwitz stability of the Metzler matrix A.

Recall that in the DT case, marginal stability corresponds to having one or more simple eigenvalues on the circle of radius $\rho(B) = 1$. Since $B \geq 0$, from Theorem 3.5, $\rho(B) = 1$ is always an eigenvalue. To obtain a complete description of all the dominant eigenvalues λ such that $|\lambda| = \rho(B) = 1$, it is convenient to distinguish the following cases (analogous splitting was done for CT compartmental systems).

- B is irreducible
  - B is Schur stable if and only if $\mathbb{1}^\top B \lneq \mathbb{1}^\top$. In fact, in order to have $\rho(B) = 1$ as an eigenvalue, it has to be $\mathbb{1}^\top B = \mathbb{1}^\top$.
  - If instead $\mathbb{1}^\top B = \mathbb{1}^\top$, then $\rho(B) = 1$ is an eigenvalue, and we have two cases
    1. If B is primitive, $\rho(B) = 1$ is strictly dominating all other eigenvalues (Theorem 3.10).
    2. If instead B is imprimitive of cyclicity index r, then B has r dominating eigenvalues, simple and radially equispaced on the unit circle as described in Section 6.5.

    In both cases B is marginally stable, see Theorem 6.36.
  - Any $u \gneq 0$ is such that the system (7.16) is inflow connected.
- B is reducible
  - B is Schur stable if and only if it B is outflow connected (i.e., it has no traps), see Theorem 6.37.
  - If instead B is not outflow connected (i.e., it has traps) then B is marginally stable. With reference to the Frobenius normal form (4.3), and denoting $B_{11}, \ldots, B_{kk}$ the terminal diagonal blocks of B, the multiplicity of the 1 eigenvalue is equal to the number of terminal diagonal blocks for which $\mathbb{1}^\top B_{ii} = \mathbb{1}^\top$. The presence of other eigenvalues on the unit circle depends on the primitivity of the blocks B_{ii}, and so does their multiplicity (each imprimitive diagonal block B_{ii} is contributing a number of radially equispaced eigenvalues equal to its cyclicity index r_i). See Theorem 6.38 for the details.

The four cases analyzed in Section 7.3.1 for CT compartmental systems are valid also for DT, after the necessary minor adjustments.

7.4 Another application of (nonlinear) CT positive systems: network epidemic models

A couple of references for this section are [7, 19].

Epidemic models form a vast class of nonlinear positive systems, widely used to describe the spread of infectious diseases in a population, but also of information, rumours, innovations, etc. in a community.

Here we are interested in investigating these spreading processes over a network. Before that however, it is convenient to introduce the various classes of models on a single compartment (i.e., a "single node", if you want).

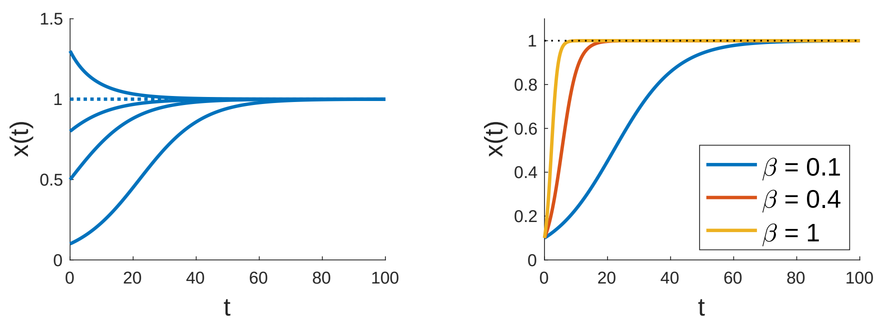
7.4.1 Scalar epidemic models

Logistic growth.

A basic model used to describe spreading in a population is the logistic curve, i.e., the scalar curve

$$x(t) = \frac{c x_o}{x_o + (c - x_o) e^{-\beta t}} \tag{7.17}$$

where $x(t) \in \mathbb{R}_+$ represents a population, $x_o = x(0)$ is the initial condition, the constant $c > 0$ is the carrying capacity of the environment and $\beta > 0$ is the intrinsic growth rate. The profile of (7.17) for different values of x_o and β is shown in Fig. 7.6.



(a) (b)

Figure 7.6: Logistic curve. (a): different initial conditions x_o. (b): different growth rates β. $c = 1$ in both plots.

The ODE of which (7.17) is the solution is known as the *logistic equation*:

$$\dot{x} = \beta x \left(1 - \frac{x}{c}\right) \tag{7.18}$$

The ODE (7.18) helps understanding and interpreting the sigmoidal behavior of (7.17).

- When $x(t) \ll c$ then (7.18) can be approximated with a growing exponential $\dot{x} \approx \beta x$, i.e., far from the carrying capacity there is exponential growth;
- When $x(t) \approx c$ then $x(t)$ is on a plateau, i.e., growth tends to vanish near the carrying capacity.

The presence of a negative quadratic term in (7.18) puts a brake on the growth, and makes the model more realistic: resources are not unlimited, and when the population approaches c its growth is limited by scarcity of resources. For some applications x_o can exceed the carrying capacity c, but it will nevertheless converge to it as t grows.

The reasoning behind the logistic equation is basic to the epidemic models which are discussed next. In these models the carrying capacity is $c = 1$, and the population $x(t)$ has also an interpretation as fraction of a total population, i.e., the carrying capacity is never exceeded.

SI model.

"SI" stays for "susceptible" and "infected", the two possible states of this model, which will be represented here by the two state variables s and x. These variables represent fractions of a population, and hence are both nonnegative and constrained in the interval [0, 1]:

$$s(t) \in [0,\, 1], \qquad x(t) \in [0,\, 1] \qquad \forall\, t \geq 0,$$

and are linked by a conservation law:

$$s(t) + x(t) = 1 \qquad \forall\, t \geq 0 \tag{7.19}$$

The model has just the parameter β, here normally called the infection rate, $\beta > 0$. The ODE for the infected fraction is of quadratic type: the rate of infection is proportional to the amount of contacts between susceptibles s and infected x (sometimes called a "mass-action" law)

$$\dot{x} = \beta s x = \beta(1 - x)x \tag{7.20}$$

As the second expression shows, the constraint (7.19) makes this mass-action principle into the logistic equation (7.18). From (7.19), we also get $\dot{s} + \dot{x} = 0$ hence the ODE for s is not required, as it can be deduced from (7.20) (it is simply $\dot{s} = -\beta s(1 - s)$). The solution of (7.20) is the logistic curve:

$$x(t) = \frac{x_o e^{\beta t}}{1 - x_o + x_o e^{\beta t}} = \frac{x_o}{x_o + (1 - x_o)e^{-\beta t}} \qquad x_o = x(0) \in [0,\, 1]$$

The states $x^* = 0$ (infection eradication) and $x^* = 1$ (full contagion) are both equilibria of (7.20). $x^* = 1$ is asymptotically stable for all trajectories with $x_o > 0$.

SIS model.

"SIS" stays for "susceptible-infected-susceptible" and, in addition to β, there is another process described, the recovery of infected individuals, represented by a term linear in x containing a second parameter, the recovery rate γ, $\gamma > 0$. The basic ODE becomes:

$$\dot{x} = \beta s x - \gamma x = \beta x(1 - x) - \gamma x = (\beta - \gamma)x - \beta x^2 \tag{7.21}$$

For $\beta \neq \gamma$ the solution can again be computed in closed form:

$$x(t) = \frac{(\beta - \gamma)x_o}{\beta x_o - (\gamma - \beta(1 - x_o))e^{-(\beta - \gamma)t}} \qquad x_o = x(0) \in [0,\, 1]$$

To describe its asymptotic behaviour, it is convenient to introduce the so-called basic reproductive number $R_o = \frac{\beta}{\gamma}$ (corresponding to the average fraction of individuals that an infected individual can infect during its infection). Using R_o, (7.21) can be rewritten as

$$\dot{x} = \gamma(R_o - 1)x - \beta x^2 = \beta x(1-x) - \gamma x \tag{7.22}$$

which admits one or two equilibrium points, depending on R_o:

1. when $R_o \leq 1$, the only admissible equilibrium point of (7.22) is $x^* = 0$, which is asymptotically stable for all $x_o \in [0, 1]$, meaning that the infection is eradicated.

2. when $R_o > 1$, then (7.22) admits a second equilibrium point $x^* = \frac{R_o - 1}{R_o} = 1 - \frac{\gamma}{\beta}$ which is asymptotically stable for all $x_o > 0$, while $x^* = 0$ becomes instead unstable. This case corresponds to endemic infection.

Notice that the bifurcation between the two cases happens exactly at $R_o = 1$ (where the two equilibria coalesce into each other in the origin). Notice further that in both cases the stability character of the linearization at $x^* = 0$ is straightforwardly seen from the Jacobian linearization of (7.22) (i.e., $\dot{x} = \gamma(R_o - 1)x$). When $R_o > 1$, to check the local stability of $x^* = 1 - \frac{\gamma}{\beta}$, it is convenient to consider the associated ODE for s:

$$\dot{s} = (\gamma - \beta s)(1 - s)$$

Computing the Jacobian linearization at $s^* = \frac{\gamma}{\beta}$, one get the linearized system $\dot{s} = \gamma(1 - R_o)s$ which is asymptotically stable for $R_o > 1$.

Bass model

All population models seen so far are homogeneous in x, meaning that they all have $x^* = 0$ as an equilibrium point. Consequently, the spreading has to be initiated by a seed, i.e., at least one infected individual. This make perfect sense for epidemic applications, but if we are considering to use contagion-like rules for describing spreading in other contexts, it may be a limitation. For instance when we consider diffusion of information in social media, or diffusion of innovation in consumer products. One model that is frequently used in the latter context and that is now adopted also for social media [17, **?**] is the so-called *Bass model* [4]:

$$\dot{x} = \gamma(1 - x) + \beta x(1 - x) \tag{7.23}$$

The variable $x \in [0, 1]$ represents the cumulative fraction of adopters of a product, and implicitly we assume that $\lim_{t\to\infty} x(t) = 1$. The parameters $\gamma \geq 0$ and $\beta \geq 0$ are called respectively the coefficient of innovation and of imitation. These names highlight the roles that are normally associated to the two terms in (7.23): the second term represents the rate of imitators, i.e., adopters under the influence of word-of-mouth (or e.g. recommender systems for online media). The first term represents instead the rate of new adopters that are innovators, i.e., that adopt a product on their own initiative, without relying on e.g. word-of-mouth or without copying some else behavior. The model is substantially similar to the SIS model (7.21) with the exception of an extra term equal to γ in the right hand side terms, which however changes its behavior and makes it more similar to that of an SI model, with the exception that the first term in (7.23) can now initiate the process of diffusion without need of an external seed.

The natural use of the Bass model is in fact in correspondence of $x(0) = 0$, case in which the explicit solution of (7.23) is

$$x(t) = \frac{1 - e^{-(\gamma+\beta)t}}{1 + \frac{\beta}{\gamma} e^{-(\gamma+\beta)t}}$$

Notice that $\lim_{t\to\infty} x(t) = 1$ for all $x(0) \in [0,\, 1]$, i.e., $x^* = 1$ is asymptotically stable for all trajectories.

7.4.2 Network epidemic models

The models described in the previous section are now replicated for multiple compartments, which we identify with the nodes of a network. So at each node we have the same state variables as in SI or SIS models, but with a subindex that identifies the node. Coupling between the nodes is also included in the dynamics, and it is provided by the adjacency matrix A of the graph, which we assume nonnegative. We consider an heterogeneous model, i.e., we assume that different compartments have different values of the parameters β_i (and, when present, γ_i etc.). The other difference between the nodes is given by the entries of the adjacency matrix A, which is weighted and in general directed.

Network SI model.

The equation for the fraction of infected population at node i is

$$\dot{x}_i = \beta_i(1 - x_i) \sum_{j=1}^{n} a_{ij} x_j, \qquad \beta_i > 0, \qquad i = 1, \dots, n \tag{7.24}$$

Eq. 7.24, which can be rewritten as $\dot{x}_i = \beta_i s_i \sum_{j=1}^{n} a_{ij} x_j$, expresses the fact that susceptibles at node i may meet infected at node i but also infected at node j. The edge weights a_{ij} express the "importance" (or the frequency) of these intra-group (a_{ii}) or inter-group (a_{ij}) contacts. We assume $A \geq 0$. The corresponding equation for susceptibles can still be obtained from the conservation law $s_i = 1 - x_i$: $\dot{s}_i = -\beta_i s_i \sum_{j=1}^{n} a_{ij}(1 - s_j)$, $i = 1, \dots, n$. In vector form, if $x = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^\top$, (7.24) becomes

$$\dot{x} = B(I - \mathrm{diag}(x))Ax = (I - \mathrm{diag}(x))BAx \tag{7.25}$$

where $\mathrm{diag}(x)$ is the diagonal matrix having $x_1, \dots, x_n$ on the diagonal, $B = \mathrm{diag}(\beta_1, \dots, \beta_n)$ and we have used the fact that diagonal matrices commute. As for the vector of susceptibles, from $s = \mathbb{1} - x$ one gets

$$\dot{s} = -\mathrm{diag}(s)BA(\mathbb{1} - s) \tag{7.26}$$

Theorem 7.8 (Network SI stability) *Assume $\mathcal{G}(A)$ strongly connected, $A \geq 0$. The system (7.25) is invariant in $[0,\, \mathbb{1}]^n$: $x(0) \in [0,\, 1]^n \implies x(t) \in [0,\, 1]^n$ for all t (and, of course, also $s(t) \in [0,\, 1]^n$ for all t). It has two equilibrium points*

1. *disease-free: $x^* = 0$ unstable;*
2. *full infection: $x^* = \mathbb{1}$ asymptotically stable.*

For all $x(0) \neq 0$, *it is* $\lim_{t\to\infty} x(t) = \mathbb{1}$.

Proof. To show forward invariance in $[0, 1]^n$ for all t, simply observe that $x(0) \in [0, 1]^n$ means $I - \text{diag}(x(0)) \geq 0$. In addition, $BA \geq 0$ irreducible. Hence (7.25) behaves like a positive system, and never escapes the positive orthant (Proposition 6.6). Similarly, when x_i approaches the upper boundary, $1-x_i \to 0$ which leads to $\dot{x}_i \to 0$, hence the set $[0, 1]^n$ is indeed forward invariant for (7.25). $x^* = 0$ and $x^* = \mathbb{1}$ are clearly equilibrium points. Showing that there is no other equilibrium point can be done by contradiction. Assume $\bar{x} \in [0, 1]^n$ is another equilibrium point, $\bar{x} \neq \{0, \mathbb{1}\}$. Then $\exists\, i$ such that $\bar{x}_i \neq \{0, 1\}$, meaning that $0 < 1 - \bar{x}_i < 1$. In order to have $\dot{\bar{x}}_i = 0$, it must be $\sum_{j=1}^n \beta_i a_{ij} \bar{x}_j = 0$, i.e., $\bar{x}_j = 0$ for all j for which $a_{ij} > 0$. The argument can now be repeated for all these indexes j, and iterated. Since BA is irreducible, the argument eventually involves all indices, including i, hence we have a contradiction. Computing the Jacobian linearization of (7.25):

$$J = BA - \text{diag}(x)BA - \text{diag}(BAx)$$

Computed around $x^* = 0$ it gives the linearized system $\dot{x} = BAx$ which is unstable, since for $BA \geq 0$ irreducible we can apply Perron-Frobenius (Theorem 3.7) and get that $\rho(BA) > 0$ is an (unstable) eigenvalue. The linearization around $x^* = \mathbb{1}$ can be easily expressed in terms of s (as linearization around $s^* = 0$ of (7.26)): $\dot{s} = -\text{diag}(BA\mathbb{1})s$. It is asymptotically stable since the vector $BA\mathbb{1}$ is positive, hence so is (locally) the nonlinear system (7.25). To show that $\lim_{t\to\infty} x(t) = \mathbb{1}$ for all $x(0) \neq 0$, consider the linear Lyapunov function $V(x) = \mathbb{1}^\top(\mathbb{1} - x)$. It is $V(x) > 0$ in $[0, 1]^n \setminus \{\mathbb{1}\}$, and $V(\mathbb{1}) = 0$. Furthermore, $\dot{V}(x) = -\mathbb{1}^\top \dot{x} = -\mathbb{1}^\top((I - \text{diag}(x))BAx \leq 0$ in $[0, 1]^n$, and $\dot{V}(x) = 0$ only in 0 and $\mathbb{1}$. Since $V(0) = n$, from $\dot{V}(x) \leq 0$, it follows that all trajectories must converge to $x^* = \mathbb{1}$. □

Network SIS model.

The heterogeneous network SIS model has parameters $\beta_i > 0$ and $\gamma_i > 0$ that change from node to node, and obeys at each node to the ODE:

$$\dot{x}_i = \beta_i(1 - x_i)\sum_{j=1}^n a_{ij}x_j - \gamma_i x_i, \qquad \beta_i > 0, \quad \gamma_i > 0, \quad i = 1, \ldots, n$$

Denoting $B = \text{diag}([\beta_1, \ldots, \beta_n])$ and $\Gamma = \text{diag}([\gamma_1, \ldots, \gamma_n])$, the network SI model can be written in vector form as

$$\dot{x} = (-\Gamma + BA - \text{diag}(x)BA)\, x \tag{7.27}$$

where $A \geq 0$. Obviously invariance in $[0, 1]^n$ follows from the previous analysis of SI models. Using $s = \mathbb{1} - x$ and the identity $\text{diag}(x)C\mathbb{1} = \text{diag}(C\mathbb{1})x$ valid for any $C \in \mathbb{R}^{n\times n}$, we obtain for s:

$$\dot{s} = \Gamma\mathbb{1} - (\Gamma + \text{diag}(BA\mathbb{1}) - \text{diag}(s)BA)\, s$$

In (7.27), $BA \geq 0$ is irreducible, while $-\Gamma$ is a Mezler Hurwitz matrix with a (negated) inverse which is nonnegative: $\Gamma^{-1} \geq 0$ (trivial). The basic reproductive number R_o for the model (7.27) is expressed as the spectral radius of the following nonnegative matrix:

$$R_o = \rho(\Gamma^{-1}BA)$$

This definition plays the same role as R_o defined for the scalar SI model (7.21). To understand why, we need to use the notion of regular splitting introduced in Proposition 6.32. The link is explained in the proof of Theorem 7.9.

It turns out that also for a network SI model R_o decides the number of equilibria and their stability character.

Theorem 7.9 (Network SIS stability) *Consider the system (7.27), with $\mathcal{G}(A)$ strongly connected and $A \geq 0$.*

1. $R_o = \rho(\Gamma^{-1}BA) \leq 1 \iff$ *the system admits a unique equilibrium point $x^* = 0$ (disease-free), which is asymptotically stable for all $x(0) \in [0, \mathbb{1}]^n$;*

2. $R_o = \rho(\Gamma^{-1}BA) > 1 \iff$ *the system admits two equilibria: $x^* = 0$ (disease-free) which is unstable, and $x^* = \bar{x} > 0$ (endemic epidemic) which is asymptotically stable. In particular, $x(0) \neq 0 \implies \lim_{t\to\infty} x(t) = \bar{x}$.*

Proof. The full proof can be found in [7], with some material taken from [16]. Here we just sketch some parts of it. Clearly $x^* = 0$ is always an equilibrium point of (7.27) regardless of R_o.

To show the local stability/instability of $x^* = 0$, notice that the Jacobian of the linearization of (7.27) at 0 is $-\Gamma + BA$. We can make use of Proposition 6.32: the matrix $-\Gamma + BA$ is Metzler by construction and, since $\gamma_i > 0$ for all i, $-\Gamma$ is Metzler Hurwitz, while $BA \geq 0$, hence the splitting of the Jacobian into $-\Gamma$ and BA is regular. Therefore $\mu(-\Gamma + BA) < 0 \iff R_o = \rho(\Gamma^{-1}BA) < 1$, i.e., $x^* = 0$ is (locally, based on the linearization) asymptotically stable if and only if $R_o < 1$.

When $R_o = 1$, then nothing can be said on the stability using the linearization, as $\mu(-\Gamma + BA) = 0$.

Let us show that when there is another equilibrium $\bar{x} \geq 0$ then it must be $\bar{x} > 0$ and also $R_o = \rho(\Gamma^{-1}BA) > 1$. Consider $\bar{x} \geq 0$ ($\bar{x} \neq 0$). By construction, at an equilibrium point

$$\bar{x} = \Gamma^{-1}BA\bar{x} - \mathrm{diag}(\bar{x})\Gamma^{-1}BA\bar{x} = \mathrm{diag}(\mathbb{1} - \bar{x})\Gamma^{-1}BA\bar{x} \tag{7.28}$$

Since $-\Gamma$ is Metzler Hurwitz (and hence nonsingular), it is $\Gamma^{-1} \geq 0$ (this is trivial, as $\gamma_i > 0$). Irreducibility of $A \geq 0$ implies irreducibility of $\Gamma^{-1}BA \geq 0$. If $\bar{x}_i = 0$, then in (7.28) it must be $\sum_{j=1}^n \frac{\beta_i}{\gamma_i} a_{ij}\bar{x}_j = 0$, and, by applying a contradiction argument similar to the one in the proof of Theorem 7.8, it can be shown that the entire $\bar{x}$ must vanish, hence $\bar{x}$ must be positive: $\bar{x}_i > 0$ for all i. Then $\Gamma^{-1}BA\bar{x} > 0$ and also $\mathrm{diag}(\bar{x})\Gamma^{-1}BA\bar{x} > 0$, which, in its turn, in (7.28) leads to $\bar{x} < \Gamma^{-1}BA\bar{x}$. Applying Proposition 6.35, we get $1 < \rho(\Gamma^{-1}BA) = R_o$.

Hence for $R_o = \rho(\Gamma^{-1}BA) < 1$ there cannot be any other equilibrium point. To show that $x^* = 0$ is asymptotically stable for all $x(0)$ (and not just locally) consider the following Lyapunov function: $V(x) = c^\top x$ where $c > 0$ is the left Perron-Frobenius eigenvector of $-\Gamma + BA$: $c^\top(-\Gamma + BA) = c^\top \mu(-\Gamma + BA)$. Differentiating along the trajectories of (7.27):

$$\begin{aligned}\dot{V}(x) =& c^\top \left(-\Gamma + BA - \mathrm{diag}(x)BA\right) x \\ =& c^\top \mu(-\Gamma + BA) - \underbrace{c^\top \mathrm{diag}(x)BAx}_{\geq 0}\end{aligned}$$

When $R_o = \rho(\Gamma^{-1}BA) < 1$ then, from Proposition 6.32, $\mu(-\Gamma + BA) < 0$, hence $\dot{V}(x) < 0$. When instead $R_o = \rho(\Gamma^{-1}BA) = 1$, it is $\mu(-\Gamma + BA) = 0$, hence $\dot{V}(x) \leq 0$. To complete the argument also in this case we can resort to the LaSalle invariance principle: the largest invariant set in the level surfaces $\{x \text{ s.t. } \mathrm{diag}(x)BAx = 0\}$ is given by $x^* = 0$, because of the irreducibility of BA (use the same contradiction argument mentioned above). Hence for $R_o = \rho(\Gamma^{-1}BA) \leq 1$, $\lim_{t\to\infty} x(t) = 0 \; \forall\, x(0) \in [0,1]^n$. The proof that for $R_o = \rho(\Gamma^{-1}BA) > 1$ there is always a second (unique) positive equilibrium point $\bar{x} > 0$ is rather complicated. In [7] it is based on a Brouwer fixed point arguments.

The intuition is the following: if $\bar{x}$ is an equilibrium point of components $0 < \bar{x}_i < 1$, then (7.28) can be written as

$$\mathcal{F}(\bar{x}) := (\mathrm{diag}(\mathbb{1} - \bar{x}))^{-1}\,\bar{x} = \begin{bmatrix} \frac{\bar{x}_1}{1-\bar{x}_1} \\ \vdots \\ \frac{\bar{x}_n}{1-\bar{x}_n} \end{bmatrix} = \Gamma^{-1}BA\bar{x}$$

Consider any $x \in (0,1)^n$. Each $\frac{x_i}{1-x_i}$ is monotonically increasing in x_i. When all components are very small, $x_i = x_i^{\text{small}}$, then $\frac{x_i^{\text{small}}}{1-x_i^{\text{small}}} \sim x_i^{\text{small}}$, and, since $\rho(\Gamma^{-1}BA) > 1$, it is plausible to assume that $\mathcal{F}(x^{\text{small}}) \sim x^{\text{small}} < \rho(\Gamma^{-1}BA)x^{\text{small}}$, while when all components $x_i = x_i^{\text{big}}$ are very close to 1, $\frac{x_i^{\text{big}}}{1-x_i^{\text{big}}} \gg x_i^{\text{big}}$, i.e., $\mathcal{F}(x^{\text{big}}) > \rho(\Gamma^{-1}BA)x^{\text{big}}$. Hence it is reasonable to assume that in between x^{small} and x^{big} there is an equilibrium point $\bar{x}$. The Brouwer fixed point argument requires that one shows that $[x^{\text{small}}, x^{\text{big}}]$ is invariant for the dynamics, which we skip.

For the endemic equilibrium point $\bar{x}$ we can show asymptotic stability with domain of attraction $[0,1]^n \setminus \{0\}$. Define $\tilde{x} = x - \bar{x}$ and consider the associated ODE

$$\dot{\tilde{x}} = \underbrace{(-\Gamma + BA - \mathrm{diag}(\bar{x})BA)}_{:=\Phi(\bar{x})}\tilde{x} - \mathrm{diag}(\tilde{x})BAx$$

The matrix $\Phi(\bar{x})$ is Metzler, irreducible and such that, at the equilibrium point $\bar{x} > 0$, $\Phi(\bar{x})\bar{x} = 0$, meaning that its spectral abscissa is $\mu(\Phi(\bar{x})) = 0$ and $\bar{x}$ is the Perron-Frobenius eigenvector of $\Phi(\bar{x})$. We can invoke Proposition 7.10, valid for singular, stable Metzler matrices. Then $\exists$ $D = \mathrm{diag}(d_1, \ldots, d_n)$, $d_i > 0$ such that

$$\Phi^\top(\bar{x})D + D\Phi(\bar{x}) \leq 0$$

For this D, consider the Lyapunov function $V(\tilde{x}) = \tilde{x}^\top D\tilde{x}$. Clearly $V(\tilde{x}) > 0 \; \forall\, \tilde{x} \neq 0$, $V(0) = 0$. Differentiating:

$$\begin{aligned} \dot{V}(\tilde{x}) &= \underbrace{\tilde{x}^\top(\Phi^\top(\bar{x})D + D\Phi(\bar{x}))\tilde{x}}_{\leq 0} - 2\tilde{x}^\top \mathrm{diag}(\tilde{x})DBAx \\ &\leq -2\tilde{x}^\top \mathrm{diag}(\tilde{x})DBAx \end{aligned}$$

Since $DBA \geq 0$ irreducible, and $\tilde{x}^\top \mathrm{diag}(\tilde{x}) = \begin{bmatrix} \tilde{x}_1^2 & \ldots \tilde{x}_n^2 \end{bmatrix}$, it is

$$\dot{V}(\tilde{x}) \leq -2\sum_{i=1}^n \tilde{x}_i^2 d_i \beta_i \sum_{j=1}^n a_{ij}x_j$$

Each term $\tilde{x}_i^2 d_i \beta_i \sum_{j=1}^n a_{ij}x_j \geq 0$, hence $\dot{V}(\tilde{x}) = 0$ only when they all vanish. Using again the same contradiction argument based on irreducibility, $\dot{V}(\tilde{x}) = 0$ only when $x = \bar{x}$ or $x = 0$, meaning that $\dot{V}(\tilde{x}) < 0$ elsewhere. Given that $V(\tilde{x}) = 0$ only in $\tilde{x}$, LaSalle invariance principle can be applied and leads to the desired conclusion. □

Proposition 7.10 *Consider the system $\dot{x} = Ax$ with A Metzler and irreducible. If $\mu(A) = 0$ then there exists a diagonal matrix $D = \mathrm{diag}(d_1, \ldots, d_n)$, $d_i > 0$ such that $A^\top D + DA$ is negative semidefinite.*

Network Bass model

An heterogeneous network Bass model can be obtained using principles similar to those used for the SI and SIS models. In particular we obtain at each node

$$\dot{x}_i = \gamma_i(1 - x_i) + \beta_i(1 - x_i)\sum_{j=1}^{n} a_{ij}x_j, \qquad \beta_i > 0, \quad \gamma_i > 0, \quad i = 1, \ldots, n$$

or in vector form,

$$\dot{x} = \Gamma(\mathbb{1} - x) + (I - \mathrm{diag}(x))\, BAx \tag{7.29}$$

where, as before, $A \geq 0$, $B = \mathrm{diag}([\beta_1, \ldots, \beta_n])$ and $\Gamma = \mathrm{diag}([\gamma_1, \ldots, \gamma_n])$.

Theorem 7.11 (Network Bass model: equilibrium and stability) *Consider the system (7.29), with $\mathcal{G}(A)$ strongly connected and $A \geq 0$. The system is invariant in $[0, \mathbb{1}]^n$: $x(0) \in [0, 1]^n \implies x(t) \in [0, 1]^n$ for all t. It has a single equilibrium point $x^* = \mathbb{1}$ asymptotically stable for all $x(0) \in [0, 1]^n$.*

Proof. Existence and uniqueness of the equilibrium point $x^* = \mathbb{1}$ are straightforward: both $\mathbb{1} - x$ and $I - \mathrm{diag}(x)$ are nonnegative and vanishing only in x^*, plus $BA \geq 0$ and irreducible. Asymptotic stability of $x^* = \mathbb{1}$ follows if we use the same Lyapunov function $V(x) = \mathbb{1}^\top(\mathbb{1} - x)$ and reasoning used in Theorem 7.8 for the network SI model. □

7.5 Yet another application of (nonlinear) CT positive systems: Lotka-Volterra system

The Lotka-Volterra model is a classical model of ecological population dynamics. Let the vector $x \geq 0$ represent the concentration of n species, interacting through a matrix A. The model is of quadratic form

$$\dot{x} = \mathrm{diag}(x)\,(Ax + r) \tag{7.30}$$

where $r \in \mathbb{R}^n$ is the vector of intrinsic growth rates. In components the model reads

$$\dot{x}_i = x_i\left(r_i + \sum_{j=1}^{n} a_{ij}x_j\right), \qquad i = 1, \ldots, n \tag{7.31}$$

which shows that what is modeled is the effect of pairwise "encounters" between species, according to a "mass-action" principle, with interaction coefficients given by the matrix A, plus linear growth/decay (depending on r_i and its sign). The coefficients of A are classified according to their sign:

1. $a_{ij} > 0$ means that species i has a benefit in encounters with species j;
2. $a_{ij} < 0$ means that species i has instead a loss from encounters with species j.

In the ecological literature, the coefficients of A are also classified in pairwise interactions as follows:

1. mutualistic interaction: $a_{ij} > 0$, $a_{ji} > 0$;
2. predator-prey interaction: $a_{ij} > 0$, $a_{ji} < 0$;
3. competitive interactions: $a_{ij} < 0$, $a_{ji} < 0$;

Notice that pairs of coefficients that are not simultaneously non-zero are also possible, i.e., $(a_{ij}, a_{ji}) = (+, 0)$, $(-, 0)$, etc. In addition, diagonal elements of A are typically nonpositive (a species consumes resources): $a_{ii} \leq 0$. When A is diagonal, the growth is logistic in each species.

From (7.31) it is evident that $x_i(0) = 0 \Longrightarrow x_i(t) = 0$ for all t, from which it follows that $x_i(0) > 0 \Longrightarrow x_i(t) \geq 0$ for all t, i.e., $\mathbb{R}^n_+$ is forward invariant for (7.30), hence its state vector is suitable to represent species concentrations.

We now investigate equilibria of (7.30). Clearly $x^* = 0$ is an equilibrium point, corresponding to extinction of all species.Furthermore, as just mentioned, $x_i(0) = 0$ implies $x_i(t) = 0$ for all t, i.e., extinct species cannot reappear (which is a reasonable properties for an ecological model). We are in particular interested in equilibria $x^* = \bar{x} > 0$, in correspondence of which all species persist in the ecosystem (i.e., no species is becoming extinct). The Jacobian linearization of (7.30) is

$$J(x) = \frac{\partial\left(\mathrm{diag}(x)(Ax + r)\right)}{\partial x} = \mathrm{diag}(x)A + \mathrm{diag}(Ax) + \mathrm{diag}(r) \tag{7.32}$$

At $x^* = 0$, $J_0 = \mathrm{diag}(r)$. The dynamics of each species is locally decoupled from the others, and for each of them $\dot{x}_i = r_i x_i$. If $r < 0$, then $x^* = 0$ is locally asymptotically stable. As soon as $r_i > 0$ for some i then $x^* = 0$ becomes unstable. Consider a positive equilibrium $x^* = \bar{x} > 0$. If A is invertible, then from $A\bar{x} + r = 0$ we obtain $\bar{x} = -A^{-1}r$ and the Jacobian at $\bar{x}$ becomes

$$J_{\bar{x}} = \mathrm{diag}(-A^{-1}r)A \tag{7.33}$$

The Jacobian matrix (7.33) is normally referred to as the community matrix of the ecosystem.

When $\bar{x} > 0$, $J_{\bar{x}}$ and A only differ for a premultiplication by a diagonal pd matrix (i.e., $\mathrm{diag}(\bar{x})$). We can therefore make use of the notion of D-stability (Definition 5.7): when A is D-stable then we are guarantee that $J_{\bar{x}}$ is Hurwitz, and hence that $\bar{x}$ is at least locally stable. A stronger result is obtained if instead of D-stability we ask for diagonal stability of A (definition 5.6): in this case we have also global stability in the "feasible region", that is for all $x > 0$. Recall that Hurwitz stability, D-stability and diagonal stability are linked by the hierarchy (5.6).

Theorem 7.12 (Stability of Lotka-Volterra systems via diagonal stability [10]) *Consider a Lotka-Volterra system (7.30). Assume that A is invertible, and that $\bar{x} > 0$ is an equilibrium point. If A is diagonally stable, then $\bar{x}$ is asymptotically stable with domain of attraction* $\mathrm{int}(\mathbb{R}^n_+)$ *(i.e. $\forall\, x(0) > 0$).*

Proof. The proof can be found in [10, 14]. It makes use of the linear-logarithmic Lyapunov function

$$V(x) = \sum_{i=1}^{n} p_i \left(x_i - \bar{x}_i - \bar{x}_i \log\left(\frac{x_i}{\bar{x}_i}\right)\right) \tag{7.34}$$

where $p_i > 0$ are the diagonal elements of the diagonal matrix $P = \mathrm{diag}(p_1, \ldots, p_n)$ that guarantees diagonal stability of A. Notice first that $V(x)$ is defined for all $x \in \mathrm{int}(\mathbb{R}^n_+)$ and that, as long as $\bar{x} > 0$, $V(\bar{x}) = 0$. Furthermore, $V(x)$ is separable in the variables $x_1, \ldots, x_n$, and can be written as

$$V(x) = \sum_{i=1}^n p_i V_i(x_i), \quad \text{where} \quad V_i(x_i) = x_i - \bar{x}_i - \bar{x}_i \log\left(\frac{x_i}{\bar{x}_i}\right).$$

The time derivative of $V_i(x_i)$ is $\dot{V}_i(x_i) = 1 - \frac{\bar{x}_i}{x_i}$. For $0 < x_i < \bar{x}_i$ it is $\dot{V}_i(x_i) < 0$, while for $x_i > \bar{x}_i$ it is $\dot{V}_i(x_i) > 0$, meaning that $V_i(x_i)$ has a unique minimum at $x_i = \bar{x}_i$ for all $x_i > 0$. Since $V(x)$ is separable in the variables x_i, its global minimum in $\mathrm{int}(\mathbb{R}^n_+)$ is at the point $\bar{x}$, where each component has its minimum. Computing the time derivative of V in $\mathrm{int}(\mathbb{R}^n_+)$,

$$\begin{aligned}\dot{V}(x) &= \sum_{i=1}^n p_i(x_i - \bar{x}_i)\frac{\dot{x}_i}{x_i} = \sum_{i=1}^n p_i(x_i - \bar{x}_i)\Big(r_i + \sum_{j=1}^n a_{ij}x_j\Big) \\ &= \sum_{i=1}^n p_i(x_i - \bar{x}_i)\sum_{j=1}^n a_{ij}(x_j - \bar{x}_j) = \frac{1}{2}(x-\bar{x})^\top(PA + A^\top P)(x - \bar{x})\end{aligned}$$

where we have used that, at $\bar{x}$, $r_i = -\sum_{j=1}^n a_{ij}\bar{x}_j$. Clearly $\dot{V}(x)$ is negative definite if $PA + A^\top P$ is negative definite, i.e., if A is diagonally stable.

□

The two complications with Theorem 7.12 are that existence of $\bar{x} > 0$ is not guaranteed, and that checking diagonal stability is not straightforward, at least analytically. It get simpler if we add extra assumption like for instance that A is normal, because in this case Hurwitz stability of A suffices, see Lemma 5.9.

Corollary 7.13 (Stability of Lotka-Volterra systems with normal A) *In the Lotka-Volterra system (7.30), assume that A is normal and Hurwitz, and that $\bar{x} > 0$ is an equilibrium point. Then $\bar{x}$ is asymptotically stable with domain of attraction* $\mathrm{int}(\mathbb{R}^n_+)$.

The situation gets even simpler in the mutualistic case discussed next, because existence of a positive equilibrium $\bar{x} > 0$ can be guaranteed for certain sign patterns of r.

7.5.1 Stability for mutualistic Lotka-Volterra

In the mutualistic case, A is a Metzler matrix. Assume that we have nonnegative intrinsic growth rates $r \geq 0$ (which is the most common case). Then if A Hurwitz, from Theorem 6.34, we know that $-A^{-1} \geq 0$, hence in (7.33) $-A^{-1}r \geq 0$, meaning that also $J_{\bar{x}}$ is Metzler. The simplest case to treat is when A is irreducible.

Theorem 7.14 (Stability of mutualistic Lotka-Volterra systems) *Consider a mutualistic Lotka-Volterra system (7.30), i.e., assume that A is Metzler. Assume further that A is irreducible. Then the following conditions are equivalent:*

1. *A is Metzler Hurwitz;*

2. *For each $r \gneq 0$ there exists a unique positive equilibrium point $\bar{x} = -A^{-1}r > 0$ and it is asymptotically stable with domain of attraction $\text{int}(\mathbb{R}^n_+)$;*

3. *the Jacobian $J_{\bar{x}}$ is Metzler Hurwitz.*

Proof. 1 $\iff$ 2: follows from Propositions 6.39 and 6.40. The extra term $\text{diag}(x)$ in (7.30) is ininfluent when $x = \bar{x} > 0$.
1 $\iff$ 3: Recall from Theorem 6.24 that for Metzler matrices D-stability, diagonal stability and Hurwitz stability are identical concepts. From Theorem 6.24, we have that A Metzler Hurwitz if and only if it is D-stable. From (7.33), $J_{\bar{x}}$ is obtained from A via a premultiplication with the positive definite diagonal matrix $D = \text{diag}(\bar{x}) = \text{diag}(-A^{-1}r)$. Since D^{-1} is also diagonal and positive definite, A and $J_{\bar{x}}$ are Metzler Hurwitz simultaneously. Attractivity for all $x(0) > 0$ follows from Theorem 7.12 and can be shown via a Lyapunov argument based on the function (7.34). □

Irreducibility of A Metzler can be relaxed, as least as long as we impose diagonal dominance. In this case, in fact, the tools developed in Section 6.6.1 and already used for compartmental systems in Section 7.3 can be applied. The following is the adaptation of Theorems 6.27 and 6.42 to Lotka-Volterra systems.

Proposition 7.15 (Stability of reducible mutualistic Lotka-Volterra systems) *Consider a mutualistic Lotka-Volterra system (7.30) in which $r \geq 0$ and the matrix A is Metzler and diagonally dominant by columns: $\mathbb{1}^\top A \leq 0$. Denote $\mathcal{V}^{\text{in}} = \{i \in \mathcal{V}$ s. .t. $r_i > 0\}$ and $\mathcal{V}^{\text{out}} = \{\in \mathcal{V}$ s. .t. $\sum_j a_{ij} < 0\}$. Then the following conditions are equivalent:*

1. *A is Metzler Hurwitz;*

2. *$\mathcal{G}(A)$ is output connected to $\mathcal{V}^{\text{out}}$.*

Furthermore, also the following conditions are equivalent.

1. *$\mathcal{G}(A)$ is input connected from $\mathcal{V}^{\text{in}}$ and output connected to $\mathcal{V}^{\text{out}}$.*

2. *The system has a unique strictly positive equilibrium point $\bar{x} = -A^{-1}r > 0$ and it is asymptotically stable with domain of attraction $\text{int}(\mathbb{R}^n_+)$.*

7.5.2 Extension to structurally balanced interaction matrices

The result can be extended to the case of $\mathcal{G}^{\text{nd}}(A)$ structurally balanced.

Theorem 7.16 (Stability of structurally balanced Lotka-Volterra systems) *Consider a Lotka-Volterra system (7.30) in which $\mathcal{G}^{\text{nd}}(A)$ is structurally balanced with signature vector s, and A is invertible and irreducible. Assume $\bar{x} > 0$ is an equilibrium point. Then the following are equivalent:*

1. *SAS is Metzler Hurwitz, where $S = \text{diag}(s)$;*

2. *$\bar{x}$ is asymptotically stable (and an attractor for all $x(0) > 0$);*

3. *The Jacobian $SJ_{\bar{x}}S$ is Metzler Hurwitz.*

Proof. Since $S^{-1} = S$, SAS is a change of basis, hence A and SAS have the same eigenvalues. Furthemore, $SJ_{\bar{x}}S = S\text{diag}(\bar{x})AS = \text{diag}(\bar{x})SAS$, i.e., $SJ_{\bar{x}}S$ is the Jacobian of SAS at $\bar{x}$. For SAS and $SJ_{\bar{x}}S$ Theorem 7.14 holds. □

Corollary 7.17 *Under the same assumptions as in Theorem 7.16, the Lotka-Volterra system* (7.30) *is a monotone system in* $\text{int}(\mathbb{R}^n_+)$*.*

Proof. Recall from Proposition 6.14 that a nonlinear system is monotone if and only if the graph $\mathcal{G}^{\text{nd}}(J(x))$ is structurally balanced for all x in the domain, where $J(x)$ is the Jacobian of the system at x. Here the system is (7.30) and we are interested only in $x > 0$ (i.e., the boundary of $\mathbb{R}^n_+$ is excluded). From the expression of the Jacobian in (7.32), we have that, when $x > 0$, A and $\text{diag}(x)A$ have the same sign pattern, while the other two terms, $\text{diag}(Ax)$ and $\text{diag}(r)$, can only affect the diagonal, which is irrelevant when computing $\mathcal{G}^{\text{nd}}(J(x))$. Hence $\mathcal{G}^{\text{nd}}(J(x))$ is structurally balance if and only if $\mathcal{G}^{\text{nd}}(A)$ is. □

The corollary is in princple valid also when A is unstable.

Notice that in Theorem 7.16 we are not specifying the sign of r, only that there exists $\bar{x} = -A^{-1}r > 0$. This is possible for r which are not positive, but also for some $r > 0$, as the following example shows.

Example 7.18 The interaction matrix

$$A = \begin{bmatrix} -1.5 & 0.5 & -0.6 & 0 \\ 0.9 & -1.3 & 0 & 0.8 \\ 0 & 0 & -1.2 & -0.3 \\ 0 & 0.5 & 0 & -1.4 \end{bmatrix}$$

is Hurwitz but not Metzler. However $\mathcal{G}^{\text{nd}}(A)$ is structurally balance, with signature vector $s = \begin{bmatrix} -1 & -1 & 1 & -1 \end{bmatrix}$, meaning that Theorem 7.16 applies. Not all growth vectors $r \geq 0$ lead to positive equilibrium point $\bar{x}$, but some do. For instance

$$r = \begin{bmatrix} 0.3 \\ 0.9 \\ 0.1 \\ 0.6 \end{bmatrix} \quad \text{gives} \quad \bar{x} = \begin{bmatrix} 0.9839 \\ 2.0984 \\ -0.2112 \\ 1.1780 \end{bmatrix}, \quad \text{while} \quad r = \begin{bmatrix} 0.6 \\ 0.7 \\ 0.5 \\ 0.8 \end{bmatrix} \quad \text{gives} \quad \bar{x} = \begin{bmatrix} 1.0579 \\ 2.0796 \\ 0.0881 \\ 1.3141 \end{bmatrix}$$

See Fig. 7.7 for a simulation. Similarly, positive equilibria can occur also for r which has some negative entries, for instance

$$r = \begin{bmatrix} 1.1 \\ 2.4 \\ 0.1 \\ -1.2 \end{bmatrix} \quad \text{gives} \quad \bar{x} = \begin{bmatrix} 1.8408 \\ 3.3236 \\ 0.0009 \\ 0.3298 \end{bmatrix}$$

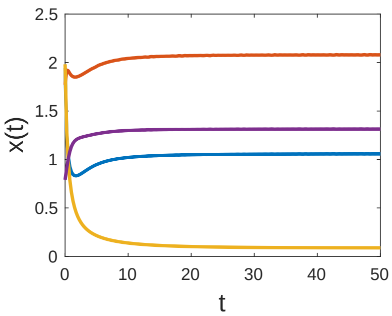


Figure 7.7: Lotka-Volterra system with structurally balanced adjacency matrix.

Chapter 8

Consensus

Consider the following distributed decision-making problem: a group of n agents wants to find an agreement on something. For instance they discuss an argument, exchange their opinions, and influence each other, until they achieve a common opinion. Mathematically, the opinion on "something" is represented as a number, and the common value that the agents want to find an agreement upon is called consensus value. By extension the problem itself, and the dynamical system that solves it, is called the *consensus problem*.

In order to find this consensus value, each agent contributes its own value, which becomes the initial condition of the state variable of that agent in the dynamical system. The group of agents has no central coordinator, and all calculations should be done in a distributed way, according to a graph of interactions among the agents. In particular, the state of each agent can be transmitted only to its first out-neighbors on the graph. Each agent does some operation with the values it receives form its in-neighbours, updates its own state accordingly, and transmits it to its out-neighbours. The procedure is iterated until a common value is achieved. A popular example of such common value is the average of the initial values contributed by the agents.

Denote $\mathcal{G}(A)$ the interaction graph, and with $A \geq 0$ its adjacency matrix. The values on which the agents rely upon to compute their common value are encoded as states at time $t = 0$:

$$x(0) = \begin{bmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{bmatrix} \in \mathbb{R}^n$$

The simplest possible dynamical system we can think of is a single integrator at each node, i.e.,

$$\dot{x}_i = u_i, \quad i = 1, \dots, n$$

where u_i is the input available to agent i.

The *consensus problem* consists then in choosing a law $u_i = u_i(x_i, x_j \text{ s. t. } j \in \mathcal{N}_i^{\text{in}})$ so that

$$\lim_{t\to\infty} |x_i(t) - x_j(t)| = 0 \tag{8.1}$$

or, in vector form,

$$x(t) \stackrel{t\to\infty}{\longrightarrow} \alpha(x(0))\mathbb{1}$$

where $\alpha(x(0))$ is a scalar function of the initial condition $x(0)$. In the common case of *average consensus*, the asymptotic value $\alpha(x(0))$ is the average of the initial conditions:

$$\alpha(x(0)) = \frac{1}{n}\sum_{j=1}^{n} x_j(0) = \frac{1}{n}\mathbb{1}^\top x(0) \tag{8.2}$$

Notice that the expression (8.1) is different from having that all $x_i(t)$ converge to the same value from all $x(0)$, for instance having $x_i(t) \overset{t\to\infty}{\longrightarrow} 0$ for all i and $\forall\ x(0)$. In fact this latter convergence occurs regardless of the value of the initial condition, while in (8.1) there is a memory of the initial condition, as indicated explicitly in the expression $\alpha(x(0))$. If we think of opinions, having all opinions converge to 0 (or to a constant x^*, independent of the initial conditions) is not very meaningful, because it implies that all agents forget their initial opinions, hence it makes little sense to interpret the value they achieve as the result of a interaction. Mathematically the difference between (8.1) and a global convergence to a constant x^* is explained as the difference between marginal stability and asymptotic stability.

In the opinion dynamics literature, the consensus problem in DT is known as the *DeGroot model* and it is studied since the Sixties. A CT equivalent is known as the Abelson model.

8.1 CT consensus problem

Let us discuss the CT consensus problem first. In this case we will make use of a well-know quantity in algebraic graph theory, the Laplacian of the graph $\mathcal{G}(A)$. Assume A has zero diagonal: $a_{ii} = 0$, $i = 1, \dots, n$.

Definition 8.1 *The* Laplacian matrix *associated to A is the following matrix $L = [\ell_{ij}] \in \mathbb{R}^{n\times n}$ of entries:*

$$\ell_{ij} = \begin{cases} \sum_{k=1}^n a_{ik} & \text{if } i = j \\ -a_{ij} & \text{if } i \neq j \end{cases} \tag{8.3}$$

In matrix form: $L = \text{diag}(A\mathbb{1}) - A$, or, in expanded form,

$$L = \begin{bmatrix} \sum_{j=1}^n a_{1j} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \sum_{j=1}^n a_{2j} & & \\ \vdots & & \ddots & \\ -a_{n1} & & & \sum_{j=1}^n a_{nj} \end{bmatrix}$$

The basic dynamical system that solves the consensus problem is the following linear system

$$\dot{x} = -Lx \tag{8.4}$$

In components (8.4) reads

$$\dot{x}_i = u_i = -\Big(\underbrace{\sum_{j=1}^n a_{ij}x_i}_{\text{diag.}} - \underbrace{\sum_{j=1}^n a_{ij}x_j}_{\text{off-diag.}}\Big) = -\sum_{j=1}^n a_{ij}(x_i - x_j) = \sum_{j=1}^n a_{ij}(x_j - x_i) \tag{8.5}$$

which shows that what is being integrated is the difference between the coordinates, i.e., $x_j - x_i$. In particular, (8.5) means that at each t each agent computes a weighted sum of the relative states, i.e., of the differences between its own state and that of its neighbors, with weights given by the corresponding edge weights a_{ij}. The law (8.5) is therefore distributed, according to the graph $\mathcal{G}(A)$. From (8.5), it can be seen that $x_i = x_j$ for all $i, j = 1, \ldots, n$, is indeed an equilibrium point, which is coherent with the requirements of the consensus problem.

Let us investigate first the case of $\mathcal{G}(A)$ undirected, then the one of $\mathcal{G}(A)$ directed.

8.1.1 Undirected graph case

Assume $\mathcal{G}(A)$ is undirected, i.e., that A is symmetric (and so is L). Most of the following properties of L are verified straightforwardly.

Proposition 8.2 *For a symmetric Laplacian matrix L it holds:*

1. *$-L$ is Metzler matrix;*
2. *$\lambda = 0$ is an eigenvalue of L;*
3. *$L\mathbb{1} = 0$ (i.e., the row sums of L are all vanishing);*
4. *$\mathbb{1}^\top L = 0$ (i.e., the column sums of L are all vanishing);*
5. *L is singular.*

Since $-L$ is a Metzler matrix, we can apply Propositions 3.15 and 3.16 and say that the spectral abscissa $\mu(-L)$ must be a real eigenvalue of $-L$. Hence if $\lambda_i(-L)$ are the eigenvalues of $-L$ (which are all real, since $-L$ is symmetric), it is

$$\lambda_n(-L) \leq \ldots \leq \lambda_2(-L) \leq \lambda_1(-L) = \mu(-L)$$

Item 3 of Proposition 8.2 implies that, by construction, $-L$ is non-strict diagonally dominant, i.e., it is a diagonally equipotent matrix in the terminology introduced in Section 6.6. It follows from this that $\lambda_1(-L) = \mu(-L) = 0$, $\lambda_i(-L) \leq 0$ $i = 2, \ldots, n$.

In the case of $\mathcal{G}(A)$ connected, (i.e., A irreducible), more can be said. From the Perron-Frobenius theorem for Metzler matrices (Theorem 3.17) we have:

Proposition 8.3 *If the undirected graph $\mathcal{G}(A)$ is connected, then the eigenvalue $\lambda_1(-L) = 0$ has multiplicity 1, and it is strictly bigger that all other eigenvalues of $-L$: $\lambda_i(-L) < 0$ $i = 2, \ldots, n$.*

Notice that the last property follows from a necessary and sufficient algebraic condition for connectivity of a graph.

Proposition 8.4 *The undirected graph $\mathcal{G}(A)$ is connected if and only if $\lambda_2(L) \neq 0$, that is, if and only if* $\operatorname{rank}(L) = n - 1$*. When instead $\mathcal{G}(A)$ is not connected, then the multiplicity of the eigenvalue $\lambda = 0$ is equal to the number of connected components of $\mathcal{G}(A)$.*

Proof. $\mathcal{G}(A)$ connected means A (and L) irreducible. Furthermore, $-L$ is Metzler and diagonally equipotent by rows, hence condition 3 of Theorem 6.26 applies and the first statement follows. As for the second statement, each connected component of $\mathcal{G}(A)$ corresponds to a diagonal block of A (possibly after a permutation) which is diagonally equipotent by rows. Theorem 6.31 applies and gives the multiplicity of $\lambda = 0$. □

Example 8.5 The adjacency matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix} \quad \Longrightarrow \quad L = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & -2 \\ 0 & -2 & 2 \end{bmatrix}$$

has the connected graph $\mathcal{G}(A)$ shown in Fig. 8.1(a), and Laplacian eigenvalues $\lambda_i(L) = \{0,\ 1.2,\ 4.7\}$. For the matrix

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix} \quad \Longrightarrow \quad L = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 2 \end{bmatrix}$$

instead, the graph $\mathcal{G}(A)$ shown in Fig. 8.1(b) is not connected, and the eigenvalues are $\lambda_i(L) = \{0,\ 0,\ 4\}$. □

(a) (b)

Figure 8.1: Graphs of Example 8.5.

It follows from Proposition 8.2 that the system (8.4) cannot be asymptotically stable, see Theorem 6.21. In particular, from $\mathbb{1}^\top L = 0$ it follows that the solutions of (8.4) have an integral of motion, i.e., they all obey a conservation law.

Proposition 8.6 *The system* (8.4) *has the conservation law*

$$\mathbb{1}^\top x(t) = \mathbb{1}^\top x(0) = \text{const} \quad \forall t \geq 0 \tag{8.6}$$

Proof. From $\mathbb{1}^\top \dot{x} = 0 \implies \mathbb{1}^\top L x(t) = 0$. Integrating

$$\mathbb{1}^\top x(t) = \mathbb{1}^\top x(0) = \text{const} \qquad \text{for all } t \geq 0.$$

□

Corollary 8.7 *If the system* (8.4) *converges, then it does so to the average of the initial conditions* $\frac{1}{n}\mathbb{1}^\top x(0)$*, i.e., it solves the average consensus problem.*

Proof. It is enough to rescale the conservation law

$$\frac{1}{n}\mathbb{1}^\top x(t) = \frac{1}{n}\mathbb{1}^\top x(0) \quad \text{for all } t \geq 0$$

which is the average consensus value given in (8.2). □

Notice that in principle the conservation law (8.6) is compatible also with asymptotic behaviors like $x_i \to +\infty$ and $x_j \to -\infty$ provided they give $x_i + x_j = \text{const}\ \forall\ t$. However, the system (8.4) has state update matrix $-L$ which is a Metzler matrix, and we will now exploit this property. As described in Chapter 6, linear systems having a state update matrix which is Metzler can be of different types (positive system or cooperative system). Here it is $x \in \mathbb{R}^n$, meaning that the consensus system is not a positive system, but rather a cooperative system.

When L is irreducible, invoking again the Perron-Frobenius theorem for Metzler matrices (Theorem 3.17), the asymptotic behavior of the system (8.4) can be fully characterized. Let us first show the following proposition, which is the CT (and symmetric) analogous of a DT spectral projection property we saw in Section 7.1.

Proposition 8.8 *Consider an undirected, connected graph $\mathcal{G}(A)$, and let L be the corresponding Laplacian. Then*

$$\lim_{t\to\infty} e^{-Lt} = \frac{1}{n}\mathbb{1}\mathbb{1}^\top.$$

Proof. The matrix $-L$ is irreducible if and only if A is irreducible (A and $-L$ differ only for the diagonal), i.e., if and only if $\mathcal{G}(A)$ is connected. $-L$ is Metzler and irreducible, and additionally, from Proposition 8.2, since $\lambda(-L) = 0$ corresponds to the positive eigenvector $\mathbb{1}$, then 0 must be also the spectral abscissa of $-L$, $\mu(-L) = 0$, as Metzler matrices admit only one positive eigenvector (Theorem 3.17). The rest of the proof is based on convergence to the dominant spectral projector. Write $-L = TJT^{-1}$, where

$$J = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & J_1 \end{array}\right], \quad T = \left[\begin{array}{c|c} \frac{\mathbb{1}}{\sqrt{n}} & T_1 \end{array}\right], \quad T^{-1} = \left[\begin{array}{c} \mathbb{1}^\top/\sqrt{n} \\ \hline T_2 \end{array}\right]$$

Then, since all eigenvalues $\lambda_i(-L)$ are real and such that $\lambda_i(-L) < \mu(-L) = 0$ whenever $\lambda_i(-L) \neq \mu(-L)$, the block J_1 is composed of asymptotically stable modes:

$$\lim_{t\to\infty} e^{-Lt} = T\left[\begin{array}{c|c} e^0 & 0 \\ \hline 0 & \lim_{t\to\infty} e^{J_1 t} \end{array}\right]T^{-1} = T\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \end{array}\right]T^{-1} = \frac{1}{n}\mathbb{1}\mathbb{1}^\top$$

□

The condition $\lambda_i(-L) < \lambda_1(-L) = \mu(-L) = 0$, $i = 2, \ldots, n$, holds whenever the graph is connected. We have the following.

Corollary 8.9 *If $\mathcal{G}(A)$ is connected, then the system (8.4) is marginally stable.*

Since $\mathbb{1}$ is the dominant eigenvector, a consequence of marginal stability is that

$$x(t) \xrightarrow{t\to\infty} \text{span}(\mathbb{1}) = \{x \in \mathbb{R}^n \ \text{ s. t. } \ x = \alpha\mathbb{1} \ \text{ with } \ \alpha \in \mathbb{R}\}.$$

Notice that $\text{span}(\mathbb{1})$ is often called the *agreement subspace*, as all its points obey to $x_i = x_j$ for all coordinates. Notice that it is a vector subspace, i.e., the consensus values in $\text{span}(\mathbb{1})$ can assume both positive and negative values.

A consequence Proposition 8.8 and Corollary 8.9 is the consensus behavior for (8.4).

Theorem 8.10 *Consider an undirected, connected graph $\mathcal{G}(A)$, and let L be the corresponding Laplacian. Then the system (8.4) converges to*

$$x^* = \lim_{t\to\infty} x(t) = \alpha \mathbb{1}, \qquad \text{where } \alpha = \frac{1}{n}\mathbb{1}^\top x(0)$$

i.e., it solves the average consensus problem.

Proof. Combining the results of Proposition 8.8 and Corollary 8.9 we get

$$x^* = \lim_{t\to\infty} x(t) = \lim_{t\to\infty} e^{-Lt}x(0) = \frac{1}{n}\mathbb{1}^\top x(0)\mathbb{1}$$

□

Example 8.11 Consider a connected undirected graph $\mathcal{G}(A)$ with $n = 100$. A typical trajectory looks like in Fig. 8.2(a). Different initial conditions leads to different cosensus points, see Fig. 8.2(b) for a 2D phase portrait.

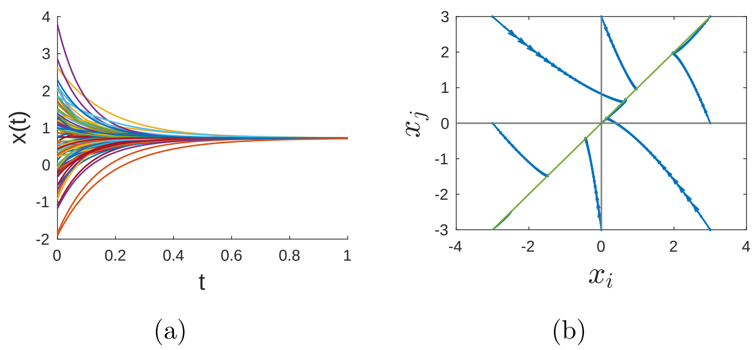


Figure 8.2: (a): a trajectory (b): a 2D phase portrait from different initial conditions.

Notice that we could have proceeded also as in Section 7.1, making use of the transversality between the conservation law $\mathbb{1}^\top x(t) = \text{const}$ and the agreement subspace $\text{span}(\mathbb{1})$. The difference w.r.t. Section 7.1 is that now the value $\sum_i x_i(0)$ is not fixed as in a Markov chain (where it was equal to 1). Furthermore, $x(0)$ can be any point in $\mathbb{R}^n$, not just a point in the positive orthant. To see it explicitly, decompose $\mathbb{R}^n$ into two orthogonal parts $\mathbb{R}^n = \text{span}(\mathbb{1}) \oplus \text{span}(\mathbb{1})^\perp$, and split $x(t)$ accordingly, as a projection on $\text{span}(\mathbb{1})$ and one on $\text{span}(\mathbb{1})^\perp$. Denoting $\Pi = I - \mathbb{1}\mathbb{1}^\top/n$ the projection matrix onto $\text{span}(\mathbb{1})^\perp$:

$$x(t) = (I - \Pi)x(t) + \Pi x(t) = \alpha\mathbb{1} + \delta(t), \tag{8.7}$$

with $\alpha\mathbb{1} \in \text{span}(\mathbb{1})$, and $\delta \in \text{span}(\mathbb{1})^\perp$. While $\alpha = \mathbb{1}^\top x(t)/n$ is the usual invariant of motion (i.e., the conservation law), the vector δ is called the *disagreement vector*, to emphasize that it is what remains once you project on the orthogonal complement of the agreement subspace (i.e., in (8.7) you subtract the invariant of motion $\delta(t) = x(t) - \alpha$). Notice that to be precise, for each solution of (8.4) δ belongs to a slice of the $(n-1)$-dimensional space $\text{span}(\mathbb{1})^\perp$, i.e., to a coset space determined by the invariant of motion. Before considering the dynamics of the projection on $\text{span}(\mathbb{1})^\perp$ let us introduce the concept of Laplacian potential. Given the Laplacian L, the associated *Laplacian potential* is the quadratic form

$$\Phi(x) = x^\top L x = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x_j - x_i)^2 \tag{8.8}$$

By construction $\Phi(x)$ is positive semidefinite (psd), as it is a sum of squares with nonnegative coefficients a_{ij}. Hence we can also say that L is psd matrix (of rank $n-1$ when $\mathcal{G}(A)$ connected). The agreement subspace $\text{span}(\mathbb{1})$ forms the 0-level surface of $\Phi(x)$: $\Phi(\alpha\mathbb{1}) = 0$. Taking the gradient of $\Phi(x)$: $\nabla\Phi(x) = Lx$, meaning that (8.4) can be written in gradient form

$$\dot{x} = -\nabla\Phi(x) = -Lx$$

In other words, (8.4) is the "best response" dynamics to the Laplacian potential (8.8), see Chapter 15. For $\mathcal{G}(A)$ connected, since L is symmetric, psd of rank $n-1$, its eigenvalues are all real and can be written as

$$0 = \lambda_1(L) < \lambda_2(L) \leq \lambda_3(L) \leq \ldots \leq \lambda_n(L)$$

We have already seen the role of $\lambda_1(L) = 0$. The second least eigenvalue of L, $\lambda_2(L)$, is called the algebraic connectivity (or Fiedler eigenvalue). Notice that it can be expressed as a function of the disagreement vector δ and of the Laplacian potential as follows:

$$\lambda_2(L) = \min_{\mathbb{1}^\top\delta=0} \frac{\delta^\top L\delta}{\|\delta\|^2} \tag{8.9}$$

where restricting the minimum to $\mathbb{1}^\top\delta = 0$ means restricting to $\delta \in \text{span}(\mathbb{1})^\perp$.

Proposition 8.12 *Consider $\mathcal{G}(A)$ connected and undirected. The dynamics of the disagreement vector δ associated to (8.4) is given by*

$$\dot{\delta} = -L\delta \tag{8.10}$$

and is globally asymptotically stable when restricted to $\text{span}(\mathbb{1})^\perp$.

Proof. Notice first that from $x = \alpha\mathbb{1} + \delta$, since α is an invariant of motion, $\frac{d(\alpha\mathbb{1})}{dt} = 0$. Hence

$$\dot{x} = \dot{\delta} = -Lx = -L(\alpha\mathbb{1} + \delta) = -L\delta$$

since $L\mathbb{1} = 0$. Once we restrict to $\text{span}(\mathbb{1})^\perp$, we can make use of a Laypunov function

$$V(\delta) = \frac{1}{2}\|\delta\|^2 \tag{8.11}$$

which is positive definite (pd) in $\text{span}(\mathbb{1})^\perp$. From (8.9), $\delta^\top L\delta \geq \lambda_2(L)\|\delta\|^2$, i.e., $-\delta^\top L\delta \leq -\lambda_2(L)\|\delta\|^2$, hence differentiating $V(\delta)$,

$$\dot{V} = -\delta^\top L\delta \leq -\lambda_2(L)\|\delta\|^2 = -\lambda_2(L)V(\delta) < 0 \tag{8.12}$$

Lyapunov stability theory guarantees that $\delta^* = 0$ is asymptotically stable in $\text{span}(\mathbb{1})^\perp$. □

Let us combine the Lyapunov arguments in the proof of Proposition 8.12 with the Laplacian potential $\Phi(L)$. Defining $V(x) = \frac{1}{2}\|x\|^2$ as Lyapunov function in $\mathbb{R}^n$, then $\dot{V}(x) = -x^\top Lx = -\Phi(x)$, which is psd. The level surfaces of this Lyapunov function correspond exactly to span($\mathbb{1}$):

$$\{x \ \text{ s.t. } \ \dot{V}(x) = 0\} = \ker(L) = \operatorname{span}(\mathbb{1}),$$

meaning that Proposition 8.12 can be interpreted as a case of LaSalle invariance principle.

Another type of Lyapunov function used for the consensus problem (8.4) is the following nonlinear function

$$V(x(t)) = \max_{i=1,\dots,n}(x_i(t)) - \min_{i=1,\dots,n}(x_i(t))$$

The meaning of this Lyapunov function is that consensus is a contraction in the convex hull of the current states. For the 1-dimensional case (i.e., $x_i \in \mathbb{R}$), the convex hull of the state $x(t)$ is given exactly by $V(x(t))$. V is continuous but not differentiable. To define properly the derivative of this Lyapunov function V one needs to use Dini derivatives, so we skip it.

8.1.2 Directed graph case

Assume now that $A \geq 0$ is not symmetric, and still without diagonal entries. Consider a digraph $\mathcal{G}(A)$, The expression (8.3) for the Laplacian associated to $\mathcal{G}(A)$ is still valid, and so are many of its properties.

Proposition 8.13 *For a non-symmetric Laplacian matrix L it holds:*

1. *$-L$ is Metzler matrix;*
2. *$\lambda = 0$ is an eigenvalue of L;*
3. *$L\mathbb{1} = 0$ (i.e., the row sums of L are all vanishing);*
4. *L is singular.*

If in addition L is irreducible, then

5. *the eigenvalue $\lambda = 0$ has multiplicity 1, and it is strictly less that the real part of all other eigenvalues of L.*
6. *$v^\top L = 0$ where $v > 0$ is the left eigenvector of L relative to 0.*

Notice that in the digraph case the eigenvalues of L, $\lambda_i(L)$, are in general complex conjugate pairs of values. What the proposition says is that

1. the eigenvalue with the least real part is real: $\lambda_1(L) = 0$;
2. In the irreducible case, $0 = \lambda_1(L) < \text{Re}[\lambda_i(L)]$, $i = 2, \dots, n$.

Both conditions follow from the Perron-Frobenius theorem for irreducible Metzler matrices (Theorem 3.17), which also implies that the left eigenvector associated to the eigenvalue $\lambda_1(L) = 0$, here called v, is still positive, but it is now typically not equal to $\mathbb{1}$. From the first point, when considering $-L$, $\mu(-L) = \lambda_1(-L) = 0$ is the spectral abscissa. When L is irreducible, since $-L$ is Metzler and diagonally equipotent, condition 3 of Theorem 6.26 applies and lead to condition 5 of Proposition 8.13.

The analogous of Theorem 8.10 is the following.

Theorem 8.14 *Consider a strongly connected digraph $\mathcal{G}(A)$, and let L be the corresponding Laplacian. Then the system (8.4) converges to consensus*

$$x^* = \lim_{t\to\infty} x(t) = \alpha \mathbb{1} \qquad \textit{where } \ \alpha = v^\top x(0)$$

and $v > 0$, rescaled s.t. $\sum_{i=1}^n v_i = 1$, is the left eigenvector of L relative to 0.

Proof. The proof is analogous to that of Theorem 8.10. From the equivalent of Proposition 8.8 for digraphs, the spectral projector onto the dominant eigenspace is

$$\lim_{t\to\infty} e^{-Lt} = \mathbb{1} v^\top$$

from which

$$x^* = \lim_{t\to\infty} x(t) = \lim_{t\to\infty} e^{-Lt} x(0) = v^\top x(0) \mathbb{1}$$

□

From Theorem 8.14, in the directed case, we still have $x(t) \overset{t\to\infty}{\longrightarrow} \text{span}(\mathbb{1})$. However, the specific value in span($\mathbb{1}$) at which $x(t)$ converges is in general not the average of the initial conditions but $\alpha = v^\top x(0)$, i.e., a weighted average of the initial conditions, with weights given by the components of v.

Weight balance leads to average consensus

The only case in which we still recover the average consensus solution is when A is weight balanced.

Definition 8.15 *A matrix $A \geq 0$ is said* weight balanced *if its weighted indegree and outdegree coincide:*

$$\text{wdeg}^{\text{in}}(i) = \sum_{j=1}^n a_{ij} = \sum_{j=1}^n a_{ji} = \text{wdeg}^{\text{out}}(i) \quad \forall\, i = 1, \ldots, n$$

The weight balance condition can be written in matrix form as

$$A\mathbb{1} = A^\top \mathbb{1}. \tag{8.13}$$

Although this condition is weaker than symmetry of A, it also leads to average consensus. In fact, it follows from (8.13) that weight balance corresponds also to $L\mathbb{1} = L^\top \mathbb{1} = 0$, i.e., the left eigenvector of L relative to 0 is equal to the right eigenvector.

Theorem 8.16 *Consider a strongly connected and weight balanced digraph $\mathcal{G}(A)$, and let L be the corresponding Laplacian. Then the system (8.4) converges to average consensus*

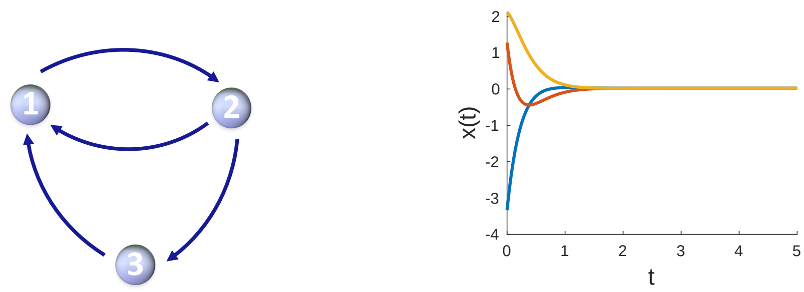
$$x^* = \lim_{t\to\infty} x(t) = \alpha \mathbb{1} \qquad \textit{where } \ \alpha = \frac{1}{n} \mathbb{1}^\top x(0)$$

The condition of Theorem 8.16 can be upgraded to a necessary and sufficient condition for average consensus: average consensus is achieved if and only if $\mathcal{G}(A)$ is strongly connected and weight balanced.

Example 8.17 The following adjacency matrix is weight balanced:

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \quad \Longrightarrow \quad L = \begin{bmatrix} 3 & -1 & -2 \\ -3 & 3 & 0 \\ 0 & -2 & 2 \end{bmatrix}$$

The matrix A corresponds to the strongly connected graph $\mathcal{G}(A)$ shown in Fig. 8.3(a), and the Laplacian eigenvalues are $\lambda_i(L) = \{0, 4 \pm 1.4i\}$. The corresponding consensus system achieves average consensus, see Fig. 8.3(b). □



(a) (b)

Figure 8.3: Graph of Example 8.17.

Lack of weight balance leads to weighted consensus

Assume that A is irreducible but not weight balanced: $A\mathbb{1} \neq A^\top \mathbb{1}$. As before, denote $v > 0$ the left eigenvector associated to $\lambda_1(L) = 0$. Theorem 8.14 holds and tells us that the consensus value at which the state vector $x(t)$ converges, $x^* = \alpha(x(0))\mathbb{1}$, is weighted by the components of v: $\alpha(x(0)) = v^\top x(0)$. Later on (Section 8.3), we will se that v has the interpretation of "social power", i.e., of relative weight that each agent has in the final outcome of the consensus process.

For non weight balanced graphs, it is interesting to seek for the equivalent of the Laplace potential introduced in the undirected case. In a quadratic form like (8.8), only the symmetric part of the matrix matters. However, when L is not weight balanced it can happen that even if $-L$ is marginally stable, $L^{\text{sym}} = (L + L^\top)/2$ is not pds.

Example 8.18 For the non-symmetric Laplacian

$$L = \begin{bmatrix} 1.3 & -0.5 & -0.8 \\ -0.9 & 1.6 & -0.7 \\ -0.8 & 0 & 0.8 \end{bmatrix}$$

we have $\lambda_i(L^{\text{sym}}) = \{-0.087, 1.565, 2.222\}$, that is, L^{sym} is not psd. □

This has an impact in considering the system (8.4) as a gradient system, and also in determining the existence of a Lyapunov function (which is of course related to the previous issue), used e.g., in Proposition 8.12.

The quadratic form to be used for L not weight balance has to take $v > 0$ into account as the following proposition says.

Proposition 8.19 *Consider a digraph $\mathcal{G}(A)$ with $A \geq 0$ irreducible. Let $\mathbb{1}$ and $v > 0$ be the right and left eigenvectors of L relative to the eigenvalue 0. Then v is the unique (up to a scalar multiplication) positive vector for which the diagonal matrix $\mathrm{diag}(v)$ is s.t. $\mathrm{diag}(v)L+L^\top \mathrm{diag}(v)$ is psd. For it, $\ker(L^\top \mathrm{diag}(v)) = \ker(L) = \mathrm{span}(\mathbb{1})$ and hence $\mathrm{rank}(\mathrm{diag}(v)L + L^\top \mathrm{diag}(v)) = n-1$.*

Proof. From Proposition 8.13, L is a singular irreducible M-matrix, and its null space $\ker(L) = \mathrm{span}(\mathbb{1})$ has dimension 1. Furthermore, since the diagonal entries of L are all positive, it is also an H_+ matrix. From Lemma 3.21-3.22 of [11] there exists a unique (up to scalar multiplication) diagonal matrix $\Xi = \mathrm{diag}(\xi_1, \ldots, \xi_n)$ such that $\ker(L^\top \Xi) = \ker(L) \subseteq \ker(\Xi L + L^\top \Xi)$, while from Theorem 3.23 of [11] L admits a unique (up to scalar multiplication) "Lyapunov scaling factor" Ξ such that $\Xi L + L^\top \Xi$ psd. Since $L^\top v = L^\top \mathrm{diag}(v)\mathbb{1}$, it must necessarily be $[\xi_1 \ \ldots \ \xi_n]^\top = \alpha v$, $\alpha \in \mathbb{R}$. From $\ker(L) = \ker(\mathrm{diag}(v)L) = \ker(L^\top \mathrm{diag}(v)) = \mathrm{span}(\mathbb{1})$, $\mathrm{diag}(v)L + L^\top \mathrm{diag}(v)$ must necessarily have rank $n-1$. □

The disagreement vector can now be defined in the same manner as for the undirected case, and still leads to the disagreement system (8.10). As Proposition 8.19 suggests, to prove its asymptotic stability in $\mathrm{span}(\mathbb{1})^\perp$, instead of the Lyapunov function (8.11), one has to use the weighted quadratic function

$$V(x) = \frac{1}{2}\delta^\top \mathrm{diag}(v)\delta \tag{8.14}$$

In fact differentiating, we get

$$\dot{V} = -\delta^\top \left(L^\top \mathrm{diag}(v) + \mathrm{diag}(v)L\right)\delta$$

where $\mathrm{diag}(v)L$ is the weight balanced equivalent of L, and, from Proposition 8.19, $L^\top \mathrm{diag}(v) + \mathrm{diag}(v)L$ is a psd matrix of rank $n-1$.

Lack of strong connectivity and rooted spanning tree

Theorem 8.14 requires strong connectivity of $\mathcal{G}(A)$. This is a sufficient but not necessary condition for consensus. (Combined with weight balance it is necessary and sufficient, but only for *average* consensus). In order to obtain a necessary and sufficient condition for (weighted) consensus, we need to resort to the notion of rooted directed spanning tree.

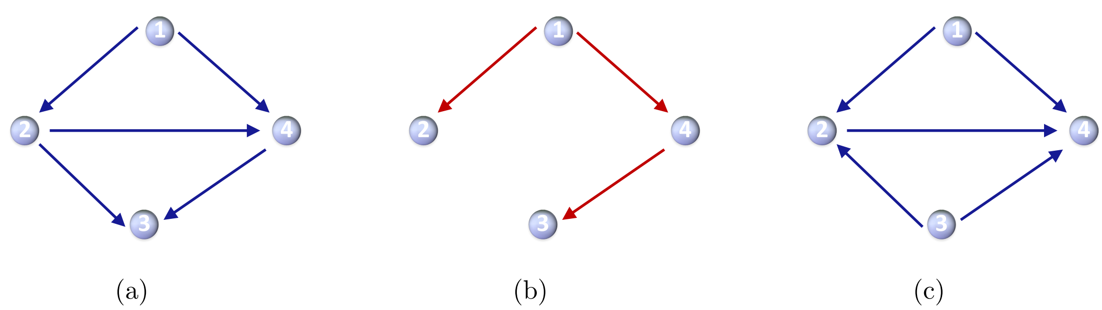


Figure 8.4: (a): A non strongly connected digraph with a rooted directed spanning tree; (b): its directed spanning tree rooted at node 1. (c): A non strongly connected graph without a directed spanning tree.

The key property of a rooted directed spanning tree is that the corresponding Laplacian has rank $n-1$.

Proposition 8.20 *Consider a directed graph $\mathcal{G}(A)$, and let L be the corresponding Laplacian. Then* $\operatorname{rank}(L) = n-1$ *if and only if $\mathcal{G}(A)$ has a rooted directed spanning tree.*

Proof. The proof is an application of Theorem 6.31: the multiplicity of $\lambda = 0$ is equal to the number of rooted strongly connected components associated to diagonally equipotent blocks. If $-L$ (which is diagonally equipotent) has a rooted spanning tree, then this multiplicity must be one. Viceversa, if $\operatorname{rank}(L) = n-1$, then necessarily $\lambda = 0$ has to have multiplicity 1 and Theorem 6.31 affirms that $\mathcal{G}(A)$ has to have a single rooted strongly connected component. $\mathcal{G}(A)$ has to contain a spanning tree rooted at one of the nodes of this strongly connected components, because otherwise $\mathcal{G}(A)$ must be disconnected (it cannot have other rooted strongly connected components), and in that case $\operatorname{rank}(L)$ cannot be $n-1$. □

Theorem 8.21 *Consider a directed graph $\mathcal{G}(A)$, and let L be the corresponding Laplacian. Then the system* (8.4) *solves the consensus problem if and only if $\mathcal{G}(A)$ has a rooted directed spanning tree.*

Proof. By construction, $L\mathbb{1} = 0$, i.e., $\lambda = 0$ is an eigenvalue of L of eigenvector $\mathbb{1}$. From the directed spanning tree assumption, all rows of L except at most one (the root) have nonzero entries. From Proposition 8.20, $\operatorname{rank}(L) = n-1$, meaning that $\lambda_1(L) = 0$ is a simple eigenvalue and its eigenspace is $\operatorname{span}(\mathbb{1})$. (Nonstrict) diagonal dominance still holds by construction, implying that it must be $\operatorname{Re}[\lambda_i(-L)] < 0$, $i = 2, \dots, n$, hence marginal stability must hold for $-L$. □

Notice that the left eigenvector of L relative to 0 need not be positive: $v \geq 0$ in general. Hence Theorem 8.21 guarantees only convergence to a weighted average in which some components of $x(0)$ may be missing.

8.1.3 Leader-follower consensus

In particular, a directed spanning tree is rooted at a node, i.e., there is a path connecting the root node to any other node. Assume without loss of generality that agent 1 is the root node. When the in-neighborhood of the root node is empty, i.e., in $\mathcal{G}(A)$ $a_{1j} = 0$ for all j, then agent 1 is not influenced by any other agent. Therefore its equation in (8.4) is simply

$$\dot{x}_1 = 0 \implies x_i(t) = x_i(0) \;\; \forall\, t.$$

On the other hand, if $\mathcal{G}(A)$ has a spanning tree Theorem 8.21 holds, meaning that the system achieves consensus. Combining these two facts, we have that all agents must converge to the value of the root node agent, $x_1(0)$. This is called *leader-follower consensus.*

Proposition 8.22 *Consider a digraph $\mathcal{G}(A)$ containing a directed spanning tree rooted at node 1. Assume $a_{1j} = 0$ for all $j = 1, \dots, n$. Then the system* (8.4) *achieves leader-follower consensus, i.e.,*

$$\lim_{t\to\infty} x(t) = x_1(0)\mathbb{1} \quad \textit{for all} \quad x(0)$$

In this case, the left eigenvector of L relative to 0 is $v = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^\top$.

Example 8.23 If the adjacency matrix of the graph in Fig. 8.4 is

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 1 & 2 & 0 & 0 \end{bmatrix} \quad \Longrightarrow \quad L = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -2 & 2 & 0 & 0 \\ 0 & -1 & 3 & -2 \\ -1 & -2 & 0 & 3 \end{bmatrix}$$

then $\mathcal{G}(A)$ has a directed spanning tree, rooted at node 1. The left eigenvector relative to 0 is $v = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^\top$, meaning that $\lim_{t\to\infty} x_i(t) = x_1(0)$, i.e., leader-follower consensus is achieved. The consensus value is the initial state of the root node which does not change over time (blue line in Fig. 8.5), while the initial conditions $x_2(0)$, $x_3(0)$ and $x_4(0)$ get lost because node 1 is not influenced by 2, 3 and 4, see Fig. 8.5. □

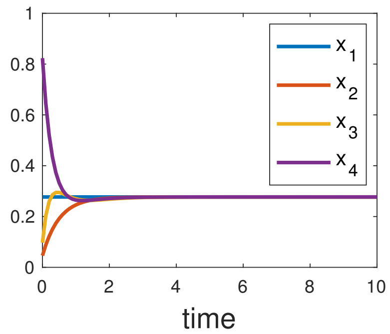


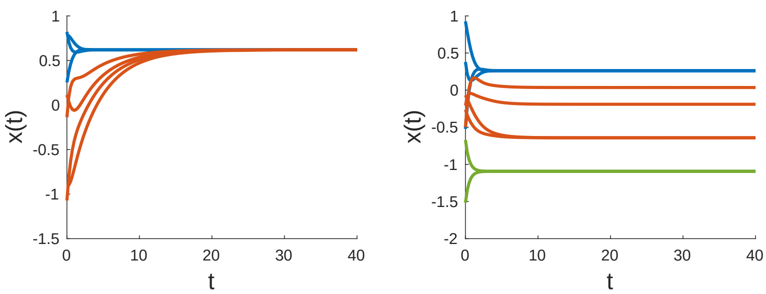
Figure 8.5: Example 8.23. Achieving leader-follower consensus.

When the root node of a directed spanning tree belongs to a strongly connected component of the graph $\mathcal{G}(A)$ then all nodes of this component can be taken as root of the directed spanning tree, see Fig. 4.6 (a). They alone decide the value of consensus to which all other nodes converge. In this case, the components of v which are different from 0 are those of the nodes that belong to the rooted strongly connected component (in the example of Fig. 4.6 (a), $v_i \neq 0$ for $i = 1, 2, 3$).

8.1.4 Cluster consensus and containment control

When instead a graph has more than one rooted strongly connected components then it cannot have a rooted directed spanning tree, as the nodes of one rooted strongly connected component are not reachable from the others, see Fig. 4.6 (b). In this case consensus cannot be achieved, or rather, each rooted strongly connected component achieves a consensus value, but these values are in general different from the others. When there are $k \geq 2$ such strongly connected components, then the agents in these components will cluster into k (normally distinct) consensus values, see Fig. 8.6(b). The agents that are not included in one of the rooted strongly connecting components do not converge to consensus, but rather to individual values. In general the specific values to which these agents converge cannot be predicted exactly. What is known is that these values are in between those of the clusters. This particular situation is sometimes called *containment control* (in the sense that one can contain all values of the extra nodes by knowing what the rooted strongly connected components do).

Example 8.24 In Fig. 4.6(a), the digraph $\mathcal{G}(A)$ has a rooted connected component consisting of the nodes 1, 2, and 3. These three agents are not influenced by the remaining agents, but influence them. A consensus dynamics (8.4) running on such graph will lead to leader-follower consensus, where the leaders are the three nodes and all others are followers. In Fig. 8.6(a) these three agents are shown in blue, all the others in red. In Fig. 4.6(b), instead, the digraph $\mathcal{G}(A)$ has two distinct rooted connected components. Consensus cannot be achieved on this graph: nodes 1-2-3 and nodes 8-9 (shown in blue and greeen in Fig. 8.6(b)) will achieve different within-cluster consensus values, while the remaining nodes will converge to some value in between. Notice that the non-rooted strongly connected component 5-6-7 does not achieve internal consensus.



(a) (b)

Figure 8.6: (a): Leader follower consensus for the network of Fig. 4.6(a). (b): Containment control (but no consensus) for the network of Fig. 4.6(b).

□

8.1.5 Other algebraic properties

Laplacian eigenvalues and Geršgorin theorem

Notice that by construction L is still non-strictly diagonally dominant by rows

$$\ell_{ii} = \sum_{j=1}^{n} a_{ij} = \sum_{j \neq i} |\ell_{ij}|.$$

In Section 5.2.2 we called this property diagonal equipotence, and we gave an interpretation in terms of the Geršgorin theorem (Theorem 5.10). In our case, Geršgorin theorem affirms that the eigenvalues of a matrix L must be contained in the union of the Geršgorin disks

$$\left\{ s \in \mathbb{C} \ \text{ s.t. } \ |s - \ell_{ii}| \leq \sum_{j \neq i} |\ell_{ij}| \right\}$$

i.e., the disks in the complex plane centered at ℓ_{ii} and of radius $\sum_{j \neq i} |\ell_{ij}|$. Since diagonal dominance is never strict, all these disks have to pass through 0, and taking $r_{\max} = \max_i \ell_{ii} = \max_i \sum_{j=1}^{n} a_{ij}$, they all must be contained inside the largest disk of radius $r_{\max}$. By looking at the disks for $-L$ (Fig. 8.7(b)), it is clear that $\mathrm{Re}[\lambda_i(-L)] \leq 0$. Combining this with Proposition 8.20, we have that any $\lambda_i(-L)$ different from $\lambda_1(-L) = 0$ must have $\mathrm{Re}[\lambda_i(-L)] < 0$.

When instead we have a matrix with strict diagonal dominance then no disk can touch the point 0, hence 0 cannot be an eigenvalue.

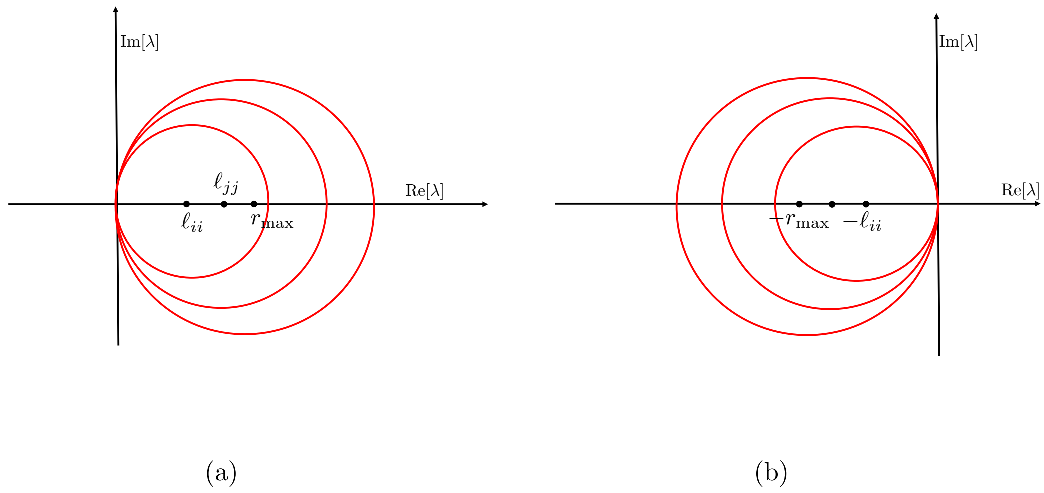


Figure 8.7: Geršgorin disks. (a): for the matrix L; (b): for $-L$.

Normalized Laplacian and Geršgorin disks

Given the Laplacian L, write it as $L = \Delta - A$ where from (8.3) $\Delta = \text{diag}(\ell_{11}, \ldots, \ell_{nn})$, with $\ell_{ii} = \sum_{j=1}^{n} a_{ij}$ the weighted in-degree of node i and A the adjacency matrix of the associated graph $\mathcal{G}(A)$. If $\mathcal{G}(A)$ is strongly connected, then $\ell_{ii} > 0$ for all i, hence Δ is a pd diagonal matrix. The *normalized Laplacian* associated to $\mathcal{G}(A)$ is defined as follows

$$\mathcal{L} := \Delta^{-1} L = I - \Delta^{-1} A$$

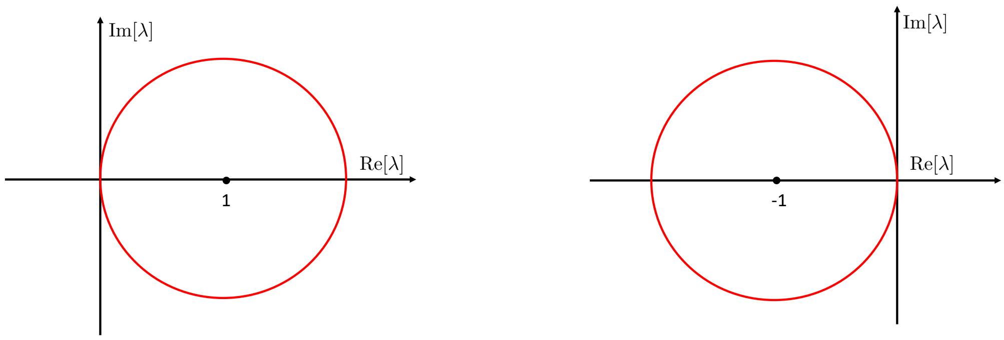
Entrywise, for the normalized Laplacian it is

$$[\mathcal{L}]_{ij} = \begin{cases} 1 & \text{if} i = j \\ \frac{-a_{ij}}{\sum_{k=1}^{n} a_{ik}} & \text{if} j \neq i \end{cases}$$

Since Δ is diagonal pd, $\mathcal{L}$ has the same algebraic properties of L: it is an irreducible and singular matrix of rank $n-1$, with $\lambda_1(\mathcal{L}) = 0$ and $\ker(\mathcal{L}) = \text{span}(\mathbb{1})$. From Theorem 5.10, the Geršgorin disks of $\mathcal{L}$ are however particularly simple, as they are all centered at $[\mathcal{L}]_{ii} = 1$, and they all have radius 1, see Fig. 8.8. The normalized Laplacian is sometimes useful because of this property. Notice that premultiplication with Δ^{-1} in the consensus problem (8.4) corresponds to rescaling the "velocity" vector $\dot{x}$. Even when A is symmetric $\mathcal{L}$ need not be symmetric.

Diagonal stability and symmetrization

As we have seen above, the weight balanced case and the strongly connected case differ slightly in the way we can construct Lyapunov functions, compare (8.11) (valid also in the directed, weight balanced case) and (8.14). In both cases, the derivative of the Lyapunov function is psd, and the rank of the associated matrix is $n-1$ (or corank 1, as it is often written). In particular, in the weight balanced case the associated matrix is $L + L^\top \succeq 0$. Proposition 8.19 affirms that for $\mathcal{G}(A)$ strongly connected, the left eigenvector relative to 0, $v > 0$, can be used as Lyapunov scaling factor: $L^\top \text{diag}(v) + \text{diag}(v) L \succeq 0$. In both cases we have that $-L$ is



(a) (b)

Figure 8.8: Geršgorin disks for a normalized Laplacian. (a): for the matrix $\mathcal{L}$; (b): for $-\mathcal{L}$.

diagonally semistable. In the weight balanced case, we also get that the symmetric part of L, L^{sym}, is psd of corank 1, while this is not true for the non weight balanced case. This fact is sometimes useful when considering quantities derived from the Laplacian matrix, like the effective resistance.

8.2 DT consensus problem

The consensus problem can be formulated also in DT, and still consists in finding a (linear) dynamical system that achieves (8.1). This is often referred to as the DeGroot model, from the original paper [?]. In particular, the DT system equivalent of (8.4) is the following

$$x(t+1) = Wx(t) \tag{8.15}$$

where $W \geq 0$ is a row stochastic matrix. If we do not assume W to be symmetric, then we are dealing with a digraph case. Several conditions for DT consensus are available. For instance, if we assume strong connectivity and acyclicity of $\mathcal{G}(W)$ then we have the following.

Theorem 8.25 *If W is a primitive row stochastic matrix, then the system (8.15) achieves consensus and*

$$x^* = \lim_{t\to\infty} x(t) = \alpha\mathbb{1} \qquad \textit{where } \alpha = v^\top x(0) \tag{8.16}$$

with $v > 0$ the left eigenvector of W relative to the eigenvalue $\rho(W) = 1$, rescaled s.t. $v^\top\mathbb{1} = 1$. If in addition W is doubly stochastic, then the system (8.15) solves the average consensus problem.

Proof. Since $W \geq 0$ is primitive, we can apply the strongest version of the Perron-Frobenius theorem (Theorem 3.10). Combined with row stochasticity of W, Theorem 3.10 implies that $\rho(W) = 1$ is a simple strictly dominating eigenvalue of W of right eigenvector $v > 0$, and that

$$\lim_{t\to\infty} W^t = \mathbb{1}v^\top \tag{8.17}$$

From Proposition 6.17 we deduce that

$$\lim_{t\to\infty} x(t) = \lim_{t\to\infty} W^t x(0) = v^\top x(0)\mathbb{1} \tag{8.18}$$

When W is doubly stochastic, then $v = \mathbb{1}/n$ and average consensus follows. □

As can be seen from (8.16), the consensus value can still be interpreted as an element in the agreement subspace span($\mathbb{1}$). The point along this vector subspace is identified by the scalar value $\alpha = v^\top x(0)$, i.e., a function of the initial conditions exactly as in the CT case. When primitivity is missing, then (8.17) fails to hold, and the limit (8.18) does not exist. Hence consensus is not achieved. Recall that an easy sufficient condition for primitivity for an irreducible W is that at least one diagonal entry of W is positive: $w_{ii} > 0$ for at least one i.

Similarly to the CT case, irreducibility of W can be weakened to existence of a rooted directed spanning tree, provided that we guarantee the aperiodicity of W. The next theorem formulates a necessary and sufficient condition for DT consensus.

Theorem 8.26 *The system (8.15) achieves consensus if and only if W has a rooted directed spanning tree and the associated rooted strongly connected component is primitive. When this happen, then $x^* = \mathbb{1}v^\top x(0)$, where $v \geq 0$ is the left eigenvector of W relative to $\rho(W) = 1$, normalized so that $v^\top \mathbb{1} = 1$.*

Proof. From Theorem 6.38, when W has a rooted directed spanning tree, then the multiplicity of the $\lambda = 1$ eigenvalue is equal to 1. The Frobenius normal form (4.2) associated to W has a single rooted strongly connected component (perhaps of dimension 1). Using, if needed, a permutation to bring W in Frobenius normal form,

$$W = \left[\begin{array}{c|c} \bar{W}_{11} & 0 \\ \hline \bar{W}_{21} & \bar{W}_{22} \end{array}\right] = \left[\begin{array}{c|cccc} W_{11} & 0 & \dots & \dots & 0 \\ \hline W_{21} & W_{22} & 0 & \dots & 0 \\ W_{31} & W_{32} & W_{33} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ W_{\ell 1} & W_{\ell 2} & \dots & W_{\ell \ell-1} & W_{\ell\ell} \end{array}\right].$$

with W_{11} row stochastic (of dimension n_1) and $W_{22}, \dots, W_{\ell\ell}$ all row substochastic (strict in at least one row for each block). We need to compute the limit of the powers of W:

$$\lim_{t\to\infty} W^t = \lim_{t\to\infty} \left[\begin{array}{c|c} \bar{W}_{11}^t & 0 \\ \hline \star & \bar{W}_{22}^t \end{array}\right]$$

If the cyclicity index of W_{11} is equal to 1, then for it (8.17) holds: $\lim_{t\to\infty} W_{11}^t = \mathbb{1}_{n_1} v_1^\top$, where v_1 is the left eigenvector associated to $\rho(W_{11}) = 1$, rescaled so that $v_1^\top \mathbb{1}_{n_1} = 1$. For the other diagonal blocks it is instead $\rho(W_{ii}) < 1$, hence $\lim_{t\to\infty} W_{ii}^t = 0$, $i = 2, \dots, \ell$. Therefore,

$$W^\infty = \lim_{t\to\infty} W^t = \left[\begin{array}{c|c} \mathbb{1}v_1^\top & 0 \\ \hline S & 0 \end{array}\right]$$

where, using the argument from [20] (p. 698) that, since $\lambda = 1$ is simple, $\hat{W}^\infty$ is the projector onto $\ker(I - \hat{W})$ along range$(I - W)$, we get range$(I - W) = \ker(W^\infty)$ but also range$(I - W) = \ker(W^\infty)$, which can be written as

$$(I - W)W^\infty = \begin{bmatrix} I - \bar{W}_{11} & 0 \\ -\bar{W}_{21} & I - \bar{W}_{22} \end{bmatrix} \begin{bmatrix} \mathbb{1}_{n_1} v_1^\top & 0 \\ S & 0 \end{bmatrix} = 0$$

from which we get $S = (I - \bar{W}_{22})^{-1}\bar{W}_{21}\mathbb{1}_{n_1}v_1^\top$. Since W is row stochastic, $\bar{W}_{21}\mathbb{1}_{n_1} + \bar{W}_{22}\mathbb{1}_{n_2} = \mathbb{1}_{n_2}$, where $n_2 = n - n_1$, which implies that $(I - \bar{W}_{22})^{-1}\bar{W}_{21}\mathbb{1}_{n_1} = \mathbb{1}_{n_2}$. Hence $S = \mathbb{1}_{n_2}v_1^\top$. Splitting x into variables in the rooted strongly connected component (x_1) and remaining variables (x_2), we therefore have:

$$x^* = W^\infty x(0) = \begin{bmatrix} \mathbb{1}_{n_1}v_1^\top x_1(0) \\ \mathbb{1}_{n_2}v_1^\top x_1(0) \end{bmatrix} = \mathbb{1}v^\top x(0)$$

where $v = \begin{bmatrix} v_1 \\ 0 \end{bmatrix}$. □

The structure of the left eigenvector v says that this consensus problem corresponds to what in Section 8.1.3 we called a leader-follower problem.

8.2.1 DT consensus and Euler discretization

It is possible to obtain the DT consensus model (8.15) from an Euler discretization of the CT consensus system (8.4). The Euler discretization of (8.5) is

$$x_i(t+1) = x_i(t) + \epsilon \sum_{j=1}^{n} a_{ij}(x_j - x_i)$$

where $\epsilon =$ step size of the discretization. In matrix form, this reads

$$x(t+1) = (I - \epsilon L)x(t) = Wx(t) \tag{8.19}$$

where $L =$ Laplacian. Denote $r_{\max} = \max_i(\ell_{ii}) = \max_i\left(\sum_{j=1}^{n} a_{ij}\right)$ the max diagonal element of L. For W expressed as in (8.19) we have the following.

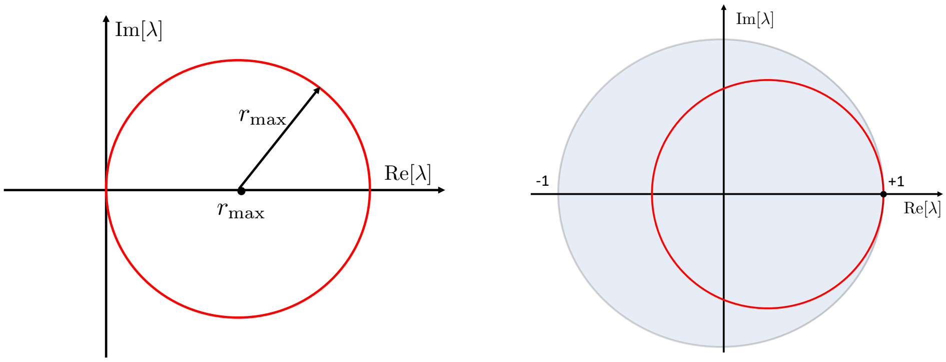
Proposition 8.27 *Consider the system (8.15), with $W = I - \epsilon L$, $\epsilon \in (0, 1/r_{\max}]$ and $L =$ Laplacian associated to $\mathcal{G}(A)$ strongly connected. Then*

1. *W is row stochastic;*
2. *$\rho(W) = 1$ is an eigenvalue of W of right eigenvector $\mathbb{1}$;*
3. *If $\mathcal{G}(A)$ weight balanced then W is doubly stochastic;*

If in addition $0 < \epsilon < 1/r_{\max}$, then

4. *W is primitive.*
5. *For all $\lambda \in \operatorname{spec}(W)$, $\lambda \neq \rho(W)$, it is $|\lambda| < \rho(W)$;*

Proof. $W\mathbb{1} = (I - \epsilon L)\mathbb{1} = \mathbb{1}$, since $L\mathbb{1} = 0$, meaning that 1 is an eigenvalue of W. From Geršgorin theorem and our discussion in Sect. 8.1.5, all eigenvalues of L are in the disk $\{s \in \mathbb{C} \text{ s. t. } |s - r_{\max}| \le r_{\max}\}$. Using the transformation $z = 1 - \epsilon s$, the disk is mapped to a disk passing through the point $z = 1$ (when $s = 0$), see Fig. 8.9. The radius of this disk depends on the choice of ϵ. In particular, if $\epsilon = 1/r_{\max}$, then the disk passes though both -1 and $+1$, hence it could happen that $I - L/r_{\max}$ is imprimitive, if some of the eigenvalues of L lie on the boundary of $\{s \in \mathbb{C} \text{ s. t. } |s - r_{\max}| \le r_{\max}\}$. If instead we choose $\epsilon < 1/r_{\max}$



(a) disk for spec(L) (b) disk for spec($I - \epsilon L$)

Figure 8.9: Spectral disks for L and $I - \epsilon L$.

then we are guaranteed that the disk in z is strictly inside the unit disk and touches only the point $+1$. Hence $\epsilon \in (0, 1/r_{\max}) \implies$ all eigenvalues $\lambda \in \text{spec}(W)$, $\lambda \neq \rho(W)$, are such that $|\lambda| < \rho(W)$, and therefore if W is irreducible, it is also primitive. If $\mathcal{G}(A)$ is weight balanced, then $L\mathbb{1} = L^\top \mathbb{1} = 0$, from which $(I - \epsilon L)\mathbb{1} = (I - \epsilon L)^\top \mathbb{1} = \mathbb{1}$, i.e., W is doubly stochastic. □

Consequently Theorem 8.25 is applicable to W obtained as Euler discretization of a CT consensus problem.

There is another reason why the expression $W = I - \epsilon L$ is useful. W represented in this way corresponds to the first order truncation of the series expansion of $e^{-\epsilon L}$, the exact solution of (8.4) computed in $x(\epsilon)$. The main reason to take this representation is that (up to the diagonal terms) the first order series expansion preserves the topology given by $\mathcal{G}(A)$, i.e., the idea that the interactions among the nodes are mediated by the edges of graph. This correspondence is instead lost if one considers the full exponential e^{-Lt}.

Extension to $\mathcal{G}(A)$ having a directed spanning tree follows from this type of reasoning. In this case, the condition $\epsilon < 1/r_{\max}$ guarantees aperiodicity.

Theorem 8.28 *Consider the system* (8.15)*, with* $W = I - \epsilon L$*,* $\epsilon \in (0, 1/r_{\max})$ *and* $L =$ *Laplacian associated to* A*. The system* (8.15) *achieves consensus if and only if* $\mathcal{G}(A)$ *has a directed spanning tree.*

8.3 Social power and weighted average

In both CT and DT consensus problems, we have seen that the asymptotic value of the opinion of the n agents is given by a weighted average of the initial conditions

$$x^* = \lim_{t \to \infty} x(t) = v^\top x(0)\mathbb{1} = \alpha \mathbb{1}$$

where we are assuming w.l.o.g. that v, the left eigenvector of the state matrix, is normalized to 1: $v^\top \mathbb{1} = 1$. The weights are decided by the components of v, and in general it is $v \geq 0$ or $v > 0$ under strong connectivity of the underlying graph.

In the context of opinion propagation (i..e, of agents that try to achieve a common opinion on a topic, by exchanging their own opinions), the left eigenvector v is normally interpreted as *social power*, i.e., the higher is the relative weight of agent i, v_i, on the final outcome $\alpha = v^\top x(0)$, the higher is the "power" that agent i has on the common decision being achieved. When in DT W is doubly stochastic (or, in CT, L is weight balanced) then all agents have the same social power, as $v = \mathbb{1}/n$, and in fact $\alpha = \frac{1}{n}\sum_i x_i(0)$ is the average of initial conditions (all taken with the same weight).

Recall from Section 7.1 the idea of a Markov chain:

$$y(t+1) = Py(t) \quad y(t) \in \mathbb{R}^n_+ \ \ \forall t \geq 0, \quad P \text{ column stochastic matrix}$$

Transpose this equation

$$y(t+1)^\top = y(t)^\top P^\top \tag{8.20}$$

and call $W = P^\top$ this row stochastic matrix. From what we have computed in Section 7.1, in the primitive case,

$$(y^*)^\top = \lim_{t\to\infty} y(t)^\top = \lim_{t\to\infty} y(0)^\top W^t = y(0)^\top \mathbb{1} v^\top \tag{8.21}$$

where v is the left eigenvector of W relative to $\rho(W) = 1$. The interpretation that can be given to (8.21) is analogous to what we have seen above for $x(t)$: $y(t) \overset{t\to\infty}{\longrightarrow} \text{span}(v)$, i.e., $y(t)$ has the interpretation of state dual to $x(t)$, converging to the asymptotic social power v. Hence also the system (8.20) can be interpreted as dual dynamics to the consensus dynamics (8.15). The two systems (8.15) and (8.20) obey to exactly the same convergence conditions. The only difference is that since we assume $y(0) \geq 0$, the system (8.20) is a positive system (i.e., $y(t) \geq 0$ for all t), while no such assumption holds for $x(t)$, which is a cooperative system instead.

An analogous argument can be set up in CT.

Chapter 9

Time-varying consensus

So far, in our consensus schemes the interaction graph was fixed for all t:

$$x(t+1) = Wx(t) \quad W \text{ row-stochastic} \tag{9.1}$$

However, it is possible that the graph of interactions among agents changes with time, for instance some agents do no participate to all parts of the discussion. Then also W = adjacency matrix of the interaction graph becomes time-varying, and instead of the static consensus (9.1) we have a time-varying consensus problem

$$x(t+1) = W(t)x(t) \qquad W(t) \text{ row-stochastic } \forall\, t \geq 0 \tag{9.2}$$

The solution of (9.2) is

$$\begin{aligned} x(1) &= W(0)x(0) \\ x(2) &= W(1)W(0)x(0) \\ &\vdots \\ x(t) &= W(t-1)\dots W(0)x(0) \end{aligned}$$

Denote $W(0:t-1) = W(t-1)\dots W(0)$. Since a product of row stochastic matrices is a row stochastic matrix, also $W(0:t-1)$ is a row stochastic matrix.

For the time invariant case (9.1) we have seen several conditions under which (9.1) converges to consensus:

- W is primitive (and hence also irreducible), see Theorem 3.10;
- W which has a directed spanning tree, see Theorem 8.26.

They all correspond to W which, after a certain power, say T, is such that for $t \geq T$, W^t contains a positive column. When W^t, for $t \geq T$, has a positive column, then it converges to a rank-1 matrix as $t \to \infty$. Notice that the condition that W^t contains a positive column for $t \geq T$ corresponds to say that on the graph $\mathcal{G}(W)$ there exists a directed path of length at most T from a root node to all other nodes, i.e., the directed spanning tree condition we saw in Theorem 8.26.

The system (9.2) is a time-varying linear system, which is an object with different properties when it comes to its asymptotic behavior. In particular, it is known that even if $W(0), \dots, W(t)$

are all "good" row stochastic matrices, in the sense that their powers converge to a rank-1 matrix: for each $W(t)$, $\lim_{k\to\infty} W(t)^k = \mathbb{1}v_t^\top$ for some $v_t > 0$, their product $W(0:t-1)$ is not necessarily "good" in the same sense, i.e., $\lim_{t\to\infty} W(0:t)$ need not converge to a rank-1 matrix.

Example 9.1 Consider the pair of stochastic matrices

$$W_1 = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad W_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

For each of them it is $\lim_{t\to\infty} W_i^t = \mathbb{1}v_i^\top$ for some $v_i > 0$, i.e., both converge to a rank-1 matrix

$$\lim_{t\to\infty} W_1^t = \begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \end{bmatrix}, \quad \lim_{t\to\infty} W_2^t = \begin{bmatrix} 1/6 & 1/3 & 1/2 \\ 1/6 & 1/3 & 1/2 \\ 1/6 & 1/3 & 1/2 \end{bmatrix}$$

Assume we have a time-varying consensus problem in which W_1 and W_2 occur according to the following switching strategy

$$W(t) = \begin{cases} W_1 & \text{if } t \text{ even} \\ W_2 & \text{if } t \text{ odd} \end{cases}$$

Hence it is

$$W(0:2t-1) = \ldots \underbrace{W_2\,W_1}_{=Q}\underbrace{W_2\,W_1}_{=Q} = Q^t$$

However

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4/9 & 4/9 & 1/9 \end{bmatrix} \quad \text{is such that} \quad \lim_{t\to\infty} Q^t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

which is not a rank-1 matrix. Therefore

$$x^* = \lim_{t\to\infty} x(t) = \lim_{t\to\infty} W(0:t-1)x(0) = \lim_{t\to\infty} Q^t x(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \\ \frac{x_1(0)}{2} + \frac{x_2(0)}{2} \end{bmatrix}$$

is not a consensus point. □

We must therefore find conditions under which a time-varying "consensus-like" system converges to consensus. A first sufficient condition is the following.

Theorem 9.2 *If the sequence of row-stochastic matrices $\{W(t)\}_{t=0}^\infty$ satisfies the following conditions for some $\beta > 0$*

1. $w_{ij}(t) \in \{0\} \cup [\beta, 1]$ $\forall\, i, j = 1, \ldots, n$, $\forall\, t \in \mathbb{N}$ *("nonvanishing couplings")*

2. *there exists a sequence of times t_k, $k = 0, 1, \ldots$ and an integer $T > 0$ such that $t_{k+1} - t_k \leq T$ $\forall\, k$ and for each k there exists a node $p_k \in \{1, \ldots, n\}$ for which*

$$[W(t_k : t_{k+1})]_{ip_k} > 0 \quad \forall\, i = 1, \ldots, n, \qquad \text{("positive column")}$$

then

$$x^* = \lim_{t\to\infty} x(t) = \lim_{t\to\infty} W(0:t-1)x(0) = \alpha\mathbb{1}$$

for some $\alpha \in \mathbb{R}$, i.e., consensus is achieved.

Proof. (only meaning) The two conditions of the theorem have the following meaning:

1. The nonzero entries of $W(t)$ never become too small (and in particular, do not vanish when $t \to \infty$).

2. For an infinite sequence of time intervals $(t_{k+1} - t_k)$ of bounded length (less that T), the corresponding matrix product $W(t_k : t_{k+1})$ has a positive column.

Having a positive column means that the graph associated to $W(t_k : t_{k+1})$ has a directed spanning tree, i.e., that some influence passes from a node p to all nodes, although this happens not in a single step, but in at most a finite number of steps ($t_{k+1} - t_k$ steps here). Since a positive column also has a positive diagonal element, then aperiodicity is guarateed as well. Then we can break time into intervals in which $W(t_{k+1} - t_k)$ all have a positive column:

$$W(0:t_k) = \underbrace{W(t_{k-1}:t_k)}_{\substack{\text{pos. col.}\\ W_k}} \cdot \ldots \cdot \underbrace{W(t_1:t:2)}_{\substack{\text{pos. col.}\\ W_2}}\ \underbrace{W(0:t_1)}_{\substack{\text{pos. col.}\\ W_1}}$$

It is a classical result that a product of stochastic matrices all having a positive column converges to a rank-1 matrix as the number of factors diverges to ∞:

$$\lim_{k\to\infty} \prod_{j=1}^{k} W_j = \mathbb{1}v^\top \quad \text{for some} \quad v > 0 \quad \text{s.t.} \quad v^\top \mathbb{1} = 1$$

This result is normally referred to as "ergodicity" in the Markov chain literature.

□

Alternative conditions can be found directly on the graphs $\mathcal{G}(W(t))$ associated to the matrices $W(t)$. What is needed are the following two conditions:

- some form of connectivity;
- aperiodicity.

An example of sufficient condition based on these two properties is the following.

Theorem 9.3 *If the sequence of row-stochastic matrices $\{W(t)\}_{t=0}^{\infty}$ satisfies the following conditions for some $\beta > 0$*

1. $w_{ij}(t) \in \{0\} \cup [\beta, 1]\ \forall\, i, j = 1, \ldots, n,\ \forall\, t \in \mathbb{N}$ *("nonvanishing couplings")*

2. $w_{ii}(t) \geq \alpha\ \forall\, i, j = 1, \ldots, n$ *and* $\forall\, t \in \mathbb{N}$; *("self-confidence")*

3. *there exists a time $T > 0$ such that $\forall\, t \in \mathbb{N}$ the union of digraphs*

$$\mathcal{G}(W(t)) \cup \mathcal{G}(W(t+1)) \cup \ldots \cup \mathcal{G}(W(t+T))$$

has a directed spanning tree ("recurrent connectivity") *then*

$$x^* = \lim_{t\to\infty} x(t) = \lim_{t\to\infty} W(0:t-1)x(0) = \mathbb{1}v^\top x(0) = \alpha\mathbb{1}$$

for $\alpha = v^\top x(0) \in \mathbb{R}$ *for some* $v > 0$*, i.e., consensus is achieved.*

Proof. We will not see the proof of this theorem but only its meaning.

1. This is the same nonvanishing weights condition of Theorem 9.2;
2. This condition guarantees primitivity of $G(W(t))$;
3. This property corresponds to a form of "uniform connectivity" on each time-window of width T, regardless of the initial condition. (Actually the connectivity is of the weakest form, i.e., a directed spanning tree).

□

Recall that for time-varying linear systems, stability á la Lyapunov is not enough. A stronger form of stability, uniform stability is required, i.e., stability regardless of the initial time t_o at which one starts the solution. The uniform connectivity condition that is used here is somewhat the analogous of this uniform stability requirement.

The conditions we are giving here are only sufficient, not necessary. It is very difficult to find necessary and sufficient condition even for time-varying linear systems. An exception is the case in which all matrices $W(t)$ have all positive diagonal entries. Then aperiodicity is always guaranteed, and only the "uniform connectivity" property must be checked. In that case necessary and sufficient conditions can be found, see e.g. [18].

Chapter 10

Synchronization

So far each agent (i.e., node of our network) is endowed with a scalar state variable, $x_i \in \mathbb{R}$ and the dynamics at a node is simply an integrator $\dot{x}_i = u_i$. In this chapter we see how to extend consensus to vectorial systems (i.e., each agent has a vector of state variables, rather that a scalar state), and also how to use consensus-related ideas to synchronize linear systems.

10.1 Consensus for higher order systems

Assume that instead of a scalar variable, each node has a vector of state variables $x_i \in \mathbb{R}^m$, and that we have "full actuation", i.e., each agent has m inputs available, so that the open loop dynamics is $\dot{x}_i = u_i$, with $u_i \in \mathbb{R}^m$. This is a somewhat trivial extension of the problem we have already investigated, and in fact to achieve consensus it can be solved trivially by an analogous law:

$$\underbrace{\dot{x}_i}_{\text{vector}} = \sum_{j=1}^{n} \underbrace{a_{ij}}_{\text{scalar}} (\underbrace{x_j}_{\text{vector}} - \underbrace{x_i}_{\text{vector}}) = \sum_{j=1}^{n} a_{ij} \left(\begin{bmatrix} x_{j,1} \\ \vdots \\ x_{j,m} \end{bmatrix} - \begin{bmatrix} x_{i,1} \\ \vdots \\ x_{i,m} \end{bmatrix} \right) \tag{10.1}$$

or, in components at each node

$$\dot{x}_{i,k} = \sum_{j=1}^{n} a_{ij}(x_{j,k} - x_{i,k})$$

The system (10.1) can be written in more compact form using Kronecker products $\otimes$: for two matrices $A \in \mathbb{R}^{p\times p}$ and $B \in \mathbb{R}^{q\times q}$, the Kronecker product of A and B is

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1p}B \\ a_{21}B & & & \\ \vdots & & & \\ a_{p1}B & \dots & & a_{pp}B \end{bmatrix} \in \mathbb{R}^{pq\times pq}$$

If $\lambda_i(A) \in \text{spec}(A)$ (resp. $\lambda_i(B) \in \text{spec}(B)$) are the p (resp. q) eigenvalues of A (resp. B), then the pq eigenvalues of $A \otimes B$ are the products of $\lambda_i(A)$ and $\lambda_i(B)$:

$$\lambda_i(A \otimes B) = \lambda_j(A)\lambda_k(B), \quad i = 1, \dots, pq, \;\; j = 1, \dots, p, \;\; k = 1, \dots, q \tag{10.2}$$

Denote $\bar{x}$ the vector obtained piling up all nm components of the x_i:

$$\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{i,1} \\ \vdots \\ x_{1,m} \\ x_{2,1} \\ \vdots \\ x_{n,m} \end{bmatrix}$$

and

$$\bar{L} = L \otimes I_m \qquad \text{where } \; I_m = m \times m \text{ identity matrix}$$

and L is the Laplacian defined in (8.3), from a given adjacency matrix A (A is $n \times n$, as the graph has n nodes). The equivalent of (8.4) for the vectorial case is then

$$\dot{\bar{x}} = -\bar{L}\bar{x} = -L \otimes I_m \bar{x} = -\left[\begin{array}{ccc|ccc|c|ccc} \ell_{11} & & & \ell_{12} & & & \cdots & \ell_{1n} & & \\ & \ddots & & & \ddots & & & & \ddots & \\ & & \ell_{11} & & & \ell_{12} & & & & \ell_{1n} \\ \hline \ell_{21} & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & \ell_{21} & & & & & & & \\ \hline \vdots & & & & & & & & & \\ \hline \ell_{n1} & & & & & & & \ell_{nn} & & \\ & \ddots & & & & & & & \ddots & \\ & & \ell_{n1} & & & & & & & \ell_{nn} \end{array}\right] \left[\begin{array}{c} x_{i,1} \\ \vdots \\ x_{1,m} \\ \hline x_{2,1} \\ \vdots \\ x_{2,m} \\ \hline \vdots \\ \hline x_{n,1} \\ \vdots \\ x_{n,m} \end{array}\right] \tag{10.3}$$

The two factors in $L \otimes I_m$ have different meanings: the second corresponds to the local dynamics at a node (here trivial dynamics, represented by the identity matrix I_m), while the first corresponds to the couplings between nodes. These couplings act in an equal way on all variables at a node. Couplings between nodes given by a Laplacian are called *diffusive couplings*. The trivial local dynamics can be replaced by more complex ones, as we do next.

10.2 Synchronization

Assume now that each agent still has a vectorial state space $x_i \in \mathbb{R}^m$, but that each agent obeys to a more complex dynamical law, for instance that at each node you have a (linear) dynamical system:

$$\dot{x}_i = F x_i + G u_i, \quad i = 1, \ldots, n, \qquad F \in \mathbb{R}^{m \times m}, \quad G \in \mathbb{R}^{m \times m} \tag{10.4}$$

The case (10.3) corresponds to $F = 0$ and $G = I$, meaning that now we also have a drift term in the dynamics of each agent, and "input" dependence more complex than just integrators.

The *synchronization problem* consists in finding a control law which is distributed according to a graph $\mathcal{G}(A)$ so that the solution of the system (10.4) follows the same trajectory for all agents. Typical case is to synchronize all states to the drift dynamics $\dot{x}_o = F x_o$, which in this case is normally called the *exosystem* (or reference generator).

Consider the simplest case, in which all agents have identical dynamics and G is invertible.

Theorem 10.1 *Assume $\mathcal{G}(A)$ is strongly connected, the input matrix G is invertible and the drift matrix F has all eigenvalues $\lambda_i(F)$ on the imaginary axis or in the left half plane:* $\mathrm{Re}[\lambda_i(F)] \leq 0$. *Then the system (10.4) with the control law*

$$u_i = G^{-1} \sum_{j=1}^{n} a_{ij}\big(x_j - x_i\big) \tag{10.5}$$

achieves exponential synchronization to a solution of $\dot{x}_o = F x_o$.

Proof. The system with the feedback law (10.5) at each node reads

$$\dot{x}_i = F x_i + \sum_{j=1}^{n} a_{ij}(x_j - x_i)$$

Consider a so-called variation of constant formula, i.e., a change of coordinates $z_i = e^{-Ft} x_i$, which corresponds to "sitting on the drift term". In this new basis, it is

$$\begin{aligned}\dot{z}_i &= -F e^{-Ft} x_i + e^{-Ft} F x_i + e^{-Ft} \sum_{j=1}^{n} a_{ij}(x_j - x_i) \\ &= \sum_{j=1}^{n} a_{ij}(z_j - z_i)\end{aligned} \tag{10.6}$$

since F and e^{-Ft} commute. This is nothing but the standard Laplacian design we considered in (10.3), for which $|z_j - z_i| \to 0$, i.e., z_i converge to consensus. Consequently, x_i synchronize to a common drift dynamics $\dot{x}_o = F x_o$. Since (10.6) is a standard linear consensus problem, convergence is exponential. □

Obviously, when $\mathrm{Re}[\lambda_i(F)] < 0$ for all i (i.e., the drift matrix is Hurwitz), then the synchronization problem trivializes to a constant steady state.

When all systems (10.4) are identical, then the overall system can be written in compact form as

$$\dot{\bar{x}} = I \otimes F \bar{x} + I \otimes G \bar{u}, \qquad \text{where} \quad \bar{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^{nm} \tag{10.7}$$

The feedback law in compact form is $\bar{u} = -L \otimes G^{-1} \bar{x}$, meaning that the closed-loop system is

$$\begin{aligned}\dot{\bar{x}} &= I_n \otimes F \bar{x} - (I_n \otimes G)(L \otimes G^{-1}) \bar{x} \\ &= I_n \otimes F \bar{x} - L \otimes I_m \bar{x} \\ &= (I_n \otimes F - L \otimes I_m) \bar{x}\end{aligned}$$

With the change of basis used in the proof, in compact notation we get

$$\dot{\bar{z}} = -L \otimes I_m \bar{z}$$

which is (10.3).

We have already noticed how in (10.7) the dynamics at a node appears in both terms in the second factor of the Kronecker product while the first factor expresses the topolgy of the graph, through the coupling given by the feedback law $\bar{u}$.

Using the same principle much more elaborated feedback laws can be developed, which relax the invertibility condition in G and do not require to make use of "full actuation" at each node. Here we report one that is rather popular in the literature

Theorem 10.2 *Assume $\mathcal{G}(A)$ is strongly connected, and at each node consider the linear system*

$$\dot{x}_i = Fx_i + Gu_i, \quad i = 1, \dots, n, \qquad F \in \mathbb{R}^{m\times m}, \quad G \in \mathbb{R}^{m\times \nu} \tag{10.8}$$

or

$$\dot{\bar{x}} = (I_n \otimes F)\bar{x} + (I_n \otimes G)\bar{u}$$

in Kronecker form. The system (10.8) *with the control law*

$$\bar{u} = -c(L \otimes K)\bar{x} \tag{10.9}$$

where $c > 0$ is a constant and $K \in \mathbb{R}^{\nu\times m}$ achieves exponential synchronization to a solution of $\dot{x}_o = Fx_o$ if and only if each matrix $F - c\lambda_i(L)GK$, $i = 2, \dots, n$, is Hurwitz.

Proof. To be added.

10.2.1 Pinned synchronization

Also the extension to a leader-follower (or "pinned") synchronization is straightforward: assume that we want an extra agent, indexed by "0", obeying the same dynamics as (10.8) but without control, to provide the reference signal (they all have the same drift dynamics, but the initial condition – and hence the trajectory – can be different). For that it is enough to modify slightly the law (10.9) adding an extra coupling part

$$\bar{u}_i = cK\left(\sum_{j=1}^{n} a_{ij}(x_j - x_i) + h_i(x_0 - x_i)\right)$$

where , $h_i \geq 0$ is the coefficient "pinning" node i to the reference node 0, of state x_0. Denoting $H = \mathrm{diag}(h_1, \dots, h_n)$ the pinning diagonal matrix, in Kronecker form, the closed loop looks like

$$\dot{\bar{x}} = (I_n \otimes F - c(L + H) \otimes GK)\,\bar{x} + c(H \otimes GK)\bar{x}_0$$

where $\bar{x}_0 = \mathbb{1} \otimes x_0$.

Chapter 11

Friedkin-Johnsen models

What makes the consensus problem peculiar when compared to the myriad of linear systems studied in the system theory literature, is that, rather than asymptotically stable, it is marginally stable, i.e., it has an entire vector space of possible equilibria, the agreement subspace. Which specific point in the agreement subspace one convergences to depends on the initial condition. This is the key property: the initial condition is not lost, as when the system is asymptotically stable, but instead remains in the asymptotic limit. This "memory" of the initial condition is what makes the problem interesting in many contexts, notably in opinion dynamics: the final opinion is not some fixed value but some combination of the initial opinions of the agents.

One can also ask if there is any other way in which such "memory" of the initial opinions can be spared in the outcome of the interaction process among the agents. The answer is yes, although nearly all such schemes are intrinsically nonlinear, and exploit their nonlinearity to achieve such goal. If we restrict to the linear world, there is just another exception, the Friedkin-Johnsen (FJ) model presented in this Section. It does so by considering inhomogeneous linear models, in which the inhomogeneous part is a function of the initial condition itself.

11.1 FJ model in DT

Recall the DT consensus problem (i.e., the DeGroot model)

$$x(t+1) = Wx(t), \quad W \geq 0 \text{ row stochastic}, \quad t = 0, 1, \ldots$$

Its solution is $x(t) \overset{t\to\infty}{\longrightarrow} \alpha \mathbb{1} \in \text{span}(\mathbb{1})$, where $\text{span}(\mathbb{1})$ is the agreement subspace, meaning that all agents eventually get the same value asymptotically. In order to achieve this:

- the agents must cooperate;
- no agent can have any special "attachment" to its own opinion (i.e., to its own $x_i(0)$);
- no agent can be subject to external influences, where "external" here means outside of the community of agents forming the graph.

As an opinion model, consensus is not realistic in most cases: experiments carried out mostly on small-scale controlled groups of individuals have shown that indeed agents can influence each other, and that the end-point opinion belongs to the convex hull of the initial conditions (for the model: $x^* \in \text{co}(x(0))$) but does not corresponds to consensus.

The *Friedkin-Johnsen* (FJ) model is meant to represent this "contraction" towards $\text{co}(x(0))$ which is however not necessarily a consensus point. In DT, the standard form of FJ model is

$$x(t+1) = (I - \Theta)\, W x(t) + \Theta x(0) \tag{11.1}$$

where

$$\Theta = \begin{bmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{bmatrix}, \quad \theta_i \in [0, 1]$$

The vector $\theta = \begin{bmatrix} \theta_1 & \dots & \theta_n \end{bmatrix}$ represents the vector of stubbornness coefficients, expressing how each agent is "attached" to its own initial condition (i.e., to its own opinion at the begin of the discussion). In components, (11.1) reads

$$x_i(t+1) = (1 - \theta_i) \sum_{j=1}^{n} w_{ij} x_j(t) + \theta_i x_i(0) \tag{11.2}$$

i.e., a convex combination is taken between the usual update rule of the DeGroot model and the initial condition of the agent. Depending on the convex combinators θ_i in the model (11.1) we have

- If $\theta_i = 1$ then the i-th agent is totally stubborn, as (11.2) becomes

  $$x_i(t+1) = 0 + x_i(0)$$

  i.e., the i-th agent never changes its mind;

- If $\theta_i = 0$ then the i-th agent has no attachment to its own opinion ("oblivious agent"), as (11.2) becomes the usual consensus update rule

  $$x_i(t+1) = \sum_{j=1}^{n} w_{ij} x_j(t)$$

- In between ($\theta_i \in (0,\, 1)$) the i-the agent is partially stubborn, i.e., it has some attachment to its own opinion.

When $\Theta = 0$ then (11.1) reduced to the usual DeGroot consensus model. When $\Theta = I$, $x(t) = x(0)$ for all t, i.e., no agent changes opinion.

Definition 11.1 *The FJ model (11.1) is said* θ-connected *if all nodes of the graph* $\mathcal{G}(W)$ *either have* $\theta_i > 0$ *or are connected via directed paths to some nodes* i *for which* $\theta_i > 0$.

The θ-connectivity condition means that all agents perceive the influence of at least one stubborn agent, i.e., that on the graph $\mathcal{G}(W)$ there exists a directed path from the agents i s.t. $\theta_i > 0$ to all other agents. This notion plays the same role as the input connectivity condition seen in previous Chapters, as explained in the following theorem.

Theorem 11.2 *Consider the FJ model (11.1) with W row stochastic, and $\theta \gneq 0$. Then the following conditions are equivalent.*

1. *The system (11.1) is θ-connected.*
2. *$\rho((I-\Theta)W) < 1$ i.e., the matrix $(I-\Theta)W$ is Schur stable;*
3. *the system has a unique asymptotically stable equilibrium point $x^* = \lim_{t\to\infty} x(t) = Vx(0)$, with*
$$V = (I-(I-\Theta)W)^{-1}\,\Theta$$

When these conditions hold, then it is $x^* \in \text{co}(x(0))$.

Proof. 1. $\iff$ 2. Since Θ and $I-\Theta$ are row substochastic matrices, also $(I-\Theta)W$ is row substochastic. In particular, denoting $\mathcal{V}^\theta = \{i \in \mathcal{V} \text{ s. t. } \theta_i > 0\}$, then $\mathcal{V}^\theta$ corresponds to both strictly substochastic rows and nonzero indices in the affine term $\Theta x(0)$, i.e., in the terminology we used in Chapter 6, $\mathcal{V}^\theta = \mathcal{V}^{\text{in}} = \mathcal{V}^{\text{out}}$. θ-connectivity then corresponds to existence in the graph $\mathcal{G}((I-\Theta)W)$ of a spanning forest rooted at $\mathcal{V}^\theta$. Hence, from Theorem 6.37, the condition is equivalent to $(I-\Theta)W$ being Schur stable. Obviously condition 2 implies and is implied by the existence, uniqueness and global asymptotic stability of the equilibrium point. To compute its value, observe that since (11.1) is a linear inhomogeneous system, its solution is

$$x(t) = \left(\big((I-\Theta)W\big)^t + \sum_{k=0}^{t-1} ((I-\Theta)W)^k \Theta \right) x(0)$$

The first factor is Schur stable, hence $((I-\Theta)W)^t \overset{t\to\infty}{\longrightarrow} 0$. The second factor is a summable series, hence we can use Neumann series expansion to express it as

$$\lim_{t\to\infty} \sum_{k=0}^{t-1} ((I-\Theta)W)^k = \big(I-(I-\Theta)W\big)^{-1}$$

from which the expression for V follows. Furthermore, since each term in the summation is nonnegative, then $\big(I-(I-\Theta)W\big)^{-1} \geq 0$. To show that V is row stochastic we use the following chain of identities:

$$\begin{aligned} W\mathbb{1} &= \mathbb{1} \\ \Theta W \mathbb{1} &= \Theta \mathbb{1} \\ (I-\Theta)W\mathbb{1} &= \mathbb{1} - \Theta\mathbb{1} \\ (I-(I-\Theta)W)\mathbb{1} &= \Theta\mathbb{1} \\ \mathbb{1} &= \big(I-(I-\Theta)W\big)^{-1}\Theta\mathbb{1} \end{aligned}$$

Since V is row stochastic, $x_i^* = \sum_{j=1}^n V_{ij} x_j(0)$ is a convex combination of $x(0)$, hence $x^* \in \text{co}(x(0))$. □

Unlike in the consensus problem, which has marginal stability, the FJ model has asymptotic stability, but the equilibrium point is not the origin. In fact the system (11.1) is an inhomogeneous linear system, with a constant inhomogeneous term (i.e., $\Theta x(0)$). In (11.1), instead of

$x(0)$ in the inhomogenoeus term, one can have a constant u, expressing an external influence, or both.

Since opinions are not constrained in sign, the FJ model (11.1) is however a cooperative system, rather than a positive system.

According to Theorem 11.2, when the θ-connectivity property fails, then $(I - \Theta)W$ fails to be Schur stable. What happens is that some groups of agents become stubbornness-free, and therefore perform an averaging protocol, which leads to marginal stability for the associated interaction matrix. More formally, in the Frobenius normal form (4.2) one or more of the rooted strongly connected component is row stochastic (rather than substochastic), hence, from Theorem 6.38, the $\lambda = 1$ eigenvalue appears in $(I - \Theta)W$. From Theorems 6.36 and 6.38, in these cases we still have marginal stability in the FJ model, but to avoid oscillatory behaviors in the solution $x(t)$ we need to impose that these rooted row stochastic strongly connected components are also primitive. Notice that marginal stability can never happen when W is irreducible (and $\theta \gneq 0$): this is condition 1 of Theorem 6.36.

Proposition 11.3 *Consider the FJ model* (11.1) *with W row stochastic and irreducible. Then for any $\theta \gneq 0$ the FJ model is θ-connected, and consequently $(I - \Theta)W$ is Schur stable.*

11.2 FJ model in CT

In CT, there are various possible versions of the FJ model. One of them is known as Taylor model [?]. The one presented here was suggested by Francesca Ceragioli.

The basic idea is the same as in DT: the agents can be stubborn i.e., “attached” to their own initial condition. This can be modeled as an affine (i.e., inhomogeneous) term in the ODE that describes the evolution of the opinions of the agents, i.e., in the model (8.4).

$$\dot{x} = -((I - \Theta)L + \Theta)x + \Theta x(0), \quad \Theta = \begin{bmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{bmatrix}, \quad \theta_i \in [0, 1] \tag{11.3}$$

or, in components,

$$\dot{x}_i(t) = (1 - \theta_i) \sum_j a_{ij} \big(x_j(t) - x_i(t)\big) - \theta_i \big(x_i(t) - x_i(0)\big)$$

The stubbornness vector θ has the same interpretation as in DT, and leads to the same extreme cases as in DT:

- $\theta_i = 0$ corresponds to no stubbornness at all (i.e., to a consensus-like update rule for the corresponding agent);
- $\theta_i = 1$ corresponds to total stubbornness: $\dot{x}_i = \Theta(x - x(0)) = 0$ since $x(0)$ is the initial condition (hence $\dot{x}|_{t=0} = 0 \implies x(t) = x(0)$ for all t.).
- In between ($\theta_i \in (0, 1)$), an agent is partially stubborn.

Theorem 11.4 *Consider the system* (11.3) *on a digraph $\mathcal{G}(A)$, with L the corresponding Laplacian, and $\theta \gneq 0$. Then the following are equivalent.*

1. *The system is θ-connected.*
2. *the matrix $-((I-\Theta)L+\Theta)$ is Metzler and Hurwitz;*
3. *the system has a unique asymptotically stable equilibrium point $x^* = lim_{t\to\infty}x(t) = Vx(0)$, with*
$$V = ((I-\Theta)L+\Theta)^{-1}\Theta$$
a row-stochastic matrix.

When these conditions hold, then $x^* \in \mathrm{co}(x(0))$.

Proof. 1. $\implies$ 2. The matrix $-((I-\Theta)L+\Theta)$ is Metzler by construction. To show that it is Hurwitz:

$$\begin{aligned} L\mathbb{1} &= 0 \\ \implies (I-\Theta)L\mathbb{1} &= 0 \\ \implies ((I-\Theta)L+\Theta)\mathbb{1} &= \Theta\mathbb{1} \geq 0 \end{aligned} \tag{11.4}$$

hence $-((I-\Theta)L+\Theta)$ is weakly diagonally dominant by rows. From (11.4), the strict diagonally dominant rows correspond to the set $\mathcal{V}^\theta = \{i \in \mathcal{V} \text{ s. t. } \theta_i > 0\}$, and θ connectivity corresponds to existence of a spanning forest rooted at $\mathcal{V}^\theta$. From Theorem 6.28, this is equivalent to $-((I-\Theta)L+\Theta)$ being Metzler Hurwitz. The equivalence 2. $\iff$ 3. is standard. From (11.4), the value of the unique globally asymptotically stable equilibrium point is $x^* = ((I-\Theta)L+\Theta)^{-1}\Theta x(0)$. To show that $V = ((I-\Theta)L+\Theta)^{-1}\Theta$ is row stochastic: from Theorem 6.34, $-((I-\Theta)L+\Theta)$ Metzler Hurwitz implies $((I-\Theta)L+\Theta)^{-1} \geq 0$, and hence so is V. From (11.4) and invertibility, it is $((I-\Theta)L+\Theta)^{-1}\Theta\mathbb{1} = \mathbb{1}$. □

Similarly for the DT case, any $\theta \gneq 0$ is enough to get asymptotic stability when $\mathcal{G}(A)$ is strongly connected.

Proposition 11.5 *Consider the FJ model (11.3) on a strongly connected digraph $\mathcal{G}(A)$, with L the corresponding Laplacian. Then for any $\theta \gneq 0$ the FJ model is θ-connected, and consequently $-((I-\Theta)L+\Theta)$ is Metzler Hurwitz.*

In the CT case, it is possible to apply a change of basis that shifts the equilibrium point to the origin. Defining $z = x - x^* = x - ((I-\Theta)L+\Theta)^{-1}\Theta x(0)$ as new variable, then

$$\begin{aligned} \dot{z} = \dot{x} &= -((I-\Theta)L+\Theta)x + \Theta x(0) \\ &= -((I-\Theta)L+\Theta)z - ((I-\Theta)L+\Theta)\left((I-\Theta)L+\Theta\right)^{-1}\Theta x(0) + \Theta x(0) \\ &= -((I-\Theta)L+\Theta)z \end{aligned}$$

meaning that indeed the system in z is homogeneous, and $z(t) \overset{t\to\infty}{\longrightarrow} 0$.

11.3 Concatenated DT Friedkin-Johnsen model

The FJ model is a linear system in which the state update matrix $(I-\Theta)W$ is Schur stable: $\rho((I-\Theta)W) < 1$. As for all linear systems, convergence to the unique steady state x^* occurs exponentially fast, meaning that even though theoretically x^* is reached only when $t \to \infty$, in practice even for t moderately small it is $x(t) \simeq x^*$. So if we think of an FJ model as a discussion event, in practice it can make sense to consider multiple such discussion events occurring one after the other. More technically, we can consider a two-time-scale setting in which on the fast time-scale we have an instance of an FJ model, representing a single discussion event, and on the slow time-scale we keep track of the discussion events occurring in sequence one after the other. If the two time scales are sufficiently well separated, then we can assume that the discussion is completed (and we compute the end-point of the discussion by solving the FJ model) in the fast time scale before moving to the next discussion event[1]. On different discussion events both the interaction graph $\mathcal{G}(W)$ and the stubbornness Θ can change: this mimics e.g. the fact that an agent may have different levels of stubbornness on different topics (if the discussions are on different topics). The scenario we are interested to investigate is when the opinions are concatenated, i.e., the opinion end-point of a discussion event serves as starting point for the next discussion. We call this model *concatenated FJ model*. More properly, the ingredients are the following:

- two time scales: s for the discussion event, t for the time axis within each discussion. Consequently the state (i.e., opinion) vector is indexed by both s and t: $x(s,t)$.
- An FJ model is used for each discussion event. We assume that both the interaction matrix W and the stubbornness matrix Θ are event-dependent: $W(s)$ and $\Theta(s)$. From (11.1), the FJ model at the sth step is therefore

$$x(s,t+1) = (I - \Theta(s))\, W(s) x(s,t) + \Theta(s) x(s,0), \quad s = 1,2,\ldots. \tag{11.5}$$

- Opinions are concatenated: the final state at event $s-1$ becomes initial condition at event s

$$x(s,0) = x(s-1,\infty).$$

From Theorem 11.2, under the assumption of θ-connectivity, we have that the end-point of the FJ model at event s, $x(s,\infty)$, lies inside the convex hull $\mathrm{co}(x(s,0))$, meaning that the dynamics during each discussion contract. Assuming concatenation, i.e., $x(s,0) = x(s-1,\infty)$, means that the contractivity property propagates through the chain of discussion events, i.e., the opinions tend to get closer to each other as s grows, see Fig. 11.1(b). Hence a natural question to ask is: can the concatenated FJ model achieve consensus? This is the problem studied next.

From Theorem 11.2, the solution of (11.5) is $x(s,\infty) = P(s)x(s,0)$, $s = 1,2,\ldots$, where

$$P(s) = \left(I - (I - \Theta(s))W(s)\right)^{-1} \Theta(s) x(s,0) \tag{11.6}$$

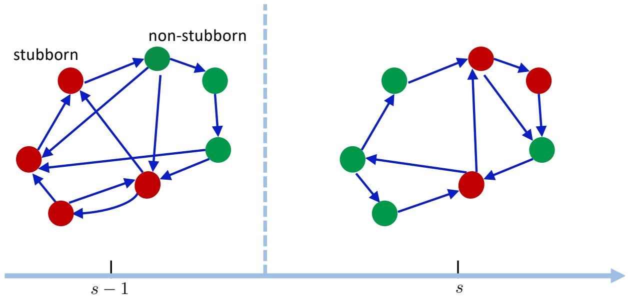
is an event-dependent row stochastic matrix. To simplify notations, let us denote

$$y(s) = x(s,\infty) \quad \text{and} \quad y(0) = x(1,0).$$

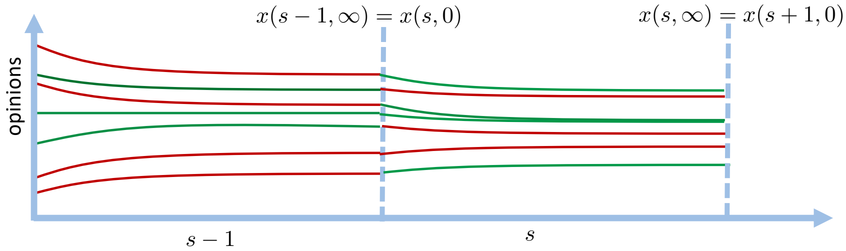
Then, the concatenated FJ model, at each event s, can be compactly written as

$$y(s) = P(s) y(s-1), \quad s = 1,2,\ldots \tag{11.7}$$

[1] Mathematicians, like Chuck Norris, can count to infinity – twice...



(a)



(b)

Figure 11.1: Concatenated FJ model. (a): at each event s, the set of stubborn agents may change and also the interaction graph $\mathcal{G}(W)$. (b): Concatenation means that final opinions at event $s-1$ become initial conditions at event s.

Since $P(s)$ is a row-stochastic matrix for all s, the system (11.7) is formally identical to a time-varying consensus problem like those we studied in Chapter 9, with the extra complication that the row stochastic matrices $P(s)$ have the structure (11.6), determined by the underlying FJ problems. As in Chapter 9, (11.7) can be rewritten as

$$y(s) = P(1:s)y(0) = P(s)P(s-1)\dots P(1)y(0).$$

From Theorem 9.2, we have the following sufficient condition for consensus.

Theorem 11.6 (Consensus in the concatenated FJ model) *Consider the concatenated FJ model (11.7) where $P(s)$ is given in (11.6). If*

1. *$W(s)$ a sequence of row stochastic irreducible matrices of entries $w_{ij}(s) \in \{0\} \cup [\beta_1, 1]$, $\forall\, i,j = 1,\dots,n$, $\forall\, s \in \mathbb{N}$ and for some $\beta_1 > 0$,*
2. *$\Theta(s) = \mathrm{diag}(\theta_1(s),\dots,\theta_n(s))$ s.t. $\theta_i(s) \in \{0\} \cup [\beta_2, \beta_3]$, $\forall\, i = 1,\dots,n$, $\forall\, s \in \mathbb{N}$ and for some $\beta_2 > 0$, $\beta_3 < 1$,*

then the model (11.7) achieves consensus:

$$y^* = \lim_{s\to\infty} y(s) = \lim_{s\to\infty} P(1:s)y(0) = \alpha\mathbb{1}$$

for some $\alpha \in \mathbb{R}$.

Proof. First of all, the irreducibility condition on $W(s)$ implies the θ-connectivity condition used in Theorem 11.2, meaning that indeed at every s we have $\rho((I-\Theta(s))W(s)) < 1$ and hence we can compute $P(s)$ in (11.6). The conditions on $w_{ij}(s)$ and on $\theta_i(s)$ are the analogous of the nonvanishing couplings condition considered in Theorem 9.2, and imply that $P_{ij}(s)$ is nonvanishing when $s \to \infty$. In fact, using the Neumann series expansion of $(I-\Theta(s))W(s)$,

$$P(s) = \lim_{t\to\infty} \sum_{k=0}^{t-1} ((I-\Theta(s))W(s))^k \Theta(s)$$

all factors $((I-\Theta(s))W(s))^k$ are nonnegative, and, using the characterization (3.1) of irreducibility and the fact that $\theta_i(s) \leq \beta_3 < 1$ (no totally stubborn agent), we get that $\sum_{k=0}^{t-1}((I-\Theta(s))W(s))^k > 0$ as soon as $t \geq n$. Without loss of generality, assume that at event s the first r agents are stubborn ($\beta_2 \leq \theta_i(s) \geq \beta_3$ for $i = 1, \ldots, r$), while the remaining $n-r$ are non-stubborn ($\theta_i(s) = 0$ for $i = r+1, \ldots, n$), then in $P(s)$ we have the following nonzero pattern

$$P(s) = \begin{bmatrix} * & \dots & \dots & * \\ \vdots & & & \vdots \\ & & & \\ \vdots & & & \vdots \\ * & \dots & \dots & * \end{bmatrix} \begin{bmatrix} * & & & & \\ & \ddots & & & \\ & & * & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} = \begin{bmatrix} * & \dots & * & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ & & & & & \\ \vdots & & \vdots & \vdots & & \vdots \\ * & \dots & * & 0 & \dots & 0 \end{bmatrix}.$$

If $r > 0$, then $P(s)$ has at least one positive column. If this is the case for all s, then $P(1:s)$ is a product of s matrices all having a positive column, hence it follows from an argument analogous to that used in the proof of Theorem 9.2 that $\lim_{s\to\infty} P(1:s) = \mathbb{1}v^\top$, i.e., $P(1:s)$ converges to a rank-1 matrix. Therefore $y(s)$ converges to consensus. If instead, at some s, $r = 0$ (no stubborn agent), then the FJ problem reduces to a consensus problem, for which consensus is achieved. It is maintained for all FJ problems afterwards. □

11.4 Multidimensional Friedkin-Johnsen model

This section is from [21].

Ideally, the FJ model (11.1) describes a discussion among a group of agents on a single issue. Assume that we have m issues to discuss, and that we want to discuss them simultaneously (not concatenated!) because they are interdependent. Each of them is modeled as an FJ model, but these models must be run simultaneously because they are interdependent. Each agent has now a vector of m opinion variables: $x_i = \begin{bmatrix} x_i^1 & \dots & x_i^m \end{bmatrix}^\top \in \mathbb{R}^m$. Denoting $C \in \mathbb{R}^{m\times m}$ the topic interdependence matrix, then a possible model is the following

$$\bar{x}(t+1) = (I-\Theta)W \otimes C\bar{x}(t) + \Theta \otimes I_m \bar{x}(0) \tag{11.8}$$

where $\otimes$ = Kronecker product and $\bar{x} = \begin{bmatrix} x_1^\top & \dots & x_n^\top \end{bmatrix}^\top \in \mathbb{R}^{nm}$.

Theorem 11.7 (Stability of multidimensional FJ model) *Consider the multidimensional FJ model (11.8).*

1. *If C is row stochastic then the evolution stays in the convex hull: $x_i(0) \in [a,\, b]^m$ for all i $\Longrightarrow x_i(t) \in [a,\, b]^m$ for all i, where $[a,\, b]$ is an interval of $\mathbb{R}$.*
2. *The model (11.8) is Schur stable if and only if $\rho(C) < \rho((I - \Theta)W)^{-1}$.*
3. *When $\rho(C) = 1$ then the model (11.8) is Schur stable if and only if the model (11.1) is θ-connected.*

Proof. Writing (11.8) in components,

$$x_i(t+1) = (1-\theta_i)C \sum_j w_{ij} x_j(t) + \theta_i x_i(0)$$

We know that in the ordinary FJ model $x(t) \in \mathrm{co}(x(0))$ always hold. When C is stochastic (i.e., $C \geq 0$and $C\mathbb{1} = \mathbb{1}$), then the presence of C adds yet another convex combination. In particular, for each component of x_i, if $\xi(t) = \sum_j w_{ij} x_j(t)$, $x_i^\ell(t+1) = (1-\theta_i)\sum_k c_{\ell k}\xi_k(t) + \theta_i x_i^\ell(0)$. These convex combinations belong to the convex hull.

The system (11.8) is a linear system, hence it is Schur stable iff its state update matrix (here $(I-\Theta)W \otimes C$) has all eigenvalues in the open unit disk. From the properties of Kronecker products, $\lambda((I-\Theta)W \otimes C) = \lambda((I-\Theta)W)\lambda(C)$. Hence a necessary and sufficient condition is that $\rho((I-\Theta)W)\rho(C) < 1$.

When $\rho(C) = 1$, the condition above becomes $\rho((I-\Theta)W) < 1$ which is the usual condition for an FJ model to be stable. The θ-connectivity condition becomes then a sufficient condition.

Example 11.8 For the case $m = 2$, let us consider the model (11.8) with the following matrices:

$$W = \begin{bmatrix} 0.22 & 0.12 & 0.36 & 0.3 \\ 0.147 & 0.215 & 0.344 & 0.294 \\ 0 & 0 & 1 & 0 \\ 0.09 & 0.178 & 0.446 & 0.286 \end{bmatrix} \quad \text{and} \quad \Theta = \mathrm{diag}(W)$$

and the initial condition

$$x_1(0) = \begin{bmatrix} 25 \\ 25 \end{bmatrix}, \quad x_2(0) = \begin{bmatrix} 25 \\ 15 \end{bmatrix}, \quad x_3(0) = \begin{bmatrix} 75 \\ -50 \end{bmatrix}, \quad x_4(0) = \begin{bmatrix} 85 \\ 5 \end{bmatrix}$$

Consider further following interdependence matrices

$$C_1 = I, \quad C_2 = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}, \quad C_3 = \begin{bmatrix} 0.8 & -0.2 \\ -0.3 & 0.7 \end{bmatrix}$$

Since $C_1 = I$, the topics are independent, hence the two FJ models can be run independently. The steady state solution for C_1 is:

$$x_1^* = \begin{bmatrix} 60 \\ -19.3 \end{bmatrix}, \quad x_2^* = \begin{bmatrix} 60 \\ -21.5 \end{bmatrix}, \quad x_3^* = \begin{bmatrix} 75 \\ -50 \end{bmatrix}, \quad x_4^* = \begin{bmatrix} 75 \\ -23.2 \end{bmatrix}$$

For C_2 and C_3, the topics are instead interdependent. The steady state solution for C_2 is:

$$x_1^* = \begin{bmatrix} 39.2 \\ 12 \end{bmatrix}, \quad x_2^* = \begin{bmatrix} 39 \\ 10 \end{bmatrix}, \quad x_3^* = \begin{bmatrix} 75 \\ -50 \end{bmatrix}, \quad x_4^* = \begin{bmatrix} 56 \\ 5.3 \end{bmatrix}$$

The steady state solution for C_3 is:

$$x_1^* = \begin{bmatrix} 52.3 \\ -30.9 \end{bmatrix}, \quad x_2^* = \begin{bmatrix} 52.1 \\ -33.3 \end{bmatrix}, \quad x_3^* = \begin{bmatrix} 75 \\ -50 \end{bmatrix}, \quad x_4^* = \begin{bmatrix} 68.3 \\ -33.2 \end{bmatrix}$$

The trajectories are shown in Fig. 11.2.

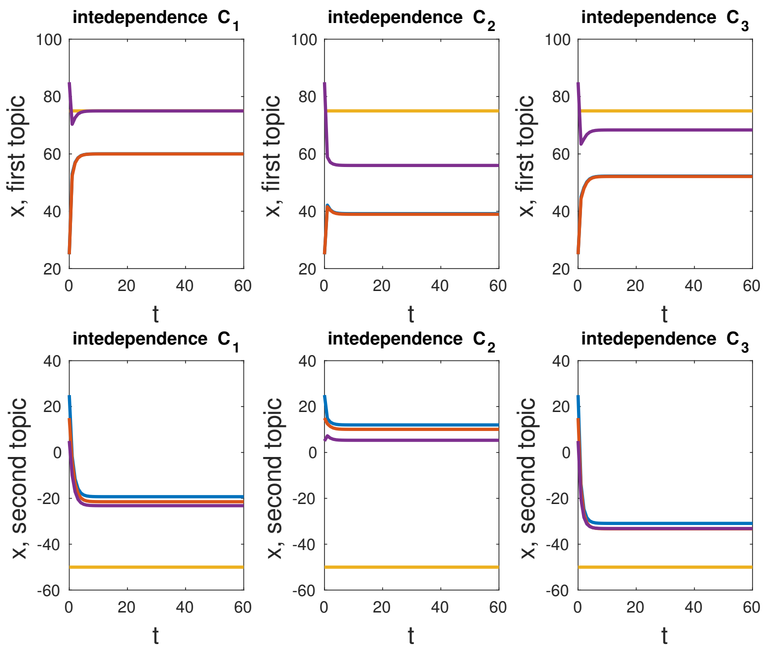


Figure 11.2: Multidimensional FJ model. Top row: opinion on the first topic, for the three interdependence matrices C_i. Bottom row: opinion on the second topic.

Chapter 12

Bipartite consensus

> *"Since universal ultimate agreement is an ubiquitous outcome of a very broad class of mathematical models, we are naturally led to inquiry what on earth one must assume in order to generate the bimodal outcomes of community cleavage studies"*
>
> R. P. Abelsson[1].

So far, all agents cooperate to achieve a common goal. Cooperation is encoded in the sign of the adjacency matrix: $A \geq 0$. However, there are situations in which, alongside cooperation, there is also some form of antagonism among the agents. This is especially true on social networks, where unfriendly interactions happen rather often. One way to represent this on graphs is to use signed graphs. In particular we can associate the following interpretation to the edge signs:

- positive edges: any type of "friendly tie" (cooperation/alliance/trust);
- negative edges: any type of "unfriendly tie" (antagonism/rivalry/mistrust).

In sociology, this way of encoding antagonism dates back to the Fifties, as a way of formalizing why situations like "the enemy of my enemy is my friend" lead to less "social tension" than situations like "the enemy of my friend is my friend", see Fig. 12.1. To understand the meaning of these social tensions, imagine a situation in which you are inviting two of your friends for dinner. Do you expect the dinner to be more successful if these two people like each other or if they "hate" each other? Conceptually, these situations can be describing using the notion of structural balance given in Section 4.2. When a graph is structurally balance, then the community of agents can be split into two disjoint subcommunities in each of which there are only friends, and all agents of one subcommunity are "enemies" of all agents in the other subcommunity, see Fig. 4.8(a).

The question that we pose ourselves in the chapter is the following: is it still possible to achieve consensus in presence of antagonistic interactions?

[1]In R. P. Abelsson. Mathematical models of the distribution of attitudes under controversy. Contribution in Mathematical Psychology. N. Frederiksen and H. Gulliksen (eds.) Holt Rinehart and Winston, 1964, pp. 142-160

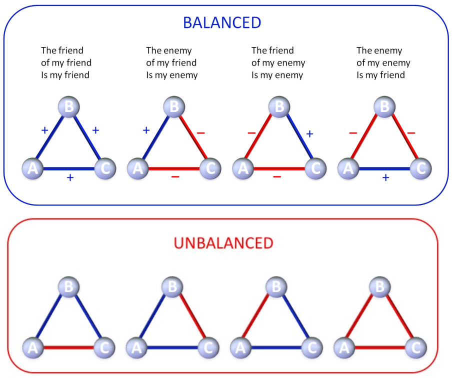


Figure 12.1: Structural balance/unbalance in triads.

12.1 CT bipartite consensus

Recall that in our consensus reasoning we use the fact that $A \geq 0$ to conclude that the corresponding negated Laplacian $-L$ is a Metzler matrix, and we rely upon the fact that Perron-Frobenius theorem guarantees the dominance of a special eigenspace span($\mathbf{1}$), to which all trajectories converge to. If A is not nonnegative, however, in general it is no longer possible to conclude on the existence of such properties for the adjacency matrix.

In the structurally balanced case a notion similar to consensus, but respecting the partition into two opposite faction still exists. We call this bipartite consensus. A group of agents is said to achieve *bipartite consensus* if

$$\lim_{t\to\infty} \big(|x_i(t)| - |x_j(t)|\big) = 0 \qquad \forall\, i, j \in \mathcal{V} \tag{12.1}$$

or, in vector form

$$|x(t)| \overset{t\to\infty}{\longrightarrow} \alpha(x(0))\mathbf{1}$$

where $\alpha(x(0))$ is a scalar depending on the initial condition $x(0)$. The presence of the absolute value means that if the node set $\mathcal{V}$ is partitioned in to $\mathcal{V}_1$ and $\mathcal{V}_2$ (the two factions in which a structurally balanced graph can be split, see Section 4.2),

$$x_i(t) \overset{t\to\infty}{\longrightarrow} \alpha(x(0)) \quad \text{on one side of the partition,} \quad i \in \mathcal{V}_1$$
$$x_j(t) \overset{t\to\infty}{\longrightarrow} -\alpha(x(0)) \quad \text{on the other side of the partition,} \quad j \in \mathcal{V}_2$$

To set up a consensus problem we need a Laplacian matrix. For signed graphs there is more than one possibility. Here we consider the following.

Definition 12.1 *Given a signed graph of adjacency matrix A_s, the* signed Laplacian $L_s = [\ell_{s,ij}]$ *is the matrix of entries*

$$\ell_{s,ij} = \begin{cases} \sum_{j=1}^{n} |a_{s,ij}| & \textit{if} \ \ i = j \\ -a_{s,ij} & \textit{if} \ \ i \neq j \end{cases} \tag{12.2}$$

With respect to (8.3), the diagonal terms now contain the sum of the absolute values of the row elements. Using the signed Laplacian (12.2), the equivalent of (8.4) becomes the *bipartite* *consensus problem*

$$\dot{x} = -L_s x \tag{12.3}$$

or, in components,

$$\dot{x}_i = \sum_{j=1}^{n} |a_{s,ij}| \big(\mathrm{sgn}(a_{s,ij}) x_j - x_i \big) \tag{12.4}$$

This Laplacian collapses into the usual Laplacian (8.3) when the adjacency matrix is nonnegative. The interpretation of the update rule (12.4) is the following:

1. When the agent j is friendly with agent i, $\mathrm{sgn}(a_{s,ij}) > 0 \implies$ the difference $x_j - x_i$ is taken in (12.4) and weighted by $a_{s,ij}$, as usually done in averaging rules;

2. When the agent j is unfriendly with agent i, $\mathrm{sgn}(a_{s,ij}) < 0 \implies -x_j - x_i$ is taken in the sum of (12.4), i.e., agent i takes the "opposite" of the information sent by its "enemy" agent j, and weights it with $|a_{s,ij}|$.

Combining the two cases one gets

$$|a_{s,ij}| \big(\mathrm{sgn}(a_{s,ij}) x_j - x_i \big) = \begin{cases} a_{s,ij} \big(x_j - x_i \big) & \text{if } j \text{ is friend of } i \\ |a_{s,ij}| \big(-x_j - x_i \big) & \text{if } j \text{ is enemy of } i \end{cases}$$

Let us investigate the properties of the signed Laplacian (12.3). By construction it is still diagonally dominant, even though in the weakest possible form, which we called diagonal equipotence in Section 5.2.2:

$$|\ell_{s,ii}| = \sum_{j=1}^{n} |a_{s,ij}| = \sum_{j=1}^{n} |\ell_{s,ij}| \tag{12.5}$$

We know from Theorem 5.12 that, under suitable assumptions like irreducibility of A_s, at least marginal stability is guaranteed. In facr, according to item 2 of Theorem 5.12, diagonally equipotent matrices can be marginally or asymptotically stable. It is easily seen in examples (we shown it in Example 12.5 below) that one or more of the row sums in L_s (without absolute value) can be different from zero, that is, $L_s \mathbf{1} \neq 0$, meaning that 0 is no longer an eigenvalue of L_s. We then expect that this leads to situation which we have not encountered in our consensus analysis in Chapter 8: in this case $-L_s$ becomes Hurwitz stable, instead of marginally stable (recall that for consensus-like properties we need marginal stability, so as to have an entire eigenspace of equilibria). The following lemma specializes item 2 of Theorem 5.12 to the case of signed Laplacian matrices, and affirms that structural balance is a necessary and sufficient condition for existence of the 0 eigenvalue in L_s, and hence for marginal stability of $-L_s$.

Lemma 12.2 *Consider a strongly connected signed graph $\mathcal{G}(A_s)$ and the associated signed Laplacian L_s. Then $\lambda_1(L_s) = 0$ is an eigenvalue of L_s if and only if $\mathcal{G}(A_s)$ is structurally balanced.*

Proof. (sketch) From Proposition 4.11, we know that a graph which is structurally balanced admits a diagonal change of basis $S = \mathrm{diag}(s_1, \ldots, s_n)$, $s_i = \pm 1$, such that SA_sS is nonnegative. It is easy to show that the same S renders also $-SL_sS$ Metzler. In addition $L_g = SL_sS$ is a "standard" Laplacian, with all properties of the Laplacian we discussed in Proposition 8.13, among them $\lambda_1(L_g) = 0$ being an eigenvalue of eigenvector $\mathrm{span}(\mathbf{1})$. Since L_s and L_g are similar matrices, it must be that $\lambda_1(L_s) = 0$ is an eigenvalue also of L_s. Showing the converse is more involved, but intuitively, we can say that whenever the signed graph is structurally unbalanced, a matrix S rendering SA_sS nonnegative fails to exist, and this implies that at least one of the inequalities in (12.5) must be strict. In turn this can be shown to clash with the hypothesis that 0 is an eigenvalue of L_s. $\square$

Combining Lemma 12.2 with Theorem 5.12, we have that $\lambda(L_s) = 0$ is the spectral abscissa of $-L_s$, from which the following stability characterization of L_s follows.

Corollary 12.3 *For the signed Laplacian L_s of a strongly connected signed graph $\mathcal{G}(A_s)$ it is*

1. *$-L_s$ is marginally stable if and only if $\mathcal{G}(A_s)$ is structurally balanced;*
2. *$-L_s$ is Hurwitz if and only if $\mathcal{G}(A_s)$ is structurally unbalanced.*

In turn, from Corollary 12.3 we get a complete characterization of the solutions of (12.3), as summarized by the following theorem.

Theorem 12.4 *Consider a strongly connected signed graph $\mathcal{G}(A_s)$ and the associated system (12.3).*

1. *If $\mathcal{G}(A_s)$ is structurally balanced, then the system (12.3) converges to bipartite consensus*

$$x^* = \lim_{t\to\infty} x(t) = \alpha S \mathbf{1} \qquad \text{where } \alpha = v^\top S x(0) \in \mathbb{R}$$

   with S the diagonal similarity transformation that renders SA_sS nonnegative, and $v > 0$ the left eigenvector of SL_sS relative to the eigenvalue 0, normalized such that $v^\top \mathbf{1} = 1$;

2. *If $\mathcal{G}(A_s)$ is structurally unbalanced then the system (12.3) converges to the origin*

$$x^* = \lim_{t\to\infty} x(t) = 0 \qquad \forall\, x(0) \in \mathbb{R}^n$$

Signed Laplacian L_s	
$\mathcal{G}(A_s)$ structurally balanced	$\mathcal{G}(A_s)$ structurally unbalanced
$\lambda(L_s) = 0$ is an eigenvalue	$\lambda(L_s) = 0$ is not an eigenvalue
$\mu(-L_s) = 0$	$\mu(-L_s) < 0$
marginal stability	Hurwitz stability
$\exists\, S$ s.t. $-SL_sS$ is Metzler	$\nexists\, S$ s.t. $-SL_sS$ is Metzler
bipartite consensus	collapse to the origin

Table 12.1: Properties of structurally balanced vs unbalanced signed Laplacians.

In words:

- If in a signed graph it is possible to have a partition like the one in Fig. 4.8(a), then bipartite consensus is possible: the agents on each faction agree on a common value which is the same in absolute value but not in sign as that of the other faction.

- If the graph cannot be partitioned into two "opposing" factions (i.e., it is like in Fig. 4.8(b)), then no clear global partition into two factions exists, and all agents "collapse" to the origin. Notice that, just like a consensus state, the origin has all equal components. However, it is improper to treat thi as a consensus point (it is asymptotically stable, hence independent of the initial conditions, which are forgotted by the dynamics). It is more a "neutralization" point, to which all opinions "collapse" and become insignificant.

Notice how the diagonal transformation S is the key to understand the bipartition of the solution of (12.3) in the structurally balanced case. Essentially, all agents on one side of the partition shown in Fig. 4.8(a) end up in the same value, while all other agents end up in the opposite value.

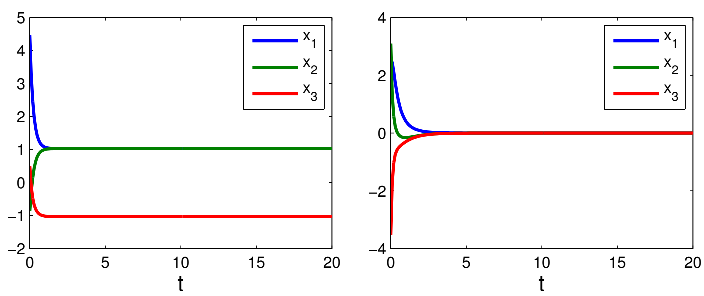
Example 12.5 Consider the signed adjacency matrices

$$A_1 = \begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & -4 \\ -2 & -4 & 0 \end{bmatrix}, \qquad \text{and} \qquad A_2 = \begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & 4 \\ -2 & 4 & 0 \end{bmatrix}.$$

The graph $\mathcal{G}(A_1)$ is structurally balance with partition $\{1, 2\}$ and $\{3\}$, while $\mathcal{G}(A_2)$ is not, see Fig. 12.2. The corresponding solutions of the system (12.3) are given in Fig. 12.3. □

(a) structurally balanced $\mathcal{G}(A_1)$ (b) structurally unbalanced $\mathcal{G}(A_2)$

Figure 12.2: Signed graphs of Example 12.5



(a) (b)

Figure 12.3: Signed graphs of Example 12.5

Similar to the standard consensus case, the strong connectivity assumption can be weakened to existence of a directed spanning tree also for bipartite consensus.

12.2 DT bipartite consensus

The idea of bipartite consensus makes sense also in DT. Consider the following DDE

$$x(t+1) = W_s x(t) \tag{12.6}$$

where W_s is a signed matrix such that the associated matrix of absolute values $W_s^{\text{abs}} = [|w_{s,ij}|]$ is a row stochastic matrix, and $w_{s,ii} \geq 0$, $i = 1, \ldots, n$. Also in this case, the properties of the associated graph $\mathcal{G}(W_s)$ decide whether (12.6) will achieve bipartite consensus or will collapse to the origin.

Theorem 12.6 *Consider a strongly connected signed graph $\mathcal{G}(W_s)$ where W_s^{abs} is row stochastic and $W_{s,ii} \geq 0$, with $W_{s,ii} > 0$ for at least one i.*

1. *If $\mathcal{G}(W_s)$ is structurally balanced, then the system (12.6) converges to bipartite consensus*

$$x^* = \lim_{t\to\infty} x(t) = \alpha S \mathbf{1} \qquad \text{where } \alpha = v^\top S x(0) \in \mathbb{R}$$

with S the diagonal similarity transformation that renders SW_sS nonnegative, and $v > 0$ the left eigenvector of SW_sS relative to the eigenvalue $\rho(W_s) = 1$, normalized such that $v^\top \mathbf{1} = 1$;

2. *If $\mathcal{G}(W_s)$ is structurally unbalanced then the system (12.6) converges to the origin*

$$x^* = \lim_{t\to\infty} x(t) = 0 \qquad \forall\, x(0) \in \mathbb{R}^n$$

Chapter 13

Bounded confidence models

This class of models is also known as Hegselmann-Krause (HK) models. At the basis of the model is the idea that the influences that an agent receive from all other agents should be somehow related to the opinions that these agents are expressing. In practice, individuals are more willing to be influenced by other individuals when the latter have similar opinions than when they have very different opinions. This phenomenon is called homophily: like-minded individuals tend to interact more than individuals with widely different opinions.

The novel aspect of a bounded confidence model is exactly to encode this preference for interaction with agents with similar opinions. Since the opinions are the states of the model, the bounded confidence model has a state-dependent interaction graph.

In particular, if the opinion of an agent is represented as $x_i \in \mathbb{R}$, then agent j exerts an influence over agent i if $|x_i - x_j| \leq d$, where $d =$ confidence range. Denote

$$\mathcal{T}_i(x(t)) = \{j = 1, \dots, n \ \text{ s.t. } \ |x_i - x_j| \leq d\}$$

the neighborhood of "trusted" agents of agent i at time t (including self), of cardinality $|\mathcal{T}_i(x(t))|$. In DT, the update rule of the bounded confidence model is

$$x_i(t+1) = \frac{1}{|\mathcal{T}_i(x(t))|} \sum_{j \in \mathcal{T}_i(x(t))} x_j(t) \tag{13.1}$$

i.e., it is an averaging of the opinions (including self), but only over the trusted agents. The trusted neighborhood $\mathcal{T}_i(x(t))$ can change over time, hence the update rule is itself time-varying. Notice that (13.1) can be written as

$$x_i(t+1) = x_i(t) + \frac{1}{|\mathcal{T}_i(x(t))|} \sum_{j \in \mathcal{T}_i(x(t))} (x_j(t) - x_i(t))$$

where the averaging process is made more evident. The ODE (13.1) can be written also as

$$x(t+1) = W(x(t))x(t)$$

where $W(x(t))$ is a row stochastic matrix of components

$$w_{ij}(x(t)) = \begin{cases} \frac{1}{|\mathcal{T}_i(x(t))|} & \text{if } \ j \in \mathcal{T}_i(x(t)) \\ 0 & \text{if } \ j \notin \mathcal{T}_i(x(t)) \end{cases}$$

The only difference w.r.t. the time-varying consensus which we already saw in Section 9 is that the time-varying part is in reality state-dependent, which makes the model behave qualitatively quite differently.

At each t we can consider the interaction graph associated to $W(x(t))$

$$\mathcal{G}(x(t)) = (\mathcal{V}, \mathcal{E}(x(t))).$$

$\mathcal{G}(x(t))$ is undirected and what changes with t is its edge sets:

$$(i,j) \in \mathcal{E}(x(t)) \iff j \in \mathcal{T}_i(x(t)) \big(\iff i \in \mathcal{T}_j(x(t))\big) \iff |x_i(t) - x_j(t)| \leq d$$

meaning that the graph itself changes with the opinions: the edges appear or disappear as the opinions evolve. Let us see what are the main properties of the HK model

Theorem 13.1 *The model (13.1) has the following properties:*

1. *order preserving:* $x_i(0) \leq x_j(0) \implies x_i(t) \leq x_j(t)\ \forall\, t$.

Assume $x_1 \leq x_2 \leq \ldots \leq x_n$. *Then*

2. $x_1(t)$ *nondecreasing:* $x_1(t+1) \geq x_1(t)\ \forall\, t$;
3. $x_n(t)$ *nonincreasing:* $x_n(t+1) \leq x_n(t)\ \forall\, t$;
4. *If* $(i, i+1) \notin \mathcal{E}(x(t)) \implies (i, i+1) \notin \mathcal{E}(x(t+\tau))\ \forall\, \tau > 0$ *(i.e., adjacent agents that are disconnected stay disconnected);*
5. *If* $\mathcal{G}(x(t))$ *disconnected* $\implies \mathcal{G}(x(t+\tau))$ *disconnected* $\forall\, \tau > 0$;
6. *The connected components of* $\mathcal{G}(x(t))$ *can split as* t *grows, but cannot merge;*
7. *Properties 2 and 3 are valid for any connected component of* $\mathcal{G}$;
8. $x_i(t)$ *converges to* x_i^* *in finite time. For any* i, j *it is either* $x_i^* = x_j^*$ *or* $|x_i^* - x_j^*| > d$.

Proof.

1. Consider agents i and j with opinions $x_i(t) \leq x_j(t)$ and neighborhood sets as shown in Fig. 13.1. Let $\mathcal{S}_i$ (the argument $x(t)$ is omitted for brevity) be the set of agents in the

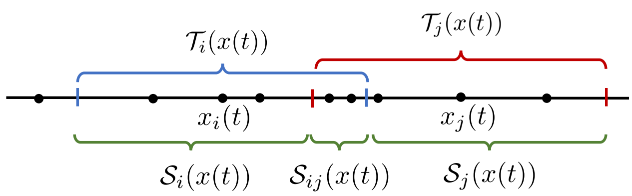


Figure 13.1: Bounded confidence

neighborhood of i but not of j, $\mathcal{S}_j$ that of agents in the neighborhood of j but not of i, and $\mathcal{S}_{ij}$ that of neighbours in common, see Fig. 13.1. Compute the mean of the opinions in these 3 sets: $\bar{x}_{S_i} = \frac{\sum_{k \in \mathcal{S}_i} x_k}{|\mathcal{S}_i|}$ and similarly for $\bar{x}_{S_j}$ and $\bar{x}_{S_{ij}}$. The update rule (13.1) can be considered as the mean of x over $\mathcal{T}_i$, and can be written as

$$x_i(t+1) = \bar{x}_{\mathcal{T}_i} = \frac{|\mathcal{S}_i|\bar{x}_{\mathcal{S}_i} + |\mathcal{S}_{ij}|\bar{x}_{\mathcal{S}_{ij}}}{|\mathcal{S}_i| + |\mathcal{S}_{ij}|} \leq \bar{x}_{\mathcal{S}_{ij}} \tag{13.2}$$

and similarly

$$x_j(t+1) = \bar{x}_{\mathcal{T}_j} \geq \bar{x}_{\mathcal{S}_{ij}}$$

(Graphically, from Fig. 13.1, the inequality (13.2) means that taking the mean of all opinions in $\mathcal{T}_i$ is certainly not bigger than taking the mean of only those in its "higher end" S_{ij}, hence $\bar{x}_{\mathcal{T}_i} \leq \bar{x}_{\mathcal{S}_{ij}}$ follows straightforwardly, and similarly for the other). From these inequalities we have

$$x_i(t+1) \leq x_j(t+1)$$

2. Obvious: if $x_1(t)$ has $\mathcal{T}_1(x(t)) = \{1\}$ then $x_1(t+1) = x_1(t)$. If instead $\mathcal{T}_1(x(t))$ contains more than just agent 1, then $x_1(t+1)$ is averaging $x_1(t)$ with other values that are bigger $\Longrightarrow$ it will increase.

3. Same reasoning.

4. $(i, i+1) \notin \mathcal{E}(x(t))$ means that agents up to i and agents from $i+1$ on are disconnected. Hence the different sets of agents can be treated distinctly. Using the argument above on each of these sets the conclusion follows.

5. Follows by the same argument.

6. Follows by the same argument.

7. Follows by the same argument.

8. Since $x_i(t)$ are sorted and bounded in $[x_1(0),\ x_n(0)]$, for all t, then the $x_i(t)$ must converge. To see the second part (informally), if $x_i^* \neq x_j^*$ but $|x_i^* - x_j^*| < d$ then x_i^* and x_j^* cannot be stationary according to the rule (13.1).

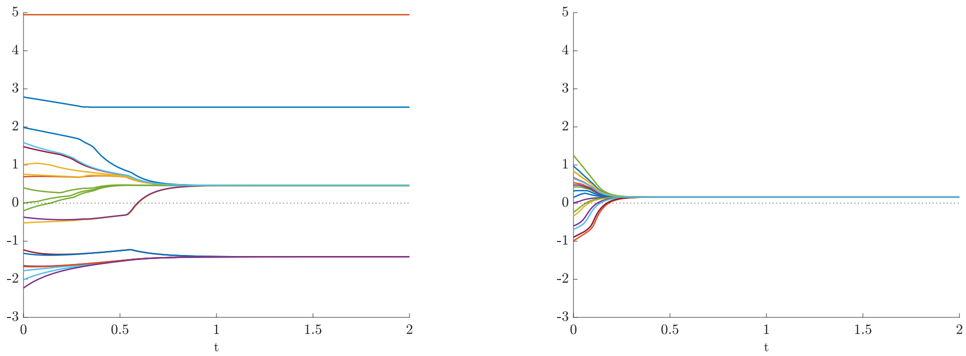
□

Putting together properties 2 and 3 of Theorem 13.1, the range of opinions i.e., the difference $v(x(t)) = \max_i x_i(t) - \min_i x_i(t)$, is nonincreasing over time and bounded below by 0. Obviously, the following "contraction" property holds:

$$x_j(t) \in [\min_i x_i(0),\ \max_i x_i(0)], \quad \forall\, j = 1, \ldots, n.$$

The last item of the previous Theorem tells us what the asymptotic behavior of (13.1) is: $x(t)$ tends to cluster into groups of opinions inside which $\mathcal{G}(x(t))|_{\text{cluster}}$ becomes fully connected, and a consensus is achieved. According to Theorem 13.1, the distance between the consensus values of different clusters is $\geq d$, but in practice it is $\geq 2d$. In particular, when the initial range of opinions $v(x(0))$ is less than d, then consensus is achieved among all agents, i.e., all $x_i(t)$ stay inside the same cluster and converge to the same value for t large enough. In practice, also $v(x(0)) < 2d$ may suffice to reach consensus, depending on $x(0)$.

Example 13.2 Consider an HK model for $n = 100$ agents. A simulation with opinions $x(0)$ uniformly distributed in $[-5, 5]$ and $d = 1$ is shown in Fig. 13.2. The resulting clusters have consensus values that different for around 2.



(a) (b)

Figure 13.2: Bounded confidence simulation. (a): clustering. (b): consensus.

13.1 Extensions: symmetry/asymmetry, homogeneity/heterogeneity

More elaborated models for bounded confidence opinion dynamics can be obtained making use of different (and possibly "individualized", i.e., different from agent to agent) sets of parameters. For instance expressing the confidence interval for agent i as $[x_i - \ell_i, x_i + u_i]$ where $\ell_i \geq 0$ and $u_i \geq 0$ are the lower and upper confidence thresholds for the i-th agent, then an influence function can be defined as

$$\phi_i(x_i, x_j) = \begin{cases} 1 & \text{if } -\ell_i \leq x_j - x_i \leq u_i \\ 0 & \text{otherwise} \end{cases}$$

With this notation, the cardinality of the trusted neighbors becomes

$$|\mathcal{T}_i(x(t))| = \sum_{j=1}^{n} \phi_i(x_i(t), x_j(t)),$$

meaning that (13.1) can be reexpressed as

$$x_i(t+1) = \frac{1}{\sum_{j=1}^{n} \phi_i(x_i, x_j)} \sum_{j=1}^{n} \phi_i(x_i, x_j) x_j$$

Various cases can then be analyzed:

- *Symmetric case*: $\ell_i = u_i$ for all i (i.e., the confidence interval is centered at the agent's opinion; both lower and upper threshold are equal to the individualized' confidence range: $d_i = \ell_i = u_i$). *Asymmetric case*: $u_i \neq \ell_i$ for some i, see Fig. 13.3.
- *Homogeneous case*: $\ell_i = \ell$ and $u_i = u$ for all i (i.e., all agents have the same confidence interval, with possibly $\ell \neq u$). *Heterogeneous case*: the thresholds ℓ_i and u_i are individual.

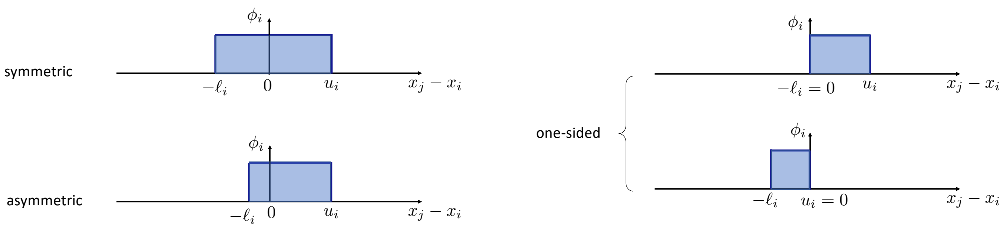


Figure 13.3: Types of confidence intervals in bounded confidence models.

The model investigated in the early part of Section 13 is the homogeneous, symmetric case. The asymmetric case is called *one sided* when $\ell_i =$ or $u_i = 0$ for some i, see Fig. 13.3. When instead for an agent i both $\ell_i = u_i = 0$, then the agent is called *(totally) stubborn*: the set of neighbors $\mathcal{T}_i(x(t))$ contains at each t only agents having an opinion identical to that of agent i, meaning that the agent never modifies its opinion, while it can influence other agents.

Not all properties of Theorem 13.1 are obeyed by the various classes of bounded confidence models. For instance an heterogeneous model cannot guarantee the ordering of the opinions (property 1) [?], and neither convergence in finite time (property 8). Similarly, also one-sided models may achieve only asymptotic convergence.

13.2 Extension to multidimensional opinions

Opinions so far are scalar variables, but it is possible to extend the whole model to opinions that are vectorial $x_i \in \mathbb{R}^m$, representing for instance the opinion of an agent on multiple topics. Bounded confidence requires then to define a distance in $\mathbb{R}^m$, for instance the Euclidean distance $\|y\|_2 = \sqrt{\sum_{k=1}^m y_k^2}$. The neighbour set of agent i is then defined as

$$\mathcal{T}_i(x(t)) = \left\{ j \ \text{ s.t. } \|x_i - x_j\|_2 = \sqrt{\sum_{k=1}^{m} (x_{i,k} - x_{j,k})^2 \leq d} \right\}$$

i.e., all agents that are in a ball of radius d around x_i in opinion space. Consequently, the neighborhood graph $\mathcal{G}(x(t))$ is still state-dependent. The behavior is still similar to the scalar case we saw above, with minor differences:

- opinion still split into clusters;
- $\mathcal{G}(x(t))$ can have components that split but also that merge together (this cannot happen in dimension 1)

- convergence is still achieved in finite time, and within each cluster all opinions converge to a consensus value.

Chapter 14

Biased assimilation

Individuals are often biased when inspecting the opinions of other people. In social sciences, this phenomenon is called *biased assimilation*, and is exemplified by the idea that people tend to accept evidence which is supportive of their own opinion while critically examining contrarian one. The bias assimilation model presented in this chapter tries to describe this phenomenon. It is a highly nonlinear model, and the nonlinearity derives from weighting the usual DeGroot averaging law with a term that depends on the agent own state raised to a power equal to the bias parameter: the higher the exponent (i.e., bias parameter b introduced below) the more biased towards the agent's opinion is the assimilation of the opinions of the neighbors. In order for the model to work properly, the opinions have to be confined to the unit cube $[0, 1]^n$.

Let x_i, $i = 1, \dots, n$, denote the opinion of agent i, and consider the DT update rule

$$x_i(t+1) = \frac{x_i(t)^b \sum_{j=1}^n w_{ij} x_j(t)}{x_i(t)^b \sum_{j=1}^n w_{ij} x_j(t) + \big(1 - x_i(t)\big)^b \big(1 - \sum_{j=1}^n w_{ij} x_j(t)\big)} \tag{14.1}$$

where $W = [w_{ij}]$ is a $n \times n$ row stochastic matrix, and, as already mentioned, $b \geq 0$ is the bias parameter, assumed equal for all agents $i = 1, \dots, n$. In the numerator of (14.1), the averaging term over the neighbours opinions (including self-opinion[1]) $\sum_{j=1}^n w_{ij} x_j$ is premultiplied by the factor x_i^b. This is the way the bias mechanisms is implemented: all opinions (including self) are weighted by the opinion of the agent, raised to a power which represents the strength of the bias. A vector version of (14.1) looks like

$$x(t+1) = \text{diag}(h(x))^{-1} \text{diag}(x)^b W x$$

where the denominator of (14.1) is

$$h(x) = \text{diag}(x)^b W x + \text{diag}(\mathbb{1} - x)^b (\mathbb{1} - W x)$$

[1] Observe that most often in the literature the self-contribution to the average is not rescaled by the bias term, i.e., in place of (14.1) one uses the following:

$$x_i(t+1) = \frac{w_{ii} x_i(t) + x_i(t)^b \sum_{j \neq i} w_{ij} x_j(t)}{w_{ii} + x_i(t)^b \sum_{j \neq i} w_{ij} x_j(t) + \big(1 - x_i(t)\big)^b \big(\sum_{j \neq i} w_{ij} - \sum_{j \neq i} w_{ij} x_j(t) \big)} \tag{14.2}$$

The behavior of (14.2) is very similar to that of (14.1) but a bit more difficult to analyze, so we omit it.

The denominator in (14.1) is meant to rescale the update law so that $x_i(t) \in [0, 1]$ for all t. Let us show that indeed the system (14.1) is invariant in the unit cube $[0, 1]^n$. For $x_i(t) = 0$ it is $x_i(t+1) = 0$, and similarly, for $x_i(t) = 1$ it is $x_i(t+1) = 1$. In between, both the numerator and the denominator are nonnegative quantities, but the numerator is smaller. Since this is true for each i regardless of the value of $x_j \in [0, 1]$, $j \neq i$, we have that indeed the dynamics remains confined to $[0, 1]^n$.

The model (14.1) is highly nonlinear and its dynamical behavior difficult to investigate. What we can observe is that it has a number of equilibria on the boundary of the cube: any corner of the cube, i.e., x^* s.t. $x_i^* \in \{0, 1\}$, is an equilibrium point. These can be considered “extremal” (or “polarized”) opinions for the model. In two of these corners, $x^* = 0$ and $x^* = \mathbb{1}$, the opinions are unanimous, which we will refer to as *unilaterally polarized* opinions. The other $2^n - 2$ corners x^* of the unit cube correspond to opinion vectors that contain both extreme opinions 0 and 1, hence we call them *bilaterally polarized* equilibria. The model has also equilibria in the interior of the cube. In particular, $x^* = \mathbb{1}/2$ is always an equilibrium point. It corresponds to the centroid of the unit cube, and can be considered an “indecision state”, as it is maximally distant from each polarized opinion. Other interior equilibria are possible, but their presence (and value) depend on W and b.

The values of the bias parameter b can be classified in the following cases:

1. no bias: $b = 0$;
2. weak bias: $0 < b \leq 1$;
3. strong bias: $b \geq 1$.

The case $b = 0$ is denoted “no bias” because we recover the usual DeGroot update law, seen in (8.15):

$$x_i(t+1) = \frac{\sum_{j=1}^n w_{ij} x_j(t)}{\cancel{\sum_{j=1}^n w_{ij} x_j(t)} + 1 - \cancel{\sum_{j=1}^n w_{ij} x_j(t)}}$$

The asymptotic behavior of the model is analyzed in the following theorem.

Theorem 14.1 *Consider the biased assimilation model (14.1), with irreducible row-stochastic matrix W.*

- *The unilaterally polarized equilibria $x^* = 0$ and $x^* = \mathbb{1}$ are locally asymptotically stable for all $b > 0$, with domain of attraction consist of at least $(0, 1/2)^n$, resp. $(1/2, 1)^n$.*
- *The indecision state $x^* = \mathbb{1}/2$ is an unstable equilibrium point for all $b > 0$.*
- *The bilaterally polarized equilibria x^* have stability properties that depend on the value of b:*
  1. *they are unstable if $0 < b < 1$;*
  2. *they are locally asymptotically stable if $b \geq 1$.*

Proof.

Denoting $s_i(t) = \sum_{j=1}^n w_{ij}x_j(t)$, (14.1) can be written as

$$x_i(t+1) = f_i(x) = \frac{x_i(t)^b s_i(t)}{x_i(t)^b s_i(t) + \big(1-x_i(t)\big)^b\big(1-s_i(t)\big)}$$

Let us consider first the stability of the unilaterally polarized equilibrium $x^* = 0$. In order to prove that $x_i(0) < 1/2$ for all i implies $\lim_{t\to\infty} x_i(t) = 0$ for all i, let us consider the ratio

$$\frac{1-x_i(t+1)}{x_i(t+1)} = \left(\frac{1-x_i(t)}{x_i(t)}\right)^b \frac{1-s_i(t)}{s_i(t)}$$

From $0 \le x_i(0) < 1/2$ it is

$$\frac{1-x_i(0)}{x_i(0)} > \frac{1-x_M(0)}{x_M(0)} > 1$$

where $x_M(0) = \max_i x_i(0)$. Since it is also $\frac{1-s_i(0)}{s_i(0)} > \frac{1-x_M(0)}{x_M(0)} > 1$, at $t=1$ we have

$$\frac{1-x_i(1)}{x_i(1)} \ge \left(\frac{1-x_M(0)}{x_M(0)}\right)^{b+1} > \frac{1-x_M(0)}{x_M(0)} > 1$$

which implies $x_i(1) < x_M(0) < 1/2$ for all i, and, consequently, $x_M(1) < x_M(0)$. Iterating the reasoning we get

$$\frac{1-x_i(1)}{x_i(1)} \ge \left(\frac{1-x_M(0)}{x_M(0)}\right)^{(b+1)^\top}$$

meaning that $x_M(t) \to 0$ exponentially as $t \to \infty$. The proof for the equilibrium point $x^* = \mathbb{1}$ is identical if one replaces $x_i(t)$ with $y_i(t) = 1 - x_i(t)$.

To show local stability/instability of the other equilibria, let us compute the Jacobian linearization of (14.1):

$$\frac{\partial f_i(x)}{\partial x_i} = \frac{\big(bx_i^{b-1}s_i + x_i^b w_{ii}\big)\, h_i(x) - x_i^b s_i \frac{\partial h_i(x)}{\partial x_i}}{h_i^2(x)}$$

$$\frac{\partial f_i(x)}{\partial x_k} = \frac{x_i^b w_{ik} h_i(x) - x_i^b s_i \frac{\partial h_i(x)}{\partial x_k}}{h_i^2(x)}$$

where

$$\frac{\partial h_i(x)}{\partial x_i} = bx_i^{b-1}s_i + x_i^b w_{ii} - b(1-x_i)^{b-1}(1-s_i) - (1-x_i)^b w_{ii}$$

$$\frac{\partial h_i(x)}{\partial x_k} = x_i^b w_{ik} - (1-x_i)^b w_{ik}$$

When computed in $x^* = \mathbb{1}/2$,

$$\left.\frac{\partial h_i(x)}{\partial x_i}\right|_{x=\mathbb{1}/2} = \left.\frac{\partial h_i(x)}{\partial x_k}\right|_{x=\mathbb{1}/2} = 0$$

and

$$\left.\frac{\partial f_i(x)}{\partial x_i}\right|_{x=\mathbb{1}/2} = b + w_{ii}, \qquad \left.\frac{\partial f_i(x)}{\partial x_k}\right|_{x=\mathbb{1}/2} = w_{ik}$$

meaning that the Jacobian matrix at $\mathbb{1}/2$ is $J(\mathbb{1}/2) = \left[\frac{\partial f_i(x)}{\partial x_j}\Big|_{x=\mathbb{1}/2} \right]_{ij} = W + bI$. Since W is row stochastic and irreducible, $\rho(W) = 1$, hence $\rho(J(\mathbb{1}/2)) > 1$, for any $b > 0$.

Consider now one of the bilaterally polarized equilibria, denote it ξ. By construction $\exists\ i$ such that $\xi_i = 0$, and j such that $\xi_j = 1$. For the i-th coordinate $0 < h_i(\xi) = 1 - s_i < 1$, while for the j-th it is $0 < h_j(\xi) = s_j < 1$. When $b < 1$, $f_i(x)$ fails to be Lipschitz at ξ and

$$\frac{\partial f_i(x)}{\partial x_i}\bigg|_{x=\xi} = \begin{cases} +\infty & \text{if } \xi_i = 0 \\ 0 & \text{if } \xi_i = 1 \end{cases}, \qquad \frac{\partial f_i(x)}{\partial x_k}\bigg|_{x=\xi} = 0$$

after applying l'Hôpital rule. Hence the Jacobian matrix at ξ, $J(\xi) = \left[\frac{\partial f_i(x)}{\partial x_j}\Big|_{x=\xi} \right]_{ij}$ is diagonal and with at least one element equal to $+\infty$, meaning that ξ is unstable. When $b = 1$,

$$\frac{\partial f_i(x)}{\partial x_i}\bigg|_{x=\xi} = \begin{cases} \frac{s_i}{1-s_i} & \text{if } \xi_i = 0 \\ 0 & \text{if } \xi_i = 1 \end{cases}, \qquad \frac{\partial f_i(x)}{\partial x_k}\bigg|_{x=\xi} = 0$$

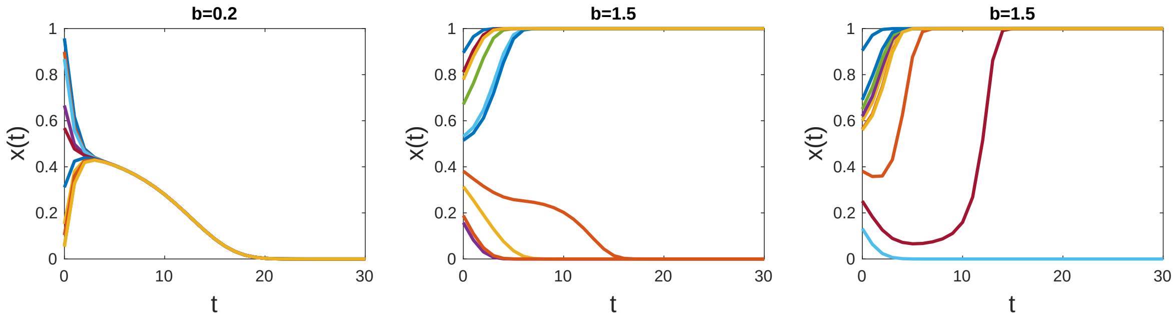
hence the Jacobian matrix $J(\xi)$ is diagonal and with at least one element and at least one equal to 0. Since, at ξ, $0 < \frac{s_i}{1-s_i} < 1$, it is $\rho(J(\xi)) < 1$, hence ξ is locally asymptotically stable. Finally, when instead $b > 1$, direct computations show that $J(\xi) = 0$. Since $\rho(J(\xi)) = 0$ the equilibria are locally asymptotically stable. □

The locally asymptotically stable equilibria $x^* = 0$ and $x^* = \mathbb{1}$ also happen to resemble consensus points, in the sense that all entries of the vector x^* are identical. Notice, however, that this form of consensus is different from that investigated in previous chapters, as here we have local asymptotic stability, not marginal stability (with continuous dependence from the initial conditions). For this reason, we prefer to call $x^* = 0$ and $x^* = \mathbb{1}$ a unilaterally polarized opinion, rather than a consensus.

Theorem 14.1 states that the region of attraction of the unilateral polarized equilibria $x^* = 0$ and $x^* = \mathbb{1}$ contains at least the subcube of unanimous opinions on the same side of the "indecision state" $\mathbb{1}/2$. These opinions mutually reinforce themselves and converge to the corresponding extreme unilateral polarized state 0 or $\mathbb{1}$. In particular such opinion trajectory differs fundamentally from one obtained from an averaging dynamics, as it escapes the convex hull of the initial conditions, hence the term "polarization".

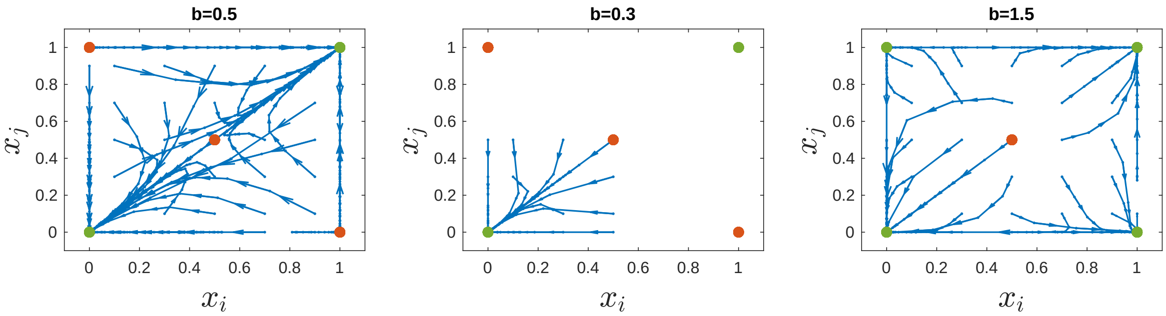
While in presence of a small bias $b < 1$ only the unilaterally polarized states 0 and $\mathbb{1}$ are normally attractors for the dynamics, and convergence to one or the other depends on the opinions of the neighbors on the graph, in presence of a large bias $b \geq 1$, the self opinion of each agent "overrides" the influence of the neighbouring agents, at least near the polarized states. So a strong bias leads to limited (or no) assimilation of the opinions of the other agents.

Example 14.2 Let us consider a fully connected graph of size $n = 10$ and adjacency matrix $W = \mathbb{1}^\top \mathbb{1}/10$. Figs. 14.1 and 14.2 show a few simulations for the model (14.1), with both small and large b. The small b cases seem to be leading to unilateral polarization. While this is predicted in Theorem 14.1 for the case of initial conditions all on the same side of the indecision state (Fig. 14.2(b)), in simulation it seems to be happening also regardless of the initial conditions, see Fig. 14.1(a) and Fig. 14.2(a). The case of large b behaves instead differently: while we still have unilateral polarization for initial conditions all on the same side of the indecision state, mixed polarization can happen for more general initial conditions Figs. 14.1 and 14.2(c). In this case the position of the initial condition of an agent is not necessarily a predictor of the side of its extremal final opinion, see Fig. 14.1(c).



(a) (b) (c)

Figure 14.1: Biased assimilation model (14.1). (a): small bias leads to unilateral polarization. (b) large bias leads to bilateral polarization. (c): for large bias, the initial condition is not always a predictor of the polarization of the final opinion.



(a) (b) (c)

Figure 14.2: Biased assimilation model (14.1). Phase portraits of 2D slices of the opinion evolution. (a): unilateral polarization occurs when $b < 1$ (the asymptotically stable equilibria are in green, while those in red are unstable). (b): the domain of attraction of 0 (resp. $\mathbb{1}$) is at least the subcube $(0, 1/2)^n$ (resp. $(1/2, 1)^n$). (c): for large bias, bilateral polarization occurs, and all corners are locally asymptotically stable equilibria.

14.1 Extension to signed graphs

The peculiar features of the model (14.1) are

1. The indecision state $\mathbb{1}/2$ is alway an unstable equilibrium point, meaning that the agents (almost) always get polarized;
2. For small bias values, polarization is always unilateral, i.e., all agents end up in an agreement state.

Both features have some value, but, depending on the context, may also represent significant limitations: why should the agents always get polarized? (Recall that averaging dynamics never lead to polarization as it is intended here, but to its contrary). Why should they always achieve unanimity?

One way to cope with these limitations is to introduce antagonism in the model, in the form of signed interactions such as those we saw in Chapter 12.

Consider an irreducible row stochastic matrix $W = [w_{ij}]$, and alongside, a "trust/mistrust" signature matrix $\Sigma = [\sigma_{ij}]$, $\sigma_{ij} \in \pm 1$, $\sigma_{ii} = 1$. Form the signed adjacency matrix $W_s = \Sigma \circ W = [\sigma_{ij} w_{ij}]$ (this is called Hadamar product, or element-wise product). For an agent i, split the incoming neighbors into friendly and unfriendly subsets: $\mathcal{N}_i^{\text{in},+} = \{j \text{ s.t. } \ w_{s,ij} > 0\}$ and $\mathcal{N}_i^{\text{in},-} = \{j \text{ s.t. } \ w_{s,ij} < 0\}$. The update rule that replaces (14.1) treats the unfriendly neighbors differently, and rather than considering their opinion, i.e., $x_j(t)$, takes the "opposite", i.e., $1 - x_j(t)$, for any $j \in \mathcal{N}_i^{\text{in},-}$. Denoting

$$s_i^+(t) = \sum_{j \in \mathcal{N}_i^{\text{in},+}} w_{ij} x_j(t) \quad \text{and} \quad s_i^-(t) = \sum_{j \in \mathcal{N}_i^{\text{in},-}} w_{ij}(1 - x_j(t))$$

the summations over $\mathcal{N}_i^{\text{in},+}$ and $\mathcal{N}_i^{\text{in},-}$, the new model is

$$x_i(t+1) = \frac{x_i(t)^b \left(s_i^+(t) + s_i^-(t)\right)}{x_i(t)^b \left(s_i^+(t) + s_i^-(t)\right) + (1 - x_i(t))^b \left(1 - s_i^+(t) - s_i^-(t)\right)} \tag{14.3}$$

14.1.1 Structural balance leads to bilateral polarization for any bias

When $\mathcal{G}(W_s)$ is structurally balanced, then we have the following straightforward generalization of Theorem 14.1.

Proposition 14.3 *Consider the model (14.3) with an irreducible row stochastic matrix W. If $\mathcal{G}(W_s)$ is structurally balanced, with node partition $\mathcal{V}_1$, $\mathcal{V}_2$, then bilateral polarization occurs, i.e., the equilibria*

$$x^* = \{\xi, \mathbb{1} - \xi\} \quad \textit{with} \quad \xi_i = \begin{cases} 1 & \textit{if } i \in \mathcal{V}_1 \\ 0 & \textit{if } i \in \mathcal{V}_2 \end{cases}$$

are locally asymptotically stable for each value of the bias $b > 0$. Furthermore, all initial conditions such that $x_i(0) > 1/2$ for $i \in \mathcal{V}_1$ and $x_j(0) < 1/2$ for $j \in \mathcal{V}_2$ converge to ξ while those with $x_i(0) < 1/2$ for $i \in \mathcal{V}_1$ and $x_j(0) > 1/2$ for $j \in \mathcal{V}_2$ converge to $\mathbb{1} - \xi$.

Proof. Considering the change of basis

$$y_i = \begin{cases} x_i & \text{if } i \in \mathcal{V}_1 \\ 1 - x_i & \text{if } i \in \mathcal{V}_2 \end{cases}$$

The proof is a straightforward combination of the properties of structurally balanced matrices and Theorem 14.1.

So, rather than $\{0, \mathbb{1}\}$ as attractors, structural balance leads to attractivity for another pair of antipolad corners ξ and $\mathbb{1} - \xi$, which are both necessarily bilaterally polarized.

14.1.2 Structural unbalance amplifies indecision for small bias

For a small bias b, the case of $\mathcal{G}(W_s)$ structurally unbalanced is significantly different from the unsigned case. In fact, for sufficiently small $b > 0$ the indecision state $x^* = \mathbb{1}/2$ becomes an attractor.

Proposition 14.4 *Consider the model (14.3) with an irreducible row stochastic matrix W and $b > 0$ sufficiently small. If $\mathcal{G}(W_s)$ is structurally unbalanced, then the indecision state $x^* = \mathbb{1}/2$ is locally asymptotically stable.*

Proof. We skip the proof, and provide instead an intuitive reasoning. The idea is that for $b = 0$ the model (14.3) corresponds to a linear DT signed consensus problem (see Sect. 12.2), after changing basis to $z := x - \mathbb{1}/2$. For it, when $\mathcal{G}(W_s)$ is unbalanced, $z^* = 0$ is a global asymptotically stable equilibrium point. When $b > 0$ but sufficiently small, the behavior of the nonlinear model (14.3) must be the same as its linear $b = 0$ counterpart, namely convergence to $z^* = 0$ occurs from all initial conditions which are not equilibria. □

Notice that, depending on the choice of W and Σ, the behavior described in Proposition 14.4 may occur also for reasonably large values of $0 < b < 1$, hence it is not a "pathological" behavior, see Example 14.6 and Fig. 14.3(a). What may happen when b grows is that other internal equilibria appear, possibly even close to $\mathbb{1}/2$, and the system bifurcates: $x^* = \mathbb{1}/2$ becomes unstable but potentially no polarization is achieved (this depends on W and Σ).

14.1.3 For large bias, opinions become polarized

Increasing b beyond 1, then the outcome of the model (14.3) is a (normally bilateral) polarization, decided by the initial conditions. In this case the presence of signed edges does not alter the behavior of the model: similarly to Theorem 14.1, the initial opinion of an agent overrides the influence of the neighbors, provided that at least one neighbor has a similar stance (i.e., has an opinion on the same side as the agent with respect to the indecision state 1/2).

Proposition 14.5 *Consider the model (14.3) with an irreducible row stochastic matrix W and b which sufficiently large ($b > 1$). Assume that, for each i, $x_i(0) \neq 1/2$ and that $\exists\, j \neq i$ such that $x_j(0)$ is lying on the "same side" as $x_i(0)$ with respect to 1/2 (i.e., $x_i(0)$ and $x_j(0)$ are both $< 1/2$ or both $> 1/2$). Then polarization occurs, i.e., $x(t)$ converges to $x^* = \xi$, of components $\xi_i = (\mathrm{sgn}(x_i(0) - 1/2) + 1)/2$.*

Proof. A proof of this proposition can be found in [24], Theorem 4. A linearization argument similar to the one used in the proof of Theorem 14.1 can also be used. □

Example 14.6 Let us consider again a fully connected graph of size $n = 10$ and adjacency matrix $W = \mathbb{1}^\top \mathbb{1}/10$. The edge sign pattern is provided by the matrix $\Sigma = [\sigma_{ij}]$, $\sigma_{ij} \in \pm 1$, which can be used to form the (Hadamar) product $W_s = \Sigma \circ W$.

Consider the case of small bias parameter: $b = 0.5$. When $\mathcal{G}(W_s)$ is structurally unbalanced, then $\lim_{t\to\infty} x(t) = \mathbb{1}/2$ for all $x(0) \in (0, \mathbb{1})^n$, see Fig. 14.3(a). The effect of the lack of balance is that it is more difficult for the agents to take a decision, and as a consequence they remain in an indecision state. When instead $\mathcal{G}(W_s)$ is structurally balanced, i.e., when $\Sigma = \sigma\sigma^\top$ for some $\sigma = \begin{bmatrix}\sigma_1 & \dots & \sigma_n\end{bmatrix}^\top$, $\sigma_i \in \pm 1$, then bilateral polarization occurs, and the opinions converge to the pair ξ, $\mathbb{1} - \xi$ of components $\xi_i = (\sigma_i + 1)/2 \in \{0, 1\}$, see Fig. 14.3(b). The indecision state is now unstable.

For large biases $b > 1$, all polarized states (i.e., all corners of the unit cube) are locally asymptotically stable, see Fig. 14.3(c).

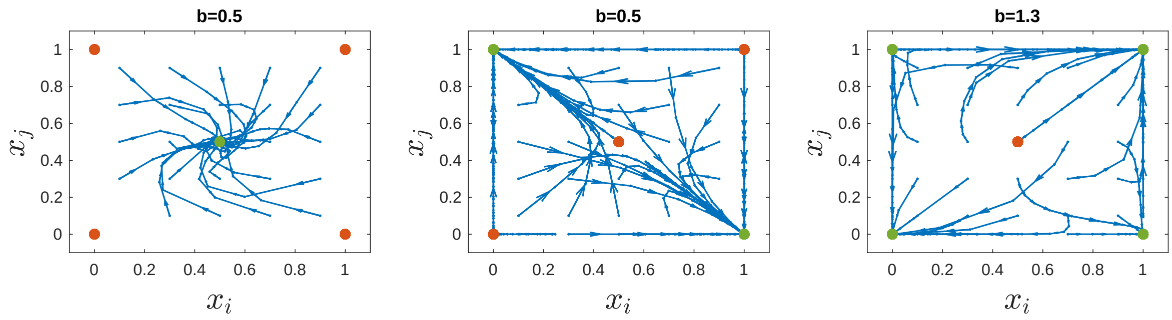


Figure 14.3: Signed biased assimilation model (14.3). Phase portraits of 2D slices of the opinion evolution. (a): small bias and $\mathcal{G}(W_s)$ structurally unbalanced. The indecision state $\mathbb{1}/2$ is the only attractor. (b): small bias and $\mathcal{G}(W_s)$ structurally balanced. Bilateral polarization to a pair of antipodal corners occurs. (c): large bias always lead to local asymptotic stability of each corner.

14.1.4 Notes and further reading

The main references for this part are: [25], [6] and [24].

Chapter 15

Opinion formation games

15.1 Consensus as best response of a network game

Let us consider a strategic game on the graph $\mathcal{G}(A)$ in which the agents $i \in \mathcal{V}$ are the players, the opinions x_i are the actions, assumed to belong to some subset (compact??) of $\mathbb{R}$, $x_i \in \Omega \subseteq \mathbb{R}$, and the aim of agent i is to minimize the cost function

$$J_i(x) = J_i(x_i, x_{-i}) = \frac{1}{2} \sum_j a_{ij}(x_i - x_j)^2 \tag{15.1}$$

where x_{-i} is the opinion (i.e., the action) of the $n-1$ other agents playing in the game. Given x_{-i}, each agent i seeks to compute its best response, i.e., the opinion x^* that minimizes the cost function (15.1):

$$x_i^* = \arg\min_{x_i \in \Omega} J_i(x_i, x_{-i}) \tag{15.2}$$

Owing to the structure of $J_i(x)$, x_i^* depends only on the actions of the agents that are first (incoming) neighbors of agent i. If $J_i(x)$ is continuously differentiable in x_i as in (15.1), this corresponds to solving

$$\nabla_{x_i} J_i(x) = \frac{\partial J_i(x)}{\partial x_i} = \sum_j a_{ij}(x_i - x_j) = L_{i,*}x = 0 \tag{15.3}$$

where $L_{i,*}$ is the ith row of the Laplacian $L = \text{diag}(A\mathbb{1}) - A$. The set of actions in which every agent is playing the best response to the other players' actions is a Nash equilibrium. More formally, a Nash equilibrium is an action profile $x^* = (x_i^*,\, x_{-i}^*)$ for which

$$J_i(x_i^*,\, x_{-i}^*) \leq J_i(x_i\, x_{-i}^*) \qquad \forall\, x_i \in \Omega \quad \text{and} \quad \forall\, i \in \mathcal{V}$$

In a Nash equilibrium, no player is able to improve its cost function by a unilateral move.

Since the expression (15.2) for the best strategy is valid for each player, the Nash equilibrium for this game is obtained by putting together the rows of L, i.e., it is given by the solution of $Lx = 0$. Therefore it corresponds to a consensus point $x^* \in \text{span}(\mathbb{1})$, at least when $\mathcal{G}(A)$ is strongly connected. (or with a rooted spanning tree?). Any consensus point in span($\mathbb{1}$) is a Nash equilibrium.

The consensus system $\dot{x} = -Lx$ can be considered as a best response dynamics associated to (15.3).

If $\mathcal{G}(A)$ is undirected, i.e., A is symmetric, then we have seen in Section 8.1.1 that we can introduce a Laplacian potential $\Phi(x) = \frac{1}{2}x^\top Lx$. In this case, since $\Phi(x) = \sum_i J_i(x)$, minimizing $J_i(x)$ with respect to x_i for all i and minimizing $\Phi(x)$ over the all unit norm vectors x gives the same result, as

$$\nabla_x \Phi(x) = Lx = \begin{bmatrix} \nabla_1 J_1 \\ \vdots \\ \nabla_n J_n \end{bmatrix}.$$

When A is not symmetric, in the cost function $\Phi(x)$ only the symmetric part of L (i.e., $(L + L^\top)/2$, corresponding to $(A + A^\top)/2$) matters, while instead $J_i(x)$ refers to the original A. Only the individual cost functions lead to a Nash equilibrium which is a consensus, as Example 15.1 shows.

Example 15.1 Consider the following adjacency matrix and associated Laplacian

$$A = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix} \qquad \Longrightarrow \qquad L = \begin{bmatrix} -a & a \\ b & -b \end{bmatrix}$$

for some a, $b > 0$. For it

$$\Phi(x) = \frac{1}{2}x^\top Lx = \frac{1}{2}\left(ax_1^2 + (a+b)x_1x_2 + bx_2^2\right)$$

while

$$J_1(x) = \frac{a}{2}(x_1 - x_2)^2, \qquad J_2(x) = \frac{b}{2}(x_1 - x_2)^2$$

meaning that, if $a \neq b$, $\Phi(x) \neq J_1(x) + J_2(x)$. Consequently,

$$\nabla\Phi(x) = \frac{1}{2}\begin{bmatrix} 2ax_1 - (a+b)x_2 \\ -(a+b)x_1 + 2bx_2 \end{bmatrix}$$

while instead

$$\nabla_{x_1} J_1(x) = a(x_1 - x_2), \qquad \nabla_{x_2} J_2(x) = b(x_1 - x_2).$$

When $a = b$, then $\nabla\Phi(x) = \begin{bmatrix} \nabla_{x_1} J_1 \\ \nabla_{x_2} J_2 \end{bmatrix}$. □

Using the cost function (15.1), it is possible to obtain also the DT version of the consensus dynamics if we start with a graph $\mathcal{G}(W)$ in which $W = [w_{ij}]$ is row stochastic. In fact, rewriting (15.1) as

$$J_i(x) = \frac{1}{2}\sum_j w_{ij}(x_i - x_j)^2$$

the gradient equation (15.3)

$$\frac{\partial J_i(x)}{\partial x_i} = \underbrace{\sum_j w_{ij}}_{=1} x_i - \sum_j w_{ij} x_j = 0$$

leads to the Nash equilibrium satisfying $x_i^* = \sum_j w_{ij} x_j^*$, or, in matrix form, $x^* = W x^*$. Notice that this expression holds even when $w_{ii} > 0$. When we have strict dominance of the spectral radius $\rho(W) = 1$ over all other eigenvalues of W, then this is the eigenvalue - eigenvector equation corresponding to consensus points $x^* \in \mathrm{span}(\mathbb{1})$. The DT consensus system (8.15) corresponds then to the best response dynamics associated to the game.

15.2 Friedkin-Johnsen model as best response of a network game with stubborn agents

Let us now consider a quadratic cost function in which the agents give a weight also to their opinions $x_{o,i} = x_i(0)$, not only to the differences in opinion with respect to the other agents. If, as in Chapter 11, we use stubbornness coefficients $\theta_i \in [0, 1)$ to express the attachment of agent i to its own initial opinion $x_{o,i}$, then the cost function we are considering is

$$J_i(x) = \frac{1}{2}\Big(\theta_i (x_i - x_{o,i})^2 + (1-\theta_i)\sum_j w_{ij}(x_i - x_j)^2\Big) \tag{15.4}$$

where $W = [w_{ij}] \geq 0$ is a row stochastic adjacency matrix for the interaction graph $\mathcal{G}(W)$. If we think of $x_i \in \Omega \subseteq \mathbb{R}$ as action, we can proceed to compute the Nash equilibrium of this game as we did before. Namely, we have the following.

Proposition 15.2 *The Nash equilibrium of the game of players $i \in \mathcal{V}$, actions $x_i \in \Omega \subseteq \mathbb{R}$ and cost function (15.4) is given by x^* such that*

$$x_i = (I - (I - \Theta)W)^{-1}\Theta x_o \tag{15.5}$$

Proof. Computing the vanishing point of the gradient

$$\begin{aligned} 0 =& \frac{\partial J_i(x)}{\partial x_i} = \theta_i (x_i - x_{o,i}) + (1-\theta_i)\sum_j w_{ij}(x_i - x_j) \\ =& \underbrace{\Big(\theta_i + (1-\theta_i)\underbrace{\sum_j w_{ij}}_{=1}\Big)}_{=1} x_i - \theta_i x_{o,i} - (1-\theta_i)\sum_j w_{ij} x_j \end{aligned}$$

i.e., $x_i^* = \theta_i x_{o,i} + (1-\theta_i)\sum_j w_{ij} x_j^*$, or in matrix form, $x^* = \Theta x_o + (I - \Theta) W x^*$, from which (15.5) follows. □

The best response dynamics associated to (15.5) is given by

$$x_i(t+1) = (1-\theta_i)\sum_j w_{ij} x_j(t) + \theta_i x_{o,i}$$

i.e., in vector form, the Friedkin-Johnsen model (11.1), of which x^* is the asymptotically stable equilibrium point.

Bibliography

[1] Claudio Altafini. Stability analysis of diagonally equipotent matrices. *Automatica*, 49(9):2780–2785, 2013.

[2] David H. Anderson. Structural properties of compartmental models. *Mathematical Biosciences*, 58(1):61 – 81, 1982.

[3] D.H. Anderson. *Compartmental Modeling and Tracer Kinetics*. Lecture Notes in Biomathematics. Springer Berlin Heidelberg, 1983.

[4] Frank M. Bass. A new product growth for model consumer durables. *Management Science*, 15(5):215–227, 1969.

[5] G.W. Cross. Three types of matrix stability. *Linear Algebra and its Applications*, 20(3):253–263, 1978.

[6] Pranav Dandekar, Ashish Goel, and David T. Lee. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110(15):5791–5796, 2013.

[7] Fall, A., Iggidr, A., Sallet, G., and Tewa, J. J. Epidemiological models and lyapunov functions. *Math. Model. Nat. Phenom.*, 2(1):62–83, 2007.

[8] L. Farina and S. Rinaldi. *Positive Linear Systems: Theory and Applications*. A Wiley-Interscience publication. Wiley, 2000.

[9] Angela Fontan and Claudio Altafini. Multiequilibria analysis for a class of collective decision-making networked systems. *IEEE Transactions on Control of Network Systems*, 5(4):1931–1940, 2018.

[10] B. S. Goh. Global stability in many-species systems. *The American Naturalist*, 111(977):135–143, 1977.

[11] Daniel Hershkowitz and Hans Schneider. Lyapunov diagonal semistability of real H-matrices. *Linear Algebra and its Applications*, 71:119 – 149, 1985.

[12] R.A. Horn and C. R. Johnson. *Matrix Analysis*. Cambdridge University Press, 1985.

[13] John A. Jacquez and Carl P. Simon. Qualitative theory of compartmental systems. *SIAM Review*, 35(1):43–79, 1993.

[14] E. Kaszkurewicz and A. Bhaya. *Matrix Diagonal Stability in Systems and Computation.* Birkhäuser Boston, 2000.

[15] H.K. Khalil. *Nonlinear Systems.* Pearson Education. Prentice Hall, 3rd edition, 2002.

[16] Ali Khanafer, Tamer Baar, and Bahman Gharesifard. Stability of epidemic models over directed graphs: A positive systems approach. *Automatica*, 74:126–134, 2016.

[17] Minkyoung Kim, David Newth, and Peter Christen. Modeling dynamics of diffusion across heterogeneous social networks: News diffusion in social media. *Entropy*, 15(10):4215–4242, 2013.

[18] Ji Liu, A. Stephen Morse, Angelia Nedi, and Tamer Baar. Internal stability of linear consensus processes. In *53rd IEEE Conference on Decision and Control*, pages 922–927, 2014.

[19] Wenjun Mei, Shadi Mohagheghi, Sandro Zampieri, and Francesco Bullo. On the dynamics of deterministic epidemic propagation over networks. *Annual Reviews in Control*, 44:116–128, 2017.

[20] C.D. Meyer. *Matrix Analysis and Applied Linear Algebra.* Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2000.

[21] Sergey E. Parsegov, Anton V. Proskurnikov, Roberto Tempo, and Noah E. Friedkin. Novel multidimensional models of opinion dynamics in social networks. *IEEE Transactions on Automatic Control*, 62(5):2270–2285, 2017.

[22] Hal L. Smith. Systems of ordinary differential equations which generate an order preserving flow. A survey of results. *SIAM Rev.*, 30(1):87–113, 1988.

[23] R.S. Varga. *Matrix Iterative Analysis.* Springer Series in Computational Mathematics. Springer Berlin Heidelberg, 1999.

[24] Lingfei Wang, Yiguang Hong, Guodong Shi, and Claudio Altafini. Signed social networks with biased assimilation. *IEEE Transactions on Automatic Control*, pages 1–1, 2022.

[25] Weiguo Xia, Mengbin Ye, Ji Liu, Ming Cao, and Xi-Ming Sun. Analysis of a nonlinear opinion dynamics model with biased assimilation. *Automatica*, 120:109113, 2020.