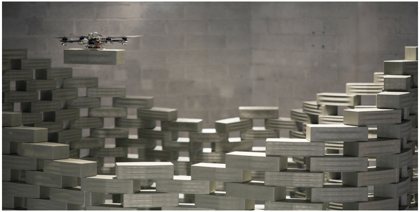
Lecture 6

From algebraic graph theory to consensus

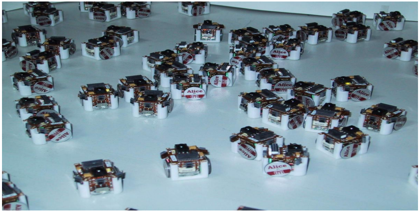
Giancarlo Ferrari Trecate[1]

[1]Dependable Control and Decision Group
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
giancarlo.ferraritrecate@epfl.ch

Opportunities offered by NCS: coordination among agents



Flight assisted architecture at ETH



Swarm of mobile robots

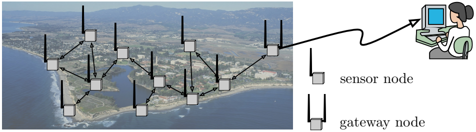
Previous lecture

- Motivating examples: agents using communication for reaching a common goal
  - dynamics captured by matrices with special properties (e.g. row-stochastic)

$$x^+ = Ax, \quad A_{ij} \geq 0, \quad \sum_{j=1}^{n} A_{ij} = 1, \ \forall i = 1, \ldots, n$$

- Basics in graph theory (as graphs capture the topology of partial communication networks)

Review: averaging in wireless sensor networks



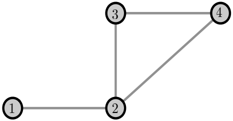
- n spatially distributed devices, each measuring the same environmental variable (temperature, light,...)
- devices exchange information over a communication network
- the operator wants to receive a single average measurement

Distributed algorithm

Sensor i computes

$$x_i^+ = \text{average}(x_i, x_j |\, j \sim i)$$

Example: $x_1^+ = \frac{x_1+x_2}{2}$, $x_2^+ = \frac{x_1+x_2+x_3+x_4}{4}$



$j \sim i \stackrel{\text{def}}{=} j$ is a neighbor of $i \stackrel{\text{def}}{=}$ the edge (i, j) exists

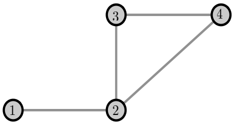
Review: collective model for the graph in the figure

Set $x = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$

$$x^+ = Ax$$

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

A is row-stochastic



$j \sim i \overset{\text{def}}{=} j$ is a neighbor of $i \overset{\text{def}}{=}$ the edge (i, j) exists

Problem

Will the sensors achieve *average consensus*, i.e.

$$x_i(k) \to \text{average}(x_i(0), i = 1, \dots, n) \text{ as } k \to +\infty, \forall i = 1, \dots, n \text{ ?}$$

Remark: communication among sensors is just partial (e.g. 1 not connected to 4)

Outline

- Algebraic graph theory
  - relevant classes of matrices (non-negative, irreducible, primitive,...) for analyzing graph connectivity properties
- Spectral properties of non-negative matrices: the Perron-Frobenious theorem
- Analysis of a simple consensus algorithm

Non-negative matrices

Let $\mathbb{1}_n = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T \in \mathbb{R}^n$. For short, $\mathbb{1} = \mathbb{1}_n$

Definitions

A matrix $A \in \mathbb{R}^{n \times n}$ is

- non-negative [positive] if $A_{ij} \geq 0$ $[A_{ij} > 0]$, $\forall i, j$
  - row-stochastic (or just stochastic, for short) if it is non-negative and $A\mathbb{1}_n = \mathbb{1}_n$ (sum of each row equal to 1)
  - column-stochastic if it is non-negative and $\mathbb{1}_n^T A = \mathbb{1}_n^T$ (sum of each column equal to 1)
  - doubly stochastic if it is both row- and column-stochastic

Notation: $A \succeq 0$, $A \succ 0$ for non-negative/positive matrices

Remarks

- "stochastic": in row i, the entries $A_{ij} \geq 0$ can be interpreted as probabilities of the event $j \in \{1, \ldots, n\}$. They sum up to 1.
- A is column-stochastic $\Leftrightarrow$ A^T is stochastic

Examples

Localization of the eigenvalues

Recall: for $A \in \mathbb{R}^{n\times n}$

- $\mathrm{Spec}(A)$ is the spectrum of A
- $\rho(A) = \max_{\lambda \in \mathrm{Spec}(A)} |\lambda|$ is the *spectral radius* of A

A general result for localizing $\mathrm{Spec}(A)$ from the elements of A

Theorem (Gershgorin disks)

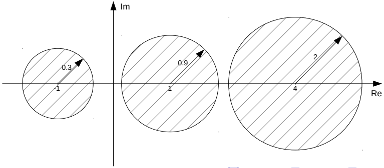
For $A \in \mathbb{R}^{n\times n}$ one has

$$\mathrm{Spec}(A) \subset \cup_{i=1}^{n} B\left(a_{ii}, \sum_{j=1, j\neq i}^{n} |a_{ij}|\right)$$

where $B(c,\gamma) \subset \mathbb{C}^n$ is the closed ball of radius γ centered in $c \in \mathbb{C}^n$

Example

$$A = \begin{bmatrix} -1 & 0.2 & 0.1 \\ 0 & 1 & -0.9 \\ 2 & 0 & 4 \end{bmatrix}$$



Spectral properties of row-stochastic matrices

Theorem

If A is stochastic, then

1. $1 \in \mathrm{Spec}(A)$
2. $\rho(A) = 1$

Sketch of the proof

Proof of point 1. A stochastic $\rightarrow A\mathbb{1} = \mathbb{1} \rightarrow \mathbb{1}$ is an eigenvector of A with eigenvalue 1

Graphical proof of point 2 using Gershgorin disks.

Powers of $A \succeq 0$

Motivation: recall the dynamics of the example systems

$$x^+ = Ax \Rightarrow x(k) = A^k x(0)$$

Remarks about boundedness of A^k

- A stochastic $\Rightarrow$ A can be stable (but not Schur) $\Rightarrow$ A^k can be bounded
- If $A \succeq 0$ is not stochastic but $\rho(A) = 1$, can A^k be unbounded ? Yes
  - $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \rightarrow A^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \ldots A^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

Powers of $A \succeq 0$

- A stochastic $\Rightarrow$ A^k convergent[1]? Not always
  - $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ convergent
  - $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \to A^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A^3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ldots$ not convergent
  - $A = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \to \lim_{k \to +\infty} A^k = \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$
    - ★ Special property: $A^k \succ 0$ for $k \geq 2$

Problem: conditions for convergence of A^k ?

Next steps:

- Associate a matrix $A \in \mathbb{R}^{n \times n}$ to a digraph G
- Analysis of A^k and related connectivity properties of G
- Analysis of convergence of A^k

[1] Matrices for which $\lim_{k \to +\infty} A^k$ exists are called *semi-convergent* in the Textbook.

Properties of A through the associated digraph

Definition

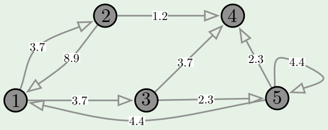
The adjacency matrix $A \in \mathbb{R}^{n \times n}$ of a weighted digraph $G = (V, E, w)$, with n nodes is given by

$$A_{ij} = \begin{cases} 0 & \text{if } (i,j) \notin E \\ w_{ij} & \text{if } (i,j) \in E \end{cases}$$

Standing assumption

All weights w_{ij} are strictly positive

Example



$$A = \begin{bmatrix} 0 & 3.7 & 3.7 & 0 & 0 \\ 8.9 & 0 & 0 & 1.2 & 0 \\ 0 & 0 & 0 & 3.7 & 2.3 \\ 0 & 0 & 0 & 0 & 0 \\ 4.4 & 0 & 0 & 2.3 & 4.4 \end{bmatrix}$$

Remark: A_{ij} means "from i to j"

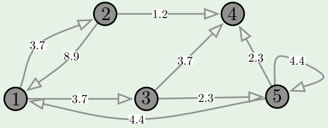
Properties of A through the associated digraph

Definition

The binary adjacency matrix $A \in \{0,1\}^{n \times n}$ of a weighted digraph $G = (V, E, w)$ is given by

$$a_{ij} = \begin{cases} 0 & \text{if } (i,j) \notin E \\ 1 & \text{if } (i,j) \in E \end{cases}$$

Example

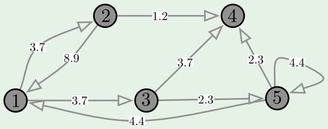


$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Remark: to any $n \times n$ matrix is possible to associate a graph capturing the zero/nonzero pattern and, viceversa, to any digraph is possible to associate an adjacency matrix capturing its topology.

Properties of A and G

Example



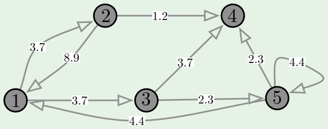
$$A = \begin{bmatrix} 0 & 3.7 & 3.7 & 0 & 0 \\ 8.9 & 0 & 0 & 1.2 & 0 \\ 0 & 0 & 0 & 3.7 & 2.3 \\ 0 & 0 & 0 & 0 & 0 \\ 4.4 & 0 & 0 & 2.3 & 4.4 \end{bmatrix}$$

Easy observations:

- v is a sink $\Leftrightarrow$ the row v of A is zero
- v is a source $\Leftrightarrow$ the column v of A is zero
- v has no self-loop $\Rightarrow A_{vv} = 0$

Properties of A and G

Example



$$A = \begin{bmatrix} 0 & 3.7 & 3.7 & 0 & 0 \\ 8.9 & 0 & 0 & 1.2 & 0 \\ 0 & 0 & 0 & 3.7 & 2.3 \\ 0 & 0 & 0 & 0 & 0 \\ 4.4 & 0 & 0 & 2.3 & 4.4 \end{bmatrix}$$

Recall: the in-/out-degree of u are

$$d^{in}(u) = \sum_{j \in \mathcal{N}^{in}(u)} w_{ju}, \quad d^{out}(u) = \sum_{j \in \mathcal{N}^{out}(u)} w_{uj}$$

- If $w_{ij} \in \{0, 1\}$ then $d^{out}(u) =$(# successors of u) and $d^{in}(u) =$(# predecessors of u)
- A is stochastic [column-stochastic] $\Leftrightarrow d^{out}(u) = 1$, $[d^{in}(u) = 1]$, $\forall u \in V$

Powers of A and graph exploration

Proposition

Let $l_{ij} = (A^k)_{ij}$. Then $l_{ij} \neq 0$ if and only if there is a path from i to j of length k. Moreover, if A is the binary adjaceny matrix, l_{ij} counts the number of directed paths from i to j of length k

Proof that "$l_{ij} \neq 0 \;\Leftrightarrow$ there is a path from i to j of length k"

For $k = 1$ recall that $A_{ij} > 0$ if and only if (i, j) is an edge.
For $k = 2$ one has

$$(A^2)_{ij} = (i\text{th row of } A) \cdot (j\text{th column of } A) = \sum_{h=1}^{n} A_{ih} A_{hj}$$

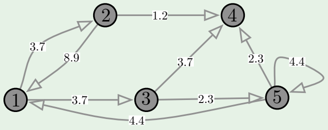
A path of length 2 exists from i to j only if, for some h, (i, h) and (h, j) are edges. This is equivalent to $A_{ih}A_{hj} > 0$. Since all elements of A are non-negative, this is equivalent to $(A^2)_{ij} > 0$.
For bigger k, one can proceed by induction.

Powers of A and graph exploration

Proposition

Let $l_{ij} = (A^k)_{ij}$. Then $l_{ij} \neq 0$ if and only if there is a path from i to j of length k. Moreover, if A is the binary adjaceny matrix, l_{ij} counts the number of directed paths from i to j of length k

Example



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

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

Remark: paths include self-loops

Irreducible matrices and strongly connected graphs

Irreducible matrices: an interesting subset of non-negative matrices.

... but first we need to introduce permutations

Definition

$P \in \{0,1\}^{n \times n}$ is a permutation matrix if it has a single 1 in each row and column.

Example

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad P \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

P describes the permutation $1 \rightarrow 3$, $2 \rightarrow 1$, $3 \rightarrow 2$

Inverse of P

- Permutations are orthogonal matrices: $P^T = P^{-1}$

Check on the example: $P^T P \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = P^T \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Meaning of $P^T AP$ (similarity transformation through P)

Example (ctd)

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

Then $AP = \begin{bmatrix} a_{12} & a_{13} & a_{11} \\ a_{22} & a_{23} & a_{21} \\ a_{32} & a_{33} & a_{31} \end{bmatrix}$ (column order swapped)

and $P^T AP = \begin{bmatrix} a_{22} & a_{23} & a_{21} \\ a_{32} & a_{33} & a_{31} \\ a_{12} & a_{13} & a_{11} \end{bmatrix}$ (row order of AP swapped)

... not very illuminating ...

Meaning of $P^T AP$

Interpretation

Consider the map $A\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$. Apply the same permutation P^T in the domain and the codomain. Example: $\tilde{x}=\begin{bmatrix}x_2\\x_3\\x_1\end{bmatrix}$ and $\tilde{y}=\begin{bmatrix}y_2\\y_3\\y_1\end{bmatrix}$

Then, $\tilde{x}$ and $\tilde{y}$ are related by $P^T AP$, i.e. $P^T AP\tilde{x}=\tilde{y}$

The same holds if $A\in\mathbb{R}^{n\times n}$, $n>3$

Irreducible matrices

Definition

A matrix $A \succeq 0$ is reducible if there is a permutation P such that $P^T AP$ is upper block triangular, i.e.

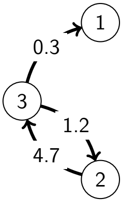
$$P^T AP = \begin{bmatrix} B & C \\ 0_{(n-r)\times r} & D \end{bmatrix} \text{ for some } 0 < r < n$$

where $B \in \mathbb{R}^{r\times r}$ $C \in \mathbb{R}^{r\times(n-r)}$, $D \in \mathbb{R}^{(n-r)\times(n-r)}$ Otherwise, it is called irreducible

- In the definition, the dimension of the zero block is important
- Why reducible matrices are useful for analyzing digraphs ?

Irreducibility $\leftrightarrow$ strong connectivity of G

Problem: is this graph strongly connected ?

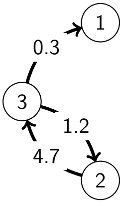


Obviously NO.

Adjacency matrix: $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4.7 \\ 0.3 & 1.2 & 0 \end{bmatrix}$

Irreducibility $\leftrightarrow$ strong connectivity of G

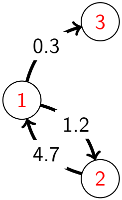
Problem: is this graph strongly connected ?



Obviously NO.

Adjacency matrix: $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4.7 \\ 0.3 & 1.2 & 0 \end{bmatrix}$

Node permutations $1 \to 3,\ 2 \to 2,\ 3 \to 1.$ $P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$



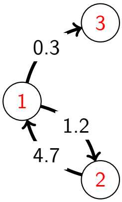
New adjacency matrix:

$$\tilde{A} = \left[\begin{array}{cc|c} 0 & 1.2 & 0.3 \\ 4.7 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right]$$

Zero block of size $(n-r) \times r$ for $r = 2,\ n = 3$

Permutation does not affect connectivity. By construction one has $\tilde{A} = P^T A P$. Moreover, $\tilde{A}$ is reducible

Irreducibility $\leftrightarrow$ strong connectivity of G



New adjacency matrix:

$$\tilde{A} = \begin{bmatrix} 0 & 1.2 & 0.3 \\ 4.7 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Check:

$$P^T A P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 4.7 \\ 0.3 & 1.2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \dots$$

Discussion: powers of $A = \begin{bmatrix} B & C \\ 0_{(n-r)\times r} & D \end{bmatrix}$

Irreducibility $\leftrightarrow$ strong connectivity of G

Theorem

Let A be the adjacency matrix of the weighted digraph G with $n \geq 2$ nodes. The following statements are equivalent

1. A is irreducible
2. $\sum_{k=0}^{n-1} A^k \succ 0$
3. G is strongly connected

Remarks

- The (i,j) element of $\sum_{k=0}^{n-1} A^k \succeq 0$ is nonzero if and only if one can reach j from i in at most $n-1$ hops
- Easy check of strong connectivity: condition in point 2

Mid-lecture summary

- digraph $G \leftrightarrow$ adjacency matrix A
  $(A_{ij}) > 0 \Leftrightarrow (i,j)$ is an edge $\quad (A^K)_{ij} > 0 \Leftrightarrow$ can go from i to j in K hops
- For consensus, study powers of $A \succeq 0$. When do they converge ?

Irreducible matrices

Theorem

Let A be the adjacency matrix of the weighted digraph G with $n \geq 2$ nodes. The following statements are equivalent

1. A is irreducible
2. $\sum_{k=0}^{n-1} A^k \succ 0$
3. G is strongly connected

Primitive matrices

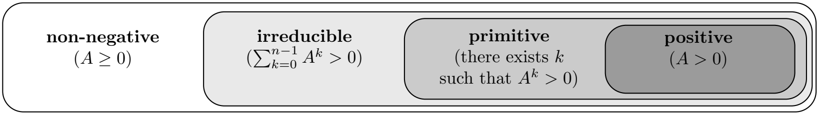
Definition

A nonegative matrix A is primitive if $A^k \succ 0$ for some $k \in \mathbb{N}$

Remark

In the graph $G = (V, E, w)$ associated to a primitive A, one can reach j from i in exactly k hops, $\forall i, j \in V \Rightarrow G$ is strongly connected $\Rightarrow A$ is irreducible

Summary of relations



- In the figure, replace $>$ with $\succ$ and $\geq$ with $\succeq$
- All inclusions are strict
- All classes include both stochastic and nonstochastic matrices

Graph interpretation of primitive matrices

Definition

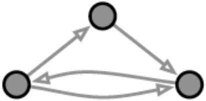
A *strongly connected* digraph G is periodic if the greatest common divisor of the length of all cycles is $k > 1$. In this case, k is called period. Otherwise G is termed aperiodic

- Periodicity is here defined only for directed graphs (the notion of cycle for undirected graphs is different)
- Cycles in digraphs are simple paths $\Rightarrow$ their number is finite $\Rightarrow$ the GCD always exists

Examples

(a) A periodic digraph with period 2

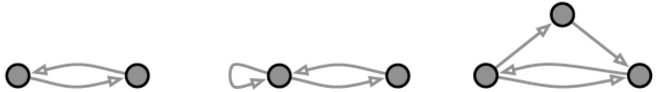
(b) An aperiodic digraph with cycles of length 1 and 2.



(c) An aperiodic digraph with cycles of length 2 and 3.

Graph interpretation of primitive matrices

Examples



(a) A periodic digraph with period 2

(b) An aperiodic digraph with cycles of length 1 and 2.

(c) An aperiodic digraph with cycles of length 2 and 3.

- Graph (a): $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \to A^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A^3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ldots$ not primitive
- Graph (b): $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \to A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ primitive!
- Graph (c): primitive (check at home)

Theorem

Let G a digraph with adjacency matrix A. The following statements are equivalent

- G is strongly connected and aperiodic
- A is primitive

Spectral properties of non-negative matrices

Definition

For $A \in \mathbb{R}^{n \times n}$ and $\lambda \in \mathbb{C}$

- $v \in \mathbb{C}^n$ is a (right) eigenvector of A if $Av = \lambda v$
- $w \in \mathbb{C}^n$ is a left eigenvector of A if $w^T A = \lambda w^T$

Remark

Left eigenvectors are the eigenvectors of A^T

Spectral properties of non-negative matrices

Theorem (Perron Frobenius)

Let $A \in \mathbb{R}^{n \times n}$, $n \geq 2$ be non-negative. Then

1. There is a real $\lambda \in \mathrm{Spec}(A)$ such that $\lambda \geq |\mu| \geq 0$, $\forall \mu \in \mathrm{Spec}(A)$
2. There are right and left eigenvectors v and w of λ that verify $w \succeq 0$ and $v \succeq 0$

If, additionally, A is irreducible (i.e. G is strongly connected), then

3. λ is > 0 and simple
4. w and v are $\succ 0$ and unique (up to rescaling)

If, additionally, A is primitive (i.e. G is strongly connected and aperiodic), then

5. λ verifies $\lambda > |\mu| \geq 0$, $\forall \mu \in \mathrm{Spec}(A)$, $\mu \neq \lambda$

Remarks

- In all cases, $\lambda = \rho(A)$. If λ verifies (1) is called *dominant*. If λ verifies (5) it is *strictly dominant*

Spectral properties of non-negative matrices

Theorem (Perron Frobenius)

Let $A \in \mathbb{R}^{n \times n}$, $n \geq 2$ be non-negative. Then

1. There is a real $\lambda \in \mathrm{Spec}(A)$ such that $\lambda \geq |\mu| \geq 0$, $\forall \mu \in \mathrm{Spec}(A)$
2. There are right and left eigenvectors v and w of λ that verify $w \succeq 0$ and $v \succeq 0$

If, additionally, A is irreducible (i.e. G is strongly connected), then

3. λ is > 0 and simple
4. w and v are $\succ 0$ and unique (up to rescaling)

If, additionally, A is primitive (i.e. G is strongly connected and aperiodic), then

5. λ verifies $\lambda > |\mu| \geq 0$, $\forall \mu \in \mathrm{Spec}(A)$, $\mu \neq \lambda$

Remarks

- (1) and (2) are about *existence* of a dominant λ and nonstrict positivity
- (3) and (4) are about *uniqueness* and strict positivity

Spectral properties of non-negative matrices

Theorem (Perron Frobenius)

Let $A \in \mathbb{R}^{n \times n}$, $n \geq 2$ be non-negative. Then

1. There is a real $\lambda \in \mathrm{Spec}(A)$ such that $\lambda \geq |\mu| \geq 0$, $\forall \mu \in \mathrm{Spec}(A)$
2. There are right and left eigenvectors v and w of λ that verify $w \succeq 0$ and $v \succeq 0$

If, additionally, A is irreducible (i.e. G is strongly connected), then

3. λ is > 0 and simple
4. w and v are $\succ 0$ and unique (up to rescaling)

If, additionally, A is primitive (i.e. G is strongly connected and aperiodic), then

5. λ verifies $\lambda > |\mu| \geq 0$, $\forall \mu \in \mathrm{Spec}(A)$, $\mu \neq \lambda$

Remarks

- Powerful if combined with stochasticity (see next) !

Examples

Examples

Examples

Primitivity and convergence of A^k

Proposition

If A is primitive with $\lambda = \rho(A)$ and v, w normalized (i.e. the chosen left and right eigenvector associated with $\rho(A)$ verify $v^T w = 1$) then

$$\lim_{k \to +\infty} \left(\frac{A}{\lambda} \right)^k = vw^T \tag{1}$$

If, in addition, A is stochastic, then $\lambda = 1$ and $v = \alpha \mathbb{1}_n$. For $\alpha = 1$ and w verifying $\mathbb{1}_n^T w = 1$,

$$\lim_{k \to +\infty} A^k = \mathbb{1}_n w^T \tag{2}$$

Remark

- All rows of $\mathbb{1}_n w^T$ are the same
- w and v verifying $Av = \lambda v$, $w^T A = \lambda w^T$ and $w^T v = 1$ are not unique, but they give the same limit in (1).
  Proof of non-uniqueness: if w and v verify $w^T v = 1$, then, for $\alpha \neq 0$, $\tilde{w} = \frac{1}{\alpha} w$ and $\tilde{v} = \alpha v$ verify $\tilde{w}^T \tilde{v} = 1$
- In (2), w^T is unique if one choses $v = \mathbb{1}_n$

Important remarks

- "v and w normalized" does not mean that $\|v\|_2 = 1$ and $\|w\|_2 = 1$. According to the proposition, it means that

$$v^T w = 1$$

  if A is primitive.
  Definition. If A is primitive and stochastic, then "w normalized" means $\mathbb{1}_n^T w = 1$ (that is $\sum_{i=1}^n w_i = 1$)
- In the sequel, unless otherwise stated, v and w refer to right and left eigenvectors of A associated to the dominant eigenvalue

A first result on consensus

Consider the DT system

$$x^+ = Ax \Rightarrow x(k) = A^k x(0) \tag{3}$$

Theorem (consensus with primitive, stochastic matrices)

If A is primitive and stochastic, the state trajectory $x(k)$ verifies

$$\lim_{k\to+\infty} x(k) = (w^T x(0))\mathbb{1}_n \tag{4}$$

where w is defined as in the previous proposition.
If, in addition, A is doubly stochastic, then $w = \frac{1}{n}\mathbb{1}_n$ and hence

$$\lim_{k\to+\infty} x(k) =< x(0) > \mathbb{1}_n \tag{5}$$

where $< x >= \frac{1}{n}\sum_{i=1}^{n} x_i$

Remarks

- (4) is consensus: all states $x_i(k)$ converge to the same value (a weighted average of $x(0)$)

A first result on consensus

Consider the DT system

$$x^+ = Ax \Rightarrow x(k) = A^k x(0) \tag{3}$$

Theorem (consensus with primitive, stochastic matrices)

If A is primitive and stochastic, the state trajectory $x(k)$ verifies

$$\lim_{k\to+\infty} x(k) = (w^T x(0))\mathbb{1}_n \tag{4}$$

where w is defined as in the previous proposition.
If, in addition, A is doubly stochastic, then $w = \frac{1}{n}\mathbb{1}_n$ and hence

$$\lim_{k\to+\infty} x(k) =< x(0) > \mathbb{1}_n \tag{5}$$

where $< x >= \frac{1}{n}\sum_{i=1}^n x_i$

Remarks

- (5) is *average* consensus: all states $x_i(k)$ converge to the average of $x(0)$

Rate of convergence to consensus

Definition

For a stochastic matrix A, the essential spectral radius $\rho_{ess}(A)$ is the modulus of the second-largest eigenvalue

$$1 = \underbrace{|\lambda_1|}_{\rho(A)} \geq \underbrace{|\lambda_2|}_{\rho_{ess}(A)}$$

Corollary (convergence rate)

In the consensus Theorem, if (4) is verified, it holds $\forall \epsilon > 0 \ \exists c_\epsilon > 0$ such that, for all initial states $x(0) \in \mathbb{R}^n$

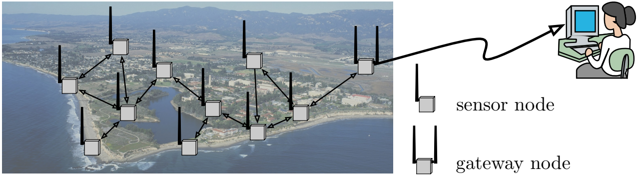
$$\|x(k) - x_{final}\|^2 \leq c_\epsilon (\rho_{ess}(A) + \epsilon)^k \|x(0) - x_{final}\|^2$$

where $x_{final} = (w^T x(0))\mathbb{1}_n$

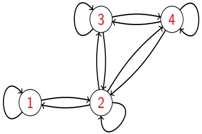
Remark

If A is primitive and stochastic, $\rho_{ess}(A) < 1$ (see point 5 of Perron Frobenius theorem). Then, there exists a sufficiently small $\epsilon > 0$ such that $\rho_{ess}(A) + \epsilon < 1$ and hence guaranteeing exponential convergence to the consensus state with rate $\log(\rho_{ess}(A) + \epsilon)$

Example averaging in wireless sensor networks



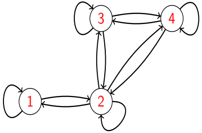
$$x^+ = Ax \quad A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$



Associated digraph

Example averaging in wireless sensor networks

$$x^+ = Ax \quad A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$



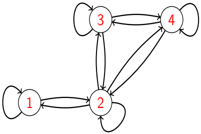
Associated digraph

- A is primitive because G is strongly connected and aperiodic. Check:

$$A^2 = \begin{bmatrix} \frac{3}{8} & \frac{3}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{3}{16} & \frac{17}{48} & \frac{11}{48} & \frac{11}{48} \\ \frac{1}{12} & \frac{11}{36} & \frac{11}{36} & \frac{11}{36} \\ \frac{1}{12} & \frac{11}{36} & \frac{11}{36} & \frac{11}{36} \end{bmatrix}$$

Example averaging in wireless sensor networks

$$x^+ = Ax \quad A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$



Associated digraph

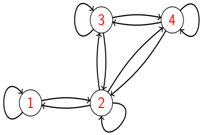
- Perron Frobenious: one strictly dominant eigenvalue with strictly positive and unique right and left eigenvectors v and w.
  Check: by direct computation, eigenvalues and associated right eigenvectors are

$$(1, \mathbb{1}_4), \quad \left(\frac{1}{24}(5+\sqrt{73}), \begin{bmatrix} -2-2\sqrt{73} \\ -11+\sqrt{73} \\ 8 \\ 8 \end{bmatrix}\right), \quad \left(\frac{1}{24}(5-\sqrt{73}), \begin{bmatrix} 2(-1+\sqrt{73}) \\ -11-\sqrt{73} \\ 8 \\ 8 \end{bmatrix}\right), \quad \left(0, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}\right)$$

  - dominant eigenvalue $\lambda = 1$ with right eigenvector $v = \mathbb{1}_4 \succ 0$
  - *left* eigenvector for $\lambda = 1$: $w = \begin{bmatrix} 1/6 & 1/3 & 1/4 & 1/4 \end{bmatrix}^T$, chosen so that $\mathbb{1}_4^T w = 1$

Example averaging in wireless sensor networks

$$x^+ = Ax \quad A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$



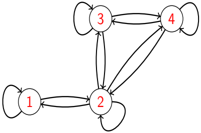
Associated digraph

- Convergence of A^k: We have that

$$\lim_{k\to+\infty} \left(\frac{A}{\lambda}\right)^k = \mathbb{1}_4 w^T = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

Example averaging in wireless sensor networks

$$x^+ = Ax \quad A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$



Associated digraph

- Theorem on consensus: $x(k)$ converges to

  $$(w^T x(0))\mathbb{1}_4 = [(1/6)x_1(0) + (1/3)x_2(0) + (1/4)x_3(0) + (1/4)x_4(0)]\, \mathbb{1}_4$$

  Since A is not doubly stochastic, average consensus is not expected. Indeed it is not reached because node 2 has more infuence than the others.

  - for $\epsilon > 0$ such that $\rho_{ess}(A) + \epsilon < 1$, the convergence rate is $\log(\rho_{ess}(A) + \epsilon) = \log(\frac{1}{24}(5 + \sqrt{73}) + \epsilon) = \log(0.5643 + \epsilon)$

Take home messages

- A non-negative/irreducible/primitive matrix A is related to the connectivity properties of the associated digraph G
  - The powers of A as well
- For non-negative matrices, the Perron-Frobenious theorem allows one to:
  - partially charachterize the eigenstructure of A
  - study convergence of A^k as $k \to \infty$
- A primitive + stochastic $\Rightarrow$ consensus !