Introduction to Discrete-time Averaging Systems

Chengcheng Zhao

Zhejiang University, Hangzhou, China
&University of Victoria, Victoria, Canada

August 2, 2021

Background

- Swarm behaviour in nature

(a) Bird flocking

(b) Ant swarming

(c) Fish Swarming

Background

- Swarm behaviour in nature

(d) Bird flocking

(e) Ant swarming

(f) Fish Swarming

- No control center
- Individual animals only interact with their neighbours
- Collective animal behavior

How can we use the idea behind in social and engineering fields?

Background

- Consensus algorithms
  - Average consensus: all states converges to average
  - Maximum consensus: all states converge to maximum value
- Wide application
  - Smart grids, VANETS, social networks, and crowd-sensing

Figure 1: Wide applications

Background

- Consensus algorithms
  - Average consensus: all states converges to average
  - Maximum consensus: all states converge to maximum value
- Wide application
  - Smart grids, VANETS, social networks, and crowd-sensing

Figure 1: Wide applications

Two application examples to show the related averaging systems

Application Examples

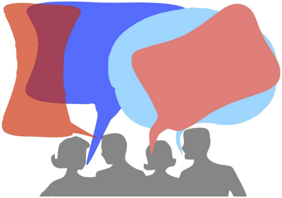


Figure 2: Interactions in a social influence network

Application Examples

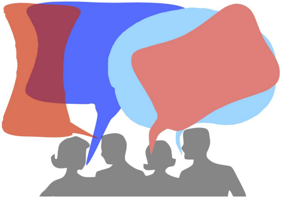


Figure 2: Interactions in a social influence network

- Social influence networks: opinion dynamics
  - A group of n individuals who must act together as a team

Application Examples

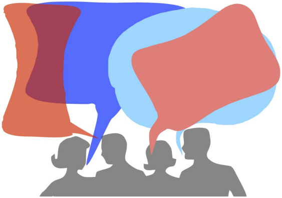


Figure 2: Interactions in a social influence network

- Social influence networks: opinion dynamics
  - A group of n individuals who must act together as a team
  - Each individual has its own subjective estimate p_i for the unknown parameters

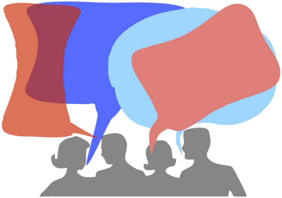


Figure 2: Interactions in a social influence network

- Social influence networks: opinion dynamics
  - A group of n individuals who must act together as a team
  - Each individual has its own subjective estimate p_i for the unknown parameters
  - Individual i is appraised of p_i of each other member $j \neq i$ of the group

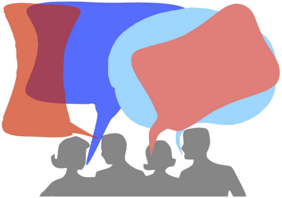


Figure 2: Interactions in a social influence network

- Social influence networks: opinion dynamics
  - A group of n individuals who must act together as a team
  - Each individual has its own subjective estimate p_i for the unknown parameters
  - Individual i is appraised of p_i of each other member $j \neq i$ of the group
  - How to model predictions that the individual will revise its estimate?

Application Examples

- Social influence networks: opinion dynamics
  - The model (French-Harary-DeGroot) predicts that the individual will revise its estimate to be

$$p_i(k+1) = \sum_{j=1}^{n} a_{ij} p_j(k) \tag{1}$$

Application Examples

- Social influence networks: opinion dynamics
  - The model (French-Harary-DeGroot) predicts that the individual will revise its estimate to be

$$p_i(k+1) = \sum_{j=1}^{n} a_{ij} p_j(k) \tag{1}$$

    - $a_{ij} \geq 0$ denotes the weight that individual i assigns to individual j;
    - $\sum_{j=1}^{n} a_{ij} = 1$ for all i;
    - a_{ii} describes the attachment of individual i to its own opinion;
    - a_{ij} is an interpersonal influence weight that individual i accords to individual j;

Application Examples

- Social influence networks: opinion dynamics
  - Scientific questions of interests
    - Is this model of human opinion dynamics believable?
      Is there empirical evidence in its support?

Application Examples

- Social influence networks: opinion dynamics
  - Scientific questions of interests
    - Is this model of human opinion dynamics believable?
      Is there empirical evidence in its support?
    - How does one measure the coefficients a_{ij}?

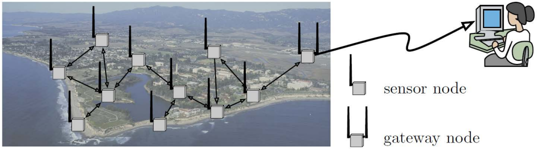
Application Examples

- Social influence networks: opinion dynamics
  - Scientific questions of interests
    - Is this model of human opinion dynamics believable?
      Is there empirical evidence in its support?
    - How does one measure the coefficients a_{ij}?
    - Under what conditions do the estimate converge to the same estimate?
      And to what final estimate?

Application Examples

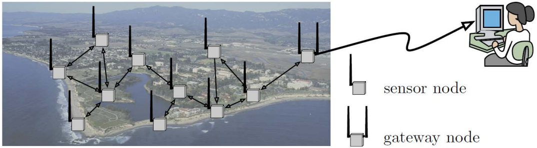
- Social influence networks: opinion dynamics
  - Scientific questions of interests
    - Is this model of human opinion dynamics believable?
      Is there empirical evidence in its support?
    - How does one measure the coefficients a_{ij}?
    - Under what conditions do the estimate converge to the same estimate?
      And to what final estimate?
    - What are more realistic, empirically-motivated models, possibly including stubborn individuals or antagonistic interactions?

Application Examples



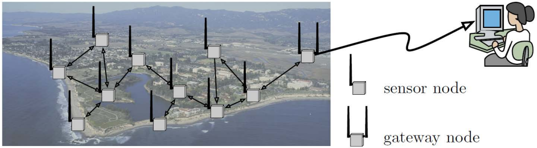
- Wireless sensor networks
  - A collection of spatially-distributed devices

Application Examples



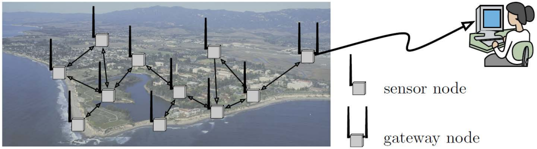
- Wireless sensor networks
  - A collection of spatially-distributed devices
  - Measure physical and environmental variables (e.g., temperature, vibrations, sound, light, etc)

Application Examples



- Wireless sensor networks
  - A collection of spatially-distributed devices
  - Measure physical and environmental variables (e.g., temperature, vibrations, sound, light, etc)
  - Perform local computations, and transmit information to neighboring device throughout the network

Application Examples



- Wireless sensor networks
  - A collection of spatially-distributed devices
  - Measure physical and environmental variables (e.g., temperature, vibrations, sound, light, etc)
  - Perform local computations, and transmit information to neighboring device throughout the network

How can all devices obtain the accurate estimate in a distributed way?

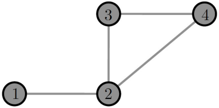


Figure 3: The communication graph for devices

- Wireless sensor networks: linear averaging
  - Each node has a measured temperature $x_i(0)$

Application Examples

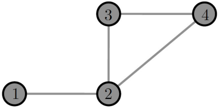


Figure 3: The communication graph for devices

- Wireless sensor networks: linear averaging
  - Each node has a measured temperature $x_i(0)$
  - Apply the following linear averaging algorithm (**Algorithm 1**)

Application Examples

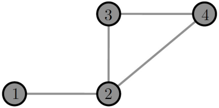


Figure 3: The communication graph for devices

- Wireless sensor networks: linear averaging
  - Each node has a measured temperature $x_i(0)$
  - Apply the following linear averaging algorithm (**Algorithm 1**)

$$x_i(k+1) = \text{mean}(x_i(k), \{x_j(k)\}), \text{for all neighbors } j \tag{2}$$

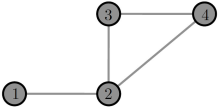


Figure 3: The communication graph for devices

- Wireless sensor networks: linear averaging
  - Each node has a measured temperature $x_i(0)$
  - Apply the following linear averaging algorithm (**Algorithm 1**)

$$x_i(k+1) = \text{mean}(x_i(k), \{x_j(k)\}), \text{for all neighbors } j \tag{2}$$

    - $x_i(k+1)$ is the value at iteration k, update example: $x_1(k+1) = x_1(k)/2 + x_2(k)/2$

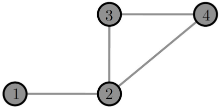


Figure 3: The communication graph for devices

- Wireless sensor networks: linear averaging
  - Each node has a measured temperature $x_i(0)$
  - Apply the following linear averaging algorithm (**Algorithm 1**)

$$x_i(k+1) = \text{mean}(x_i(k), \{x_j(k)\}), \text{for all neighbors } j \tag{2}$$

    - $x_i(k+1)$ is the value at iteration k, update example: $x_1(k+1) = x_1(k)/2 + x_2(k)/2$
  - update rule $x(k+1) = Ax(k)$

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \tag{3}$$

Application Examples

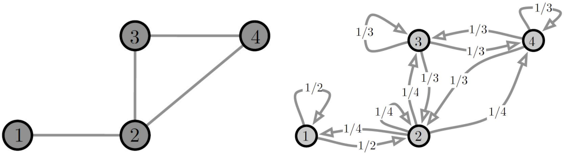
- Apply Algorithm 1

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \quad (4)$$

Application Examples

- Apply Algorithm 1

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \tag{4}$$



(a) (b)

Figure 4: The original communication graph and the weighted graph

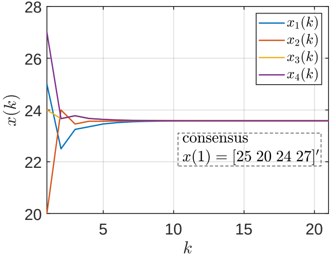
Application Examples

- Apply Algorithm 1

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \tag{5}$$

- Apply Algorithm 1

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \tag{5}$$



As the average is 24, average consensus cannot be achieved.

Application Examples

- Apply a new weight strategy (**Algorithm 2**)

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \tag{6}$$

- Apply a new weight strategy (**Algorithm 2**)

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \tag{6}$$

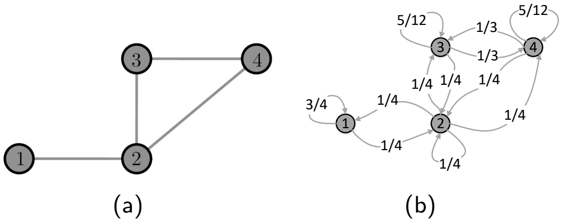


Figure 5: The original communication graph and the weighted graph

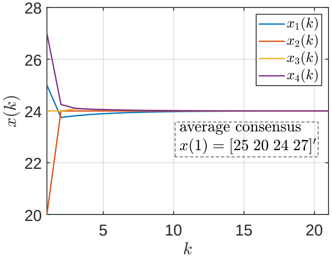
Application Examples

- Apply a new weight strategy (Algorithm 2)

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \tag{7}$$

- Apply a new weight strategy (Algorithm 2)

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix} = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix} \tag{7}$$



As the average is 24, average consensus can be achieved.

Application Examples

- Interesting findings

$$A = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix}, A = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \quad (8)$$

  - A is a non-negative matrix

Application Examples

- Interesting findings

$$A = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix}, A = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \quad (8)$$

  - A is a non-negative matrix
  - A is a row stochastic matrix

Application Examples

- Interesting findings

$$A = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix}, A = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \tag{8}$$

  - A is a non-negative matrix
  - A is a row stochastic matrix
  - The associated graph of A is strongly connected

Application Examples

- Interesting findings

$$A = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix}, A = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \tag{8}$$

  - A is a non-negative matrix
  - A is a row stochastic matrix
  - The associated graph of A is strongly connected
  - When can we achieve average consensus?

- Interesting findings

$$A = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix}, A = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \tag{8}$$

  - A is a non-negative matrix
  - A is a row stochastic matrix
  - The associated graph of A is strongly connected
  - When can we achieve average consensus?
    In Algorithm 2, A is symmetric

Averaging Systems

- Dynamic model

$$x(k+1) = Ax(k) \tag{9}$$

Averaging Systems

- Dynamic model

$$x(k+1) = Ax(k) \tag{9}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdots \\ x_n(k) \end{bmatrix} \tag{10}$$

Averaging Systems

- Dynamic model

$$x(k+1) = Ax(k) \tag{9}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdots \\ x_n(k) \end{bmatrix} \tag{10}$$

  - $A \in \mathbb{R}^{n \times n}$ has non-negative entries and unit row sums
  - $x(k) \in \mathbb{R}^n$, $k \geq 0$
    $x_i(0)$ is the initial scalar state (temperature, vibrations, sound, light)
    $x_i(k)$ is the updated state at iteration k

Averaging Systems

- Dynamic model

$$x(k+1) = Ax(k) \tag{9}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdots \\ x_n(k) \end{bmatrix} \tag{10}$$

  - $A \in \mathbb{R}^{n \times n}$ has non-negative entries and unit row sums
  - $x(k) \in \mathbb{R}^n$, $k \geq 0$
    $x_i(0)$ is the initial scalar state (temperature, vibrations, sound, light)
    $x_i(k)$ is the updated state at iteration k
- Interesting problems for the averaging model
  - Does each node converge to a value?
    Is this value the same for all nodes (consensus)?

Averaging Systems

- Dynamic model

$$x(k+1) = Ax(k) \tag{9}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdots \\ x_n(k) \end{bmatrix} \tag{10}$$

  - $A \in \mathbb{R}^{n \times n}$ has non-negative entries and unit row sums
  - $x(k) \in \mathbb{R}^n$, $k \geq 0$
    $x_i(0)$ is the initial scalar state (temperature, vibrations, sound, light)
    $x_i(k)$ is the updated state at iteration k
- Interesting problems for the averaging model
  - Does each node converge to a value?
    Is this value the same for all nodes (consensus)?
  - Is this value equal to the average of the initial conditions?
    When do the agents achieve average consensus?

Averaging Systems

- Dynamic model

$$x(k+1) = Ax(k) \quad (9)$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdots \\ x_n(k) \end{bmatrix} \quad (10)$$

  - $A \in \mathbb{R}^{n \times n}$ has non-negative entries and unit row sums
  - $x(k) \in \mathbb{R}^n$, $k \geq 0$
    $x_i(0)$ is the initial scalar state (temperature, vibrations, sound, light)
    $x_i(k)$ is the updated state at iteration k
- Interesting problems for the averaging model
  - Does each node converge to a value?
    Is this value the same for all nodes (consensus)?
  - Is this value equal to the average of the initial conditions?
    When do the agents achieve average consensus?
  - What properties do the graph and the corresponding matrix need to have in order for the algorithm to converge?

Averaging Systems

- Dynamic model

$$\begin{aligned} x(k+1) = Ax(k) \Rightarrow x(k) &= Ax(k-1) \\ &= A \times Ax(k-1) \\ &= \vdots \\ &= A^{(k+1)}x(0) \end{aligned} \tag{11}$$

Averaging Systems

- Dynamic model

$$
\begin{aligned}
x(k+1) = Ax(k) \Rightarrow x(k) &= Ax(k-1) \\
&= A \times Ax(k-1) \\
&= \vdots \\
&= A^{(k+1)}x(0)
\end{aligned} \tag{11}
$$

- Jordan normal form

$$
\begin{aligned}
A = PJP^{-1} \Rightarrow x(k+1) &= A^{(k+1)}x(0) \\
&= (PJP^{-1})^{k+1}x(0)
\end{aligned} \tag{12}
$$

Averaging Systems

- Dynamic model

$$\begin{aligned} x(k+1) = Ax(k) \Rightarrow x(k) &= Ax(k-1) \\ &= A \times Ax(k-1) \\ &= \vdots \\ &= A^{(k+1)}x(0) \end{aligned} \quad (11)$$

- Jordan normal form

$$\begin{aligned} A = PJP^{-1} \Rightarrow x(k+1) &= A^{(k+1)}x(0) \\ &= (PJP^{-1})^{k+1}x(0) \end{aligned} \quad (12)$$

$$\begin{aligned} &A^2 = PJP^{-1}PJP^{-1} = PJ^2P^{-1} \\ &A^k = PJP^{-1}PJP^{-1} \cdots PJP^{-1} = PJ^kP^{-1} \\ &\Rightarrow x(k+1) = PJ^{k+1}P^{-1}x(0) \end{aligned} \quad (13)$$

Averaging Systems

- Dynamic model

$$\begin{aligned} x(k+1) = Ax(k) \Rightarrow x(k) &= Ax(k-1) \\ &= A \times Ax(k-1) \\ &= \vdots \\ &= A^{(k+1)}x(0) \end{aligned} \tag{11}$$

- Jordan normal form

$$\begin{aligned} A = PJP^{-1} \Rightarrow x(k+1) &= A^{(k+1)}x(0) \\ &= (PJP^{-1})^{k+1}x(0) \end{aligned} \tag{12}$$

$$\begin{aligned} & A^2 = PJP^{-1}PJP^{-1} = PJ^2P^{-1} \\ & A^k = PJP^{-1}PJP^{-1} \cdots PJP^{-1} = PJ^kP^{-1} \\ & \Rightarrow x(k+1) = PJ^{k+1}P^{-1}x(0) \end{aligned} \tag{13}$$

Averaging Systems

- Suppose $A \in \mathbb{R}^{n \times n}$ is symmetric

Averaging Systems

- Suppose $A \in \mathbb{R}^{n \times n}$ is symmetric

$$x(k+1) = A^{(k+1)}x(0) = PJ^{k+1}P^{-1}x(0) \tag{14}$$

Averaging Systems

- Suppose $A \in \mathbb{R}^{n\times n}$ is symmetric

$$x(k+1) = A^{(k+1)}x(0) = PJ^{k+1}P^{-1}x(0) \tag{14}$$

  - Transformation

$$J = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

$$\Rightarrow J^{k+1} = \begin{bmatrix} \lambda_1^{k+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^{k+1} \end{bmatrix}$$

Averaging Systems

- Suppose $A \in \mathbb{R}^{n \times n}$ is symmetric
  - Take limitations on both sides of the equation

$$\begin{aligned}
\lim_{k\to\infty} x(k+1) &= \lim_{k\to\infty} PJ^{k+1}P^{-1}x(0) \\
&= \lim_{k\to\infty} P \begin{bmatrix} \lambda_1^{k+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^{k+1} \end{bmatrix} P^{-1}x(0) \\
&= P \begin{bmatrix} \lim_{k\to\infty} \lambda_1^{k+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lim_{k\to\infty} \lambda_n^{k+1} \end{bmatrix} P^{-1}x(0)
\end{aligned} \tag{15}$$

- Suppose $A \in \mathbb{R}^{n\times n}$ is symmetric
  - Take limitations on both sides of the equation

$$
\begin{aligned}
\lim_{k\to\infty} x(k+1) &= \lim_{k\to\infty} PJ^{k+1}P^{-1}x(0) \\
&= \lim_{k\to\infty} P \begin{bmatrix} \lambda_1^{k+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^{k+1} \end{bmatrix} P^{-1}x(0) \\
&= P \begin{bmatrix} \lim\limits_{k\to\infty} \lambda_1^{k+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lim\limits_{k\to\infty} \lambda_n^{k+1} \end{bmatrix} P^{-1}x(0)
\end{aligned} \tag{15}
$$

  - Consensus is correlated to the eigenvalues of the matrix A
  - Limitation exists if $\lim\limits_{k\to\infty} \lambda_i^{k+1}$ exists, i.e., $\lambda_i \leq 1$

Connectivity of the Associated Graph

- The power of matrix A

$$A^2 = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 0.3750 & 0.3750 & 0.1250 & 0.1250 \\ 0.1875 & 0.3542 & 0.2292 & 0.2292 \\ 0.0833 & 0.3056 & 0.3056 & 0.30560 \\ 0.0833 & 0.3056 & 0.3056 & 0.3056 \end{bmatrix} \tag{16}$$

- The power of matrix A

$$A^2 = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 0.3750 & 0.3750 & 0.1250 & 0.1250 \\ 0.1875 & 0.3542 & 0.2292 & 0.2292 \\ 0.0833 & 0.3056 & 0.3056 & 0.30560 \\ 0.0833 & 0.3056 & 0.3056 & 0.3056 \end{bmatrix} \tag{16}$$

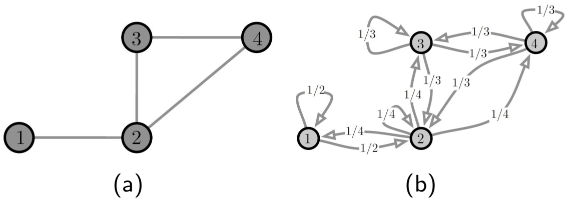


Figure 6: The original communication graph and the weighted graph

Connectivity of the Associated Graph

- The power of matrix A
  - Nonzero elements of A^2: the directed path with a length of 2 in the associated graph
    $[A^2]_{ij} > 0$, there is a directed path between node i and node j
  - The information flow between different nodes
    $[A^2]_{ij} > 0$, node i can obtain the information of node j through two hops interaction

Row-stochastic matrices and their spectral radius

- For any row-stochastic matrix $A \in \mathbb{R}^{n\times n}$
  1) 1 is an eigenvalue $\Leftarrow$ definition $A1_n = 1_n$
  2) $\mathrm{spec}(A)$ is a subset of the unit disk and $\rho(A) = 1$
- Gershgorin Disk Theorem

Theorem

For any square matrix $A \in \mathbb{R}^{n\times n}$,

$$\mathrm{spec}(A) \subset \cup_{i=\{1,\cdots,n\}}\{z||z-a_{ii}| \leq \sum_{j=1,j\neq i}^{n} |a_{ij}|\} \tag{17}$$

Proof.

$Ax = \lambda x$, $x \neq 0_n$, $|x_i| = \max_{j\{1,\cdots,n\}} |x_j| > 0 \Rightarrow \lambda x_i = \sum_{j=1}^{n} a_{ij}x_j$

$\Rightarrow \lambda - a_{ii} = \sum_{j=1,j\neq i}^{n} a_{ij}x_j/x_i$

$\Rightarrow |\lambda - a_{ii}| = |\sum_{j=1,j\neq i}^{n} a_{ij}x_j/x_i| \leq \sum_{j=1,j\neq i}^{n} |a_{ij}||x_j|/|x_i| \leq \sum_{j=1,j\neq i}^{n} |a_{ij}|$ □

Perron-Frobenius Theory

- Irreducible and primitive matrices
  - $A \in \mathbb{R}^{n\times n}$, $n \geq 2$ has non-negative entries and is
    - irreducible if $\sum_{k=0}^{n-1} A^k > 0$ (G is strongly connected)
    - primitive if there exists a positive integer k such that $A^k > 0$ (G is strongly connected and aperiodic)
    - a primitive matrix is irreducible

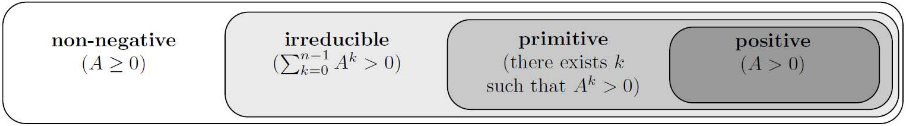


Figure 7: The set of non-negative square matrices and its subsets of irreducible, primitive and positive matrices

Perron-Frobenius Theory

- Irreducible and primitive matrices

$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$: $\text{spec}(A_1) = \{1, 1\}$, the zero/nonzero pattern in A_1^k is constant, and $\lim_{k\to\infty} A_1^k = I_2$,

$A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$: $\text{spec}(A_2) = \{1, -1\}$, the zero/nonzero pattern in A_2^k is periodic, and $\lim_{k\to\infty} A_2^k$ does not exist,

$A_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$: $\text{spec}(A_3) = \{0, 0\}$, the zero/nonzero pattern is $A_3^k = 0$ for all $k \geq 2$, and $\lim_{k\to\infty} A_3^k = 0$,

$A_4 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$: $\text{spec}(A_4) = \{1, -1/2\}$, the zero/nonzero pattern is $A_4^k > 0$ for all $k \geq 2$, and $\lim_{k\to\infty} A_4^k = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$, and

$A_5 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$: $\text{spec}(A_5) = \{1, 1\}$, the zero/nonzero pattern in A_5^k is constant and $\lim_{k\to\infty} A_5^k$ is unbounded.

Figure 8: Example 2-dimensional non-negative matrices and their properties

- Perron-Frobenius Theorem

Theorem

Let $A \in \mathbb{R}^{n\times n}$, $n \geq 2$. If A is non-negative, then
1) there exists a real eigenvalue $\lambda \geq |\mu| \geq 0$ for all other eigenvalues μ;
2) the right and left eigenvectors v and w of λ can be selected non-negative.
If additionally A is irreducible, then
3) the eigenvalue λ is strictly positive and simple;
4) the right and left eigenvectors v and w of λ are unique and positive.
If additionally A is primitive, then
5) the eigenvalue $\lambda > |\mu|$ for all other eigenvalues μ

- Proof: analyze properties of positive matrices and then use "limit"

Perron-Frobenius Theory

- Lemma for positive matrices

Lemma

Let $A \in \mathbb{R}^{n\times n}$, $n \geq 2$. *If* A *is positive, then*
Lem-1) there exists an eigenvalue $\lambda = \rho(A) > |\mu| \geq 0$ *for all other eigenvalues* μ*;*
Lem-2) λ *is simple, i.e.,* $\mathrm{algmulti}_A(\lambda) = 1$*;*
Lem-3) the right and left eigenvectors v *and* w *of* λ *are positive.*

- Proof is omitted and can be found in the reference below
- $\rho(A)$ is the only one eigenvalue on the spectral circle
- Algebraic and geometric multiples are equal to 1

C. D. Meyer, "Matrix analysis and applied linear algebra," *SIAM*, 2000.

Perron-Frobenius Theory

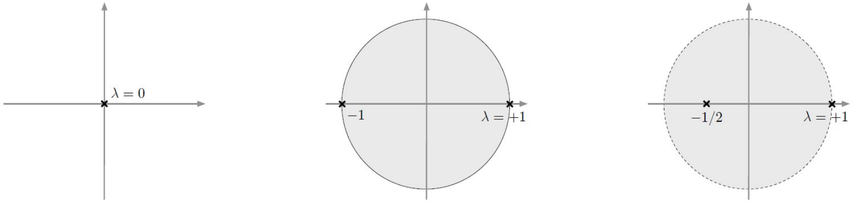
- Proof of 1) and 2) of non-negative matrix A
  - Key idea: positive matrices $\Rightarrow$ sequence convergence
    - Construct a positive matrix $A_k = A + (1/k)1_n 1_n^\top$
      $\Rightarrow A_k > 0$ and let (r_k, p_k) $(r_k = \rho(A_k),\ p_k > 0,\ \|p_k\| = 1)$ eigenpair
      $\Rightarrow \{p_k\}_{k=1}^{\infty}$ is a bounded set as contained in the unit 1-sphere in $\mathbb{R}^n$
    - Convergence: each bounded sequence in $\mathbb{R}^n$ has a convergent subsequence
      $\Rightarrow \{p_k\}_{k=1}^{\infty}$ has a convergent subsequence, $p_{k_i} > 0$ and $\|p_{k_i}\| = 1$
      $\Rightarrow \{p_{k_i}\}_{k_i=1}^{\infty} \to z$ where $z \geq 0$
    - Take limitations: $r_k = \lim_{t\to\infty} \|A_k^t\|^{1/t} \Rightarrow 0 \leq A < A_1,\ \rho(A) \leq \rho(A_1)$
      $\Rightarrow A_1 > A_2 > \cdots > A \Rightarrow r_1 > r_2 > \cdots > r$ $(r = \rho(A))$, $\{r_k\}_{k=1}^{\infty}$ is a monotonic sequence of positive numbers bounded by r
      $\Rightarrow \lim_{k\to\infty} r_k = r^*$ exists and $r^* \geq r$, $\lim_{k_i\to\infty} r_{k_i} = r^* \geq r$
      $\Rightarrow \lim_{k\to\infty} A_k = A$ implies $\lim_{k_i\to\infty} A_{k_i} = A$
      $\Rightarrow Az = \lim_{k_i\to\infty} A_{k_i} p_{k_i} = \lim_{i\to\infty} r_{k_i} p_{k_i} = r^* z \Rightarrow r^* \in \mathrm{spec}(A) \Rightarrow r^* \leq r$
      $\Rightarrow r^* = r$ and $Az = rz$ with $z \geq 0$ and $z \neq 0$

- Proof of 3) and 4) for irreducible matrices
  - $\rho(A)$ is simple: $r = \rho(A)$, let $B = (I + A)^{n-1} > 0$ and $\nu = \rho(B)$
    $\Rightarrow \lambda \in \mathrm{spec}(A) \Leftrightarrow (1+\lambda)^{n-1} \in \mathrm{spec}(B)$,
    $\mathrm{algmulti}_A(\lambda) = \mathrm{algmulti}_B((1+\lambda)^{n-1})$
    $\Rightarrow \nu = \max_{\lambda \in \mathrm{spec}(A)} |1+\lambda|^{n-1} = \{\max_{\lambda \in \mathrm{spec}(A)} |1+\lambda|\}^{n-1} = (1+r)^{n-1}$
    $\Rightarrow \mathrm{algmulti}_A(r) = 1 \Leftrightarrow \mathrm{algmulti}_B(\nu) = 1$.
  - Positive eigenvector: (r, x) is eigenpair of $A \Leftrightarrow (\nu, x)$ is eigenpair of B
    $\Rightarrow x > 0$
    $\Rightarrow r > 0$; otherwise $Ax = 0$ impossible $\Leftarrow A \geq 0$, $x > 0 \Rightarrow Ax > 0$

- Proof of 5)
  - By definition of primitive matrix
    $B = A^k > 0 \Rightarrow \lambda \in \mathrm{spec}(A) \Leftrightarrow \lambda^k \in \mathrm{spec}(B)$
    - Suppose $|\lambda_1| = 1$ and $\lambda_1 \neq \rho(A) \Rightarrow \lambda_1^k \in \mathrm{spec}(B)$
      $\Rightarrow |\lambda_1^k| = 1$ and $\mathrm{spec}(B)$ has two eigenvalues on the spectral circle
      contradict with the result for positive matrix
      only one eigenvalue $\rho(A)$ on the spectral circle
  - $\Rightarrow$ eigenvalue $\rho(A) > |\mu|$ for all other eigenvalues μ

Perron-Frobenius Theory-Perron

- Frobenius Theorem
  - $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $A_4 = \begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix}$



(a) The matrix A_3 is reducible: its dominant eigenvalue is 0 and so is its other eigenvalue.

(b) The matrix A_2 is irreducible but not primitive: its dominant eigenvalues $+1$ is not strictly larger, in magnitude, than the other eigenvalue -1.

(c) The matrix A_4 is primitive: its dominant eigenvalue $+1$ is strictly larger than the other eigenvalue $-1/2$.

Figure 9: Example 2-dimensional non-negative matrices and their properties

Perron-Frobenius Theory

- Powers of non-negative matrices

Theorem

Let $A \in \mathbb{R}^{n \times n}$, $n \geq 2$ be non-negative with dominant eigenvalue λ and the right and left eigenvectors are denoted by v and w of λ, $v^\top w = 1$. If λ is simple and strictly larger in magnitude than all other eigenvalues, then we have

$$\lim_{k \to \infty} \frac{A^k}{\lambda^k} = vw^\top \tag{18}$$

Proof.

λ is simple and strictly larger $\Rightarrow A = T \begin{bmatrix} \lambda & 0_{1 \times n-1} \\ 0_{n-1 \times 1} & B \end{bmatrix} T^{-1}$ and $\rho(B/\lambda) < 1$

$\Rightarrow \lim_{k \Rightarrow \infty} B^k/\lambda^k = 0 \Rightarrow \lim_{k \Rightarrow \infty} (\frac{A}{\lambda})^k = T(\lim_{k \Rightarrow \infty} \begin{bmatrix} 1^k & 0_{1 \times n-1} \\ 0_{n-1 \times 1} & B^k \end{bmatrix})T^{-1} =$

$T(\lim_{k \Rightarrow \infty} \begin{bmatrix} 1 & 0_{1 \times n-1} \\ 0_{n-1 \times 1} & 0_{n-1 \times n-1} \end{bmatrix})T^{-1} = vw^\top$, v is the first column of T and w is the first row of T^{-1}. □

- Row-stochastic matrices (Let A be a row-stochastic matrix and let G be its associated digraph)
  - the eigenvalue 1 is simple and all other eigenvalues $|\mu| < 1$
  - $\lim_{k\to\infty} A^k = 1_n w^\top$ for $w > 0$ and $1_n^\top w = 1$
  - G is an aperiodic strongly-connected graph

- Row-stochastic matrices (Let A be a row-stochastic matrix and let G be its associated digraph)
  - the eigenvalue 1 is simple and all other eigenvalues $|\mu| < 1$
  - $\lim_{k\to\infty} A^k = 1_n w^\top$ for $w > 0$ and $1_n^\top w = 1$
  - G is an aperiodic strongly-connected graph
- If the previous conditions are satisfied, then
  - the solution of $x(k+1) = Ax(k)$ satisfies $\lim_{k\Rightarrow\infty} x(k) = w^\top x(0) 1_n$
  - if additionally A is doubly-stochastic, then $w = \frac{1}{n} 1_n$ so that

$$\lim_{k\to\infty} x(k) = \frac{1_n^\top x(0)}{n} 1_n = \text{average}(x(0)) 1_n \tag{19}$$

Summary

- Discrete-time averaging systems
  - Background and application examples
  - Analysis intuition for convergence (consensus)
  - Conditions to ensure consensus and average consensus
- References

F. Bullo, "Lectures on network systems," *Kindle Direct Publishing*, 2019.

C. Godsil, G. F. Royle, "Algebraic graph theory," *Springer Science & Business Media*, 2013.

C. D. Meyer, "Matrix analysis and applied linear algebra," *SIAM*, 2000.

R. Olfati-Saber, J. A. Fax, and R. M. Murray. "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, 95(1): 215-233, 2007.

R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, 49(9): 1520-1533, 2004.

Thank You!
Q&A