Distributed Systems & Control

Advanced Topics in Control 2022

Lecture 2: Elements of Matrix Theory

ETH zürich

AUTOMATIC CONTROL LABORATORY

Brief announcements

227-0690-12L Advanced Topics in Control (Spring 2022) FS2022

Dashboard / Meine Kurse / 227-0690-12L Advanced Topics in Control (Spring 2022) FS2022

General Information

Lecturers: Prof. Florian Dörfler, Dr. Mathias Hudoba de Badyn, Dr. Vahid Mamduhi

Assistants: Andrea Martinelli, Dr. Dominic Liao-McPherson, Alberto Padoan, Carlo Cenedese

Student assistants: Joudi Hajar, Aristomenis Sfetsos

	When	Where	Video link
Lectures:	Mondays, 16:00-18:00	HG D1.1 (also streamed online)	link
Tutorials:	Fridays, 10:00-12:00	HG D1.1 (also streamed online)	link

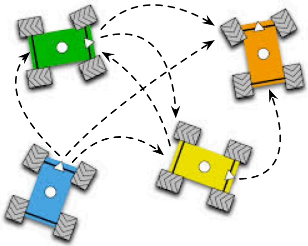
Grading: based on 3 homework assignments (50%) and a final project (50%). Here is the information on grading, homework, and the final project.

- lecture videos are protected with login `doe-22s` and password `6srK7L2`

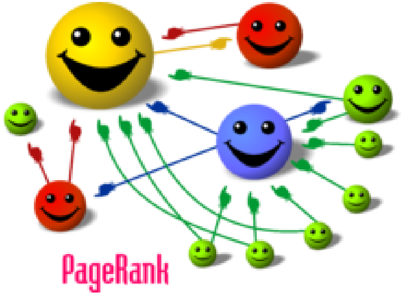
recap of motivating examples

(last lecture, exercise session, & chapter 1)

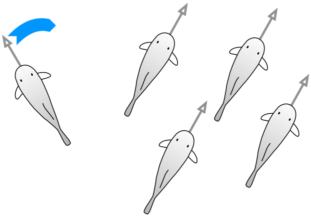
Motivating examples from last lecture & exercise



cooperative control of mobile robots



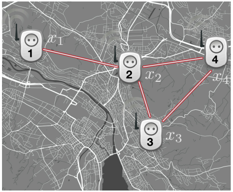
PageRank algorithm



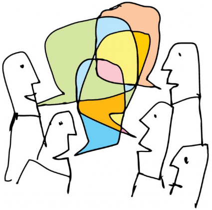
flocking

many more to come

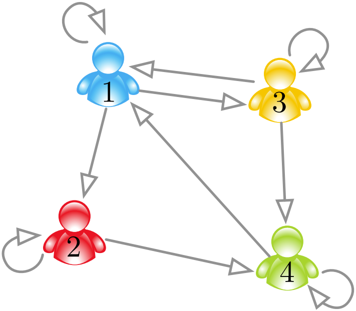
Motivating examples . . . all led to averaging algorithms



distributed sensor networks



social opinion dynamics

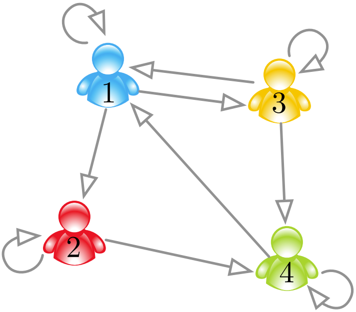


Distributed averaging algorithm:

$$\begin{bmatrix} x_1^+ \\ x_2^+ \\ x_3^+ \\ x_4^+ \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 & \\ & 1/2 & & 1/2 \\ 1/3 & & 1/3 & 1/3 \\ 1/2 & & & 1/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Convergence? Convergence to avg?

Problem posed more formally



Distributed averaging algorithm:

Consider the iteration

$$x^+ = Ax$$

Properties:

- sparsity pattern induced by graph
- nonnegative coefficients: $a_{ij} \geq 0$
- unit row sums: $\sum_j a_{ij} = 1$

Questions:

- is such an iteration convergent?
- if so, what is it converging to?
- can we design A so that we converge to the average?
- how are these properties related to the underlying graph?

Today: matrix theory approach **Next week:** graph theory approach

review of discrete-time linear dynamic systems

(last exercise session & chapter 2.1)

Solutions of linear discrete-time systems

Definition: A square matrix A defines a *discrete-time linear system* by

$$x(k+1) = Ax(k), \quad x(0) = x_0.$$

The sequence $x(k) = A^k x_0$ is called the *solution* or *evolution* of the system.

Modal expansion of the solution for a *diagonalizable* matrix:

$$x(k) = (w_1^T x_0) \cdot \lambda_1^k v_1 + \cdots + (w_n^T x_0) \cdot \lambda_n^k v_n$$

(λ_i, v_i, w_i) are eigenvalues & eigenvectors: $Av_i = \lambda_i v_i$ & $w_i^T A = \lambda_i w_i^T$

$\Rightarrow$ **matrix powers**: $A^k = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1^k & & \\ & \cdots & \\ & & \lambda_n^k \end{bmatrix} \begin{bmatrix} w_1^T \\ \vdots \\ w_n^T \end{bmatrix}$

The general *non-diagonalizable* case

Eigenvalue multiplicities:

- *algebraic multiplicity* = multiplicity as a root of $\det(\lambda I - A) = 0$
- *geometric multiplicity* = number of linearly-independent eigenvectors

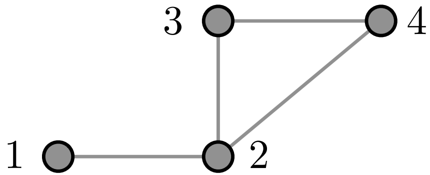
Eigenvalue simplicities: an eigenvalue is

- *simple* if it has algebraic and geometric multiplicity equal to 1
- *semi-simple* if algebraic and geometric multiplicity coincide

$\Rightarrow$ **matrix powers** for a matrix with non-semisimple eigenvalues:

$$A = T \begin{bmatrix} \ddots & & \\ \cdots & \boxed{\begin{matrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \lambda_i & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{matrix}} & \cdots \\ & & \ddots \end{bmatrix} T^{-1} \Rightarrow A^k = T \begin{bmatrix} \ddots & & \\ \cdots & \boxed{\begin{matrix} \lambda_i^k & k\lambda_i^{k-1} & \cdots & \frac{k!\lambda_i^{k-n_i+1}}{(k-n_i+1)!(n_i-1)!} \\ 0 & \lambda_i^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & k\lambda_i^{k-1} \\ 0 & \cdots & 0 & \lambda_i^k \end{matrix}} & \cdots \\ & & \ddots \end{bmatrix} T^{-1}$$

Revisiting the wireless sensor network example



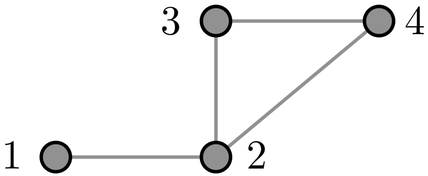
$$x^+ = \underbrace{\begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}}_{=A} x$$

Jordan form:

$$J = T^{-1}AT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{24}(5-\sqrt{73}) & 0 \\ 0 & 0 & 0 & \frac{1}{24}(5+\sqrt{73}) \end{bmatrix}$$

Spectrum: $\mathrm{spec}(= A) = \{1\ ,\ 0\ ,\ \underbrace{1/24 \cdot (5-\sqrt{73})}_{\approx -0.14}\ ,\ \underbrace{1/24 \cdot (5+\sqrt{73})}_{\approx 0.56}\}$

$$\Rightarrow \lim_{k\to\infty} x(k) = \lim_{k\to\infty} A^k x_0 = \lim_{k\to\infty} TJ^kT^{-1}x_0 = T \left[\begin{smallmatrix} 1 & 0 \\ 0 & \mathbb{0}_{3\times 3} \end{smallmatrix}\right] T^{-1}x_0$$



$$\lim_{k\to\infty} x(k) = T \begin{bmatrix} 1 & 0 \\ 0 & \mathbb{0}_{3\times 3} \end{bmatrix} T^{-1} x_0$$

$$= T_1 \cdot 1 \cdot T_1^{-1} x_0$$

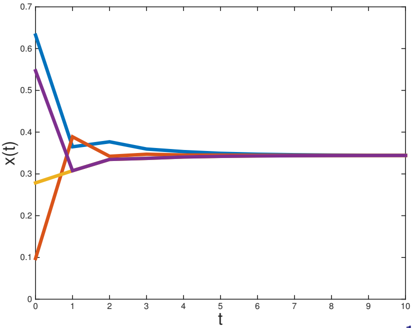
where $T_1 = T\mathbf{e}_1 = [1\ 1\ 1\ 1]^\top = \mathbb{1}$

and $T_1^{-1} = T^{-1}\mathbf{e}_1 = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}^\top$

are the right and left eigenvectors associated to the eigenvalue 1

$\Rightarrow$ asymptotic consensus (but not quite to the average):

$$\lim_{k\to\infty} x(k) = \underbrace{\begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}^\top x_0}_{\text{convex combination of initial values}} \ \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{\textit{consensus}}$$



Convergence of linear discrete-time systems

Definition (Convergent matrix). *The matrix A is* convergent *if* $\lim_{\ell\to+\infty} A^\ell$ *exists and it is equal to the zero matrix, that is:*

$$\lim_{\ell\to+\infty} A^\ell = \mathbb{0}_{n\times n}.$$

Definition (Spectrum and spectral radius of a matrix). *Given a square matrix A,*

(i) *the spectrum of A, denoted* $\mathrm{spec}(A)$*, is the set of eigenvalues of A; and*

(ii) *the spectral radius of A is radius of the smallest disk centered at the origin containing the spectrum of A:*

$$\rho(A) = \max\{\|\lambda\|_{\mathbb{C}} \mid \lambda \in \mathrm{spec}(A)\}.$$

Theorem (Convergence and spectral radius). *For a square matrix A, the following two statements are equivalent:*

(i) *A is convergent; and*

(ii) $\rho(A) < 1$.

Averaging matrices are not convergent due to an eigenvalue at 1: $A\mathbb{1} = \mathbb{1}$

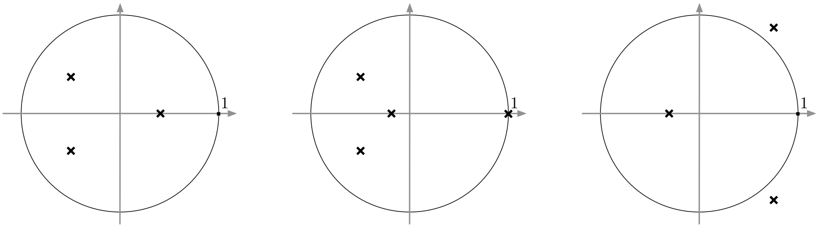
Semi-convergence of linear discrete-time systems

Definition (Semi-convergent matrix). *The matrix A is* semi-convergent *if* $\lim_{\ell\to+\infty} A^{\ell}$ *exists.*

Theorem (Semi-convergence and spectral radius). *For a square matrix A, the following two statements are equivalent:*

(i) *A is* semi-*convergent; and*

(ii) $\rho(A) \le 1$ *; and the only possible unit-norm eigenvalue is the number 1 and it must be semi-simple.*



(a) The spectrum of a convergent matrix

(b) The spectrum of a semiconvergent matrix, provided the eigenvalue 1 is semisimple.

(c) The spectrum of a matrix that is not semiconvergent.

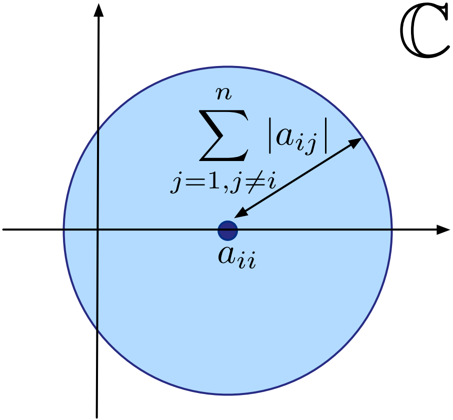
row-stochastic matrices

(Chapter 2.2)

Geršgorin's disk theorem

Theorem (Geršgorin disks). *For any square matrix $A \in \mathbb{R}^{n \times n}$,*

$$\operatorname{spec}(A) \subset \bigcup_{i \in \{1,\dots,n\}} \underbrace{\left\{ z \in \mathbb{C} \mid \|z - a_{ii}\|_{\mathbb{C}} \leq \sum_{j=1, j \neq i}^{n} |a_{ij}| \right\}}_{\textit{disk centered at } a_{ii} \textit{ in the complex plane with radius } \sum_{j=1,j\neq i}^{n} |a_{ij}|} .$$



Proof of Geršgorin's disk theorem

Theorem (Geršgorin disks). *For any square matrix* $A \in \mathbb{R}^{n\times n}$*,*

$$\operatorname{spec}(A) \subset \bigcup_{i\in\{1,\ldots,n\}} \underbrace{\left\{ z \in \mathbb{C} \;\middle|\; \|z - a_{ii}\|_{\mathbb{C}} \leq \sum_{j=1, j\neq i}^{n} |a_{ij}| \right\}}_{\textit{disk centered at } a_{ii} \textit{ in the complex plane with radius } \sum_{j=1,j\neq i}^{n} |a_{ij}|} .$$

Jacobi relaxation in parallel computation (see also E2.16 & E2.17)

$$\underbrace{\begin{bmatrix} \star & \star & & \star \\ \star & \star & & \\ \star & & \star & \star \\ & \star & & \star \end{bmatrix}}_{=A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

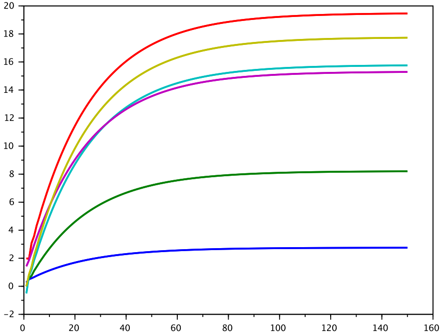
1. **setup:** n distributed processors want to solve the equation $Ax = b$
2. **assume:** $A \in \mathbb{R}^{n\times n}$ is invertible & diagonal elements a_{ii} are nonzero
3. **info:** processor i knows b_i & $a_{ij} \neq 0$ for row-neighbors $j \in \{1, \dots, n\}$
4. **initialization:** processor i has an initial guess $x_j(0)$ of its neighbors
5. **update** at time $k \in \mathbb{N}$: processor i stores a variable $x_i(k)$ & computes

$$x_i(k+1) = \frac{1}{a_{ii}}\Big(b_i - \sum\nolimits_{j=1, j\neq i}^{n} a_{ij}x_j(k)\Big)$$

6. **communicate:** i sends $x_i(k+1)$ to column-neighbors with $a_{ji} \neq 0$

Example for Jacobi relaxation

$$\underbrace{\begin{bmatrix} 3.7 & 0 & 0 & 0 & -0.6 & 0 \\ 0 & 4.5 & -0.8 & 0 & -0.7 & -0.6 \\ -1 & 0 & 1.5 & -0.7 & 0 & -0.7 \\ -0.7 & -0.9 & -0.8 & 2 & -0.5 & 0 \\ 0 & 0 & -0.8 & -0.8 & 2.8 & -0.6 \\ 0 & 0 & -0.8 & -0.7 & 0 & 1.5 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ 3 \\ -1 \\ 4 \\ 0 \end{bmatrix}}_{b}$$



$$x(k \to \infty) = \begin{bmatrix} 2.7536672 \\ 8.2157827 \\ 19.491865 \\ 15.786202 \\ 15.314281 \\ 17.762556 \end{bmatrix}$$

$$= A^{-1} b \quad \checkmark$$

Convergence analysis via Geršgorin's theorem

- **error variable:** $y(k) = x(k) - x^\star$ where $x^\star$ solves $Ax^\star = b$
- **error dynamics:** $y_i(k+1) = -\sum_{j=1, j\neq i}^{n} \frac{a_{ij}}{a_{ii}} y_j(k)$

$\Rightarrow$ matrix form: $y(k+1) = \tilde{A} y(k)$ where $\tilde{A}_{ij} = \begin{cases} 0 & \text{for } i = j \\ -a_{ij}/a_{ii} & \text{else} \end{cases}$

- **Geršgorin's disk theorem:**

$$\operatorname{spec}(\tilde{A}) \subset \bigcup_{i\in\{1,\dots,n\}} \underbrace{\left\{ z \in \mathbb{C} \;\middle|\; \|z\|_{\mathbb{C}} \leq \sum_{j=1, j\neq i}^{n} |a_{ij}/a_{ii}| \right\}}_{\text{disk centered at 0 in the complex plane with radius } \sum_{j=1, j\neq i}^{n} |a_{ij}/a_{ii}|}$$

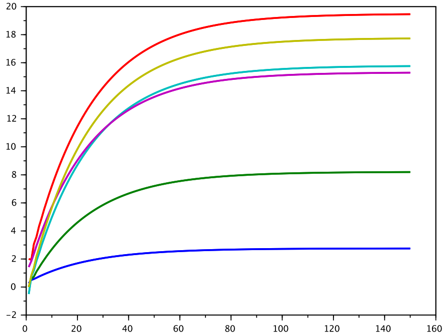
- assume A to be **strictly diagonal dominant:** $|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|$

$\Rightarrow$ Geršgorin disks of $\tilde{A}$ strictly inside the unit disk: $\sum_{j=1, j\neq i}^{n} |a_{ij}/a_{ii}| < 1$

$\Rightarrow$ **convergence** of error dynamics: $\rho(\tilde{A}) < 1$

Example for Jacobi relaxation re-visited

$$\underbrace{\begin{bmatrix} 3.7 & 0 & 0 & 0 & -0.6 & 0 \\ 0 & 4.5 & -0.8 & 0 & -0.7 & -0.6 \\ -1 & 0 & 1.5 & -0.7 & 0 & -0.7 \\ -0.7 & -0.9 & -0.8 & 2 & -0.5 & 0 \\ 0 & 0 & -0.8 & -0.8 & 2.8 & -0.6 \\ 0 & 0 & -0.8 & -0.7 & 0 & 1.5 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ 3 \\ -1 \\ 4 \\ 0 \end{bmatrix}}_{b}$$



- $\boxed{\rho(\tilde{A}) = 0.9584448}$

  $\Longrightarrow$ convergence ✓

  (but no diagonal dominance)

$\Rightarrow$ we got lucky !

- what if $\rho(\tilde{A}) > 1$?

$\Rightarrow$ Jacobi over-relaxation (E2.17)

Stochastic matrices & their spectral properties

Definition: The square matrix $A \in \mathbb{R}^{n \times n}$ is

1. *non-negative* if $a_{ij} \geq 0$ for all $i, j \in \{1, \dots, n\}$;
2. *positive* if $a_{ij} > 0$ for all $i, j \in \{1, \dots, n\}$;
3. *row-stochastic* if non-negative & $A\mathbb{1} = \mathbb{1}$;
4. *column-stochastic* if non-negative & $A^\top \mathbb{1} = \mathbb{1}$; and
5. *doubly-stochastic* if it is row- column-stochastic.

Lemma (Spectral properties):
For a row-stochastic matrix A,

1. 1 is an eigenvalue,
2. and $\rho(A) = 1$.

Perron-Frobenius theory

(Chapter 2.3)

Example: DeGroot model in opinion dynamics

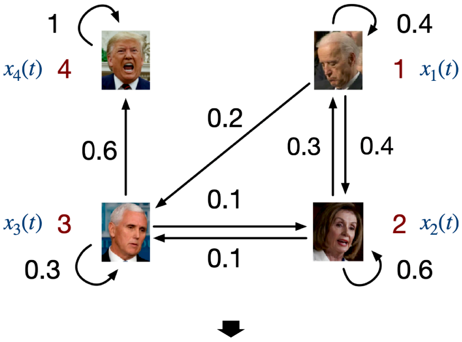
- n individuals $i \in \{1, \dots, n\}$
- individual opinions $x_i(k) \in \mathbb{R}$
- weighted averaging opinion update

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

where $a_{ij} \geq 0$ signifies how much i trusts j / how much j influences i

convex coeff: $a_{ij} \geq 0$ & $\sum_j a_{ij} = 1$

- compactly: $x(k+1) = Ax$, where A is row-stochastic: $A \geq 0$ & $A\mathbb{1} = \mathbb{1}$



$$A = \begin{bmatrix} 0.4 & 0.4 & 0.2 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0.1 & 0.3 & 0.6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

asymptotic behavior: $\lim_{k\to\infty} x(k) = \lim_{k\to\infty} A^k x_0 = ???$

ETH EduApp examples — convergent ?

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$A_4 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$$

$$A_5 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Primitivity & Irreducibility

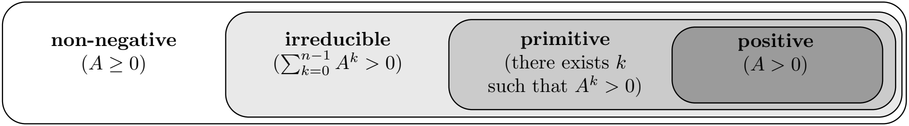
Definition: For $n \geq 2$, an $n \times n$ non-negative matrix A is

1. *irreducible* if $\sum_{k=0}^{n-1} A^k$ is positive,
2. *primitive* if there exists $k \in \mathbb{N}$ such that A^k positive.

A matrix that is not irreducible is said to be *reducible*.

Facts: For a $n \times n$ non-negative matrix A (with $n \geq 2$),

1. $A^k > 0 \implies A^\ell > 0$ for all $\ell \geq k$
2. A primitive $\implies$ A irreducibile



EduApp examples — revisited

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \lim_{k\to\infty} A_1^k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$ **reducible**

$$A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \lim_{k\to\infty} A_2^k = \text{not existent (oscillating)}$$ **irreducible**

$$A_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \lim_{k\to\infty} A_3^k = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$ **reducible**

$$A_4 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix} \qquad \lim_{k\to\infty} A_4^k = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$ **irreducible & primitive**

$$A_5 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \lim_{k\to\infty} A_5^k = \text{not existent (diverging)}$$ **reducible**

Perron-Frobenius theorem

non-negative ($A \geq 0$)

irreducible ($\sum_{k=0}^{n-1} A^k > 0$)

primitive (there exists k such that $A^k > 0$)

positive ($A > 0$)

if A **is non-negative**

1. there exists an eigenvalue $\lambda \geq |\mu| \geq 0$ for all other eigenvalues μ
2. the associated right and left eigenvectors $v \geq 0$ and $w \geq 0$

if A **is irreducible**

3. $\lambda > 0$ and λ is simple
4. $v > 0$ and $w > 0$ are unique

if A **is primitive**

5. $\lambda > |\mu|$ for all other eigenvalues μ

q: what is Perron-Frobenius useful for?

a: convergence of dynamical systems

(chapter 2.3.3)

Perron-Frobenius & semi-convergence

Theorem (Semi-convergence) For $A \geq 0$ with dominant simple eigenvalue $\rho(A)$ & right/left eigenvectors $v \geq 0$ & $w \geq 0$ normalized to $v^\top w = 1$:

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

if A **is non-negative**

1. there exists an eigenvalue $\lambda \geq |\mu| \geq 0$ for all other eigenvalues μ
2. the associated right and left eigenvectors $v \geq 0$ and $w \geq 0$

if A **is irreducible**

3. $\lambda > 0$ and λ is simple
4. $v > 0$ and $w > 0$ are unique

if A **is primitive**

5. $\lambda > |\mu|$ for all other eigenvalues μ
6. $\lim_{k\to\infty} A^k/\lambda^k = vw^T$, with normalization $v^T w = 1$

EduApp example revisited

$$A_4 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$$

- **primitivity**: $A_4^2 = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} > 0$

- **Perron-Frobenius:** $\operatorname{spec}(A_4) = \{1, -\frac{1}{2}\}$

  $\Rightarrow \rho(A_4) = 1 > 0$ is dominant eigenvalue

  $\Rightarrow$ associated dominant right/left eigenvectors $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ & $w = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

- **semi-convergence**:

$$A_4 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix} \quad \lim_{k\to\infty} A_4^k = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \underbrace{1}_{=\rho(A_4)} \cdot \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{=v} \cdot \underbrace{\begin{bmatrix} 2 & 1 \end{bmatrix}}_{=w^T} \cdot \underbrace{\frac{1}{3}}_{1/v^T w}$$

Our first result on averaging systems

Recall for A row-stochastic: $\rho(A) = 1$ is eigenvalue with eigenvector $\mathbb{1}$

Corollary (Consensus for primitive row-stochastic). *For a primitive row-stochastic matrix A,*

(i) the simple eigenvalue $\rho(A) = 1$ is strictly larger than the magnitude of all other eigenvalues, hence A is semi-convergent;

(ii) $\lim_{k\to\infty} A^k = \mathbb{1}_n w^\top$, where w is the left positive eigenvector of A with eigenvalue 1 satisfying $w_1 + \dots + w_n = 1$;

(iii) the solution to $x(k+1) = Ax(k)$ satisfies

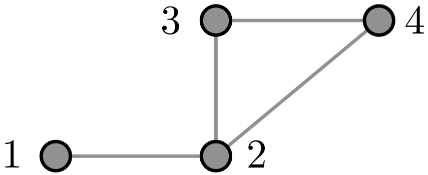
$$\lim_{k\to\infty} x(k) = \big(w^\top x(0)\big)\mathbb{1}_n;$$

(iv) if additionally A is doubly-stochastic, then $w = \frac{1}{n}\mathbb{1}_n$ (because $A^\top \mathbb{1}_n = \mathbb{1}_n$ and $\frac{1}{n}\mathbb{1}_n^\top \mathbb{1}_n = 1$) so that

$$\lim_{k\to\infty} x(k) = \frac{\mathbb{1}_n^\top x(0)}{n}\mathbb{1}_n = \operatorname{average}\big(x(0)\big)\mathbb{1}_n.$$

In this case we say that the dynamical system achieves average consensus.

Revisiting the wireless sensor network example



$$x^+ = \underbrace{\begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}}_{A} x$$

A is **primitive**:

$$A^2 = \begin{bmatrix} 3/8 & 3/8 & 1/8 & 1/8 \\ 3/16 & 17/48 & 11/48 & 11/48 \\ 1/12 & 11/36 & 11/36 & 11/36 \\ 1/12 & 11/36 & 11/36 & 11/36 \end{bmatrix} > 0$$

P-F-Theorem: $\rho(A) = 1$ with $v = \mathbb{1}$ & $w = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}^T$

$\Rightarrow \lim_{k\to\infty} x(k) = \mathbb{1}\, w^T x_0 = \mathbb{1}\left(\frac{1}{6}x_{10} + \frac{1}{3}x_{20} + \frac{1}{4}x_{30} + \frac{1}{4}x_{40}\right)$

Revisiting the DeGroot model in opinion dynamics

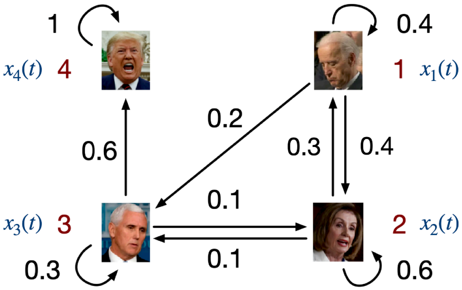
- n individuals $i \in \{1, \dots, n\}$
- individual opinions $x_i(k) \in \mathbb{R}$
- weighted averaging opinion update

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

where $a_{ij} \geq 0$ signifies how much i trusts j / how much j influences i

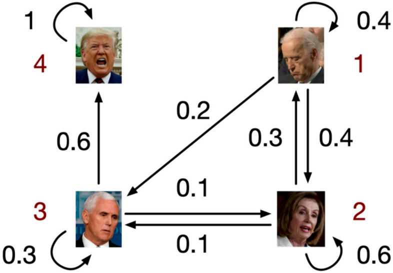
convex coeff: $a_{ij} \geq 0$ & $\sum_j a_{ij} = 1$

- compactly: $x(k+1) = Ax$, where A is row-stochastic: $A \geq 0$ & $A\mathbb{1} = \mathbb{1}$



$$A = \begin{bmatrix} 0.4 & 0.4 & 0.2 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0.1 & 0.3 & 0.6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

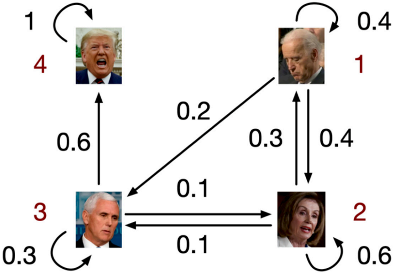
asymptotic behavior: $\lim_{k\to\infty} x(k) = \lim_{k\to\infty} A^k x_0 = ???$



$$\Rightarrow \quad A = \begin{bmatrix} 0.4 & 0.4 & 0.2 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0.1 & 0.3 & 0.6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \quad I + A + A^2 + A^3 = \begin{bmatrix} 1.92 & 1.20 & 0.53 & 0.35 \\ 0.87 & 2.52 & 0.40 & 0.21 \\ 0.07 & 0.27 & 1.45 & 2.22 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

$\Rightarrow$ **reducible & not primitive !**



$$\Rightarrow \quad A = \begin{bmatrix} 0.4 & 0.4 & 0.2 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0.1 & 0.3 & 0.6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- A row-stochastic, reducible, & **spectrum**$(A) = \{1, 0.885, 0.215, 0.2\}$
- $\rho(A) = 1$ is **dominant eigenvalue** with $v = \mathbb{1}$ & $w = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$
- despited lack of reducibility $\lim_{k\to\infty} x(k)$
  $= \lim_{k\to\infty} A^k x_0 = \lim_{k\to\infty} \frac{A^k}{\rho(A)^k} x_0$
  $= v w^T x_0 = x_4(0) \cdot \mathbb{1}$

More cliffhangers

Example 1: $\begin{bmatrix} 0 & \star \\ \star & 0 \end{bmatrix}$ is irreducible & not primitive, but $\begin{bmatrix} \star & \star \\ \star & 0 \end{bmatrix}$ is primitive

Q1: what makes irreducible matrices primitive?

Example 2: $\begin{bmatrix} x_1^+ \\ x_2^+ \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ with reducible & non-primitive matrix

Nevertheless, averaging converges: $x_1(k \to \infty) = x_2(k \to \infty) = x_1(0)$

Q2: what are the necessary and sufficient convergence conditions?

Q3: how does all of this relate to underlying graph structure?

Answers: will be revealed in next lecture(s) after studying graph theory

Reading assignment (lecture notes):

Chapter 2: Elements of Matrix Theory

Exercise session (Friday):

- led by Andrea Martinelli
- review of take-home messages
- examples & additional facts
- exercises & illustrations