Lecture:Perrorn-Frobenius Theorem

Shaoshuai Mou

PURDUE
UNIVERSITY

Review

Lemma (Convergence):

Suppose $A \in \mathbb{R}^{n \times n}$ is such that

- 1 is a simple eigenvalue
- 1 is the largest eigenvalue in magnitude.
- all the other eigenvalues are with magnitude strictly less than 1

Then $A^t \to vw'$ as fast as $|\lambda_2|^t \to 0$

where v, w are right and left eigenvectors of *A* corresponding to 1 and $w'v = 1$

λ_2 denotes the 2nd largest eigenvalue of A in magnitude.

Convergence to Consensus $\lim_{t \to \infty} A^t = \mathbf{1}w'$

Eigenvalue/Eigenvector Properties *of A*

$A\mathbf{1} = \mathbf{1}$

1 is the largest eigenvalue in magnitude. $|\lambda(A)| \le 1$

1 is a simple eigenvalue

all the other eigenvalues are with magnitude strictly less than 1

Gershgorin Circle Theorem

A is row stochastic

Analysis on **Non-Negative Matrices**

Perron-Frobenius Theorem

Any $A \in \mathbb{R}^{n\times n}$ with $n \geq 2$ is a

$\geq$, $>$: entriwisely non-negative or strictly positive.

Non-negative Matrix if $A \geq 0$; $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Irreducible Matrix if for each pair of i and j, there exists a positive integer *k* such that $(A^k)ij > 0$

$$A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad A_2^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Primitive Matrix if there exists a positive integer *k* such that $A^k > 0$

$$A_3 = \begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix} \qquad A_3^2 = \begin{bmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{bmatrix}$$

Positive Matrix if $A > 0$.

$$A_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

- The state matrix *A* in the four-agent consensus example is primitive. (*Verify by Matlab*)

Consensus

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

Consensus for global average

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

Consensus for convex combination

$$A = \begin{bmatrix} \frac{0.39}{0.44} & \frac{0.05}{0.44} & 0 & 0 \\ \frac{0.05}{0.2} & \frac{0.06}{0.2} & \frac{0.05}{0.2} & \frac{0.04}{0.2} \\ 0 & \frac{0.05}{0.24} & \frac{0.15}{0.24} & \frac{0.04}{0.24} \\ 0 & \frac{0.04}{0.12} & \frac{0.04}{0.12} & \frac{0.04}{0.12} \end{bmatrix}$$

Perron-Frobenius Theorem

proved by Oskar Perron (1907) and Georg Frobenius (1912)

The most important theorem for convergence of **non-negative matrices**!

For any $A \in \mathbb{R}^{n \times n}$ with $n \geq 2$ with spectral radius denoted by ρ

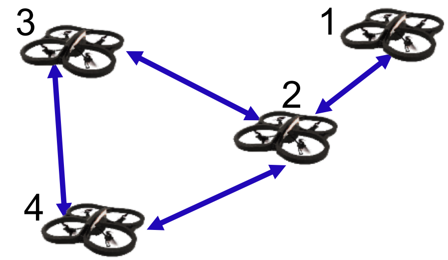
- If A is nonnegative, then ρ is one eigenvalue,
  whose eigenvector can be selected nonnegative;
- If A is irreducible, then the eigenvalue $\rho > 0$, and is simple,
  whose eigenvector is unique and positive, up to scaling;
- If A is primitive, all other eigenvalues are with magnitude strictly less than ρ.

Simple: Single Jordan block of size 1

ρ is also called Perron-Frobenius Eigenvalue

- Non-negative matrix: $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
  $\{1, 1\}$
- Irreducible matrix: $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
  $\{1, -1\}$
- Primitive matrix: $A_3 = \begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix}$
  $\{1, -1/2\}$
- Positive matrix: $A_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
  $\{2, 0\}$

Application of PF Theorem: Consensus Algorithms



For examples of consensus algorithms, one has

$$x(t+1) = Ax(t)$$

Consensus

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

Consensus for global average

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

Consensus for convex combination

$$A = \begin{bmatrix} \frac{0.39}{0.44} & \frac{0.05}{0.44} & 0 & 0 \\ \frac{0.05}{0.2} & \frac{0.06}{0.2} & \frac{0.05}{0.2} & \frac{0.04}{0.2} \\ 0 & \frac{0.05}{0.24} & \frac{0.15}{0.24} & \frac{0.04}{0.24} \\ 0 & \frac{0.04}{0.12} & \frac{0.04}{0.12} & \frac{0.04}{0.12} \end{bmatrix}$$

where A is **row stochastic** and **primitive**

$$A\mathbf{1} = \mathbf{1} \qquad |\lambda(A)| \leq 1$$

1 is a simple eigenvalue

all the other eigenvalues are with magnitude strictly less than 1

$$\lim_{t \to \infty} A^t = \mathbf{1}w' \quad w'A = w' \quad w'\mathbf{1} = 1$$

$$\lim_{t \to \infty} x(t) = \mathbf{1}w'x(0) = \begin{bmatrix} w'x(0) \\ w'x(0) \\ \vdots \\ w'x(0) \end{bmatrix}$$

$$\lim_{t\to\infty} x(t) = \mathbf{1}w'x(0) \qquad w'A = w' \qquad w'\mathbf{1} = 1$$

- For just a consensus, we do not care about the consensus value and *w* is unknown.

- For consensus to the global average $\frac{1}{n}\mathbf{1}'x(0)$, one wants $w = \frac{1}{n}\mathbf{1}$ *Is this true?*

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

$$w_{ij} = \begin{cases} \min\{\frac{1}{d_i}, \frac{1}{d_j}\} & j \in \mathcal{N}_i,\ j \neq i; \\ 1 - \sum_{j\in\mathcal{N}_i,\ j\neq i} w_{ij} & j = i \end{cases}$$

$$\left.\begin{array}{l} A = A' \\ A\mathbf{1} = \mathbf{1} \end{array}\right\} \frac{1}{n}\mathbf{1}'A = \frac{1}{n}\mathbf{1}'A$$

$$w_{ij} = w_{ji}$$

$$\gamma'\mathbf{1} = \mathbf{1}$$

- For consensus to a specific convex combination $\gamma'x(0)$, one wants $w = \gamma$

$$A = \begin{bmatrix} \frac{0.39}{0.44} & \frac{0.05}{0.44} & 0 & 0 \\ \frac{0.05}{0.2} & \frac{0.06}{0.2} & \frac{0.05}{0.2} & \frac{0.04}{0.2} \\ 0 & \frac{0.05}{0.24} & \frac{0.15}{0.24} & \frac{0.04}{0.24} \\ 0 & \frac{0.04}{0.12} & \frac{0.04}{0.12} & \frac{0.04}{0.12} \end{bmatrix} \quad \gamma = \begin{bmatrix} 0.44 \\ 0.2 \\ 0.24 \\ 0.12 \end{bmatrix}$$

$$w_{ij} = \begin{cases} \frac{1}{\gamma_i}\min\{\frac{\gamma_i}{d_i}, \frac{\gamma_j}{d_j}\} & j \in \mathcal{N}_i,\ j \neq i; \\ 1 - \sum_{j\in\mathcal{N}_i,\ j\neq i} w_{ij} & j = i \end{cases}$$

$$\gamma'A = \gamma'$$

$$\gamma'A = \gamma'$$

row sum is 1

$$\sum_{j=1}^{n} w_{ij} = 1 \Rightarrow \gamma_i \sum_{j=1}^{n} w_{ij} = \gamma_i$$

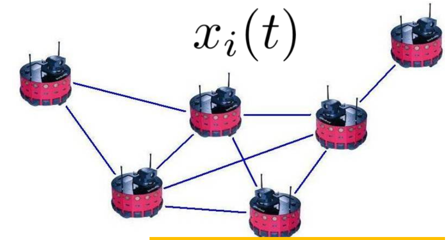
$$\left.\begin{array}{l} \gamma_i w_{ij} = \gamma_j w_{ji} \\ \gamma_i \sum_{j=1}^{n} w_{ij} = \gamma_i \end{array}\right\} \sum_{j=1}^{n} \gamma_j w_{ij} = \gamma_i \Rightarrow$$

The *i*th column sum is γ_i

Distributed Algorithm for Consensus

$$x_i(t+1) = \sum_{j \in \mathcal{N}_i} w_{ij} x_j(t)$$

$$w_{ij} = \begin{cases} > 0, & j \in \mathcal{N}_i \\ 0, & \text{otherwise.} \end{cases} \qquad \sum_{j=1}^{m} w_{ij} = 1$$



Network connectivity

$$\boxed{x(t+1) = Ax(t)}$$

Graph Theory

A is row stochastic

Gershgorin Circle Theorem

$$A\mathbf{1} = \mathbf{1}$$

1 is the largest eigenvalue in magnitude.

?

If A is also Primitive

Perron - Frobenius Theorem

1 is a simple eigenvalue

all the other eigenvalues are with magnitude strictly less than 1

$$\lim_{t \to \infty} A^t = \mathbf{1}w'$$

$$\lim_{t \to \infty} x(t) = \mathbf{1}w'x(0)$$

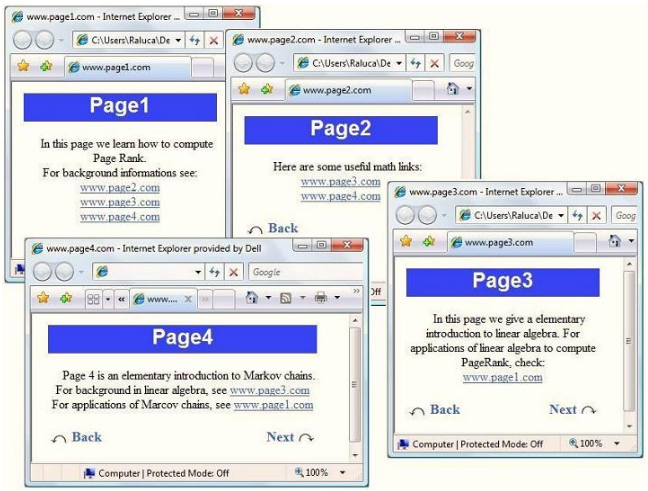
Application of PF Theorem into **PageRank** Algorithm

- **Old Method:** Text-based Ranking (1990s)

- **Google search**: relevance-based ranking (Larry Page & Sergey Brin, 1998, Stanford)

  - Key Idea: A webpage is important (relevant) if other pages *think* it is.

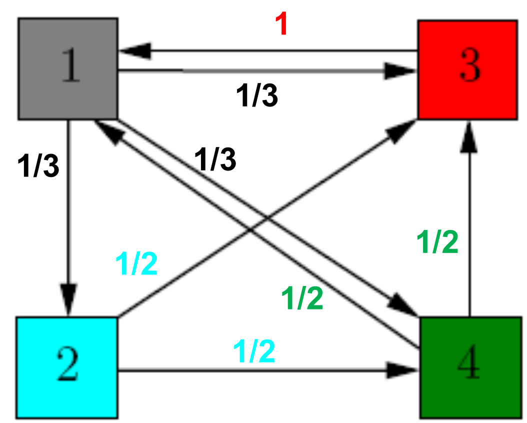
Webpage *i* thinks webpage *j* is important if webpage *i* contains a **hyperlink** to webpage *j*



a hyperlink

$$i \to j$$

a directed edge



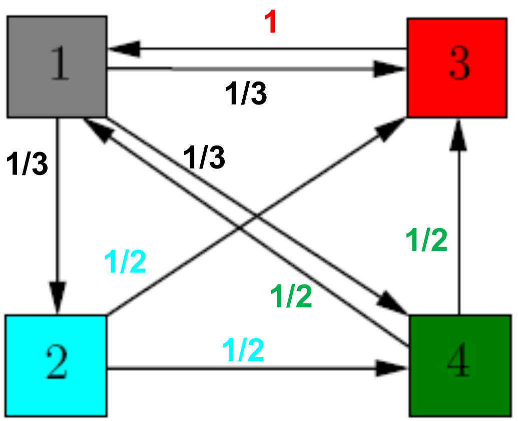
- Suppose each page transfer its relevance (importance) **evenly** to its outgoing neighbors.

Page 1 thinks Page 2, 3, 4 are important (but no preference of one over another)

Page 1 distributes its importance $x_1 \in \mathbb{R}$ evenly to Page 2, 3, 4.

[1]. S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. 1998
[2]. M. Franceschet. Pagerank: Standing on the shoulders of giants. 2011
[3]. K. Bryan and T. Leise. The $25,000,000,000 eigenvector: The linear algebra beyond google. SIAM Review, 2006
[4]. A. N. Langville and C. D. Meyer. Google's Pagerank and beyond: The science of search engine rankings. 2011

$x_i(k)$: the relevance/importance of Page i.



- Suppose each page transfer its relevance (importance) **evenly** to its outgoing neighbors.
- Each page updates its relevance (importance) based on transferred importance from its incoming neighbors.

$$x_1(k+1) = 1 \cdot x_3(k) + \frac{1}{2} \cdot x_4(k)$$

$$x_2(k+1) = \frac{1}{3} \cdot x_1(k)$$

compact form??

$$x_3(k+1) = \frac{1}{3}x_1(k) + \frac{1}{2}x_2(k) + \frac{1}{2}x_4(k)$$

$$x_4(k+1) = \frac{1}{3}x_1(k) + \frac{1}{2}x_2(k)$$

$$x(k+1) = Ax(k)$$

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

What is special about A?

- **A is column stochastic,** $\mathbf{1}'A = \mathbf{1}'$

  1 is the largest eigenvalue in magnitude.

  1 is a simple eigenvalue

- **A is primitive** (verify this by Matlab)

  all the other eigenvalues are with magnitude strictly less than 1

A is column stochastic, primitive

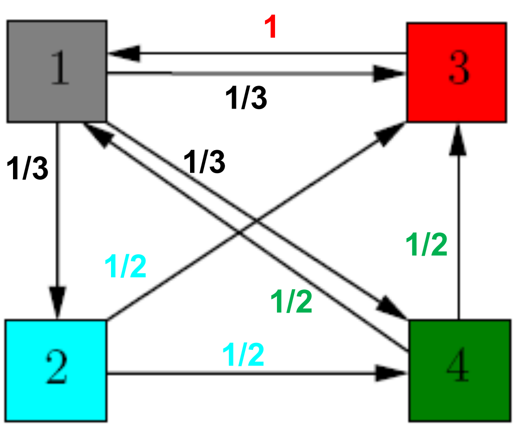
$$\lim_{t\to\infty} A^t = v\mathbf{1}' \qquad Av = v, \mathbf{1}'A = \mathbf{1}, \mathbf{1}'v = 1 \qquad v = \frac{1}{31}\begin{bmatrix}12\\4\\9\\6\end{bmatrix}$$

$$\lim_{t\to\infty} x(t) = v\,\mathbf{1}'x(0)$$

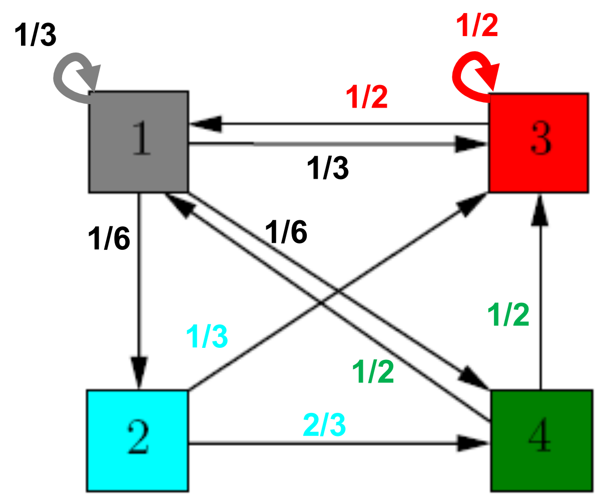
scalar

❖ The page rank is achieved by A's **right eigenvector** corresponding to 1

- Suppose each page transfer its relevance (importance) **evenly** to its outgoing neighbors.



- More generally, each page could assign **different weights** of relevance transfer to different outgoing neighbors. Or could save some weights for itself by adding self-arcs.



Research Topic: Distributed Algorithms for PageRank??

[1]. H. Ishii and R. Tempo. Distributed randomized algorithms for the pagerank computation, IEEE TAC, 2010
[2]. W. Zhao, H. F. Chen and H. T. Fang. Convergence of distributed randomized pagerank algorithms. IEEE TAC, 2013
[3]. H. Ishii, R. Tempo and E. W. Bai. A web aggregation approach for distributed randomized pagerank algorithms. IEEE TAC 2012
[4]. H. Ishii and R. Tempo. The pagerank problem, multi-agent consensus, and web aggregation: A system and control viewpoint, IEEE Control System Letters, 2012.
[5]. H. Ishii, R. Tempo and E. W. Bai. Pagerank computation via a distributed randomized approach with lossy communication. Systems and Control Letters, 2012.
[6]. J. Lei and H. F. Chen. Distributed randomized pagerank algorithm based on stochastic approximation. IEEE TAC, 2015

Summary

Perron-Frobenius Theorem

The most important theorem for convergence of **non-negative matrices**!

For any $A \in \mathbb{R}^{n\times n}$ with $n \geq 2$ with spectral radius denoted by ρ

- If A is nonnegative, then ρ is one eigenvalue,
  whose eigenvector can be selected nonnegative;
- If A is irreducible, then the eigenvalue $\rho > 0,$ and is simple,
  whose eigenvector is unique and positive, up to scaling;
- If A is primitive, all other eigenvalues are with magnitude strictly less than ρ.

Simple: Single Jordan block of size 1

ρ **is also called Perron-Frobenius Eigenvalue**

- **Paper Reading** (Due on Feb. 19, Wed): J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, C. Yu. Deterministic Gossiping, 99 (9), 2011.
- **Seminar Report for Extra Credit**. (Due on Feb. 19, Wed)