

Lecture 10: Fundamental Theorem of Markov Chains, Page Rank

Rafael Oliveira

University of Waterloo
Cheriton School of Computer Science

rafael.oliveira.teaching@gmail.com

May 26, 2025

Overview

- Main Tools
 - Linear Algebra Background
 - Perron-Frobenius
- Main Applications
 - Fundamental Theorem of Markov Chains
 - Page Rank
- Acknowledgements

Eigenvalues, Eigenvectors and Spectral Radius

- Given a square matrix $A \in \mathbb{R}^{n \times n}$, we say that $\lambda \in \mathbb{C}$ is an *eigenvalue* of A if there is a vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$.

Eigenvalues, Eigenvectors and Spectral Radius

- Given a square matrix $A \in \mathbb{R}^{n \times n}$, we say that $\lambda \in \mathbb{C}$ is an *eigenvalue* of A if there is a vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$.
- The *spectral radius* of a matrix A , denoted $\rho(A)$, is the maximum absolute value of the eigenvalues of A

Eigenvalues, Eigenvectors and Spectral Radius

- Given a square matrix $A \in \mathbb{R}^{n \times n}$, we say that $\lambda \in \mathbb{C}$ is an *eigenvalue* of A if there is a vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$.
- The *spectral radius* of a matrix A , denoted $\rho(A)$, is the maximum absolute value of the eigenvalues of A
- *Gelfand's formula*

$$\rho(A) = \lim_{t \rightarrow \infty} \|A^t\|_F^{1/t}$$

Eigenvalues, Eigenvectors and Spectral Radius

- Given a square matrix $A \in \mathbb{R}^{n \times n}$, we say that $\lambda \in \mathbb{C}$ is an *eigenvalue* of A if there is a vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$.
- The *spectral radius* of a matrix A , denoted $\rho(A)$, is the maximum absolute value of the eigenvalues of A

- Gelfand's formula*

$$\rho(A) = \lim_{t \rightarrow \infty} \|A^t\|_F^{1/t}$$

- Geometric multiplicity*: an eigenvalue λ of A has geometric multiplicity k if the space of eigenvectors of A with eigenvalue λ has dimension k . That is, if dimension of null space of $A - \lambda I$ is k .

Eigenvalues, Eigenvectors and Spectral Radius

- Given a square matrix $A \in \mathbb{R}^{n \times n}$, we say that $\lambda \in \mathbb{C}$ is an *eigenvalue* of A if there is a vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$.
- The *spectral radius* of a matrix A , denoted $\rho(A)$, is the maximum absolute value of the eigenvalues of A

- Gelfand's formula*

$$\rho(A) = \lim_{t \rightarrow \infty} \|A^t\|_F^{1/t}$$

- Geometric multiplicity*: an eigenvalue λ of A has geometric multiplicity k if the space of eigenvectors of A with eigenvalue λ has dimension k . That is, if dimension of null space of $A - \lambda I$ is k .
- Algebraic multiplicity*: an eigenvalue λ of A has algebraic multiplicity k if $(t - \lambda)^k$ is the highest power of $t - \lambda$ dividing $\det(tI - A)$

Eigenvalues, Eigenvectors and Spectral Radius

- Given a square matrix $A \in \mathbb{R}^{n \times n}$, we say that $\lambda \in \mathbb{C}$ is an *eigenvalue* of A if there is a vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$.
- The *spectral radius* of a matrix A , denoted $\rho(A)$, is the maximum absolute value of the eigenvalues of A

- Gelfand's formula*

$$\rho(A) = \lim_{t \rightarrow \infty} \|A^t\|_F^{1/t}$$

- Geometric multiplicity*: an eigenvalue λ of A has geometric multiplicity k if the space of eigenvectors of A with eigenvalue λ has dimension k . That is, if dimension of null space of $A - \lambda I$ is k .
- Algebraic multiplicity*: an eigenvalue λ of A has algebraic multiplicity k if $(t - \lambda)^k$ is the highest power of $t - \lambda$ dividing $\det(tI - A)$
- Example:

Positivity Lemma

Lemma (Positivity Lemma)

If $A \in \mathbb{R}^{n \times n}$ is a positive matrix and $u, v \in \mathbb{R}^n$ are distinct vectors such that $u \geq v$, then $Au > Av$. Moreover, there exists $\varepsilon > 0$ such that $Au > (1 + \varepsilon)Av$.

Positivity Lemma

Lemma (Positivity Lemma)

If $A \in \mathbb{R}^{n \times n}$ is a positive matrix and $u, v \in \mathbb{R}^n$ are distinct vectors such that $u \geq v$, then $Au > Av$. Moreover, there exists $\varepsilon > 0$ such that $Au > (1 + \varepsilon)Av$.

- Note that

$$(A(u - v))_i = \sum_j A_{ij}(u_j - v_j) \geq (\min_{i,j} A_{ij}) \cdot \sum_j (u_j - v_j)$$

Positivity Lemma

Lemma (Positivity Lemma)

If $A \in \mathbb{R}^{n \times n}$ is a positive matrix and $u, v \in \mathbb{R}^n$ are distinct vectors such that $u \geq v$, then $Au > Av$. Moreover, there exists $\varepsilon > 0$ such that $Au > (1 + \varepsilon)Av$.

- Note that

$$(A(u - v))_i = \sum_j A_{ij}(u_j - v_j) \geq (\min_{i,j} A_{ij}) \cdot \sum_j (u_j - v_j)$$

- Since $u_j \geq v_j$ for all j and u, v distinct implies that there is one index k such that $u_k > v_k$, we have

$$\sum_j (u_j - v_j) \geq u_k - v_k > 0$$

Positivity Lemma

Lemma (Positivity Lemma)

If $A \in \mathbb{R}^{n \times n}$ is a positive matrix and $u, v \in \mathbb{R}^n$ are distinct vectors such that $u \geq v$, then $Au > Av$. Moreover, there exists $\varepsilon > 0$ such that $Au > (1 + \varepsilon)Av$.

- Note that

$$(A(u - v))_i = \sum_j A_{ij}(u_j - v_j) \geq (\min_{i,j} A_{ij}) \cdot \sum_j (u_j - v_j)$$

- Since $u_j \geq v_j$ for all j and u, v distinct implies that there is one index k such that $u_k > v_k$, we have

$$\sum_j (u_j - v_j) \geq u_k - v_k > 0$$

- the moreover part just follows from taking small enough ε .

- Main Tools

- Linear Algebra Background
- Perron-Frobenius

- Main Applications

- Fundamental Theorem of Markov Chains
- Page Rank

- Acknowledgements

Perron's Theorem

Theorem (Perron's Theorem)

Let $A \in \mathbb{R}^{n \times n}$ be a positive matrix (i.e., all its entries are positive). Then, the following hold:

- ① $\rho(A)$ is an eigenvalue, and it has a positive eigenvector
- ② $\rho(A)$ is the only eigenvalue in the complex circumference $|\lambda| = \rho(A)$
- ③ $\rho(A)$ has geometric multiplicity 1
- ④ $\rho(A)$ has algebraic multiplicity 1

Perron's Theorem

Theorem (Perron's Theorem)

Let $A \in \mathbb{R}^{n \times n}$ be a positive matrix (i.e., all its entries are positive). Then, the following hold:

- ① $\rho(A)$ is an eigenvalue, and it has a positive eigenvector
 - ② $\rho(A)$ is the only eigenvalue in the complex circumference $|\lambda| = \rho(A)$
 - ③ $\rho(A)$ has geometric multiplicity 1
 - ④ $\rho(A)$ has algebraic multiplicity 1
- By the definition of $\rho(A)$, there is an eigenvalue $\lambda \in \mathbb{C}$ such that $|\lambda| = \rho(A)$. Let v the a corresponding eigenvector.

Perron's Theorem

Theorem (Perron's Theorem)

Let $A \in \mathbb{R}^{n \times n}$ be a positive matrix (i.e., all its entries are positive). Then, the following hold:

- 1 $\rho(A)$ is an eigenvalue, and it has a positive eigenvector
- 2 $\rho(A)$ is the only eigenvalue in the complex circumference $|\lambda| = \rho(A)$
- 3 $\rho(A)$ has geometric multiplicity 1
- 4 $\rho(A)$ has algebraic multiplicity 1

- By the definition of $\rho(A)$, there is an eigenvalue $\lambda \in \mathbb{C}$ such that $|\lambda| = \rho(A)$. Let v be a corresponding eigenvector.
- Let u be the vector defined by $u_i = |v_i|$. Then, we have

$$(Au)_i = \sum_j A_{ij} u_j \geq \left| \sum_j A_{ij} v_j \right| = |\lambda v_i| = \rho(A) \cdot u_i$$

so $Au \geq \rho(A)u$.

Perron's Theorem - item 1

- We proved $Au \geq \rho(A)u$.
- If inequality strict, then we have

$$A^2u > \rho(A) \cdot Au$$

and there is some positive $\varepsilon > 0$ such that

$$A^2u \geq (1 + \varepsilon)\rho(A)Au$$

Perron's Theorem - item 1

- We proved $Au \geq \rho(A)u$.
- If inequality strict, then we have

$$A^2u > \rho(A) \cdot Au$$

and there is some positive $\varepsilon > 0$ such that

$$A^2u \geq (1 + \varepsilon)\rho(A)Au$$

- By induction, we would have

$$A^{n+1}u \geq (1 + \varepsilon)^n \cdot \rho(A)^n \cdot Au$$

Perron's Theorem - item 1

- We proved $Au \geq \rho(A)u$.
- If inequality strict, then we have

$$A^2u > \rho(A) \cdot Au$$

and there is some positive $\varepsilon > 0$ such that

$$A^2u \geq (1 + \varepsilon)\rho(A)Au$$

- By induction, we would have

$$A^{n+1}u \geq (1 + \varepsilon)^n \cdot \rho(A)^n \cdot Au$$

- By Gelfand's formula we would have

$$\rho(A) = \lim_{n \rightarrow \infty} \|A^n\|_F^{1/n} \geq (1 + \varepsilon)\rho(A)$$

which is a contradiction. So equality must hold.

Perron's theorem - items 1 and 2

- We just proved that $\rho(A)$ is an eigenvalue, with eigenvector $u \geq 0$.

Perron's theorem - items 1 and 2

- We just proved that $\rho(A)$ is an eigenvalue, with eigenvector $u \geq 0$.
- Note that $u > 0$ since $\rho(A)u_i = (Au)_i > 0$

Perron's theorem - items 1 and 2

- We just proved that $\rho(A)$ is an eigenvalue, with eigenvector $u \geq 0$.
- Note that $u > 0$ since $\rho(A)u_i = (Au)_i > 0$
- Now we are ready for item 2: the only eigenvalue on the complex circumference $|\mu| = \rho(A)$ is $\rho(A)$

Perron's theorem - items 1 and 2

- We just proved that $\rho(A)$ is an eigenvalue, with eigenvector $u \geq 0$.
- Note that $u > 0$ since $\rho(A)u_i = (Au)_i > 0$
- Now we are ready for item 2: the only eigenvalue on the complex circumference $|\mu| = \rho(A)$ is $\rho(A)$
- If we had another eigenvalue $\lambda \neq \rho(A)$ in the circumference $|\mu| = \rho(A)$, where z is the eigenvector corresponding to λ , by the previous slide, we know that w defined as $w_i = |z_i|$ satisfies

$$Aw = \rho(A)w \Leftrightarrow \sum_j A_{ij}w_j = \rho(A) \cdot |z_i| = |\lambda z_i| = \left| \sum_j A_{ij}z_j \right|$$

for every $1 \leq i \leq n$

Perron's theorem - items 1 and 2

- We just proved that $\rho(A)$ is an eigenvalue, with eigenvector $u \geq 0$.
- Note that $u > 0$ since $\rho(A)u_i = (Au)_i > 0$
- Now we are ready for item 2: the only eigenvalue on the complex circumference $|\mu| = \rho(A)$ is $\rho(A)$
- If we had another eigenvalue $\lambda \neq \rho(A)$ in the circumference $|\mu| = \rho(A)$, where z is the eigenvector corresponding to λ , by the previous slide, we know that w defined as $w_i = |z_i|$ satisfies

$$Aw = \rho(A)w \Leftrightarrow \sum_j A_{ij}w_j = \rho(A) \cdot |z_i| = |\lambda z_i| = \left| \sum_j A_{ij}z_j \right|$$

for every $1 \leq i \leq n$

- **Lemma:** if the conditions above hold, then there is $\alpha \in \mathbb{C}$ nonzero such that $\alpha z \geq 0$

Proof by squaring both sides and using complex conjugates.

Perron's theorem - items 2 and 3

- But if $\alpha z \geq 0$ and a nonzero vector, we have

$$\lambda(\alpha z) = \alpha \cdot (\lambda z) = \alpha(Az) = A(\alpha z) \geq 0$$

Perron's theorem - items 2 and 3

- But if $\alpha z \geq 0$ and a nonzero vector, we have

$$\lambda(\alpha z) = \alpha \cdot (\lambda z) = \alpha(Az) = A(\alpha z) \geq 0$$

- Thus we know that λ is a non-negative number. However, $\rho(A)$ is the only non-negative number in the circle $|\mu| = \rho(A)$. This concludes item 2.

Perron's theorem - items 2 and 3

- But if $\alpha z \geq 0$ and a nonzero vector, we have

$$\lambda(\alpha z) = \alpha \cdot (\lambda z) = \alpha(Az) = A(\alpha z) \geq 0$$

- Thus we know that λ is a non-negative number. However, $\rho(A)$ is the only non-negative number in the circle $|\mu| = \rho(A)$. This concludes item 2.
- Now we are ready to prove item 3: the geometric multiplicity of $\rho(A)$ is 1.
- Suppose not, and let u, v be two linearly independent eigenvectors for $\rho(A)$. We can assume that both u, v are real vectors (why?).

Perron's theorem - items 2 and 3

- But if $\alpha z \geq 0$ and a nonzero vector, we have

$$\lambda(\alpha z) = \alpha \cdot (\lambda z) = \alpha(Az) = A(\alpha z) \geq 0$$

- Thus we know that λ is a non-negative number. However, $\rho(A)$ is the only non-negative number in the circle $|\mu| = \rho(A)$. This concludes item 2.
- Now we are ready to prove item 3: the geometric multiplicity of $\rho(A)$ is 1.
- Suppose not, and let u, v be two linearly independent eigenvectors for $\rho(A)$. We can assume that both u, v are real vectors (why?).
- Let $\beta > 0$ be such that $u - \beta v \geq 0$ and at least one entry is zero.

Perron's theorem - items 2 and 3

- But if $\alpha z \geq 0$ and a nonzero vector, we have

$$\lambda(\alpha z) = \alpha \cdot (\lambda z) = \alpha(Az) = A(\alpha z) \geq 0$$

- Thus we know that λ is a non-negative number. However, $\rho(A)$ is the only non-negative number in the circle $|\mu| = \rho(A)$. This concludes item 2.
- Now we are ready to prove item 3: the geometric multiplicity of $\rho(A)$ is 1.
- Suppose not, and let u, v be two linearly independent eigenvectors for $\rho(A)$. We can assume that both u, v are real vectors (why?).
- Let $\beta > 0$ be such that $u - \beta v \geq 0$ and at least one entry is zero.
- $u - \beta v \neq 0$ since the vectors are linearly independent

Perron's theorem - items 2 and 3

- But if $\alpha z \geq 0$ and a nonzero vector, we have

$$\lambda(\alpha z) = \alpha \cdot (\lambda z) = \alpha(Az) = A(\alpha z) \geq 0$$

- Thus we know that λ is a non-negative number. However, $\rho(A)$ is the only non-negative number in the circle $|\mu| = \rho(A)$. This concludes item 2.
- Now we are ready to prove item 3: the geometric multiplicity of $\rho(A)$ is 1.
- Suppose not, and let u, v be two linearly independent eigenvectors for $\rho(A)$. We can assume that both u, v are real vectors (why?).
- Let $\beta > 0$ be such that $u - \beta v \geq 0$ and at least one entry is zero.
- $u - \beta v \neq 0$ since the vectors are linearly independent
- But for each $1 \leq i \leq n$

$$\rho(A) \cdot (u - \beta v)_i = (A(u - \beta v))_i > 0$$

which contradicts our choice of β . Thus, there cannot be two linearly independent eigenvectors.

Perron-Frobenius

Theorem (Perron-Frobenius)

If a non-negative matrix $A \in \mathbb{R}^{n \times n}$ is *aperiodic* and *irreducible*, then the following hold:

- 1 $\rho(A)$ is an eigenvalue, and it has a positive eigenvector
- 2 $\rho(A)$ is the only eigenvalue in the complex circle $|\lambda| = \rho(A)$
- 3 $\rho(A)$ has geometric multiplicity 1
- 4 $\rho(A)$ has algebraic multiplicity 1

Perron-Frobenius

Theorem (Perron-Frobenius)

If a non-negative matrix $A \in \mathbb{R}^{n \times n}$ is *aperiodic* and *irreducible*, then the following hold:

- ① $\rho(A)$ is an eigenvalue, and it has a positive eigenvector
 - ② $\rho(A)$ is the only eigenvalue in the complex circle $|\lambda| = \rho(A)$
 - ③ $\rho(A)$ has geometric multiplicity 1
 - ④ $\rho(A)$ has algebraic multiplicity 1
- By previous lecture, we saw that A being aperiodic and irreducible implies that there is $m > 0$ such that A^m has all positive entries.

Theorem (Perron-Frobenius)

If a non-negative matrix $A \in \mathbb{R}^{n \times n}$ is *aperiodic* and *irreducible*, then the following hold:

- ① $\rho(A)$ is an eigenvalue, and it has a positive eigenvector
 - ② $\rho(A)$ is the only eigenvalue in the complex circle $|\lambda| = \rho(A)$
 - ③ $\rho(A)$ has geometric multiplicity 1
 - ④ $\rho(A)$ has algebraic multiplicity 1
- By previous lecture, we saw that A being aperiodic and irreducible implies that there is $m > 0$ such that A^m has all positive entries.
 - Apply Perron's theorem to A^m and note that the eigenvalues of A^m are λ_i^m , where λ_i are the eigenvalues of A

- Main Tools
 - Linear Algebra Background
 - Perron-Frobenius
- Main Applications
 - Fundamental Theorem of Markov Chains
 - Page Rank
- Acknowledgements

Fundamental Theorem of Markov Chains

- The *return time* from state i to itself is defined as

$$T_{i,i} := \min\{t \geq 1 \mid X_t = i, X_0 = i\}$$

Fundamental Theorem of Markov Chains

- The *return time* from state i to itself is defined as

$$T_{i,i} := \min\{t \geq 1 \mid X_t = i, X_0 = i\}$$

- *Expected return time*: defined as $\tau_{i,i} := \mathbb{E}[T_{i,i}]$.

Fundamental Theorem of Markov Chains

- The *return time* from state i to itself is defined as

$$T_{i,i} := \min\{t \geq 1 \mid X_t = i, X_0 = i\}$$

- Expected return time*: defined as $\tau_{i,i} := \mathbb{E}[T_{i,i}]$.

Theorem (Fundamental Theorem of Markov Chains)

Any *finite*, *irreducible* and *aperiodic* Markov Chain has the following properties:

- 1 There exists a *unique* stationary distribution π , where $\pi_i > 0$ for all $i \in [n]$
- 2 The sequence of distributions $\{p_t\}_{t \geq 0}$ will converge to π , no matter what the initial distribution is

3

$$\pi_i = \lim_{t \rightarrow \infty} P_{i,i}^t = \frac{1}{\tau_{i,i}}$$

Fundamental Theorem of Markov Chains

Theorem (Fundamental Theorem of Markov Chains)

Any *finite*, *irreducible* and *aperiodic* Markov Chain has the following properties:

- 1 There is *unique* stationary distribution π , where $\pi_i > 0$ for all $i \in [n]$
- 2 For every distribution $p_0 \in \mathbb{R}_{\geq 0}^n$,

$$\lim_{t \rightarrow \infty} P^t \cdot p_0 = \pi$$

3

$$\pi_i = \lim_{t \rightarrow \infty} P_{i,i}^t = \frac{1}{\tau_{i,i}}$$

Fundamental Theorem of Markov Chains

Theorem (Fundamental Theorem of Markov Chains)

Any *finite*, *irreducible* and *aperiodic* Markov Chain has the following properties:

- 1 There is *unique* stationary distribution π , where $\pi_i > 0$ for all $i \in [n]$
- 2 For every distribution $p_0 \in \mathbb{R}_{\geq 0}^n$,

$$\lim_{t \rightarrow \infty} P^t \cdot p_0 = \pi$$

3

$$\pi_i = \lim_{t \rightarrow \infty} P_{i,i}^t = \frac{1}{\tau_{i,i}}$$

- The transition matrix P is non-negative, irreducible and aperiodic. So we can apply Perron-Frobenius and prove items 1 and 2.

Fundamental Theorem of Markov Chains

Theorem (Fundamental Theorem of Markov Chains)

Any *finite, irreducible* and *aperiodic* Markov Chain has the following properties:

- 1 There is *unique* stationary distribution π , where $\pi_i > 0$ for all $i \in [n]$
- 2 For every distribution $p_0 \in \mathbb{R}_{\geq 0}^n$, $\lim_{t \rightarrow \infty} P^t \cdot p_0 = \pi$

3

$$\pi_i = \lim_{t \rightarrow \infty} P_{i,i}^t = \frac{1}{\tau_{i,i}}$$

If our underlying graph is undirected:

Fundamental Theorem of Markov Chains

Theorem (Fundamental Theorem of Markov Chains)

Any *finite*, *irreducible* and *aperiodic* Markov Chain has the following properties:

- 1 There is *unique* stationary distribution π , where $\pi_i > 0$ for all $i \in [n]$
- 2 For every distribution $p_0 \in \mathbb{R}_{\geq 0}^n$, $\lim_{t \rightarrow \infty} P^t \cdot p_0 = \pi$

3

$$\pi_i = \lim_{t \rightarrow \infty} P_{i,i}^t = \frac{1}{\tau_{i,i}}$$

If our underlying graph is undirected:

- If A_G adjacency matrix of $G(V, E)$ and $D = \text{diag}(d_1, d_2, \dots, d_n)$, transition matrix:

$$P = A_G D^{-1}$$

Fundamental Theorem of Markov Chains

Theorem (Fundamental Theorem of Markov Chains)

Any *finite*, *irreducible* and *aperiodic* Markov Chain has the following properties:

- 1 There is *unique* stationary distribution π , where $\pi_i > 0$ for all $i \in [n]$
- 2 For every distribution $p_0 \in \mathbb{R}_{\geq 0}^n$, $\lim_{t \rightarrow \infty} P^t \cdot p_0 = \pi$

3

$$\pi_i = \lim_{t \rightarrow \infty} P_{i,i}^t = \frac{1}{\tau_{i,i}}$$

If our underlying graph is undirected:

- If A_G adjacency matrix of $G(V, E)$ and $D = \text{diag}(d_1, d_2, \dots, d_n)$, transition matrix:

$$P = A_G D^{-1}$$

- Note that in this case, easy to guess stationary distribution:

$$\pi_i = \frac{d_i}{2m}, \quad m = |E|$$

Fundamental Theorem of Markov Chains

If our underlying graph is undirected:

Fundamental Theorem of Markov Chains

If our underlying graph is undirected:

- In this case, easy to guess stationary distribution:

$$\pi_i = \frac{d_i}{2m}, \quad m = |E|$$

Fundamental Theorem of Markov Chains

If our underlying graph is undirected:

- In this case, easy to guess stationary distribution:

$$\pi_i = \frac{d_i}{2m}, \quad m = |E|$$

- If A_G adjacency matrix of $G(V, E)$ and $D = \text{diag}(d_1, d_2, \dots, d_n)$, transition matrix:

$$P = A_G D^{-1}$$

Fundamental Theorem of Markov Chains

If our underlying graph is undirected:

- In this case, easy to guess stationary distribution:

$$\pi_i = \frac{d_i}{2m}, \quad m = |E|$$

- If A_G adjacency matrix of $G(V, E)$ and $D = \text{diag}(d_1, d_2, \dots, d_n)$, transition matrix:

$$P = A_G D^{-1}$$

- P not symmetric, but *similar* to a symmetric matrix:

$$D^{-1/2} P D^{1/2} = D^{-1/2} A_G D^{-1} D^{1/2} = D^{-1/2} A_G D^{-1/2} = P'$$

Fundamental Theorem of Markov Chains

If our underlying graph is undirected:

- If A_G adjacency matrix of $G(V, E)$ and $D = \text{diag}(d_1, d_2, \dots, d_n)$, transition matrix:

$$P = A_G D^{-1}$$

- P not symmetric, but *similar* to a symmetric matrix:

$$D^{-1/2} P D^{1/2} = D^{-1/2} A_G D^{-1} D^{1/2} = D^{-1/2} A_G D^{-1/2} = P'$$

- P and P' has *same eigenvalues*! And P' has only *real eigenvalues*!

Fundamental Theorem of Markov Chains

If our underlying graph is undirected:

- If A_G adjacency matrix of $G(V, E)$ and $D = \text{diag}(d_1, d_2, \dots, d_n)$, transition matrix:

$$P = A_G D^{-1}$$

- P not symmetric, but *similar* to a symmetric matrix:

$$D^{-1/2} P D^{1/2} = D^{-1/2} A_G D^{-1} D^{1/2} = D^{-1/2} A_G D^{-1/2} = P'$$

- P and P' has *same eigenvalues*! And P' has only *real eigenvalues*!
- Eigenvectors of P are $D^{1/2} v_i$ where v_i are eigenvectors of P' . And v_i can be taken to form *orthonormal basis*.

Fundamental Theorem of Markov Chains

- Stationary distribution: $\pi_i = \frac{d_i}{2m}$, $m = |E|$
- Transition matrix: $P = D^{-1} \cdot A_G$
 - P *similar* to a symmetric matrix: $P' = D^{-1/2} A_G D^{-1/2}$
 - P and P' has *same eigenvalues*! And P' has only *real eigenvalues*!
 - Eigenvectors of P are $D^{-1/2} v_i$ where v_i are eigenvectors of P' . And v_i can be taken to form *orthonormal basis*.

Fundamental Theorem of Markov Chains

- Stationary distribution: $\pi_i = \frac{d_i}{2m}$, $m = |E|$
- Transition matrix: $P = D^{-1} \cdot A_G$
 - P *similar* to a symmetric matrix: $P' = D^{-1/2} A_G D^{-1/2}$
 - P and P' has *same eigenvalues*! And P' has only *real eigenvalues*!
 - Eigenvectors of P are $D^{-1/2} v_i$ where v_i are eigenvectors of P' . And v_i can be taken to form *orthonormal basis*.
 - Graph strongly connected \Rightarrow *Perron-Frobenius* for irreducible non-negative matrices

Fundamental Theorem of Markov Chains

- Stationary distribution: $\pi_i = \frac{d_i}{2m}$, $m = |E|$
- Transition matrix: $P = D^{-1} \cdot A_G$
 - P *similar* to a symmetric matrix: $P' = D^{-1/2} A_G D^{-1/2}$
 - P and P' has *same eigenvalues*! And P' has only *real eigenvalues*!
 - Eigenvectors of P are $D^{-1/2} v_i$ where v_i are eigenvectors of P' . And v_i can be taken to form *orthonormal basis*.
 - Graph strongly connected \Rightarrow *Perron-Frobenius* for irreducible non-negative matrices
 - *unique* eigenvector whose eigenvalue has *max absolute value*
 - eigenvector has *all positive coordinates*
 - eigenvalue is *positive*

Fundamental Theorem of Markov Chains

- Stationary distribution: $\pi_i = \frac{d_i}{2m}$, $m = |E|$
- Transition matrix: $P = D^{-1} \cdot A_G$
 - P *similar* to a symmetric matrix: $P' = D^{-1/2} A_G D^{-1/2}$
 - P and P' has *same eigenvalues*! And P' has only *real eigenvalues*!
 - Eigenvectors of P are $D^{-1/2} v_i$ where v_i are eigenvectors of P' . And v_i can be taken to form *orthonormal basis*.
 - Graph strongly connected \Rightarrow *Perron-Frobenius* for irreducible non-negative matrices
 - *unique* eigenvector whose eigenvalue has *max absolute value*
 - eigenvector has *all positive coordinates*
 - eigenvalue is *positive*
 - This eigenvector is π !

Fundamental Theorem of Markov Chains

- Stationary distribution: $\pi_i = \frac{d_i}{2m}$, $m = |E|$
- Transition matrix: $P = D^{-1} \cdot A_G$
 - P *similar* to a symmetric matrix: $P' = D^{-1/2} A_G D^{-1/2}$
 - P and P' has *same eigenvalues*! And P' has only *real eigenvalues*!
 - Eigenvectors of P are $D^{-1/2} v_i$ where v_i are eigenvectors of P' . And v_i can be taken to form *orthonormal basis*.
 - Graph strongly connected \Rightarrow *Perron-Frobenius* for irreducible non-negative matrices
 - *unique* eigenvector whose eigenvalue has *max absolute value*
 - eigenvector has *all positive coordinates*
 - eigenvalue is *positive*
 - This eigenvector is π !
 - All random walks converge to π , as we wanted to show.

- Main Tools

- Linear Algebra Background
- Perron-Frobenius

- Main Applications

- Fundamental Theorem of Markov Chains
- Page Rank

- Acknowledgements

Page Rank

- **Setting:** we have a directed graph describing relationships between set of webpages.

Page Rank

- **Setting:** we have a directed graph describing relationships between set of webpages.

There is a directed edge (i, j) if there is a link from page i to page j .

Page Rank

- **Setting:** we have a directed graph describing relationships between set of webpages.

There is a directed edge (i, j) if there is a link from page i to page j .

- **Goal:** want algorithm to “rank” how important a page is.

Page Rank

- **Setting:** we have a directed graph describing relationships between set of webpages.

There is a directed edge (i, j) if there is a link from page i to page j .

- **Goal:** want algorithm to “rank” how important a page is.
- *Intuition:* if many other pages link to a particular page, then the linked page must be important!

Page Rank

- **Setting:** we have a directed graph describing relationships between set of webpages.

There is a directed edge (i, j) if there is a link from page i to page j .

- **Goal:** want algorithm to “rank” how important a page is.
- *Intuition:* if many other pages link to a particular page, then the linked page must be important!

Algorithm (Page Rank Algorithm)

- 1 *Initially, each page has pagerank value $\frac{1}{n}$*

Page Rank

- **Setting:** we have a directed graph describing relationships between set of webpages.

There is a directed edge (i, j) if there is a link from page i to page j .

- **Goal:** want algorithm to “rank” how important a page is.
- *Intuition:* if many other pages link to a particular page, then the linked page must be important!

Algorithm (Page Rank Algorithm)

- 1 Initially, each page has pagerank value $\frac{1}{n}$
- 2 In each step, each page:

Page Rank

- **Setting:** we have a directed graph describing relationships between set of webpages.

There is a directed edge (i, j) if there is a link from page i to page j .

- **Goal:** want algorithm to “rank” how important a page is.
- *Intuition:* if many other pages link to a particular page, then the linked page must be important!

Algorithm (Page Rank Algorithm)

- 1 Initially, each page has pagerank value $\frac{1}{n}$
- 2 In each step, each page:
 - 1 divides its pagerank value equally to its outgoing link,

Page Rank

- **Setting:** we have a directed graph describing relationships between set of webpages.

There is a directed edge (i, j) if there is a link from page i to page j .

- **Goal:** want algorithm to “rank” how important a page is.
- *Intuition:* if many other pages link to a particular page, then the linked page must be important!

Algorithm (Page Rank Algorithm)

- 1 Initially, each page has pagerank value $\frac{1}{n}$
- 2 In each step, each page:
 - 1 *divides* its pagerank value equally to its outgoing link,
 - 2 *sends* these equal shares to the pages it points to,

Page Rank

- **Setting:** we have a directed graph describing relationships between set of webpages.

There is a directed edge (i, j) if there is a link from page i to page j .

- **Goal:** want algorithm to “rank” how important a page is.
- *Intuition:* if many other pages link to a particular page, then the linked page must be important!

Algorithm (Page Rank Algorithm)

- 1 Initially, each page has pagerank value $\frac{1}{n}$
- 2 In each step, each page:
 - 1 *divides* its pagerank value equally to its outgoing link,
 - 2 *sends* these equal shares to the pages it points to,
 - 3 *updates* its new pagerank value to be the sum of shares it receives.

Page Rank - Example

Page Rank

Algorithm (Page Rank Algorithm)

- ① *Initially, each page has pagerank value $\frac{1}{n}$*
- ② *In each step, each page:*
 - ① *divides* its pagerank value equally to its outgoing link,
 - ② *sends* these equal shares to the pages it points to,
 - ③ *updates* its new pagerank value to be the sum of shares it receives.

Page Rank

Algorithm (Page Rank Algorithm)

- ① Initially, each page has pagerank value $\frac{1}{n}$
- ② In each step, each page:
 - ① *divides* its pagerank value equally to its outgoing link,
 - ② *sends* these equal shares to the pages it points to,
 - ③ *updates* its new pagerank value to be the sum of shares it receives.
- *Equilibrium* of pagerank values equal to probabilities of *stationary distribution* of random walk

$$P \in \mathbb{R}^{n \times n}, \quad P_{i,j} = \frac{1}{\delta^{\text{out}}(j)}$$

Page Rank

Algorithm (Page Rank Algorithm)

- ① Initially, each page has pagerank value $\frac{1}{n}$
- ② In each step, each page:
 - ① *divides* its pagerank value equally to its outgoing link,
 - ② *sends* these equal shares to the pages it points to,
 - ③ *updates* its new pagerank value to be the sum of shares it receives.
- *Equilibrium* of pagerank values equal to probabilities of *stationary distribution* of random walk

$$P \in \mathbb{R}^{n \times n}, \quad P_{i,j} = \frac{1}{\delta^{\text{out}}(j)}$$

- Pagerank values and transition probabilities satisfy same equations:

$$p_{t+1}(j) = \sum_{i:(i,j) \in E} \frac{p_t(i)}{\delta^{\text{out}}(i)} \Rightarrow p_{t+1} = P \cdot p_t$$

Page Rank

Algorithm (Page Rank Algorithm)

- ① Initially, each page has pagerank value $\frac{1}{n}$
- ② In each step, each page:
 - ① *divides* its pagerank value equally to its outgoing link,
 - ② *sends* these equal shares to the pages it points to,
 - ③ *updates* its new pagerank value to be the sum of shares it receives.
- *Equilibrium* of pagerank values equal to probabilities of *stationary distribution* of random walk

$$P \in \mathbb{R}^{n \times n}, \quad P_{i,j} = \frac{1}{\delta^{\text{out}}(j)}$$

- Pagerank values and transition probabilities satisfy same equations:

$$p_{t+1}(j) = \sum_{i:(i,j) \in E} \frac{p_t(i)}{\delta^{\text{out}}(i)} \Rightarrow p_{t+1} = P \cdot p_t$$

- If graph finite, irreducible and aperiodic, fundamental theorem guarantees stationary distribution.

Page Rank

- In practice, directed graph may not satisfy fundamental theorem's conditions

Page Rank

- In practice, directed graph may not satisfy fundamental theorem's conditions
- Modify original graph as follows:

Page Rank

- In practice, directed graph may not satisfy fundamental theorem's conditions
- Modify original graph as follows:
 - Fix number $0 < s < 1$
 - Divide s fraction of its pagerank value to its neighbors,
 - $1 - s$ fraction of its pagerank value to all nodes evenly

Page Rank

- In practice, directed graph may not satisfy fundamental theorem's conditions
- Modify original graph as follows:
 - Fix number $0 < s < 1$
 - Divide s fraction of its pagerank value to its neighbors,
 - $1 - s$ fraction of its pagerank value to all nodes evenly
- Equivalent to the random walk:

Page Rank

- In practice, directed graph may not satisfy fundamental theorem's conditions
- Modify original graph as follows:
 - Fix number $0 < s < 1$
 - Divide s fraction of its pagerank value to its neighbors,
 - $1 - s$ fraction of its pagerank value to all nodes evenly
- Equivalent to the random walk:
 - With probability s go to one of its neighbors (uniformly at random),
 - With probability $1 - s$ go to random page (uniformly at random)

Page Rank

- In practice, directed graph may not satisfy fundamental theorem's conditions
- Modify original graph as follows:
 - Fix number $0 < s < 1$
 - Divide s fraction of its pagerank value to its neighbors,
 - $1 - s$ fraction of its pagerank value to all nodes evenly
- Equivalent to the random walk:
 - With probability s go to one of its neighbors (uniformly at random),
 - With probability $1 - s$ go to random page (uniformly at random)
- Now resulting graph is *strongly connected* and *aperiodic* \Rightarrow unique stationary distribution

Page Rank

- In practice, directed graph may not satisfy fundamental theorem's conditions
- Modify original graph as follows:
 - Fix number $0 < s < 1$
 - Divide s fraction of its pagerank value to its neighbors,
 - $1 - s$ fraction of its pagerank value to all nodes evenly
- Equivalent to the random walk:
 - With probability s go to one of its neighbors (uniformly at random),
 - With probability $1 - s$ go to random page (uniformly at random)
- Now resulting graph is *strongly connected* and *aperiodic* \Rightarrow unique stationary distribution
- This modification does not change “relative importance” of vertices

Acknowledgement

- Lecture based largely on:
 - Hannah Cairns notes on Perron-Frobenius (see link in course webpage)
 - Lap Chi's notes
 - [Motwani & Raghavan 2007, Chapter 6]
- See Lap Chi's notes at
<https://cs.uwaterloo.ca/~lapchi/cs466/notes/L11.pdf>
- Also see Lap Chi's notes
<https://cs.uwaterloo.ca/~lapchi/cs466/notes/L14.pdf> for a proof of fundamental theorem of Markov chains for undirected graphs.

References I

 Motwani, Rajeev and Raghavan, Prabhakar (2007)

Randomized Algorithms

 Karp, R.M. and Luby, M. and Madras, N. (1989)

Monte-Carlo approximation algorithms for enumeration problems.

Journal of algorithms, 10(3), pp.429-448.

 Jerrum, M. and Sinclair, A. and Vigoda, E. (2004)

A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries.

Journal of the ACM (JACM), 51(4), pp.671-697.