Markov Chains and Page Rank

Anil Maheshwari

anil@scs.carleton.ca
School of Computer Science
Carleton University
Canada

Matrices

Matrices

1. A Rectangular Array
2. Operations: Addition; Multiplication; Diagonalization; Transpose; Inverse; Determinant
3. Row Operations; Linear Equations; Gaussian Elimination
4. Types: Identity; Symmetric; Diagonal; Upper/Lower Traingular; Orthogonal; Orthonormal
5. Transformations - Eigenvalues and Eigenvectors
6. Rank; Column and Row Space; Null Space
7. Applications: Page Rank, Dimensionality Reduction, Recommender Systems, . . .

Matrix Vector Product

Matrix-vector product: $Ax = b$

$$\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \times 4 + 1 \times -2 \\ 3 \times 4 + 4 \times -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Matrix Vector Product

$Ax = b$ as linear combination of columns:

$$\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \end{bmatrix} = 4 \begin{bmatrix} 2 \\ 3 \end{bmatrix} -2 \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

Eigenvalues and Eigenvectors

Given an $n \times n$ matrix A.
A non-zero vector v is an eigenvector of A, if $Av = \lambda v$ for some scalar λ.
λ is the eigenvalue corresponding to vector v.

Example

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

Observe that

$$\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 3 \end{bmatrix} \text{ and } \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Thus, $\lambda_1 = 5$ and $\lambda_2 = 1$ are the eigenvalues of A.
Corresponding eigenvectors are $v_1 = [1, 3]$ and $v_2 = [1, -1]$, as $Av_1 = \lambda_1 v_1$ and $Av_2 = \lambda_2 v_2$.

Computation of Eigenvalues and Eigenvectors

Given an $n \times n$ matrix A, we want to find eigenvalues λ's and the corresponding eigenvectors that satisfy $Av = \lambda v$.

We can express $Av = \lambda v$ as $(A - \lambda I)v = 0$, where I is $n \times n$ identity matrix.

Suppose $B = A - \lambda I$.

If B is invertible, than the only solution of $Bv = 0$ is $v = 0$, as $B^{-1}Bv = B^{-1}0$ or $v = 0$.

Thus B isn't invertible and hence the determinant of B is 0.

We solve the equation $det(A - \lambda I) = 0$ to obtain eigenvalues λ.

Once we know an eigenvalue λ_i, we can solve $Av_i = \lambda_i v_i$ to obtain the corresponding eigenvector v_i.

Computation of Eigenvalues and Eigenvectors

Let us find the eigenvalues and eigenvectors of $A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$

$det(A - \lambda I) = \begin{bmatrix} 2-\lambda & 1 \\ 3 & 4-\lambda \end{bmatrix} = 0$

$(2-\lambda)(4-\lambda) - 3 = 0$

$\lambda^2 - 6\lambda + 5 = 0$, and the two roots are $\lambda_1 = 5$ and $\lambda_2 = 1$.

To find the eigenvector $v_1 = [a, b]$, we can solve $\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 5 \begin{bmatrix} a \\ b \end{bmatrix}$.

This gives: $2a + b = 5a$ and $b = 3a$. Thus $v_1 = [1, 3]$ is an eigenvector corresponding to $\lambda_1 = 5$.

Similarly, for v_2, we have $\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 1 \begin{bmatrix} a \\ b \end{bmatrix}$.

This gives $2a + b = a$, or $a = -b$. Thus, $v_2 = [1, -1]$ is an eigenvector corresponding to $\lambda_2 = 1$.

Eigenvalues of A^k

Let $Av_i = \lambda_i v_i$

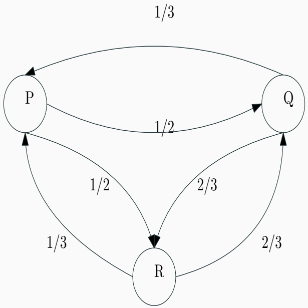
Consider: $A^2 v_i = A(Av_i) = A(\lambda_i v_i) = \lambda_i (Av_i) = \lambda_i(\lambda_i v_i) = \lambda_i^2 v_i$
$\implies A^2 v_i = \lambda_i^2 v_i$

Eigenvalues of A^k

For an integer $k > 0$, A^k has the same eigenvectors as A, but the eigenvalues are λ^k.

Markov Matrices

Markov Matrices



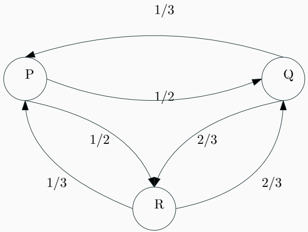
	P	Q	R
P	0	1/3	1/3
Q	1/2	0	2/3
R	1/2	2/3	0

Markov Chain

- $X_0, X_1, \ldots$ be a sequence of r. v. that evolve over time.
- At time 0, we have X_0, followed by X_1 at time 1, $\ldots$
- Assume each X_i takes value from the set $\{1, \ldots, n\}$ that represents the set of states.
- This sequence is a **Markov chain** if the probability that X_{m+1} equals a particular state $\alpha_{m+1} \in \{1, \ldots, n\}$ only depends on what is the state of X_m and is completely independent of the states of $X_0, \ldots, X_{m-1}$.

Memoryless property:
$P[X_{m+1} = \alpha_{m+1} | X_m = \alpha_m, X_{m-1} = \alpha_{m-1}, \ldots, X_0 = \alpha_0] = P[X_{m+1} = \alpha_{m+1} | X_m = \alpha_m]$, where $\alpha_0, \ldots, \alpha_{m+1}, \cdots \in \{1, \ldots, n\}$

Memoryless Property



	P	Q	R
P	0	1/3	1/3
Q	1/2	0	2/3
R	1/2	2/3	0

Markov Matrices

What is a Markov Matrix?

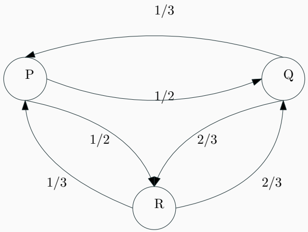
A square matrix A is a Markovian Matrix if

1. $A[i, j] =$ probability of transition from the state j to state i.
2. Sum of the values within any column is 1 (= probability of leaving from a state to any of the possible states).

State Transitions

Start in an initial state and in each successive step make a transition from the current state to the next state respecting the probabilities.

1. What is the probability of reaching the state j after taking n steps starting from the state i?
2. Given an initial probability vector representing the probabilities of starting in various states, what is the steady state? After traversing the chain for a large number of steps, what is the probability of landing in various states?



Types of States

Recurrent State: A state i is *recurrent* if starting from state i, with probability 1, we can return to the state i after making finitely many transitions.

Transient State: A state i is transient, i.e. there is a non-zero probability of not returning to the state i.

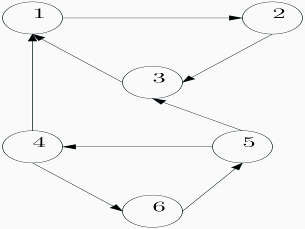
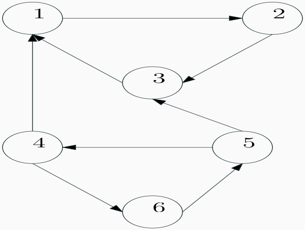


Figure 1: Recurrent States={1,2,3}. Transient States={4,5,6}

Irreducible Markov Chains

A Markov chain is **irreducible** if it is possible to go between any pair of states in a finite number of steps. Otherwise it is called **reducible**.

Observation: If the graph is strongly connected then it is irreducible.



Aperiodic Markov Chains

Period of a state

Period of a state i is the greatest common divisor (GCD) of all possible number of steps it takes the chain to return to the state i starting from i.

Note: If there is no way to return to i starting from i, then its period is undefined.

Aperiodic Markov Chain

A Markov chain is *aperiodic* if the periods of each of its states is 1.

Eigenvalues of Markov Matrices

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

Eigenvalues of A are the roots of $\det(A - \lambda I) = 0$

Eigenvalue	Eigenvector
$\lambda_1 = 1$	$v_1 = (2/3, 1, 1)$
$\lambda_2 = -2/3$	$v_2 = (0, -1, 1)$
$\lambda_3 = -1/3$	$v_3 = (-2, 1, 1)$

Observe: Largest (principal) eigenvalue is 1 and the corresponding (principal) eigenvector is $(2/3, 1, 1)$. Note that $Av_i = \lambda_i v_i$, for $i = 1, \ldots, 3$. Any vector v can be converted to a unit vector: $\frac{v}{||v||}$.

For example, for $v_1 = (\frac{2}{3}, 1, 1)$, the unit vector $\frac{v_1}{||v_1||}$ is $\frac{3}{\sqrt{22}}(\frac{2}{3}, 1, 1)$.

The vector $\frac{1}{2/3+1+1}(2/3, 1, 1) = (2/8, 3/8, 3/8)$ has the property that all its components add to 1 and it points in the same direction as v_1.

Principal Eigenvalue of Markov Matrices

Principal Eigenvalue

The largest eigenvalue of a Markovian matrix is 1

See Notes on Algorithm Design for the proof.

Idea: Let $B = A^T$
$\overrightarrow{1}$ is an Eigenvector of B, as $B\overrightarrow{1} = 1\overrightarrow{1}$
$\implies 1$ is an Eigenvalue of A.

Using contradiction, show that B cannot have any eigenvalue > 1

Eigenvalues of Powers of A

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

Note that all the entries in A^2 are > 0 and all the entries within a column still adds to 1.

$$A^2 = \begin{bmatrix} 1/3 & 2/9 & 2/9 \\ 1/3 & 11/17 & 1/6 \\ 1/3 & 1/6 & 11/17 \end{bmatrix}$$

A^k is Markovian

If the entries within each column of A adds to 1, then entries within each column of A^k, for any integer $k > 0$, will add to 1.

Random Surfer Model

Initial: Surfer with probability vector $u_0 = (1/3, 1/3, 1/3)$

$$u_1 = Au_0 = \begin{bmatrix} 0 & 1/3 & 1/3 \\ 1/2 & 0 & 2/3 \\ 1/2 & 2/3 & 0 \end{bmatrix} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 4/18 \\ 7/18 \\ 7/18 \end{bmatrix}$$

$$u_2 = Au_1 = \begin{bmatrix} 0 & 1/3 & 1/3 \\ 1/2 & 0 & 2/3 \\ 1/2 & 2/3 & 0 \end{bmatrix} \begin{bmatrix} 4/18 \\ 7/18 \\ 7/18 \end{bmatrix} = \begin{bmatrix} 7/27 \\ 10/27 \\ 10/27 \end{bmatrix}$$

Likewise, we compute $u_3 = Au_2 = [20/81, 61/162, 61/162]$,
$u_4 = Au_3 = [61/243, 91/243, 91/243]$,
$u_5 = Au_4 = [182/729, 547/1458, 547/1458]$,
$\ldots$
$u_\infty = [0.25, 0.375, 0.375] = [2/8, 3/8, 3/8]$

Linear Combination of Eigenvectors

$$u_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = c_1 \begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{aligned} u_1 &= Au_0 \\ &= c_1 Av_1 + c_2 Av_2 + c_3 Av_3 \\ &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + c_3 \lambda_3 v_3 \text{ (as } Av_i = \lambda_i v_i) \end{aligned}$$

Thus,

$$u_1 = A \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = c_1 \lambda_1 \begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix} + c_2 \lambda_2 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + c_3 \lambda_3 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

Linear Combination of Eigenvectors(contd.)

$$u_2 = Au_1 = A^2 u_0 = c_1 \lambda_1^2 \begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix} + c_2 \lambda_2^2 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + c_3 \lambda_3^2 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

In general, for integer $k > 0$, $u_k = A^k u_0 = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + c_3 \lambda_3^k v_3$, i.e.

$$u_k = A^k \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = c_1 \lambda_1^k \begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix} + c_2 \lambda_2^k \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + c_3 \lambda_3^k \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

and that equals

$$u_k = c_1 1^k \begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix} + c_2 (-\frac{2}{3})^k \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + c_3 (-\frac{1}{3})^k \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

Linear Combination of Eigenvectors(contd.)

For large values of k, $(\frac{2}{3})^k \to 0$ and $(\frac{1}{3})^k \to 0$. The above expression reduces to

$$u_k \approx c_1 \begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix} = \frac{3}{8} \begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/8 \\ 3/8 \\ 3/8 \end{bmatrix}$$

Note that the value of c_1 is derived by solving the equation for $u_0 = c_1 v_1 + c_2 v_2 + c_3 v_3$ for $u_0 = [1/3, 1/3, 1/3]$

Linear Combination of Eigenvectors(contd.)

Suppose $u_0 = [1/4, 1/4, 1/2]$
$u_1 = Au_0 = [1/4, 11/24, 7/24]$
$u_2 = Au_1 = [1/4, 23/72, 31/72]$
$u_3 = Au_2 = [1/4, 89/216, 73/216]$
$\dots$
$u_\infty = [2/8, 3/8, 3/8]$

Convergence?

Entries in A^k

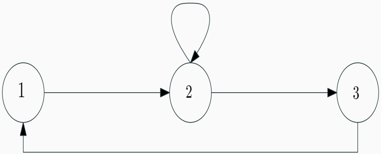
Assume that all the entries of a Markov matrix A, or of some finite power of A, i.e. A^k for some integer $k > 0$, are strictly > 0. A corresponds to an irreducible aperiodic Markov chain.

Irreducible: for any pair of states i and j, it is always possible to go from state i to state j in finite number of steps with positive probability.

Period of a state i: GCD of all possible number of steps it takes the chain to return to the state i starting from i.

Aperiodic: M is aperiodic if the GCD is 1 for the period of each of the states in M.

Properties of Markov Matrix A, when $A^k > 0$



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

$$A^3 = \begin{bmatrix} 1/2 & 1/4 & 0 \\ 1/4 & 5/8 & 1/2 \\ 1/4 & 1/8 & 1/2 \end{bmatrix} \; A^4 = \begin{bmatrix} 1/4 & 1/8 & 1/2 \\ 5/8 & 9/16 & 1/4 \\ 1/8 & 5/16 & 1/4 \end{bmatrix}$$

$A^4 > 0$ and for $k \geq 4$, $A^k > 0$.

A corresponds to irreducible aperiodic Markov chain.

Perron-Frobenius Theorem

Assume A corresponds to an irreducible aperiodic Markov chain M.

Perron-Frobenius Theorem from linear algebra states that

1. Largest eigenvalue 1 of A is unique
2. All other eigenvalues of A have magnitude strictly smaller than 1
3. All the coordinates of the eigenvector v_1 corresponding to the eigenvalue 1 are > 0
4. The steady state corresponds to the eigenvector v_1

Pagerank

Pagerank Algorithm

Problem: How to rank the web-pages?

Ranking assigns a real number to each web-page.
The higher the number, the more important the page is.
Needs to be automated, as the web is extremely large.

We will study the Page Rank algorithm.

Source: Page, Brin, Motwani, Winograd, The PageRank citation ranking: Bringing order to the Web published as a technical report in1998).

Web as a Graph

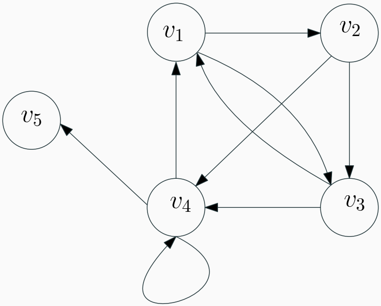
- $G = (V, E)$ is a positively weighted directed graph
- Each web-page is a vertex of G
- If a web-page u points (links) to the web-page v, there is a directed edge from u to v
- The weight of an edge uv is $\frac{1}{\text{out-degree}(u)}$

Assume $V = \{v_1, \ldots, v_n\}$
$n \times n$ adjacency matrix M of G is:

$$M(i,j) = \begin{cases} \frac{1}{\text{out-degree}(v_j)}, & \text{if } v_j v_i \in E \\ 0 & \text{otherwise} \end{cases}$$

Assumption: A surfer will make a random transition from a web-page to what it points to.

An Example

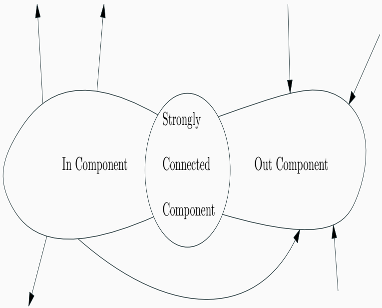


$$M = \begin{bmatrix} 0 & 0 & 1/2 & 1/3 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 & 0 \end{bmatrix}$$

Remarks

1. Assumes users will visit useful pages rather than useless pages.
2. Random Surfer Model - Assume initially a web-surfer is equally likely to be at any node of G, given by the vector $v_0 = (1/|V|, \ldots, 1/|V|)$.
3. In each step it makes a transition: $v_1 = Mv$, $v_2 = Mv_1 = M^2 v_0$, ..., $v_k = Mv_{k-1} = M^k v_0$.
4. Need to worry about sink nodes/dead ends; circling within same set of nodes; and whether we will reach a steady state?

Abstract representation of a web graph

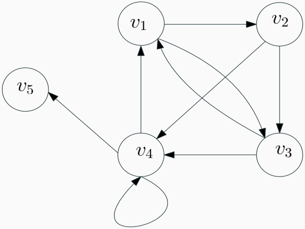


- In-Component: Nodes that can reach strongly-connected component
- Out-component: Nodes that can be reached from strongly-connected component
- Possibly multiple copies of above configuration

Avoiding Sink Nodes

Idea: Make sink nodes point to all other nodes.

$$M = \begin{bmatrix} 0 & 0 & 1/2 & 1/3 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 1/3 & 0 \\ 0 & 0 & 0 & 1/3 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 1/2 & 1/3 & 1/5 \\ 1/2 & 0 & 0 & 0 & 1/5 \\ 1/2 & 1/2 & 0 & 0 & 1/5 \\ 0 & 1/2 & 1/2 & 1/3 & 1/5 \\ 0 & 0 & 0 & 1/3 & 1/5 \end{bmatrix} = Q$$



Teleportation - Key Idea

Define $K = \alpha Q + \frac{1-\alpha}{n} E$

Teleportation Parameter: $0 < \alpha < 1$, e.g $\alpha = 0.9$

E is a $n \times n$ matrix of all 1s.

Observations on K:

1. Each entry of K is > 0
2. The entries within each column sums to 1
3. K satisfies the requirements of irreducible aperiodic Markov chain
4. Its largest eigenvalue is 1
5. By Perron-Frobenius Theorem, the **steady state (=page ranks) correspond to the principal eigenvector**

Conclusions

Computational Issues: $K = \alpha Q + \frac{1-\alpha}{n} E$
Q is sparse and E is special.

Favors: Teleport to specific pages. Teleport to topic-sensitive pages (Sports, Business, Science, News, ...) based on the profile of the user.

Caution: Real story is not that simple

References

1. Link Analysis Chapter in mmds.org
2. Chapter on Matrices in CS in my notes on algorithm design
3. Page, Brin, Motwani, Winograd, The PageRank citation ranking: Bringing order to the Web published as a technical report in1998.
4. Brin and Page, The Anatomy of a Large-Scale Hypertextual Web Search Engine, Computer Networks 56 (18): 3825-3833, Reprinted in 2012.