BACK TO GRAPHS: PATHS, CENTRALITY, PAGERANK

- *Back to grah models of sparse matrices*
- *Paths and powers of matrices*
- *Perron Frobenius theorem*
- *Application: Markov chains*
- *PageRank*
- *Notions of centrality*

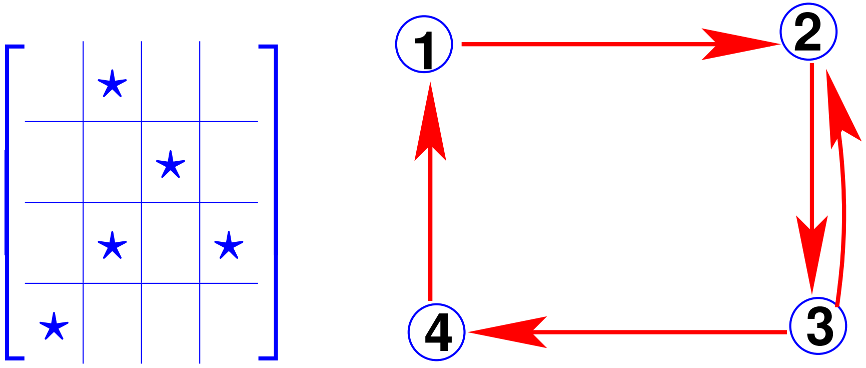
Graph Representations of Sparse Matrices. Recall:

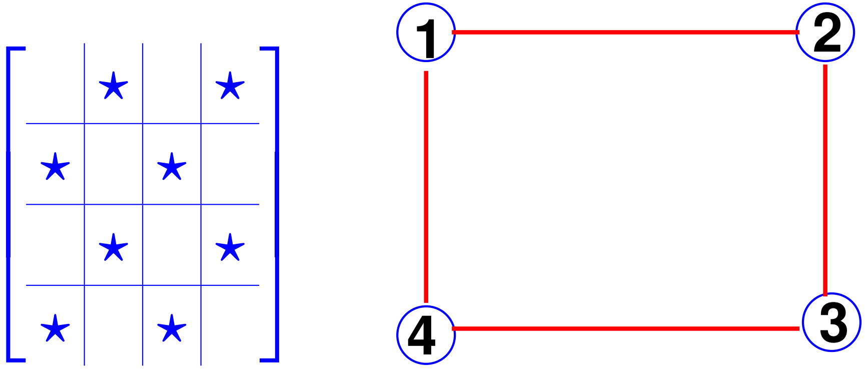
Adjacency Graph $G = (V, E)$ of an $n \times n$ matrix A :

$$V = \{1, 2, ..., N\} \qquad E = \{(i, j) | a_{ij} \neq 0\}$$

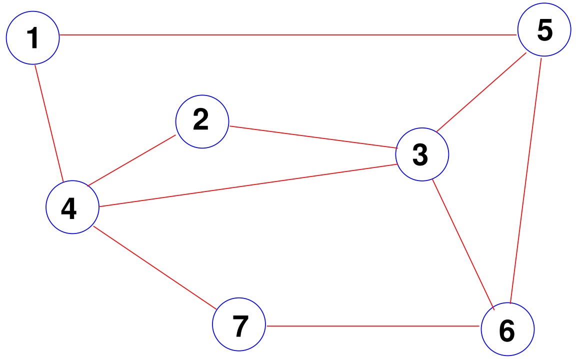
- G == undirected if A has a symmetric pattern

Example:





✍1 Show the matrix pattern for the graph on the right and give an interpretation of the path v_4, v_2, v_3, v_5, v_1 on the matrix



Example: Adjacency graph of:

$$A = \begin{bmatrix} & \star & & \star & & \\ \star & & \star & & & \\ & \star & & \star & \star & \star \\ \star & & \star & & & \\ & & \star & & & \star \\ & & \star & & \star & \end{bmatrix}.$$

Example: For any adjacency matrix A, what is the graph of A^2? [interpret in terms of paths in the graph of A]

Interpretation of graphs of matrices

✍2 What is the graph of $A + B$ (for two $n \times n$ matrices)?

✍3 What is the graph of A^T ?

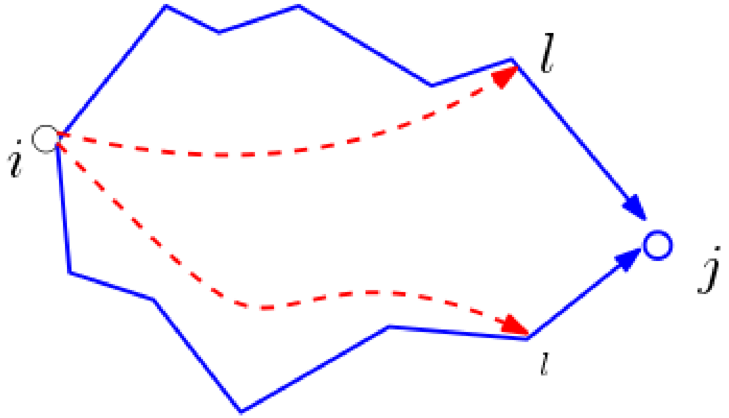
✍4 What is the graph of $A.B$?

Paths in graphs

✍5 What is the graph of A^k?

Theorem Let A be the adjacency matrix of a graph $G = (V, E)$. Then for $k \geq 0$ and vertices u and v of G, the number of paths of length k starting at u and ending at v is equal to $(A^k)_{u,v}$.

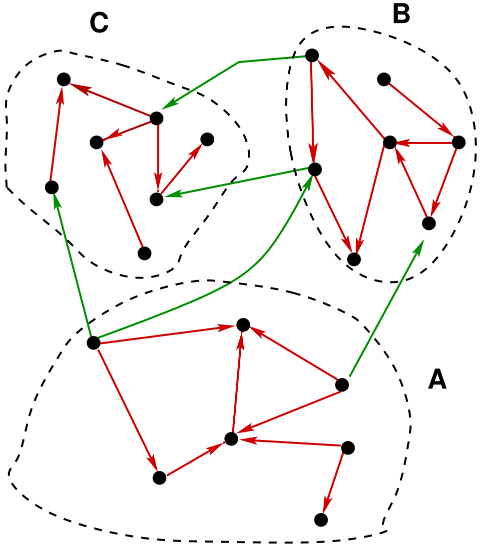
Proof: Proof is by induction. ■

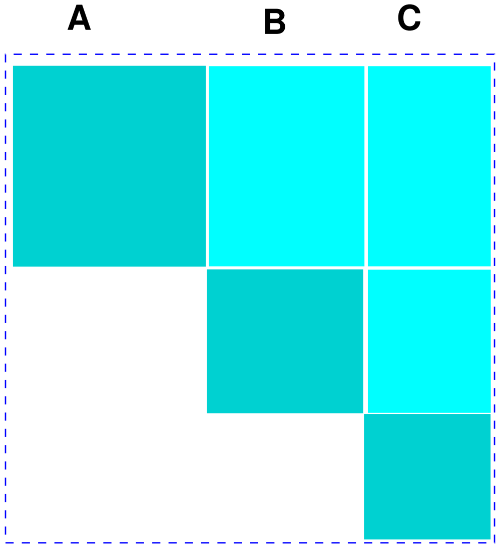


If $C = BA$ then $c_{ij} = \sum_l b_{il} a_{lj}$. Take $B = A^{k-1}$ and use induction. Any path of length k is formed as a path of length $k-1$ to some node l completed by an edge from l to j. Because a_{lj} is one for that last edge, c_{ij} is just the sum of all possible paths of length k from i to j

- Recall (definition): A matrix is *reducible* if it can be permuted into a block upper triangular matrix.

- Note: A matrix is reducible iff its adjacency graph is not (strongly) connected, i.e., iff it has more than one connected component.





- No edges from C to A or B. No edges from B to A.

Theorem: Perron-Frobenius An irreducible, nonnegative $n \times n$ matrix A has a real, positive eigenvalue λ_1 such that:
(i) λ_1 is a simple eigenvalue of A;
(ii) λ_1 admits a positive eigenvector u_1 ; and
(iii)$|\lambda_i| \leq \lambda_1$ for all other eigenvalues λ_i where $i > 1$.

- The spectral radius is equal to the eigenvalue λ_1

➤ Definition : a graph is d regular if each vertex has the same degree d.

Proposition: The spectral radius of a d regular graph is equal to d.

Proof: The vector e of all ones is an eigenvector of A associated with the eigenvalue $\lambda = d$. In addition this eigenvalue is the largest possible (consider the infinity norm of A). Therefore e is the Perron-Frobenius vector u_1. ∎

Application: Markov Chains

- Read about Markov Chains in Sect. 10.9 of:

  https://www-users.cs.umn.edu/~saad/eig_book_2ndEd.pdf

- Let $\pi \equiv$ row vector of stationary probabilities
- Then π satisfies the equation $\rightarrow$ $\pi P = \pi$

- P is the probabilty transition matrix and it is 'stochastic':

A matrix P is said to be *stochastic* if :

(i) $p_{ij} \geq 0$ for all i, j

(ii) $\sum_{j=1}^{n} p_{ij} = 1$ for $i = 1, \cdots, n$

(iii) No column of P is a zero column.

- Spectral radius is ≤ 1

✍6 Why?

- Assume P is irreducible. Then:

- Perron Frobenius $\rightarrow \rho(P) = 1$ is an eigenvalue and associated eigenvector has positive entries.

- Probabilities are obtained by scaling π by its sum.

- Example: One of the 2 models used for page rank.

Example: A college Fraternity has 50 students at various stages of college (Freshman, Sophomore, Junior, Senior). There are 6 potential stages for the following year: Freshman, Sophomore, Junior, Senior, graduated, or left-without degree. Following table gives probability of transitions from one stage to next

To From	Fr	So.	Ju.	Sr.	Grad	lwd
Fr.	.2	0	0	0	0	0
So.	.6	.1	0	0	0	0
Ju.	0	.7	.1	0	0	0
Sr.	0	0	.8	.1	0	0
Grad	0	0	0	.75	1	0
lwd	.2	.2	.1	.15	0	1

✍7 What is P? Assume initial population is $x_0 = [10, 16, 12, 12, 0, 0]$ and follow the stages of the students for a few years. What is the probability that a student will graduate? What is the probability that s/he leaves without a degree?

Page-rank

➤ Can be viewed as an application of Markov Chains

Main point: A page is important if it is pointed to by other important pages.

➤ Importance of your page (its PageRank) is determined by summing the page ranks of all pages which point to it.

➤ Weighting: If a page points to several other pages, then the weighting should be distributed proportionally.

➤ Imagine many tokens doing a random walk on this graph:

- (δ/n) chance to follow one of the n links on a page,
- $(1 - \delta)$ chance to jump to a random page.
- What's the chance a token will land on each page?

Page-Rank - definitions

If T_1, ..., T_n point to page T_i then

$$\rho(T_i) \; = \; 1 - \delta + \delta \left[\frac{\rho(T_1)}{|T_1|} + \frac{\rho(T_2)}{|T_2|} + \cdots \frac{\rho(T_n)}{|T_n|} \right]$$

- $|T_j|$ = count of links going out of Page T_j. So the 'vote' $\rho(T_j)$ is spread evenly among $|T_j|$ links.
- Sum of all PageRanks == 1: $\Sigma_T \rho(T) = 1$
- δ is a 'damping' parameter close to 1 – e.g. 0.85
- Defines a (possibly huge) Hyperlink matrix H

$$h_{ij} = \begin{cases} \frac{1}{|T_i|} & \text{if} \quad i \text{ points to } j \\ 0 & \text{otherwise} \end{cases}$$

✍8 4 Nodes

A points to B and D

B points to A, C, and D

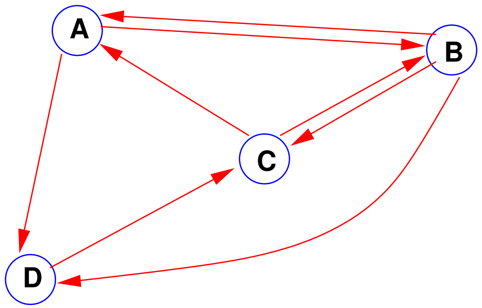
C points to A and B

D points to C

---

1) What is the H matrix?

2) the graph?



	A	B	C	D
A		$1/2$		$1/2$
B	$1/3$		$1/3$	$1/3$
C	$1/2$	$1/2$		
D			1	

- Row- sums of H are = 1.
- Sum of all PageRanks will be one: $$\sum_{\text{All-Pages}_A} \rho(A) = 1.$$
- H is a stochastic matrix [actually it is forced to be by changing zero rows]

Algorithm (PageRank)

1. Select initial row vector v ($v \geq 0$)
2. For i=1:maxitr
3 $\quad v := (1 - \delta)e^T + \delta v H$
4. end

✍9 Do a few steps of this algorithm for previous example with $\delta = 0.85$.

➤ This is a row iteration..

$$v = (1-\delta)e^T + v \, . \, \delta H$$

A few properties:

- v will remain ≥ 0. [combines non-negative vectors]
- More general iteration is of the form

$$v := v\underbrace{[(1-\delta)E + \delta H]}_{G} \quad \textbf{with} \quad E = ez^T$$

where z is a probability vector $e^T z = 1$ [Ex. $z = \frac{1}{n}e$]

- A variant of the power method.
- e is a right-eigenvector of G associated with $\lambda = 1$. We are interested in the left eigenvector.

✍10 Run *test_pr* + other drivers in matlab page

Kleinberg's Hubs and Authorities

➤ Idea is to put order into the web by ranking pages by their degree of Authority or "Hubness".

➤ An Authority is a page pointed to by many important pages.

- Authority Weight = sum of Hub Weights from In-Links.

➤ A Hub is a page that points to many important pages:

- Hub Weight = sum of Authority Weights from Out-Links.

➤ Source:

http://www.cs.cornell.edu/home/kleinber/auth.pdf

Computation of Hubs and Authorities

- Simplify computation by forcing sum of squares of weights to be 1.
- $\text{Auth}_j = x_j = \sum_{i:(i,j)\in\text{Edges}} \text{Hub}_i$.
- $\text{Hub}_i = y_i = \sum_{j:(i,j)\in\text{Edges}} \text{Auth}_j$.
- Let A = Adjacency matrix: $a_{ij} = 1$ if $(i,j) \in \text{Edges}$.
- $\mathrm{y} = A\mathrm{x}$, $\mathrm{x} = A^T\mathrm{y}$.
- Iterate . . . to leading eigenvectors of A^TA & AA^T.
- Answer: Leading Singular Vectors!

GRAPH CENTRALITY

Centrality in graphs

- Goal: measure importance of a node, edge, subgraph, .. in a graph
- Many measures introduced over the years
- Early Work: Freeman ’77 [introduced 3 measures] – based on ‘paths in graph’
- Many different ways of defininf centrality! We will just see a few

Degree centrality: (simplest) 'Nodes with high degree are important'
(note: scaling $n-1$ is unimportant)

$$C_D(v) = \deg(v)$$

Closeness centrality: 'Nodes that are close to many other nodes are important'

$$C_C(v) = \frac{1}{\sum_{w \neq v} d(v,w)}$$

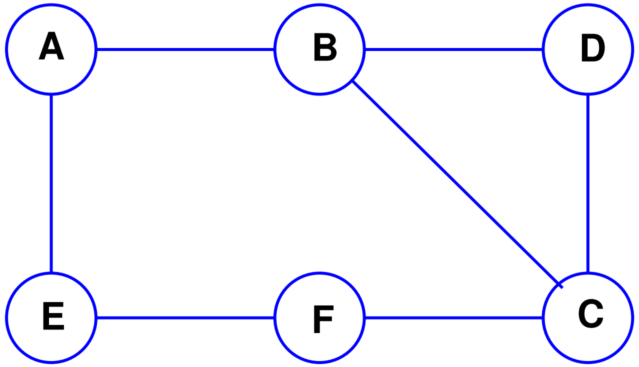
Betweenness centrality: (Freeman '77)

$$C_B(v) = \sum_{u \neq v, w \neq v} \frac{\sigma_{uw}(v)}{\sigma_{uw}}$$

- σ_{uw} = total # shortest paths from u to w
- $\sigma_{uw}(v)$ = total # shortest paths from u to w passing through v

➤ 'Nodes that are on many shortest paths are important'

Example: Find $C_D(v)$; $C_C(v)$; $C_B(v)$ when $v = C$



(u,w)	$\sigma_{uw}(v)$	σ_{uw}	/	(u,w)	$\sigma_{uw}(v)$	σ_{uw}	/
(A,B)	0	1	0	(B,E)	0	1	0
(A,D)	0	1	0	(B,F)	1	1	1
(A,E)	0	1	0	(D,E)	1	2	.5
(A,F)	0	1	0	(D,F)	1	1	1
(B,D)	0	1	0	(E,F)	0	1	0

- $C_D(v) = 3$;
- $C_C(v) = 1/[d_{CA} + d_{CB} + d_{CD} + d_{CE} + d_{CF}]$
  $= 1/[2 + 1 + 1 + 2 + 1] = 1/7$
- $C_B(v) = 2.5$ (add all ratios in table)

✍11 Redo this for $v = B$

Eigenvector centrality:

- Supppose we have n nodes v_j, $j = 1, \cdots, n$– each with a measure of importance ('prestige') p_j
- Principle: prestige of i depends on that of its neighbors.
- Prestige x_i = multiple of sum of prestiges of neighbors pointing to it

$$\lambda x_i = \sum_{j \in \mathcal{N}(i)} x_j = \sum_{j=1}^{n} a_{ji} x_j$$

- x_i = component of eigenvector associated with λ.
- Perron Frobenius theorem at play again: take largest eigenvalue $\rightarrow$ x_i's nonnegative