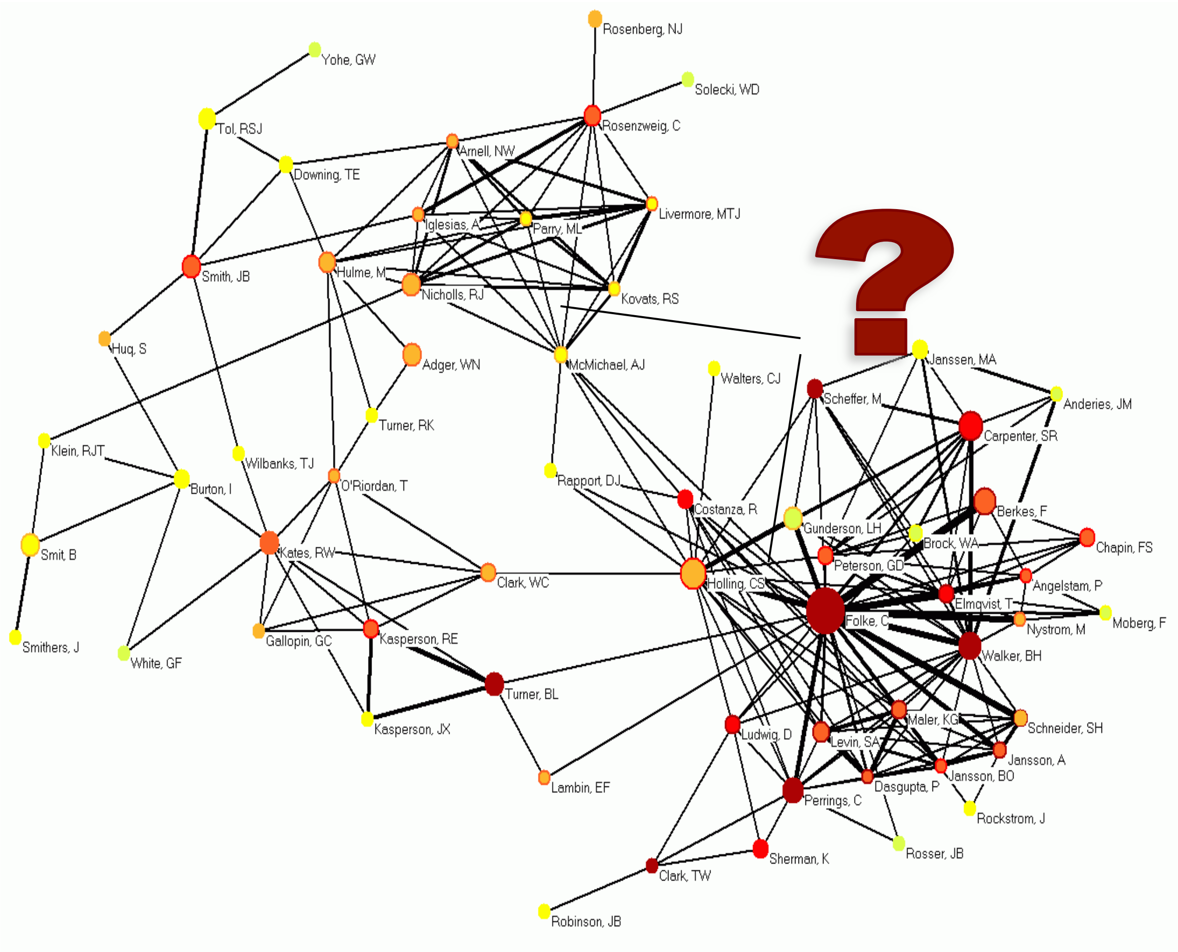
Paths and Random Walks on Graphs

Based on materials by Lala Adamic and Purnamrita Sarkar

Motivation: Link prediction in social networks



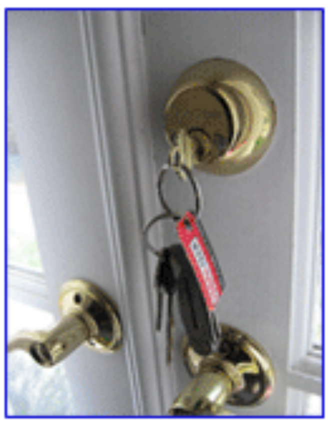
Motivation: Basis for recommendation

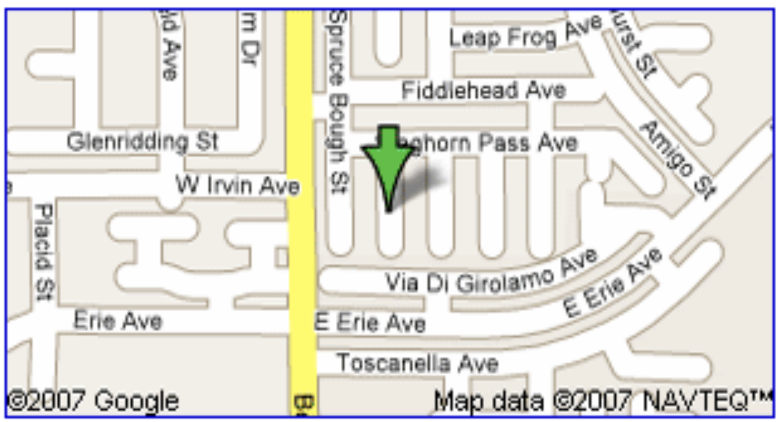
Motivation: Personalized search

Web Images Video News Maps Gmail more ▼

Google™ my car keys Search Advanced Search Preferences

Web





In the front door, where you left them last night.

Where Are **My Car Keys**?

Why graphs?

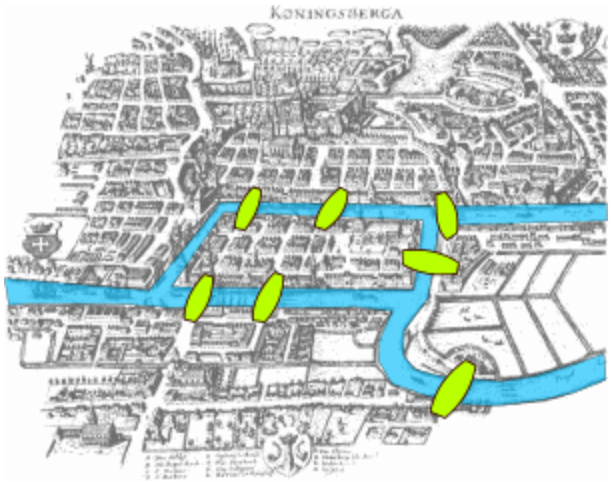
- The underlying data is naturally a graph
  - Papers linked by citation
  - Authors linked by co-authorship
  - Bipartite graph of customers and products
  - Web-graph
  - Friendship networks: who knows whom

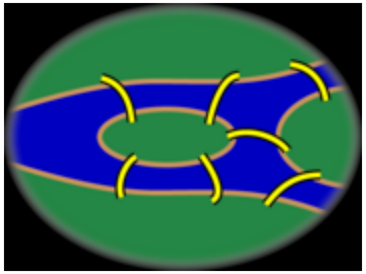
What are we looking for

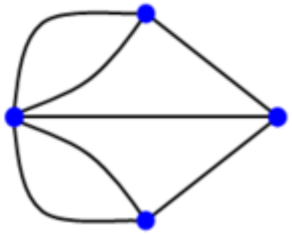
- Rank nodes for a particular query
  - Top k matches for "Random Walks" from Citeseer
  - Who are the most likely co-authors of "Manuel Blum".
  - Top k book recommendations for Jen from Amazon
  - Top k websites matching "Sound of Music"
  - Top k friend recommendations for Bob when he joins "Facebook"

History: Graph theory

- Euler's **Seven Bridges of Königsberg** – one of the first problems in graph theory
- Is there a route that crosses each bridge only once and returns to the starting point?



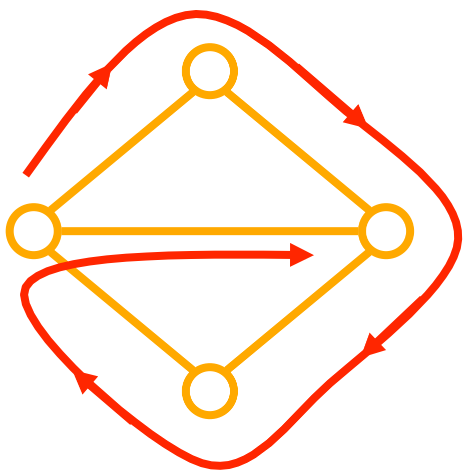




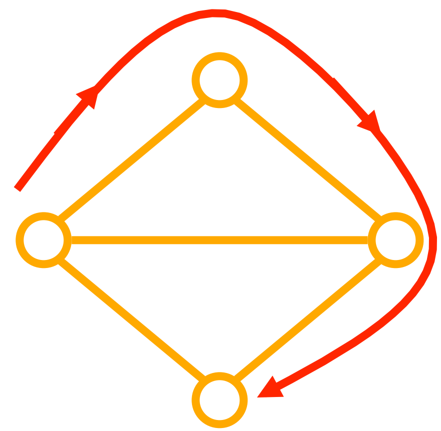
Source: http://en.wikipedia.org/wiki/Seven_Bridges_of_Königsberg

Eulerian paths

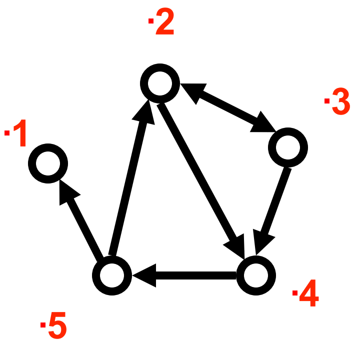
- If starting point and end point are the same:
  - only possible if no nodes have an odd degree
    - each path must visit and leave each shore
- If don’t need to return to starting point
  - can have 0 or 2 nodes with an odd degree



·Eulerian path: traverse each
·edge exactly once



·Hamiltonian path: visit
·each vertex exactly once



Node degree from matrix values

- Outdegree = $\sum_{j=1}^{n} A_{ij}$

·A = $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$

·example: outdegree for node 3 is 2, which we obtain by summing the number of non-zero entries in the 3rd *row*

$$\sum_{j=1}^{n} A_{3j}$$

- Indegree = $\sum_{i=1}^{n} A_{ij}$

·A = $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$

·example: the indegree for node 3 is 1, which we obtain by summing the number of non-zero entries in the 3rd *column*

$$\sum_{i=1}^{n} A_{i3}$$

Definitions

- n x n Adjacency matrix A.
  - A(i,j) = weight on edge from *i* to *j*
  - If the graph is undirected A(i,j)=A(j,i), i.e. A is symmetric

- n x n Transition matrix P.
  - P is row stochastic
  - P(i,j) = probability of stepping on node j from node i

    $= A(i,j)/\sum_i A(i,j)$

    i's outdegree

- n x n Laplacian Matrix L.
  - i's outdegree

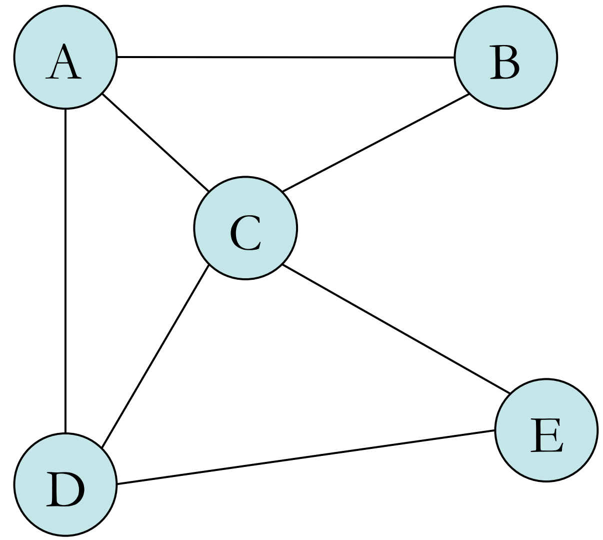
    $L(i,j)=\sum_i A(i,j) - A(i,j)$
  - Symmetric positive semi-definite for undirected graphs
  - Singular

Definitions

Adjacency Matrix

A=

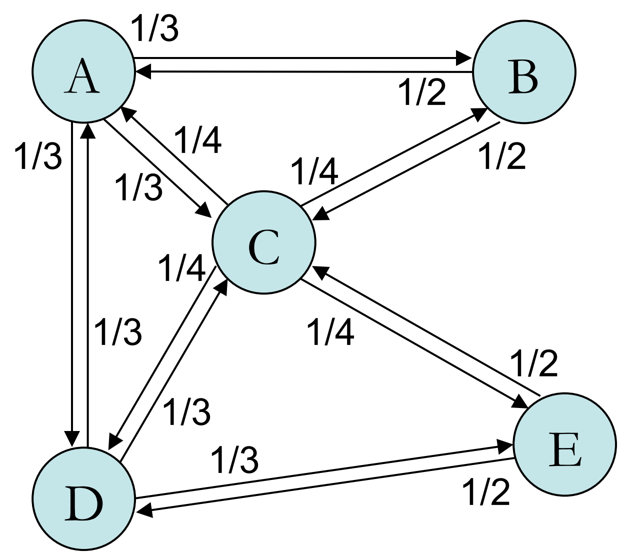
	1	1	1	
1		1		
1	1		1	1
1		1		1
		1	1	



Transition Matrix

P=

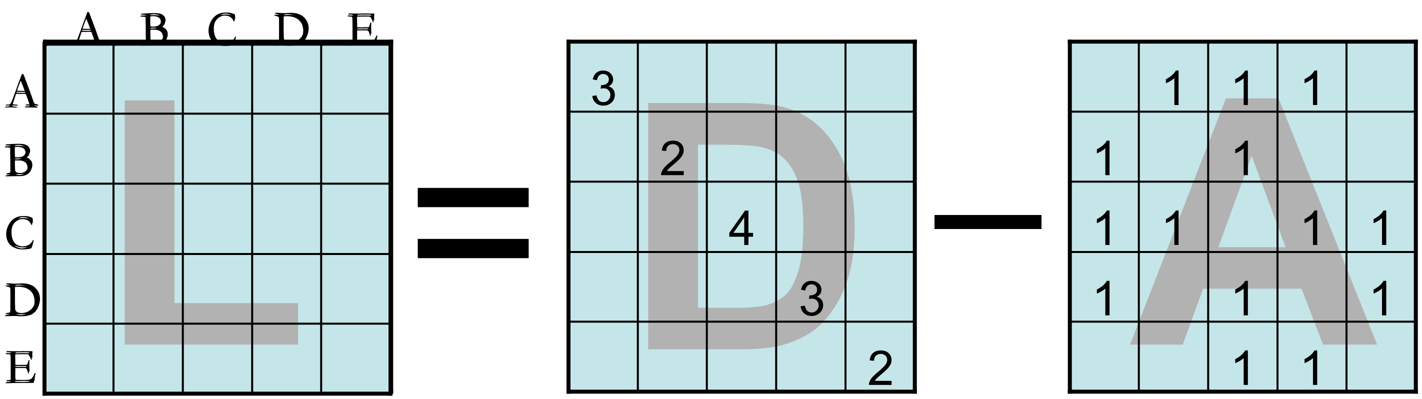
	1/3	1/3	1/3	
1/2		1/2		
1/4	1/4		1/4	1/4
1/3		1/3		1/3
		1/2	1/2	

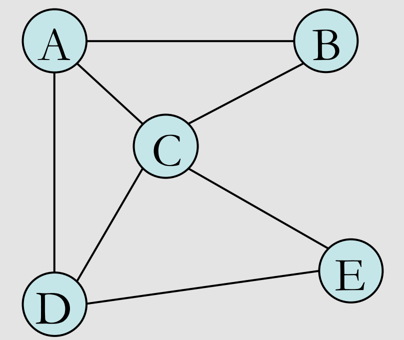


Definitions

Graph Laplacian

$$L = D - A \qquad D = \mathrm{diag}(d)$$





A =

	1	1	1	
1		1		
1	1		1	1
1		1		1
		1	1	

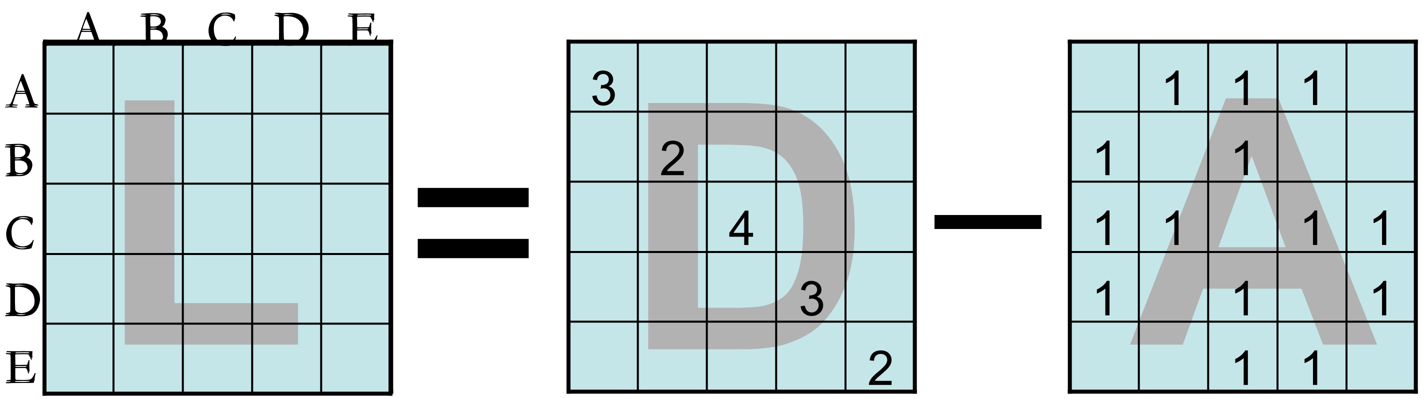
d =

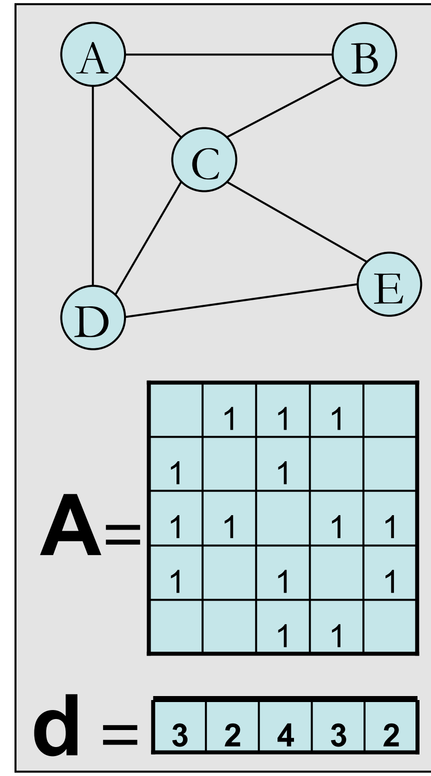
3	**2**	**4**	**3**	**2**

Spectral Graph Analysis

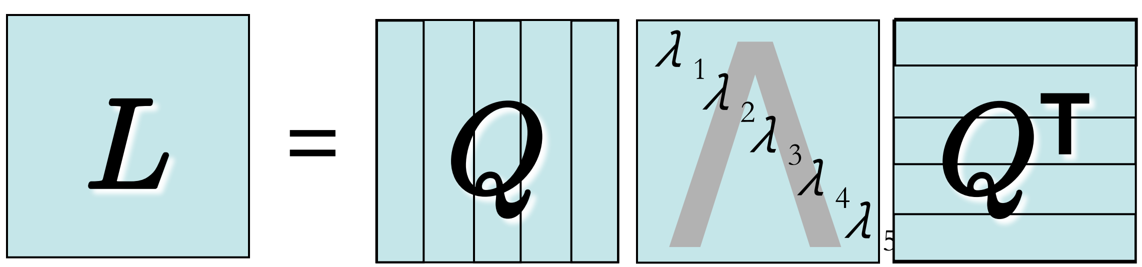
Graph Laplacian

$$L = D - A \qquad D = \mathrm{diag}(d)$$





Take the *eigendecomposition* of L



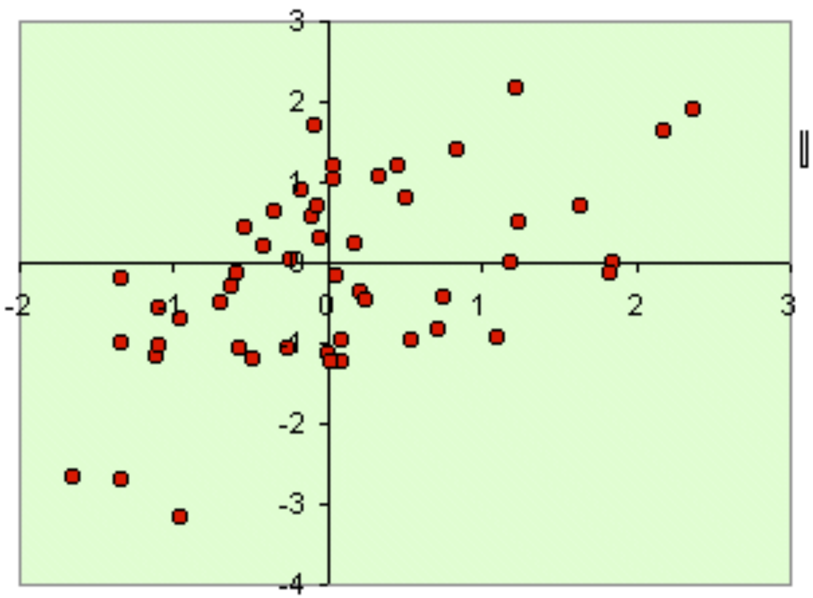
Eigenvectors

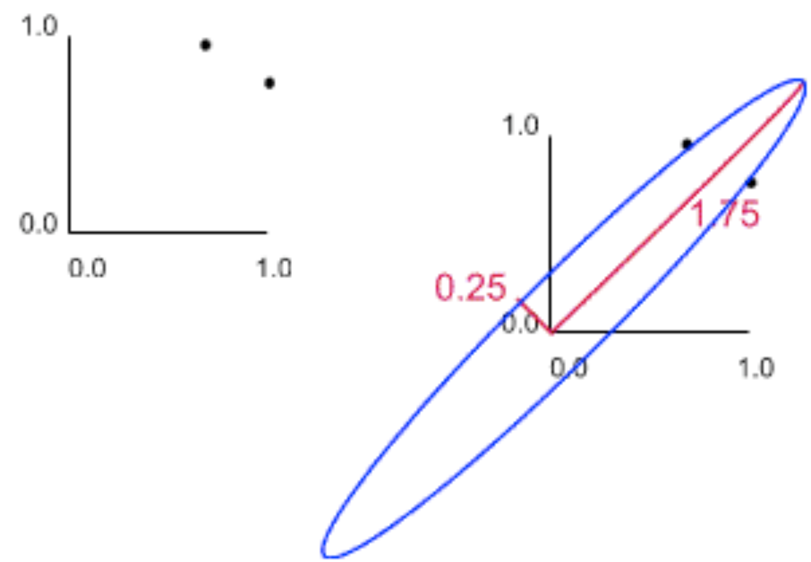
- Intuitive definition: An eigenvector is a direction for a matrix
- An eigenvector of an *n x n* matrix A is a vector such that $Av = \lambda v$, where *v* is the eigenvector and λ is the corresponding eigenvalue
  - Multiplying vector v by the scalar λ effectively stretches or shrinks the vector
- An *n x n* matrix should have *n* linearly independent eigenvectors

Eigenvectors Illustrated

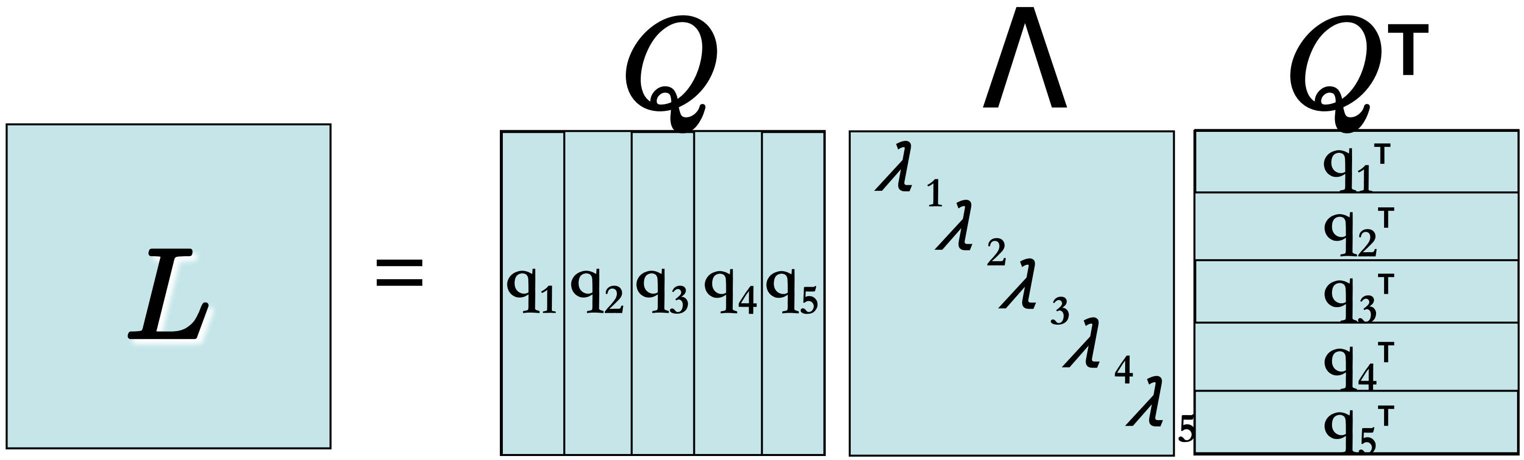
- Consider an elliptical data cloud. The eigenvectors are then the major and minor axes of the ellipse

1.00	0.75
0.75	1.00





Spectral Graph Analysis



3	-1	-1	-1	
-1	2	-1		
-1	-1	4	-1	-1
-1		-1	3	-1
		-1	-1	2

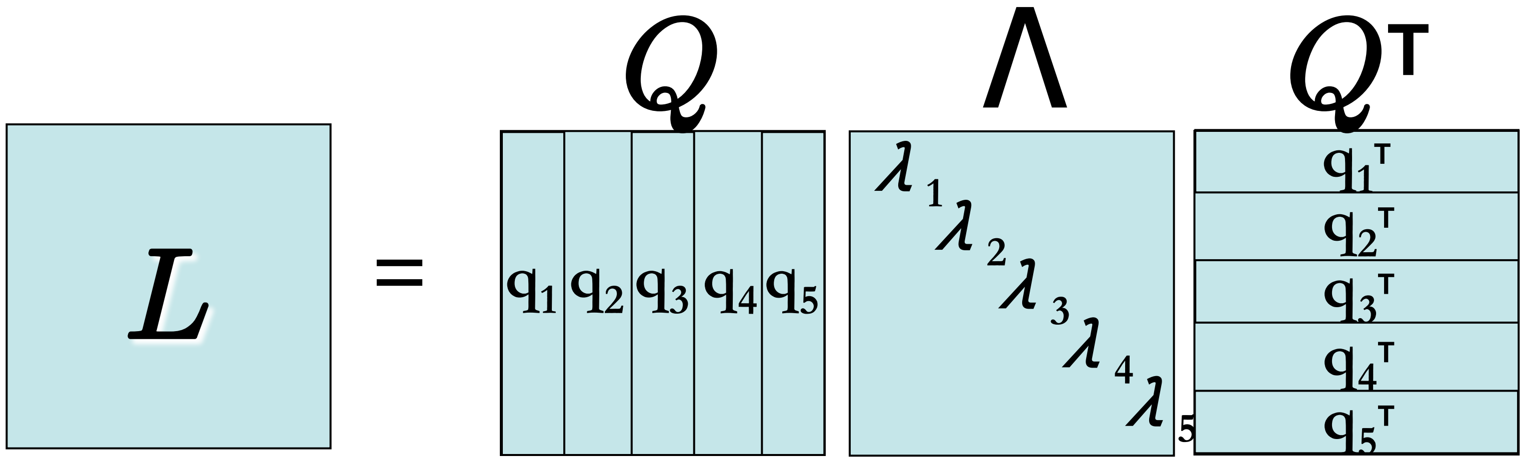
Eigenvector q_1 is constant

	q_1	q_2	q_3	q_4	q_5
A	0.45	-0.27	-0.5	-0.65	0.22
B	0.45	-0.65	0.5	0.27	0.22
C	0.45	-0.00	0.00	0.00	-0.89
D	0.45	0.27	-0.5	0.65	0.22
E	0.45	0.65	0.5	-0.27	0.22

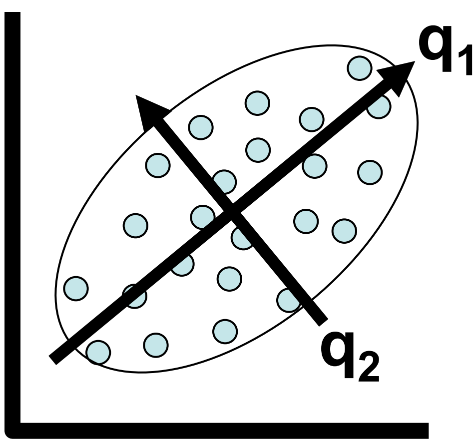
Eigenvalue $\lambda_1 = 0$

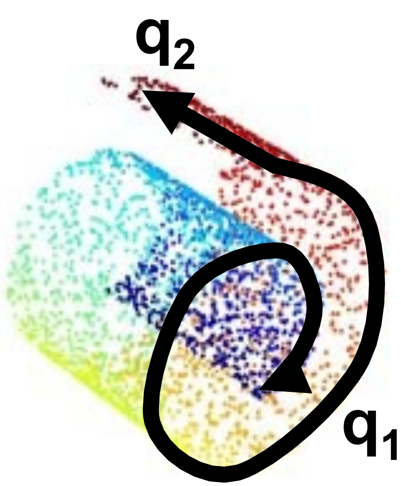
1	2	3	4	5
0.00	0	0	0	0
0	1.59	0	0	0
0	0	3.00	0	0
0	0	0	4.41	0
0	0	0	0	5.00

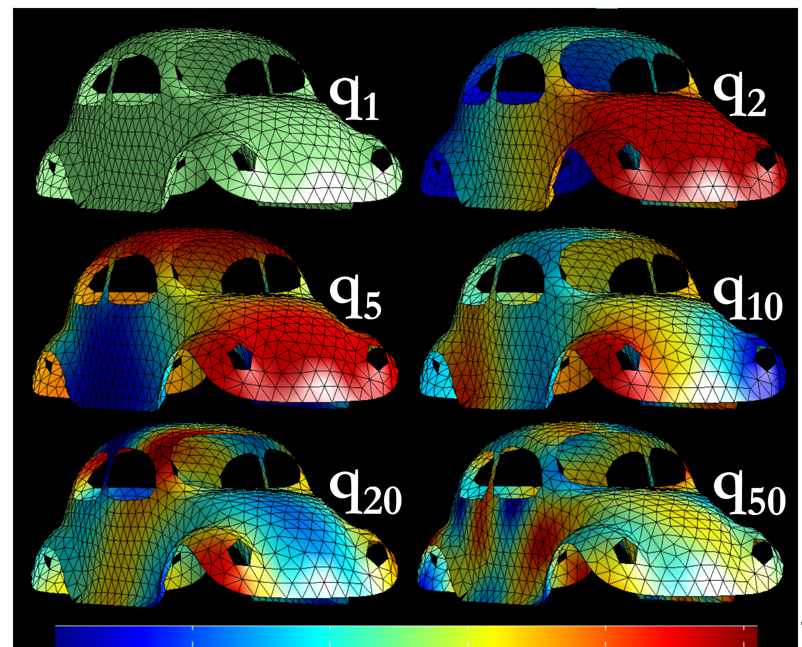
Spectral Graph Analysis



$L = Q \Lambda Q^T$







Meshes provided by Gabriel Peyré

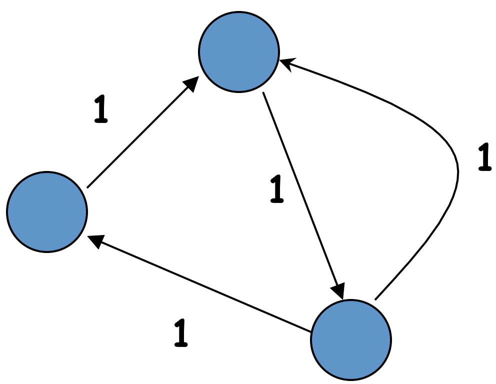
Random Walks

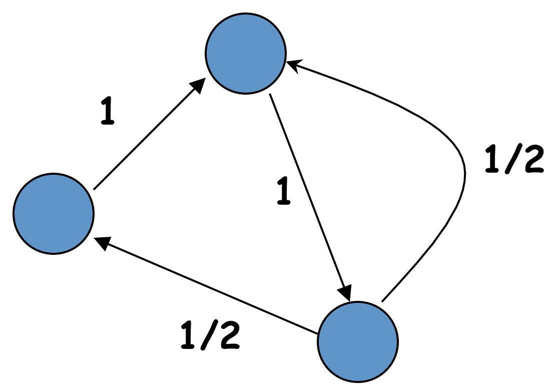
0	1	0
0	0	1
1	1	0

Adjacency matrix A

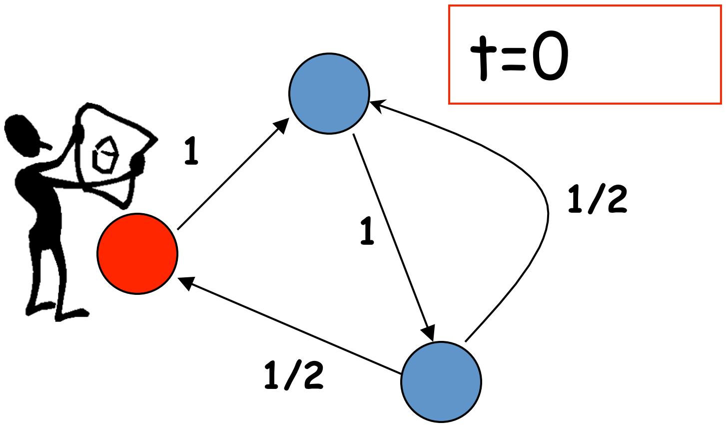
0	1	0
0	0	1
1/2	1/2	0

Transition matrix P

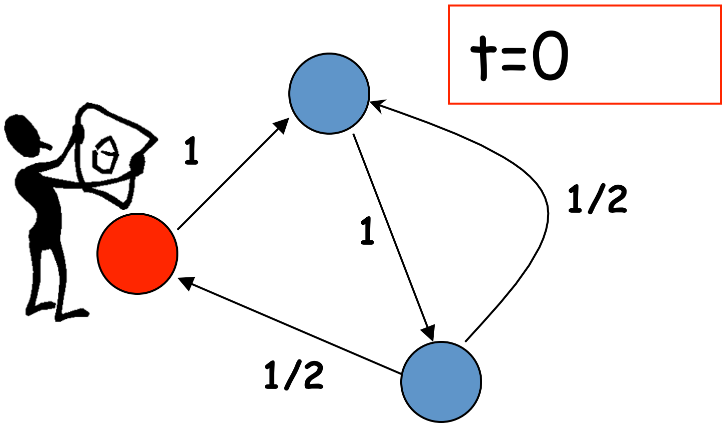


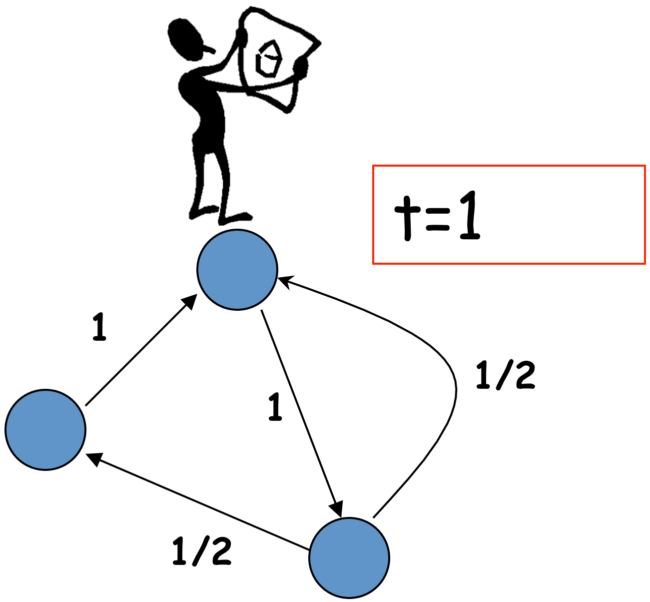


What is a random walk

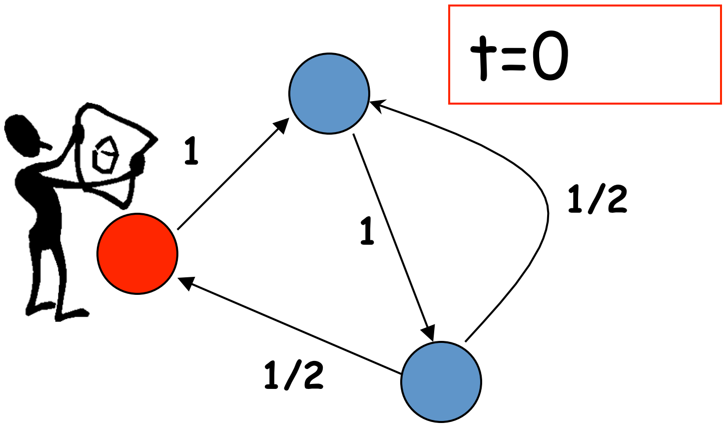


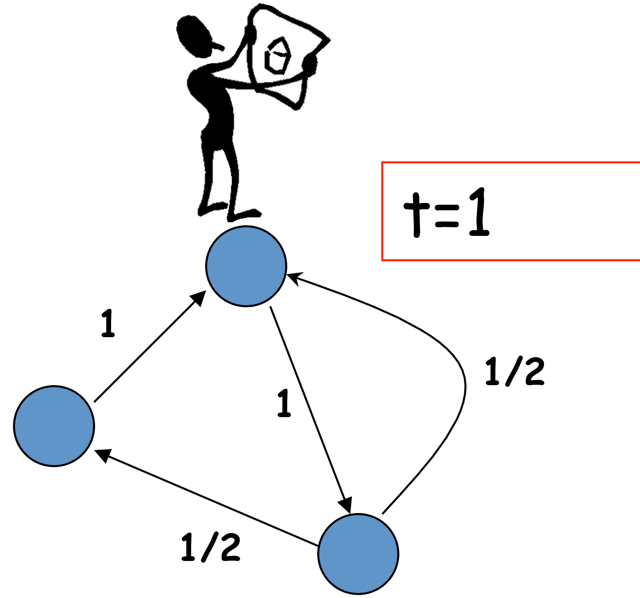
What is a random walk

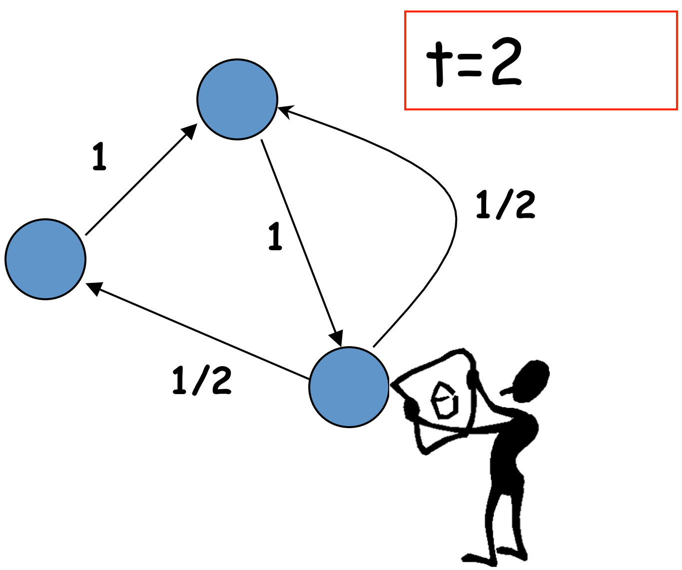




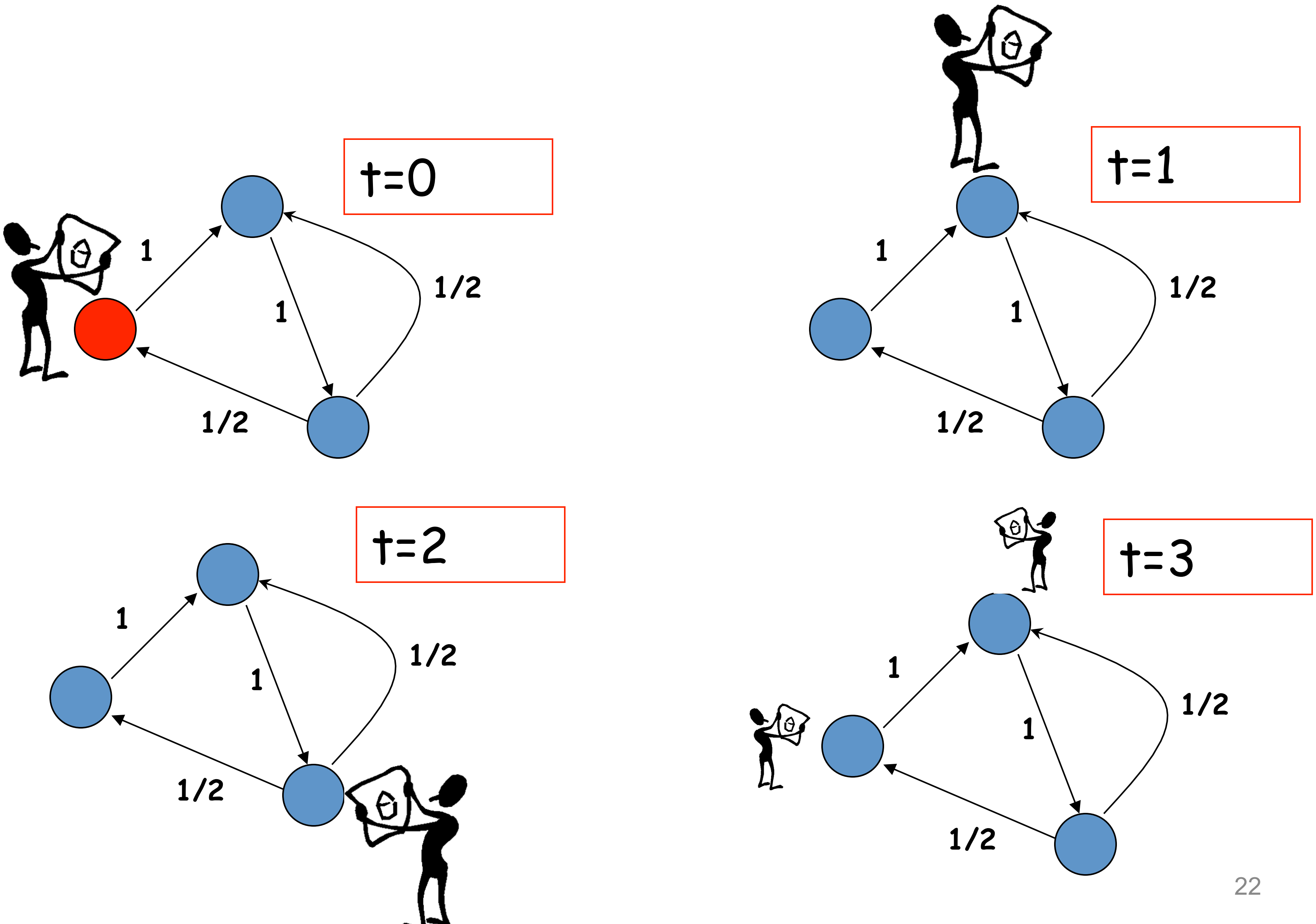
What is a random walk







What is a random walk



Probability Distributions

- $\phi_i^{(t)}$ = probability that the surfer is at node *i* at time *t*
- $$\phi_i^{(t+1)} = \sum_j \phi_i^{(t)} \times \Pr(j \to i)$$
- $$\begin{aligned}\phi_i^{(t+1)} &= \phi_i^{(t)} \times P \\ &= \phi_i^{(t-1)} \times P \times P \\ &= \phi_i^{(t-2)} \times P \times P \times P \\ &\ldots \\ &= \phi_i^{(0)} \times P^t\end{aligned}$$
- What happens when the surfer walks for a long time?

Stationary Distribution

- When the surfer keeps walking for a long time

- When the distribution does not change anymore
  - i.e. $\phi^{(t+1)} = \phi^{(t)}$

- For “well-behaved” graphs this does not depend on the start distribution!!

What is a stationary distribution? Intuitively and Mathematically

What is a stationary distribution? Intuitively and Mathematically

- The stationary distribution at a node is related to the amount of time a random walker spends visiting that node.

What is a stationary distribution? Intuitively and Mathematically

- The stationary distribution at a node is related to the amount of time a random walker spends visiting that node.
- Remember that we can write the probability distribution as

$$\phi^{(t+1)} = \phi^{(t)} \times P$$

What is a stationary distribution? Intuitively and Mathematically

- The stationary distribution at a node is related to the amount of time a random walker spends visiting that node.
- Remember that we can write the probability distribution as
$$\phi^{(t+1)} = \phi^{(t)} \times P$$
- For the stationary distribution $\phi^{(\infty)}$ we have
$$\phi^{(\infty)} = \phi^{(\infty)} \times P$$

What is a stationary distribution? Intuitively and Mathematically

- The stationary distribution at a node is related to the amount of time a random walker spends visiting that node.

- Remember that we can write the probability distribution as

$$\phi^{(t+1)} = \phi^{(t)} \times P$$

- For the stationary distribution $\phi^{(\infty)}$ we have

$$\phi^{(\infty)} = \phi^{(\infty)} \times P$$

- Whoa! that's just the left eigenvector of the transition matrix!

Power Method

(Horn & Johnson, 1985)

- P has a unique left eigenvector $\phi^{(\infty)}$
  - Called the Perron vector

Power method to compute $\phi^{(\infty)}$

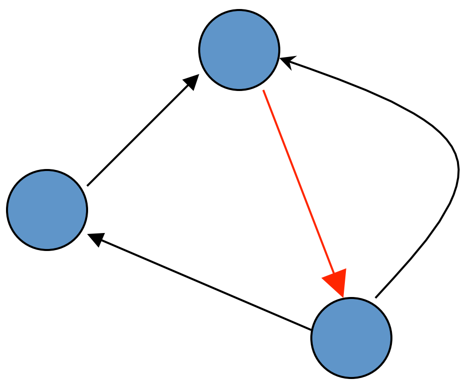
set $\phi^{(0)}$ to be a normalized nonnegative random vector
set i = 0
loop until $\phi^{(0)}, \phi^{(1)}, \ldots, \phi^{(i-1)}, \phi^{(i)}$ converges
    set $\phi^{(i+1)} = P\phi^{(i)}$
    normalize $\phi^{(i+1)}$
    i++
end loop
return $\phi^{(i)}$

Interesting Questions

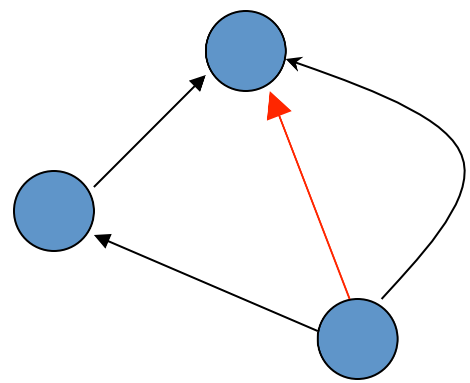
- Does a stationary distribution always exist? Is it unique?
  - Yes, if the graph is "well-behaved".

Well-behaved graphs

- Irreducible: There is a path from every node to every other node.



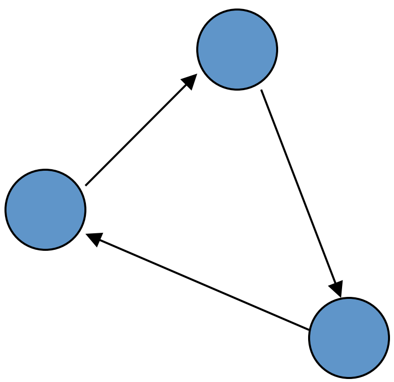
Irreducible



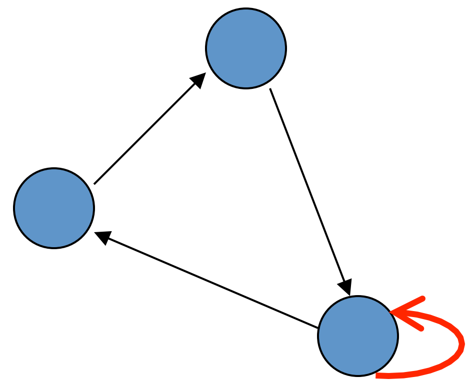
Not irreducible

Well-behaved graphs

- Aperiodic: The GCD of all cycle lengths is *1.* The GCD is also called period.



Periodicity is 3



Aperiodic

Implications of the Perron Frobenius Theorem

- If a Markov chain is irreducible and aperiodic then the largest eigenvalue of the transition matrix will be equal to **1** and all the other eigenvalues will be strictly less than **1**.
  - Let the eigenvalues of P be $\{\sigma_i \mid i=0:n-1\}$ in non-increasing order of σ_i .
  - $\sigma_0 = 1$ **>** $\sigma_1 > \sigma_2 >= \ldots\ldots >= \sigma_n$

Implications of the Perron Frobenius Theorem

- If a Markov chain is irreducible and aperiodic then the largest eigenvalue of the transition matrix will be equal to **1** and all the other eigenvalues will be strictly less than **1**.
  - Let the eigenvalues of P be $\{\sigma_i | i=0:n-1\}$ in non-increasing order of σ_i .
  - $\sigma_0 = 1 \mathbf{>} \sigma_1 > \sigma_2 >= \ldots\ldots >= \sigma_n$

- These results imply that for a well-behaved graph there exists an unique stationary distribution.

Google's PageRank

- PageRank is a "vote" by all other webpages about the importance of a page
- A link to a page counts as a vote of support
- PageRank uses a random surfer model
  - Occasionally, the surfer gets bored and jumps to a random other page

- "The 25,000,000,000 Eigenvector: the Linear Algebra Behind Google"

Random Walk on Web Graph

- Probability transition matrix given by

$$P_{u,v} = \frac{A_{u,v}}{d_u^{out}} = \frac{A_{u,v}}{\sum_{v=1}^{n} A_{u,v}} \qquad P = D^{-1}A$$

- Use a teleporting random walk (Page et al., 1998) to ensure that the graph is strongly connected and aperiodic:

$$P_{teleport} = \eta P + (1-\eta)\frac{\mathbf{1}\mathbf{1}^T - I}{|V|}$$

PageRank