Multivariate analysis and complex networks

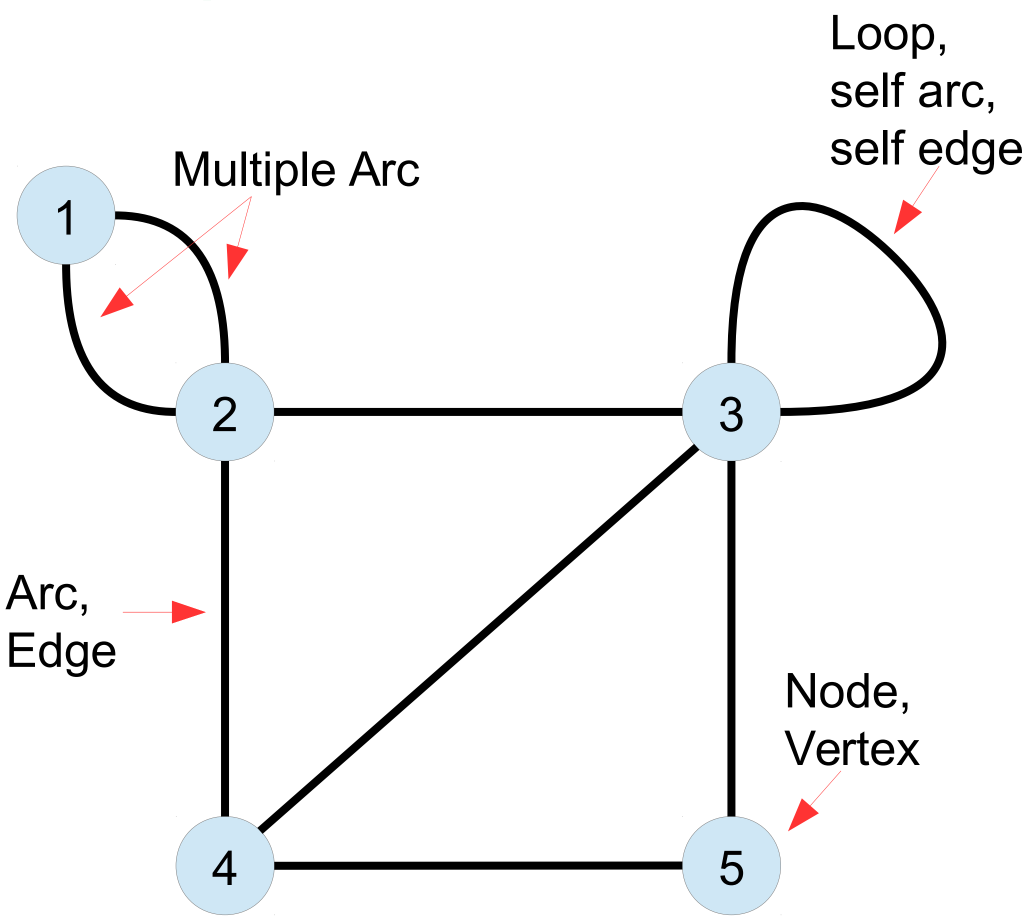
1: Graphs & Matrices

Antonio Scala

CNR-ISC, the Institute for Complex Systems

INFN School of Statistics
Ischia, May 2017

Graphs



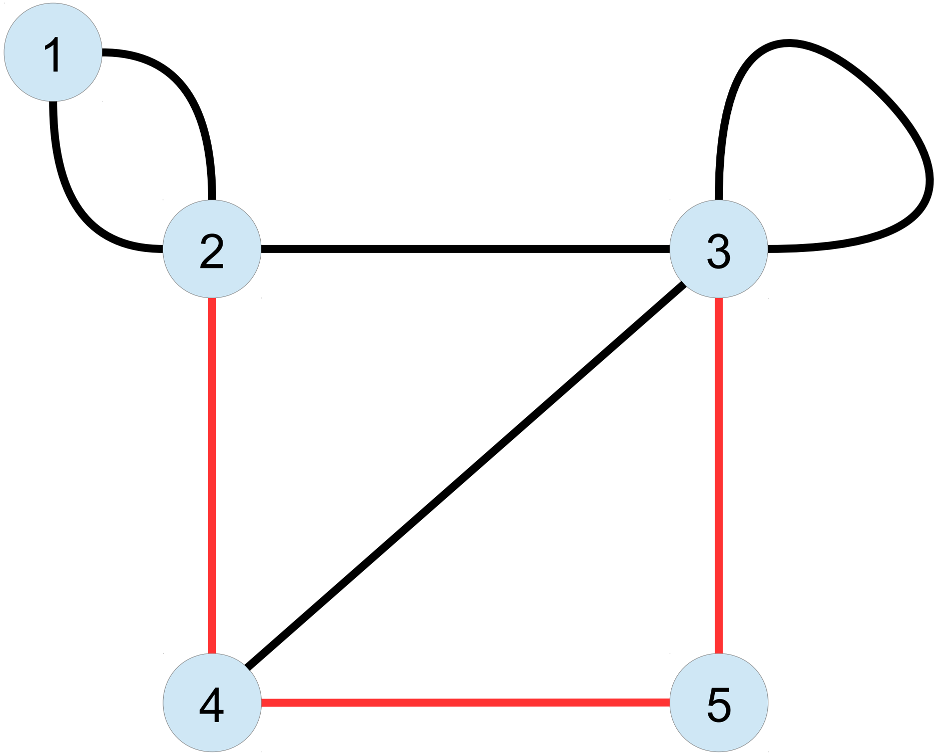
G = (V,E)

V = set of nodes

E =set of edges

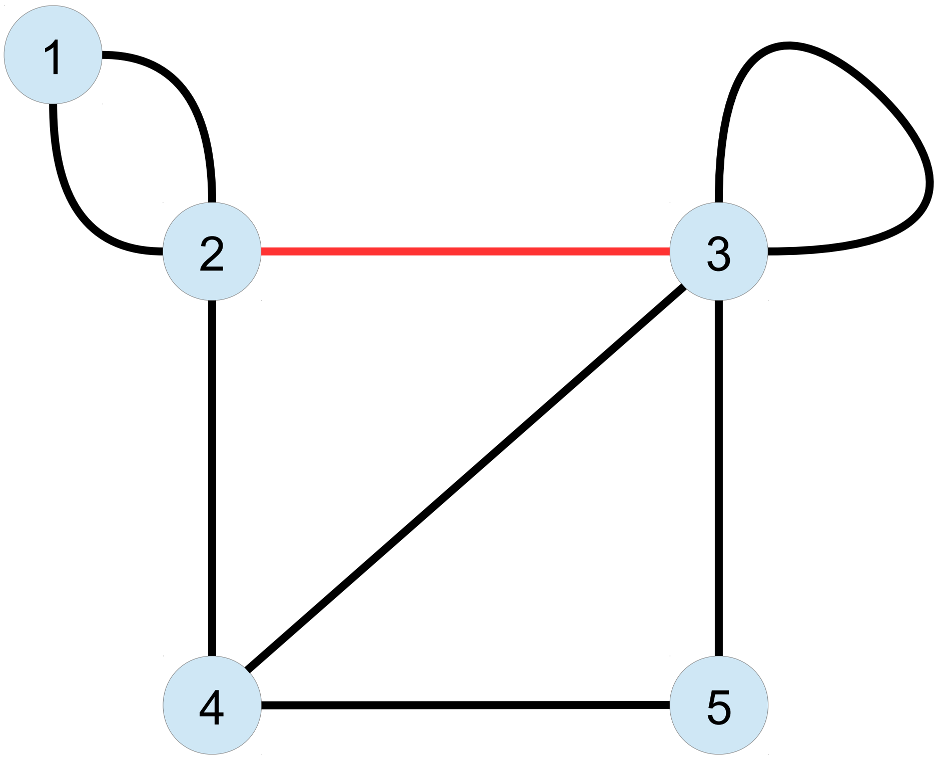
- 2 is a node
- {1,2} is an edge
- Node 2 has **degree** 4

Paths



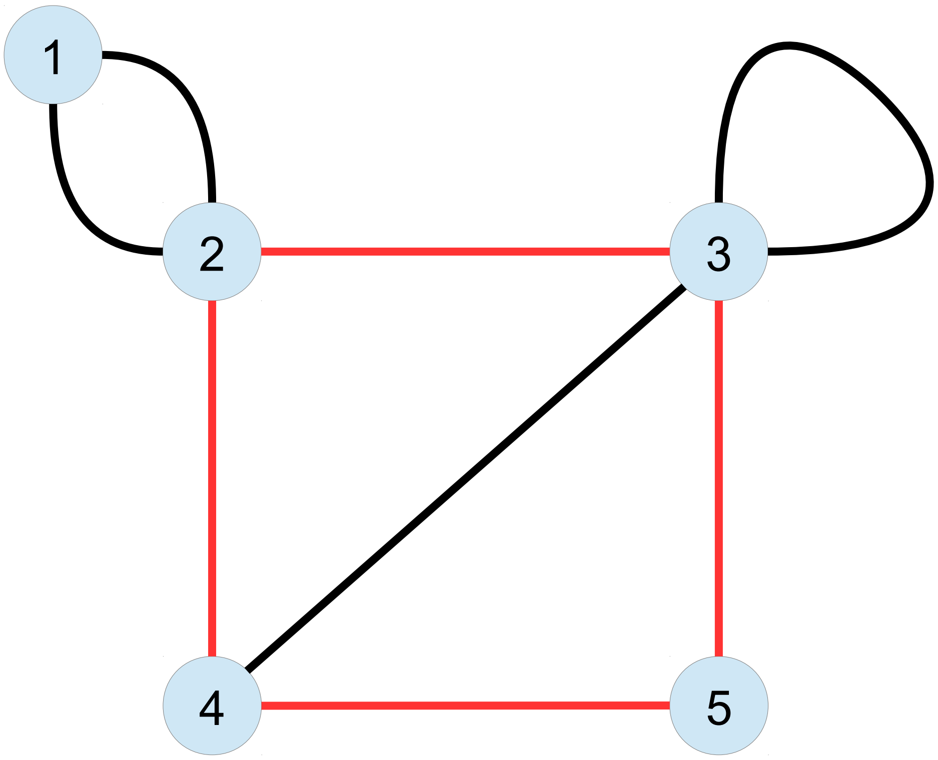
{2,3} , {3,5} ,
{5,4} , {4,2}
is a path among
node 2 and node 3

Shortest Paths (geodesics)



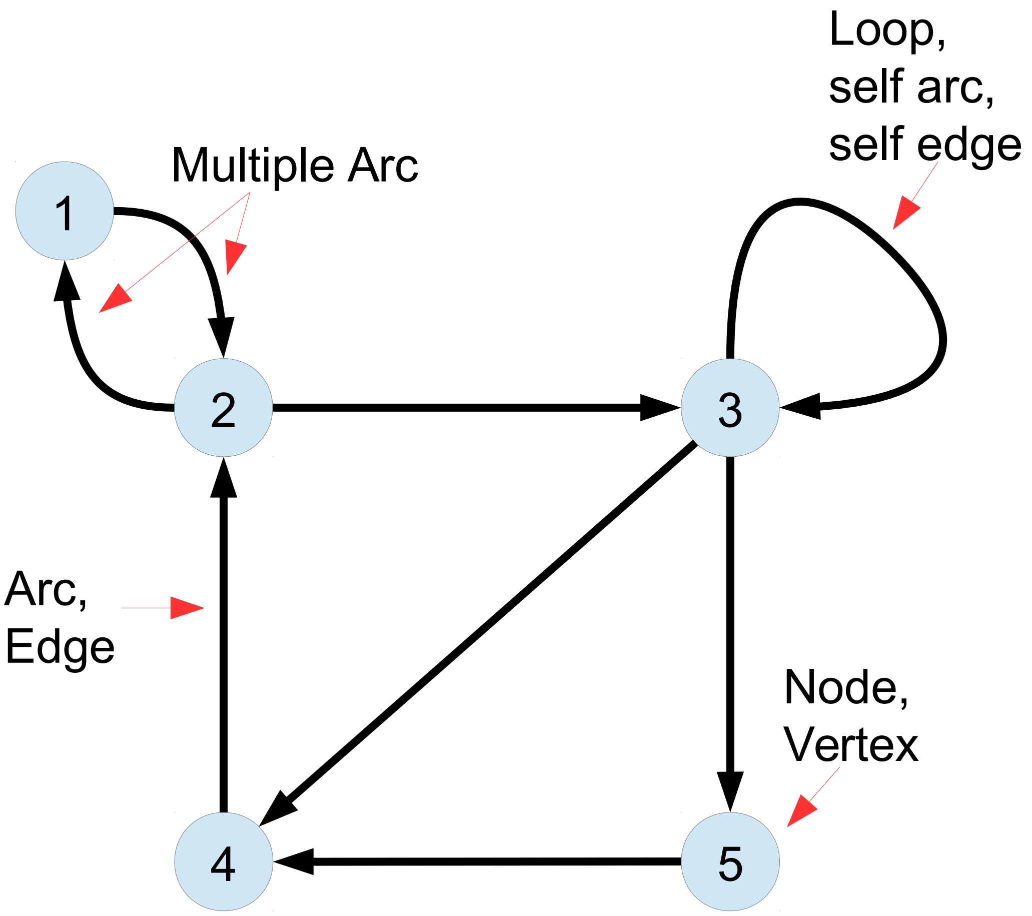
{2,3} is the shortest path among node 2 and node 3

Cycles



{2,4} , {4,5} ,
{5,3} , {3,2}
is a cycle

Digraphs



G = (V,E)

V = vertex set

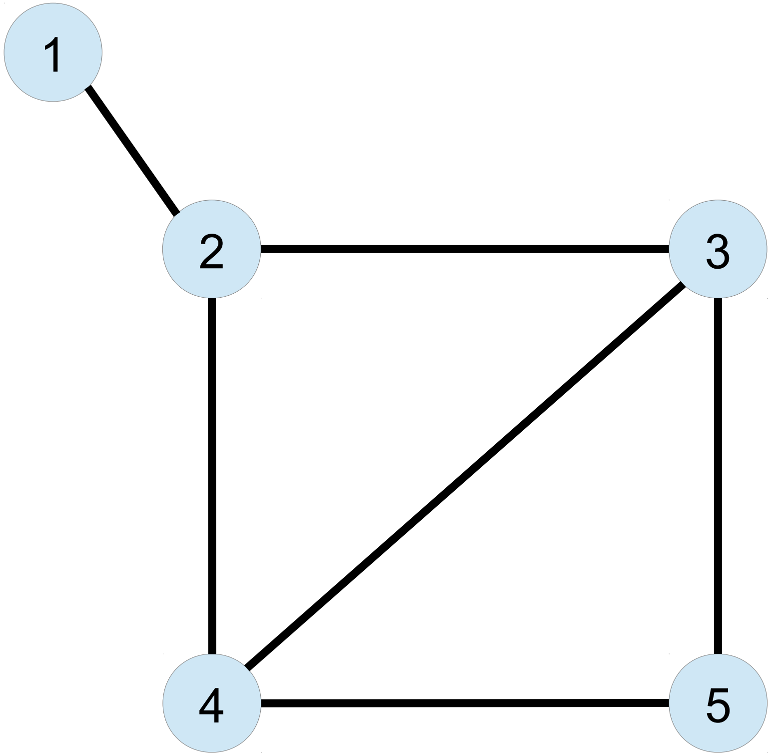
E = edge set

- (4,5) is an edge
- (5,4) is NOT an edge
- Node 4 has **indegree** 2
- Node 4 has **outdegree** 1

Kinds of graphs

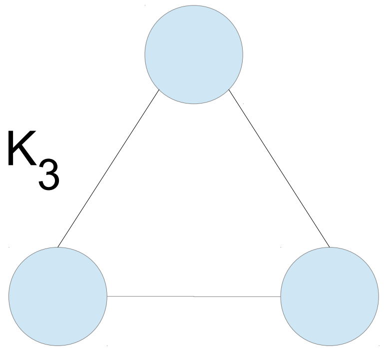
- **simple** : contains no multi–edges and no self–loops,
- **bipartite** : there exists a non–trivial partition $V = V^+ \cup V^-$ of the vertex set with $V^+ \cap V^- = \emptyset$ such that for all edges $(u, v) \in E$ (or $\{u, v\} \in E$) $u \in V^+$ and $v \in V^-$
- **complete** : there exists one and only one edge between every pair of distinct nodes,
- the **underlying graph** is the undirected version of a directed graph
- An **undirected graph** is **connected** if every node can be reached from every other node
- A **digraph** is **strongly connected** if every node can be reached from every other node; if it only holds for its underlying **undirected graph**, it is said to be **weakly connected**
- **k–regular** if all its nodes have the same degree (k)
- $G' = (V', E')$ is a **subgraph** of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$

Adjacency Matrix – Simple Graphs

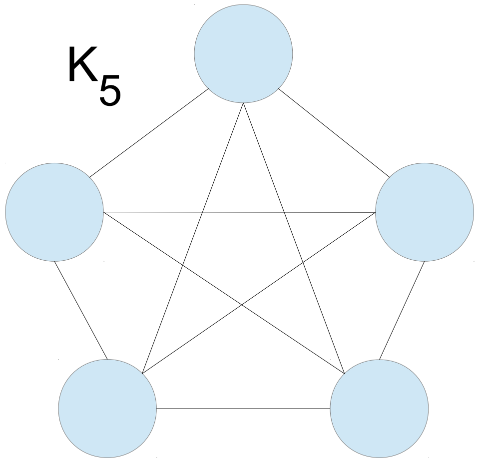


$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Complete graphs / Cliques

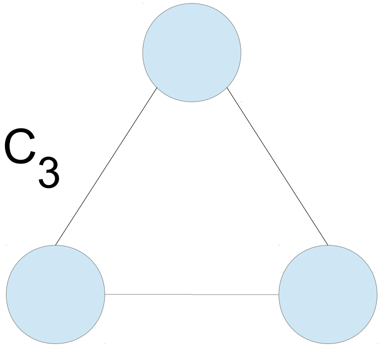


K_n is the **complete** graph of order **n**

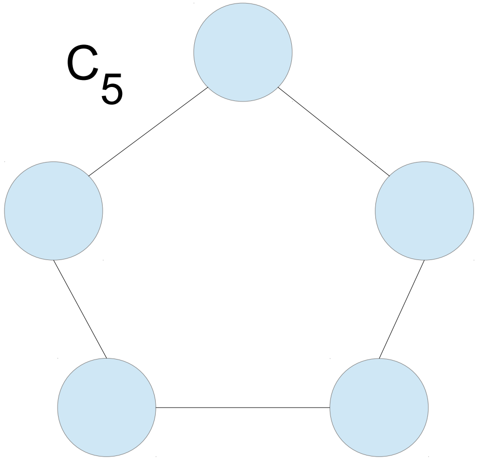


$$A(K_5) = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Cycle graphs

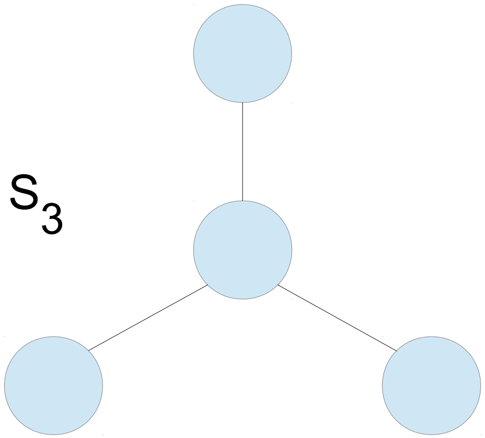


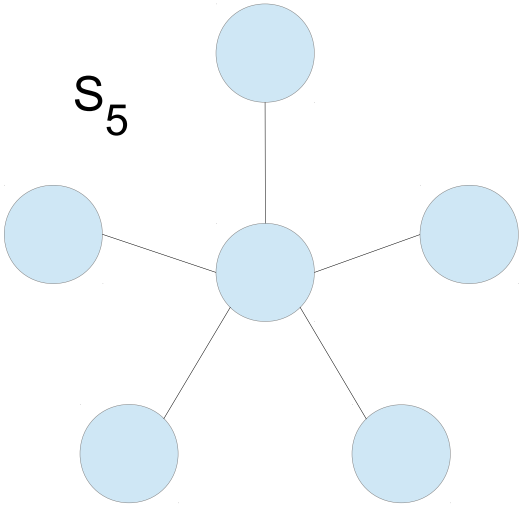
C_n is the **cycle** graph of order **n** (also called **n-cycle**)



$$A(C_5) = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Star graphs

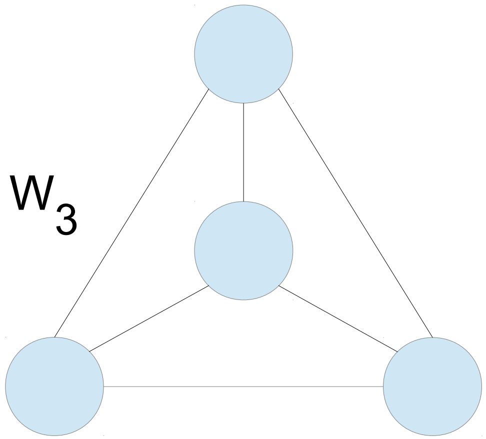


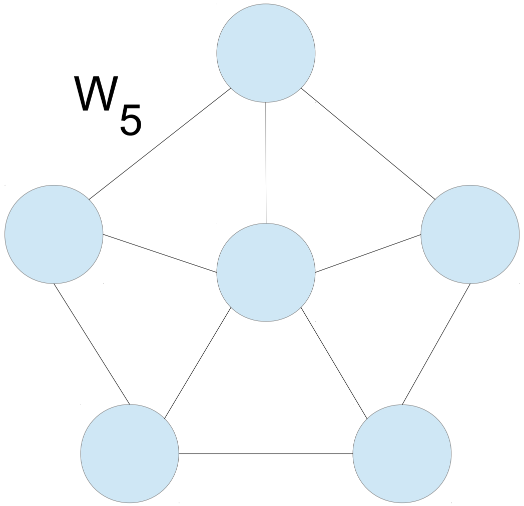


$\mathbf{S_n}$ is the **star** graph of order **n** (also called **n-star**)

$$A(S_5) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Wheel graphs

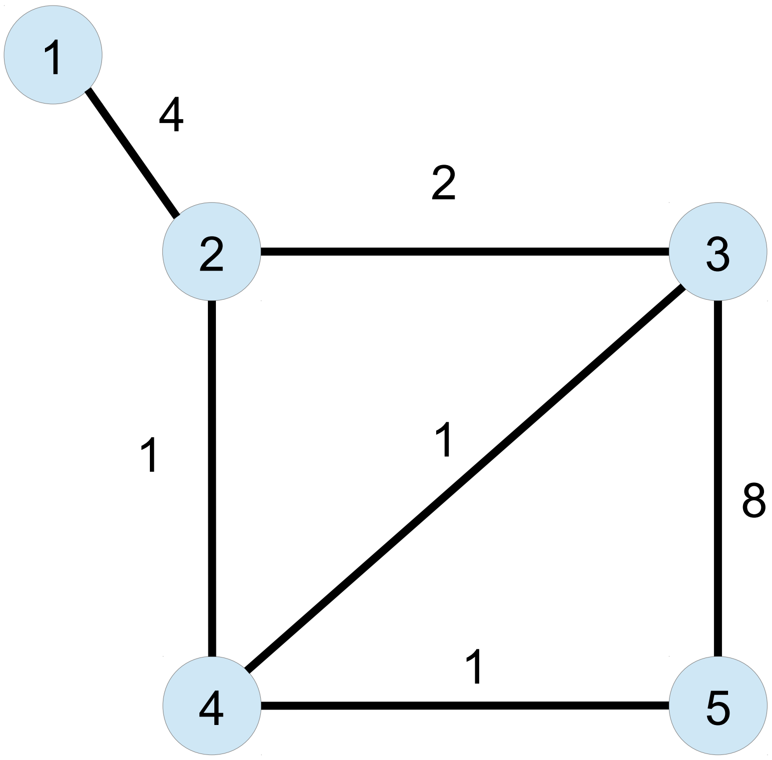




W_n is the **wheel** graph of order **n** (also called **n-wheel**)

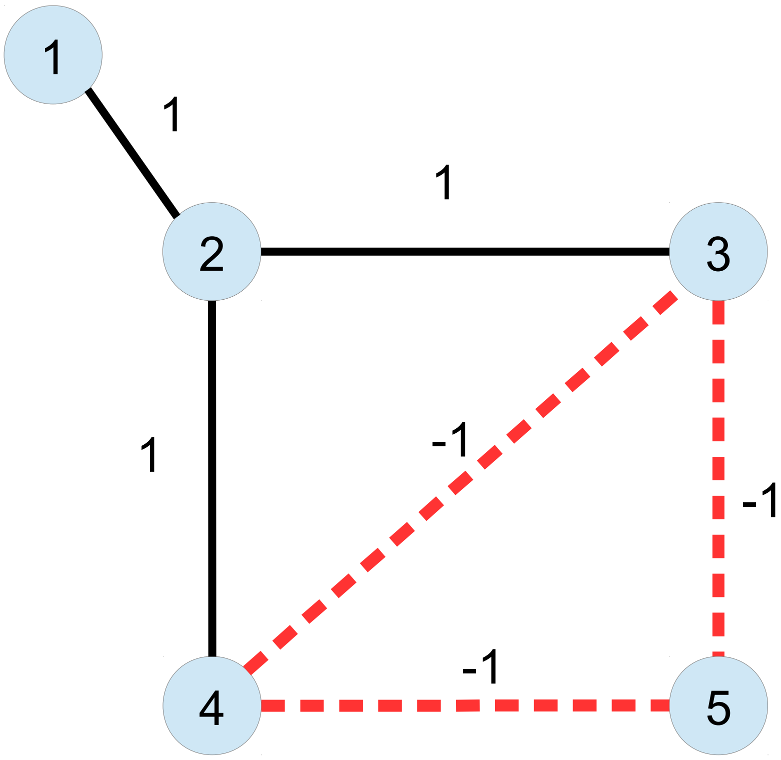
$$A(W_5) = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Adjacency Matrix – Weighted Graphs



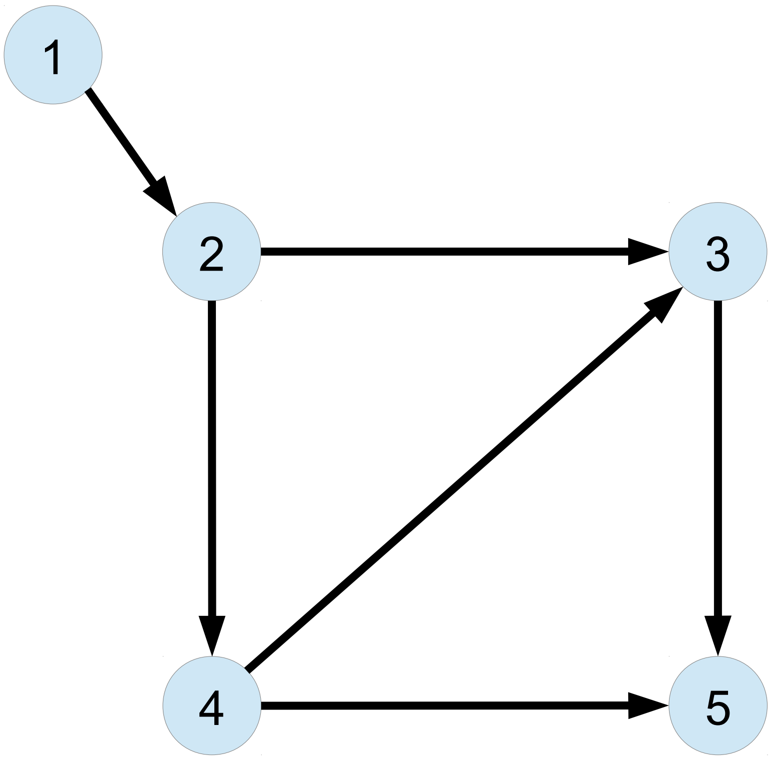
$$\begin{bmatrix} 0 & 4 & 0 & 0 & 0 \\ 4 & 0 & 2 & 1 & 0 \\ 0 & 2 & 0 & 1 & 8 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 8 & 1 & 0 \end{bmatrix}$$

Adjacency Matrix – Signed Graphs



$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & -1 & -1 & 0 \end{bmatrix}$$

Adjacency Matrix – Simple Digraphs



$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

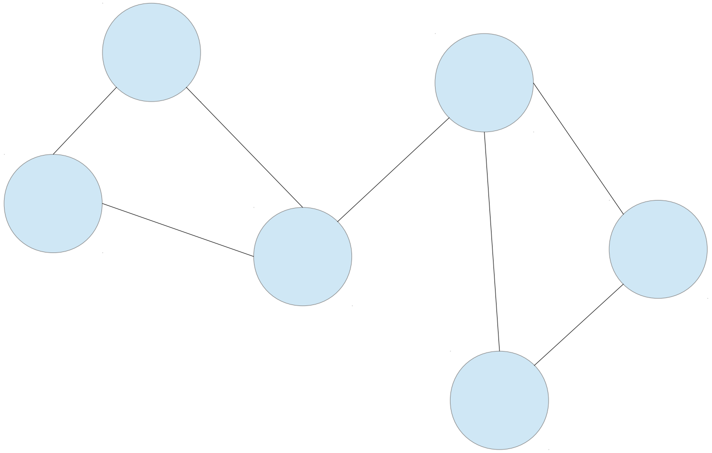
Adjacency matrix - properties

- $(A^k)_{ij}$ counts the number of paths of length k among node i and node j
- The (i,j) element of $(1+A)^k$ is non-zero if there exit a path of length <= k among node i and node j
- A graph is irreducible IFF its adjacency matrix A is irreducible IFF exists a k for which $(1+A)^k$ is **strictly positive**

Adjacency matrix - connectedness

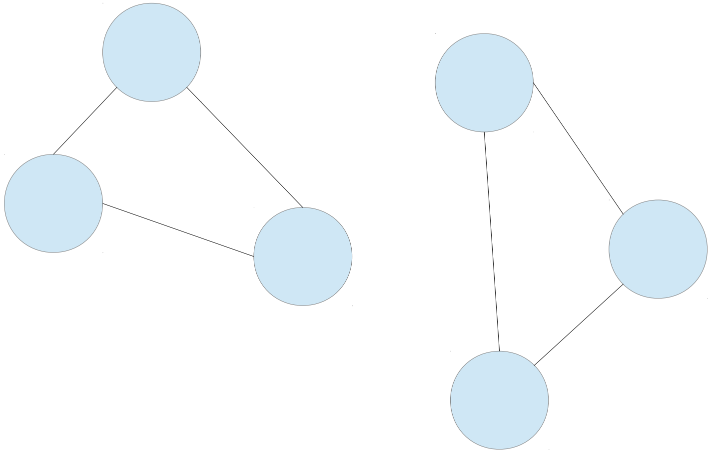
Irreducible Graph

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$



Reducible Graph

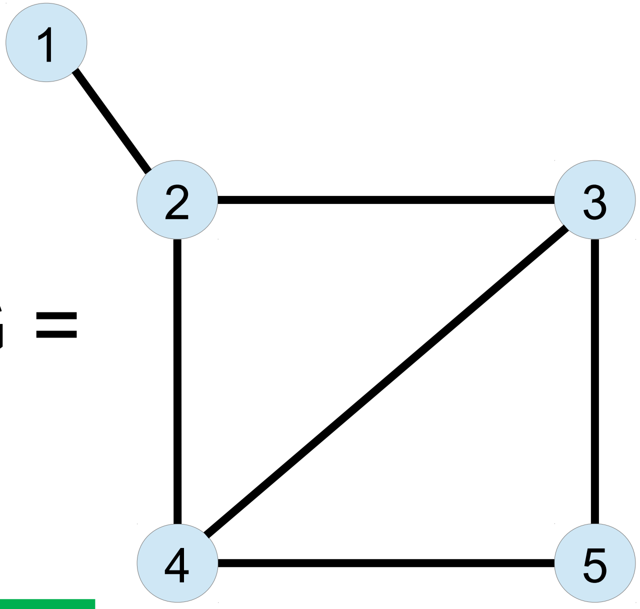
$$A = \begin{bmatrix} B & 0 \\ 0 & D \end{bmatrix}$$



Laplacian

- D degree matrix
  - $D_{ii} = d_i$ = degree of node i
  - $D_{ij} = 0$
- $\mathfrak{L} = D - A$ is the **Laplacian Matrix**

$G =$

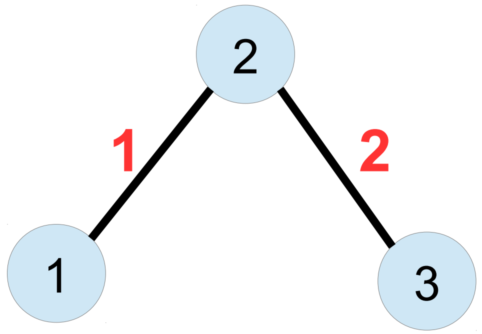


$$L = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 3 & -1 & -1 & 0 \\ 0 & -1 & 3 & -1 & -1 \\ 0 & -1 & -1 & 3 & -1 \\ 0 & 0 & -1 & -1 & 2 \end{bmatrix}$$

Incidence Matrix

- V = {1,2,3}
- E = {**1**,**2**} with
  - edge **1** = {1,2}
  - Edge **2** = {2,3}
- The **incidence matrix** B is defined on V x E
  - Let **e** = {i,j}
  - $B_{i\mathbf{e}} = 1$
  - $B_{j\mathbf{e}} = -1$
  - $B_{k\mathbf{e}} = 0$

G =



$$B = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 1 & -1 & 0 \\ 2 & 1 & 0 & -1 \end{array}$$

Incidence matrix and Laplacian

- B^T is the equivalent of the nabla operator:

$$\text{if } \mathbf{e} = \{i,j\} \text{ , } (B^T f)_{\mathbf{e}} = f_i - f_j$$

- $\mathfrak{L} = B\,B^T$
- $\mathfrak{L}$ is the equivalent of the Laplacian operator
- $\mathfrak{L}$ acts as a diffusion operator

$$(\mathfrak{L} f)_i = d_i\, f_i - \Sigma_j\, A_{ij}\, f_j$$

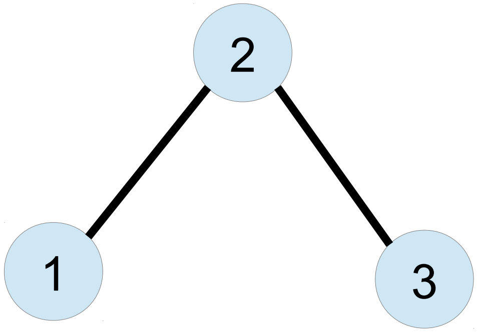
- $\mathfrak{L}$ is the equivalent of an elastic energy operator

$$f^T \mathfrak{L}\, f = \Sigma_{ij}\, A_{ij}\, (f_i - f_j)^2$$

Distance Matrix

- The distance matrix δ of a graph G of n nodes is an n x n matrix with elements

G =



δ_{ij} = minimum distance among node i and node j

$$\delta = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$$