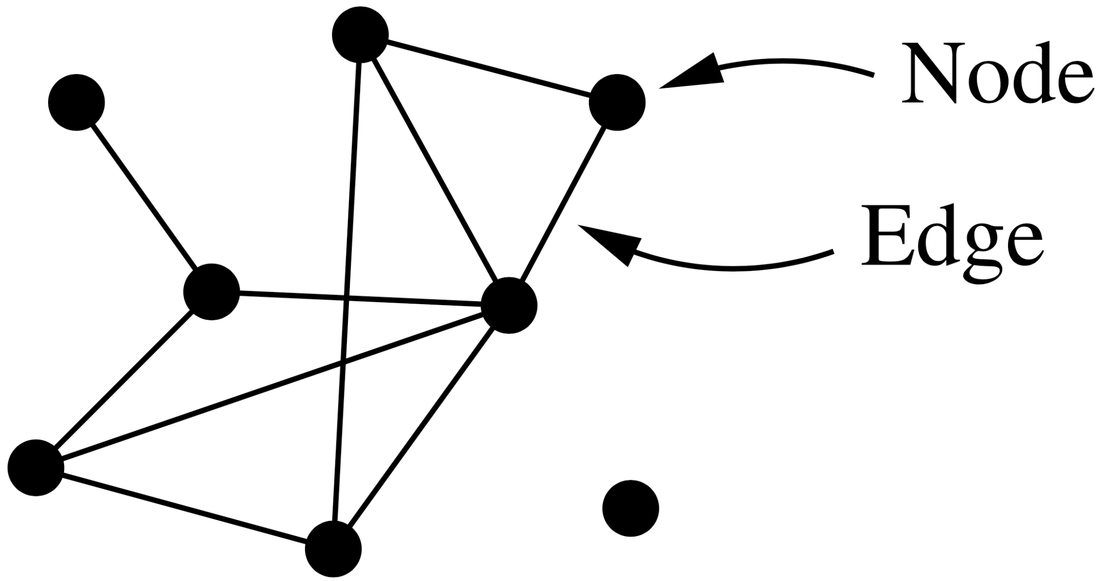
Mathematics of Networks

Networks and their representation



A network—also called a graph in the mathematical literature—is a collection of nodes (or vertices) joined by edges. Nodes and edges are also called *sites* and *bonds* (e.g., in physics and biology) or *actors* and *ties* (e.g., in sociology).

Types of Networks

Generally, the common notation in the mathematical literature for the number of nodes in a network is n and the number of edges is m.

The simplest networks have at most *a* **single edge between any pair of nodes**. When there are more than one edge between the same nodes, the set of edges between two notes are called a **multi-edge**. Usually, nodes do not have edges to themselves, but when this happens, those nodes are called **self-edges** or **self-loops**.

Networks that **neither have self-edges nor multi-edges** are called a **simple network** or **simple graph**.

A network **with multi-edges** is called a **multigraph**.

The Adjacency Matrix

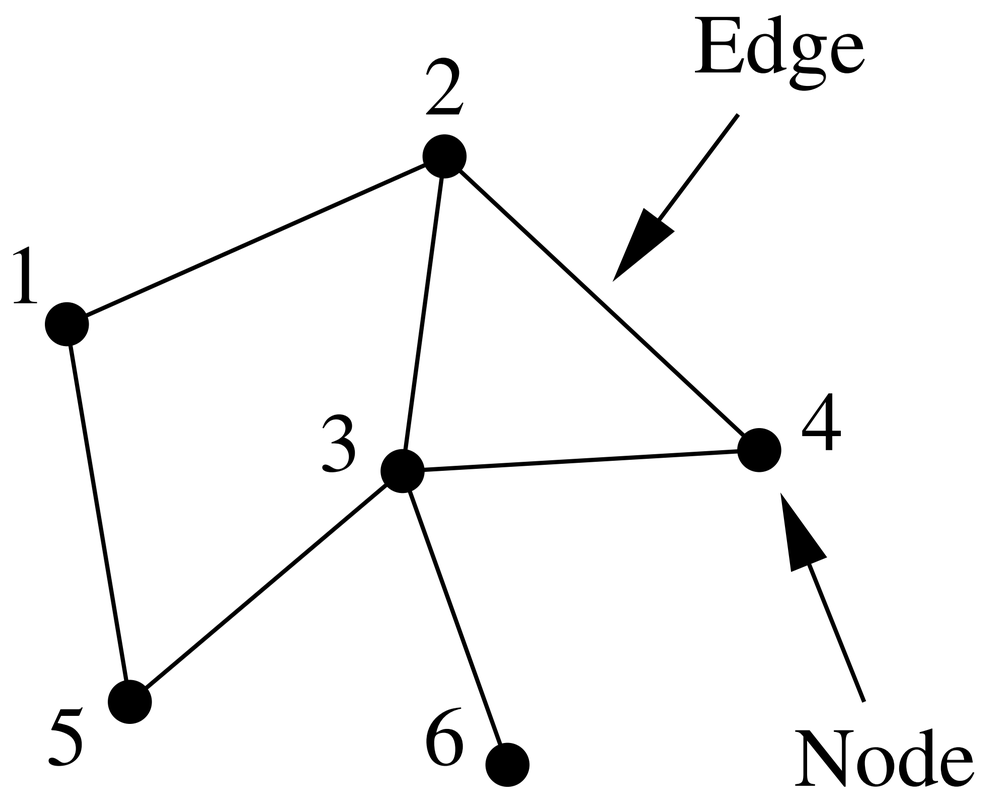
The fundamental mathematical representation of a network is the **adjacency matrix**.

Consider an undirected simple network with n nodes, labelled $1 \ldots n$ —so that labels refer to the nodes unambiguously.

The adjacency matrix A of the network is defined to be the $n \times n$ **matrix** with elements A_{ij} such that

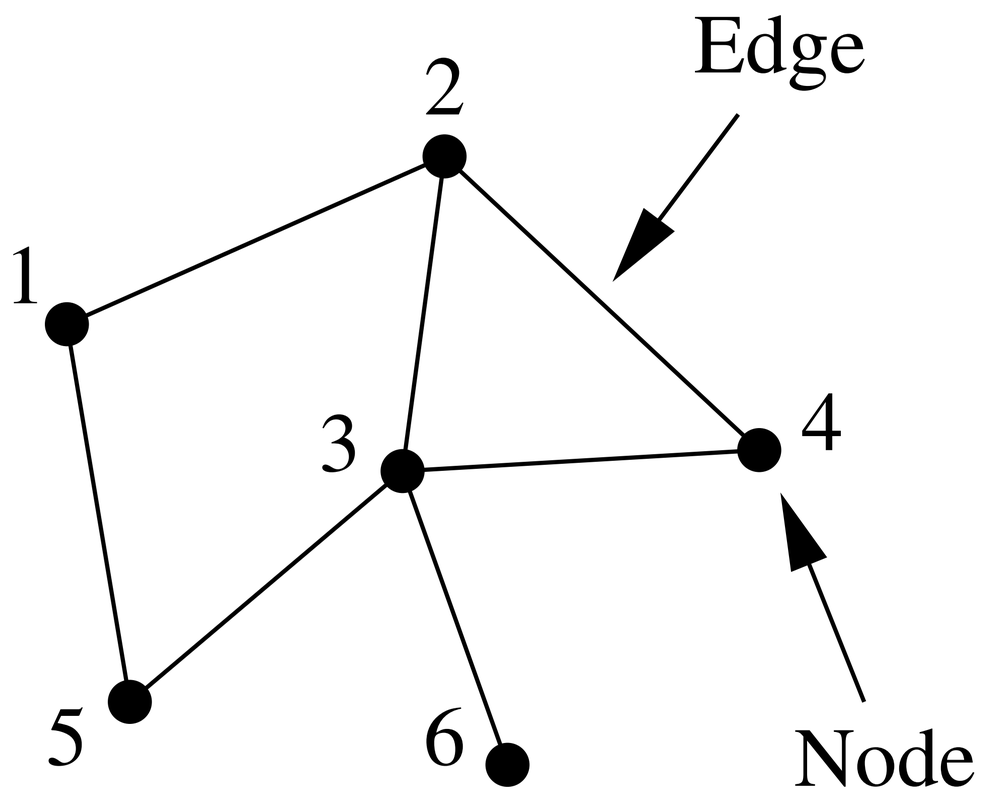
$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between nodes } i \text{ and } j \\ 0 & \text{otherwise} \end{cases}$$

The Adjacency Matrix



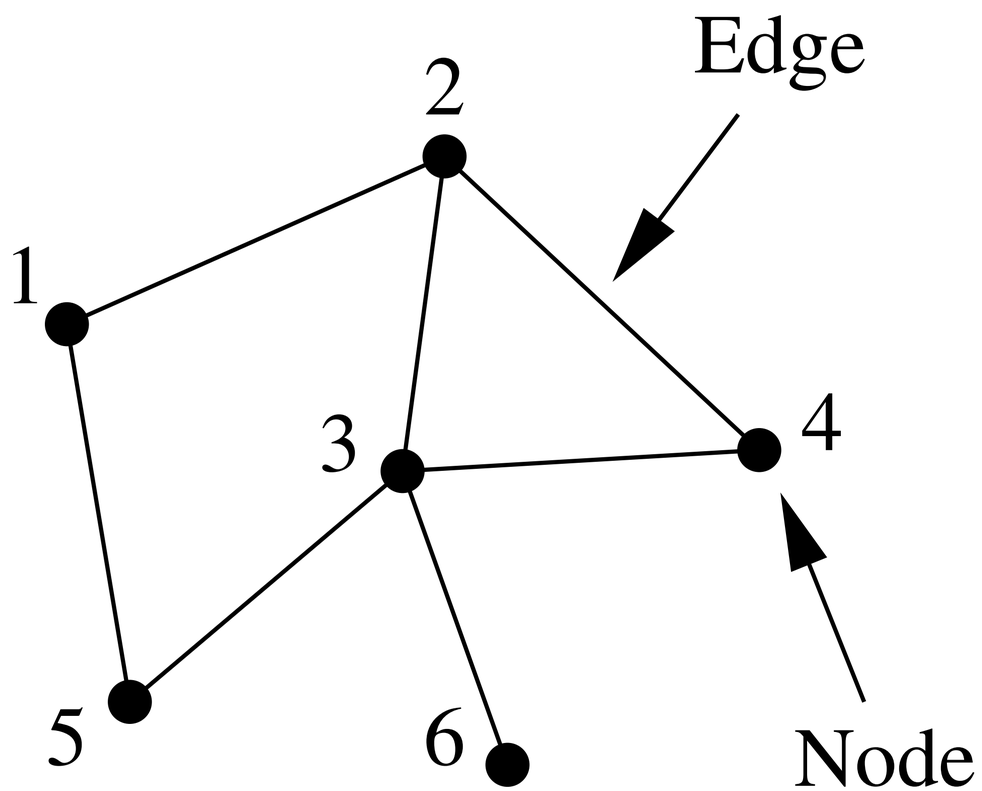
$$\mathbf{A} = \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array}$$

The Adjacency Matrix



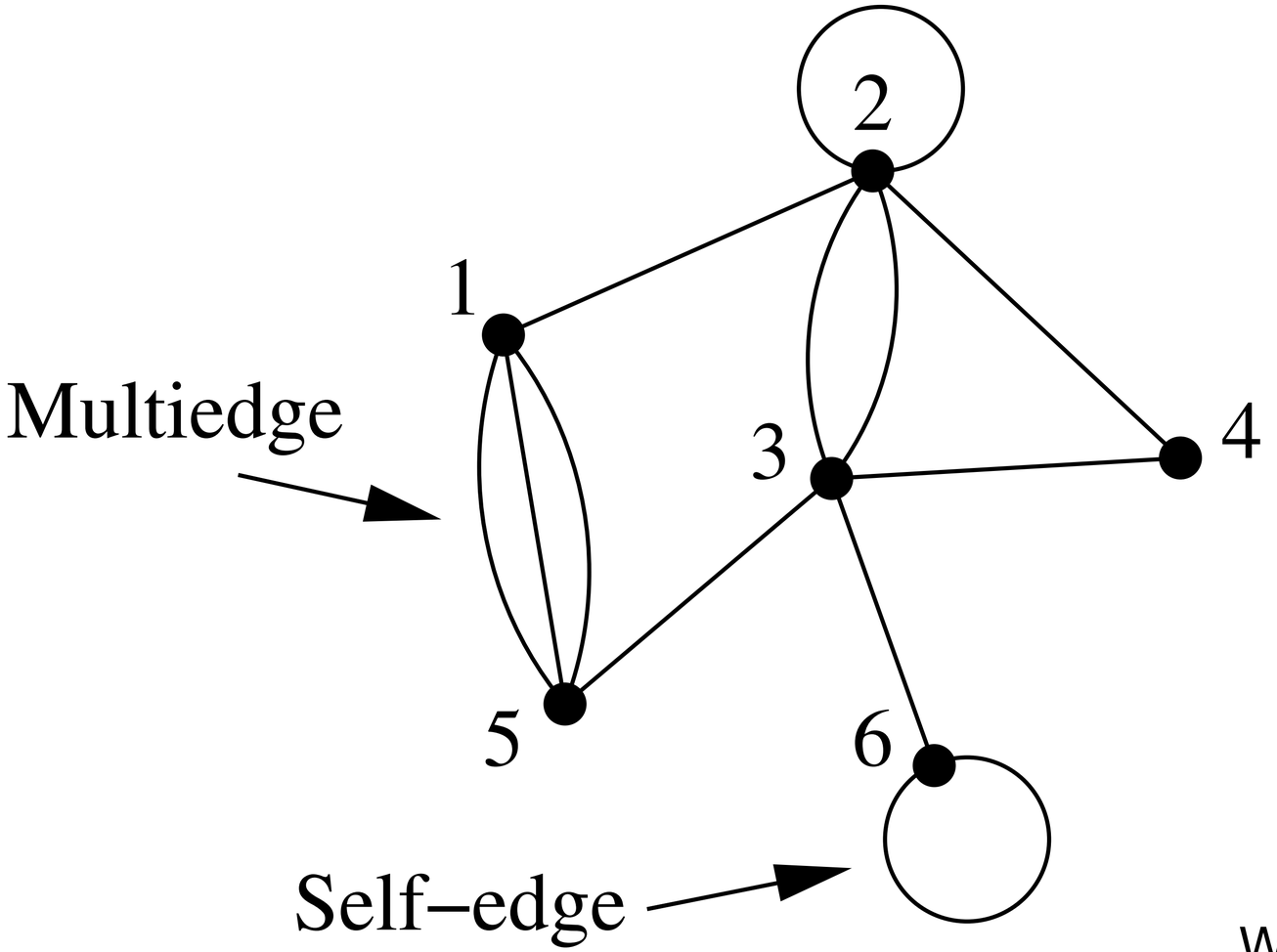
1	2	3	4	5	6	
0	1	0	0	1	0	1
1	0	1	1	0	0	2
0	1	0	1	1	1	3
0	1	1	0	0	0	4
1	0	1	0	0	0	5
0	0	1	0	0	0	6

The Adjacency Matrix



	1	2	3	4	5	6	
	0	1	0	0	1	0	1
	1	0	1	1	0	0	2
	0	1	0	1	1	1	3
	0	1	1	0	0	0	4
	1	0	1	0	0	0	5
	0	0	1	0	0	0	6

The Adjacency Matrix



	1	2	3	4	5	6	
	0	1	0	0	3	0	1
	1	2	2	1	0	0	2
	0	2	0	1	1	1	3
	0	1	1	0	0	0	4
	3	0	1	0	0	0	5
	0	0	1	0	0	2	6

Why 2 and not 1?
(hint: degenerate case of symmetry on the diagonal)

Weighted Networks

In some situations it is useful to represent edges as having a strength, weight, or value, usually a real number, e.g.,

- edges on the **Internet** where weights represent the **amount of data flowing** along them or their bandwidth;
- in a **food web**, where predator–prey interactions might have **weights measuring total energy flow** between prey and predator;
- in a **social network** where connections might have weights representing **frequency** of contact between actors.

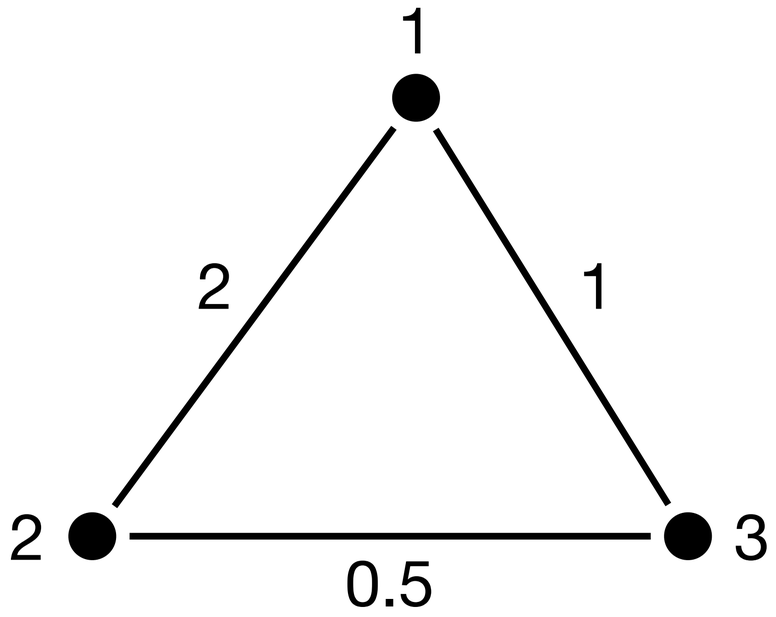
Weighted Networks

Such ***weighted*** or ***valued networks*** can be represented mathematically by an adjacency matrix with the elements A_{ij} equal to the weights of the corresponding connections

$$\begin{array}{cc} \begin{array}{ccc} 1 & 2 & 3 \end{array} & \\ \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 0.5 \\ 1 & 0.5 & 0 \end{pmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \end{array}$$

Weighted Networks

Such ***weighted*** or ***valued networks*** can be represented mathematically by an adjacency matrix with the elements A_{ij} equal to the weights of the corresponding connections



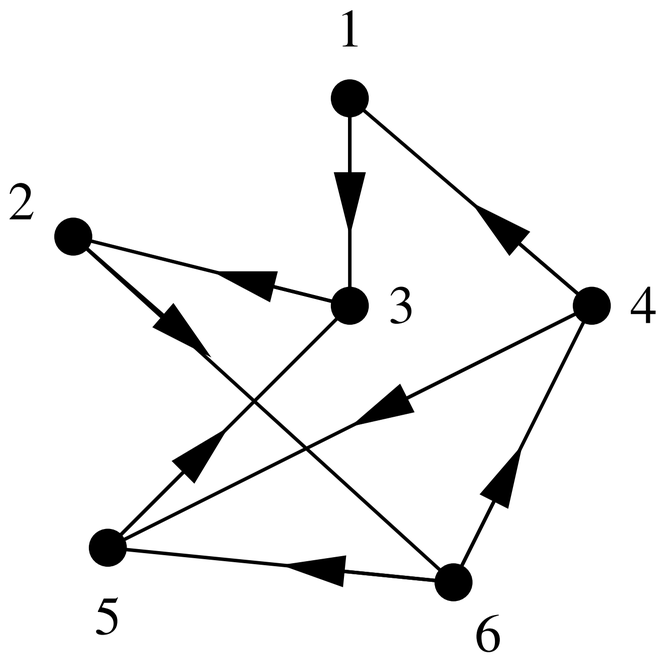
$$\begin{array}{c} \begin{array}{ccc} \color{green}{1} & \color{green}{2} & \color{green}{3} \end{array} \\ \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 0.5 \\ 1 & 0.5 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \color{green}{1} \\ \color{green}{2} \\ \color{green}{3} \end{array}$$

Types of Networks · Directed Graphs

Directed networks or ***directed graphs*** are networks in which each edge has a **direction**, pointing ***from*** **one node** ***to*** **another**—thus called **directed edges** or sometimes ***arcs***.

Examples of directed graphs are web-pages, where hyperlinks run in one direction from one web page to another; food webs, where energy flows from prey to predator; and citation networks, where citations point from one work to another.

Types of Networks • Directed Graphs

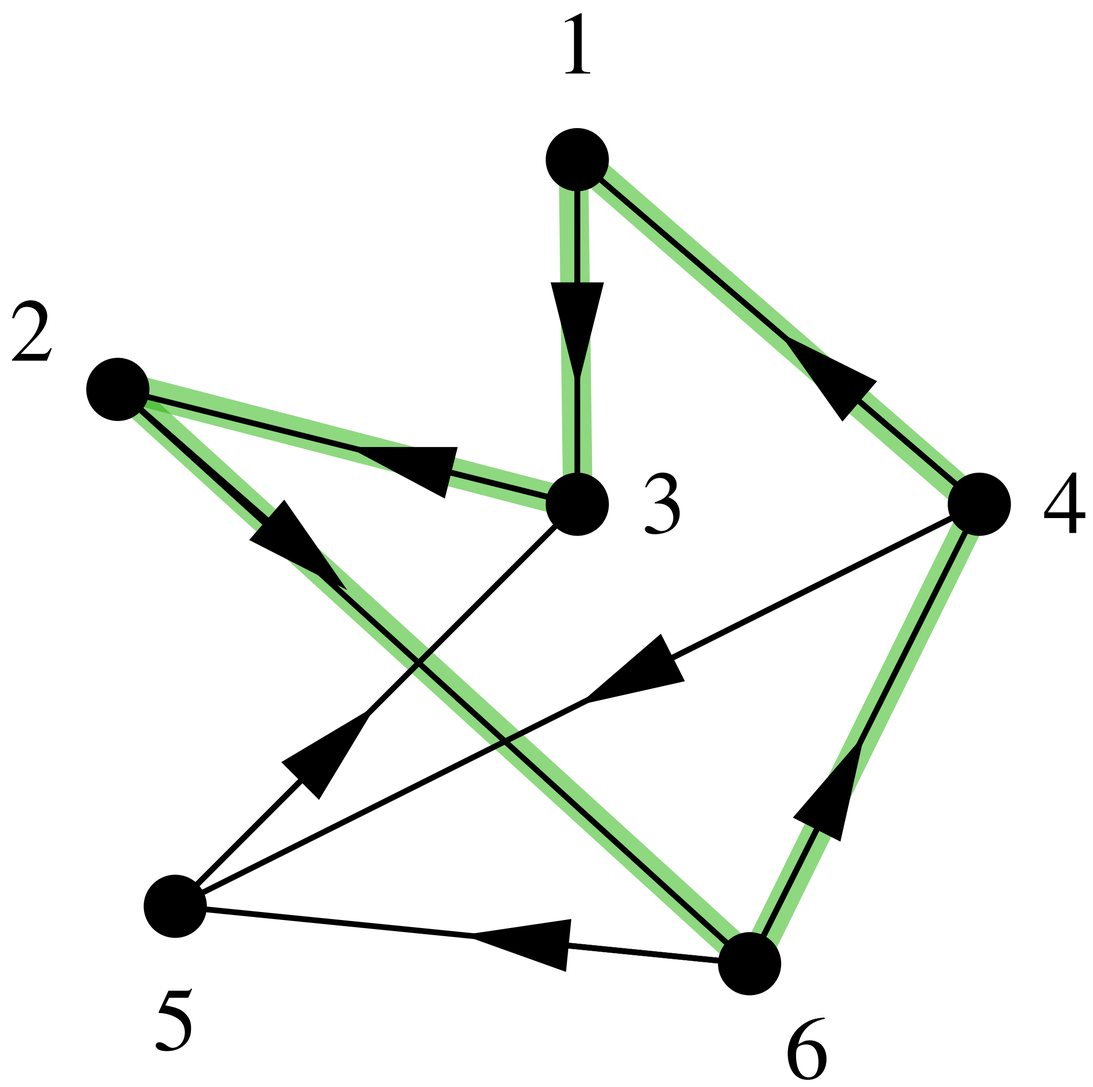


$$A_{ij} = \begin{cases} 1 & \text{if there is an edge } \textit{from } j \textit{ to } i, \\ 0 & \text{otherwise.} \end{cases}$$

$$A_{14} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \overset{\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array}}{\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}}$$

j (columns) i (rows)

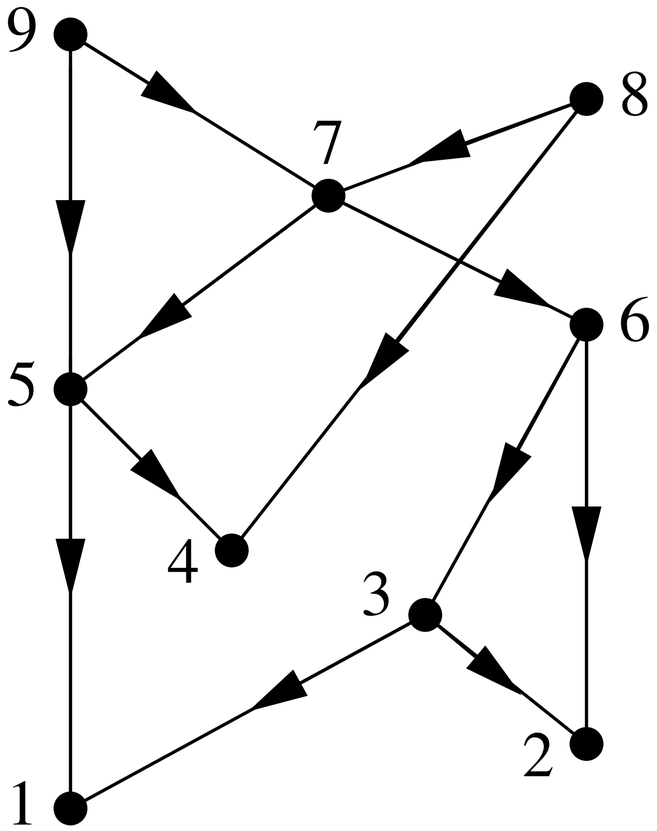
Types of Networks · Directed Graphs



A ***cycle*** in a directed network is a **closed loop of edges** with the *arrows on each of the edges pointing the same way around the loop*. A self-edge—an edge connecting a node to itself—counts as a cycle.

Types of Networks • Directed Acyclic Graphs

Directed networks that have no cycles (including self-edges) are called ***acyclic*** networks. They are frequently abbreviated with the acronym DAG.

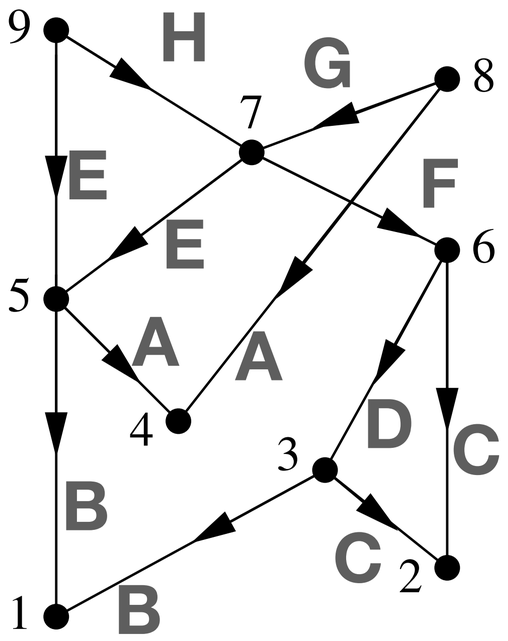


If a network is acyclic, it can be drawn with all edges pointing downward.

The proof that this is true, provides us with a method for determining whether a given network is acyclic!

Types of Networks • Directed Acyclic Graphs

Directed networks that have no cycles (including self-edges) are called ***acyclic*** networks.

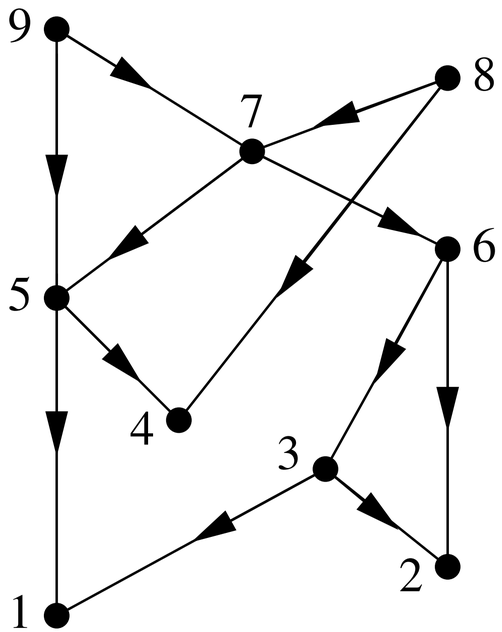


Lemma. In a DAG, with n nodes, there exists at least one node with no outgoing edges. *Proof***.** By contradiction, we assume that there are none of such nodes. Then, we have the absurdum because we can construct a cycle: we can always find an outgoing edge from any node and thus, even if we walked n edges, at the $n + 1^{th}$ step we can only select one of the nodes we already visited, creating a cycle.

Theorem. *If a network is acyclic, it can be drawn with all edges pointing downward*. **Proof.** The proof is by induction on the number of nodes. From the *Lemma* above, we know we can find at least one of those nodes that have only ingoing edges. First, we remove one of those nodes and its ingoing edges from the network, and draw them. Then, we iteratively apply the previous step $n - 1$ times, removing one of the outgoing-less nodes and drawing it in the following way: if in the original network the node had some outgoing edges to some of the nodes already drawn, we position it on top of those nodes and draw its outgoing edges, otherwise we draw it on top of the current top level.

Types of Networks • Directed Acyclic Graphs

Directed networks that have no cycles (including self-edges) are called ***acyclic*** networks.



The adjacency matrix A of a downward labelled DAG (whose element A_{ij} records the presence of an edge *from j to i*) **has all its non-zero elements above the diagonal**—also called "upper triangular".

	1	2	3	4	5	6	7	8	9	
	0	0	1	0	1	0	0	0	0	1
	0	0	1	0	0	1	0	0	0	2
	0	0	0	0	0	1	0	0	0	3
	0	0	0	0	1	0	0	1	0	4
	0	0	0	0	0	0	1	0	1	5
	0	0	0	0	0	0	1	0	0	6
	0	0	0	0	0	0	0	1	1	7
	0	0	0	0	0	0	0	0	0	8
	0	0	0	0	0	0	0	0	0	9

Types of Networks • Bi- and multi-modal Graphs

A ***bimodal network***—*a*lso called *bipartite*—is a network with **two kinds of nodes and edges** that **run only between nodes of different kinds**.

Bipartite networks are e.g., used to represent the membership of a set of people or objects in groups of some kind: people are represented by one set of nodes, groups by the other, and edges join people to groups in which they belong. For example, a bipartite network can represent actors and films, where the edges connect actors to the films in which they appear. Edges cannot connect actors to other actors or films to other films.

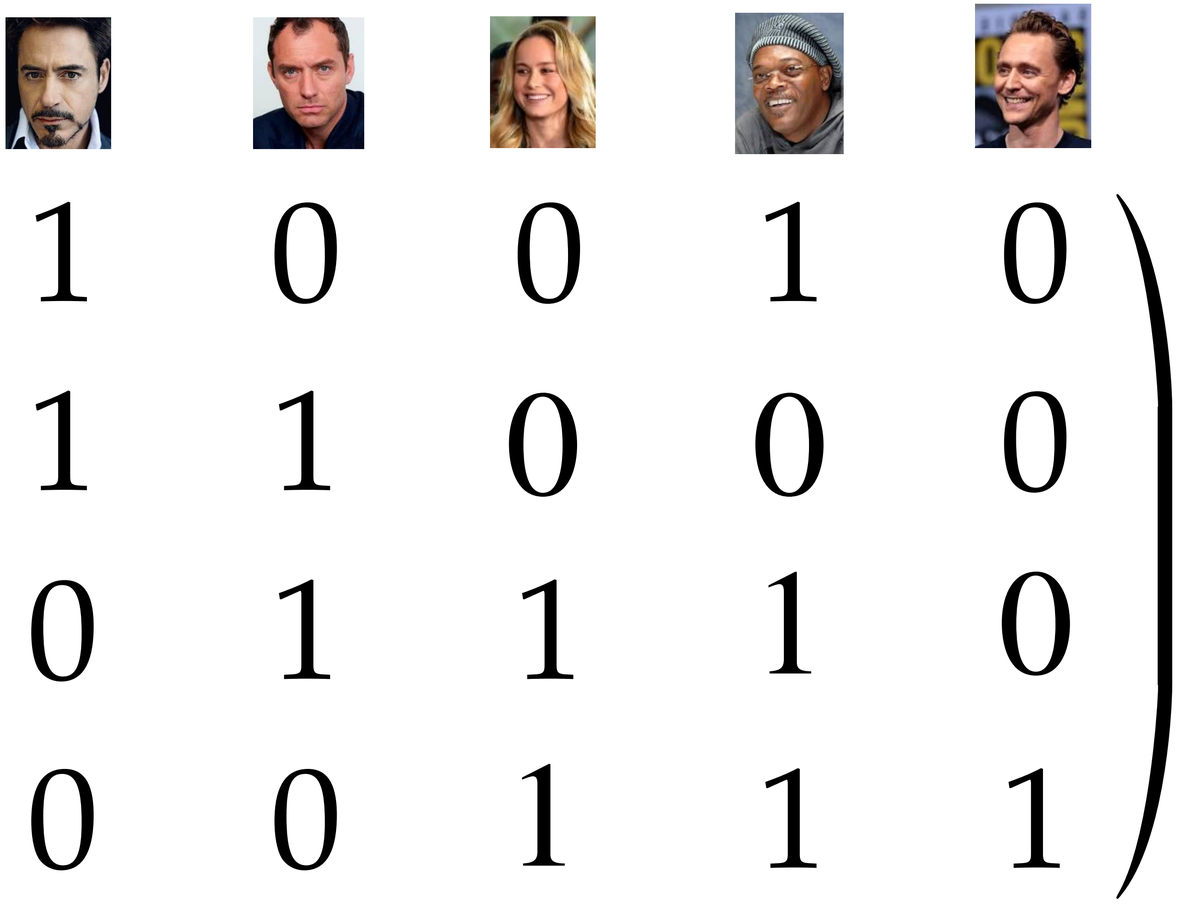
Types of Networks • Bi- and multi-modal Graphs

The equivalent of the adjacency matrix for an (undirected unweighted) **bipartite graph** is a rectangular matrix called the *incidence matrix*. In the incidence matrix, we count as n the number of nodes in the network and g the numbers of groups, then the incidence matrix has dimension $n \times g$ where its elements are described as

$$B_{ij} = \begin{cases} 1 & \text{if item } j \text{ belongs in group } i \\ 0 & \text{otherwise} \end{cases}$$

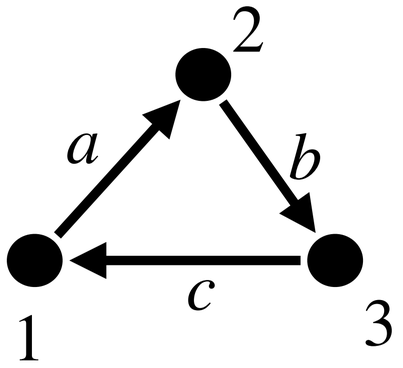
	1	2	3	4	5
A	1	0	0	1	0
B	1	1	1	1	0
C	0	1	1	0	1
D	0	0	1	1	1

Types of Networks • Bi- and multi-modal Graphs

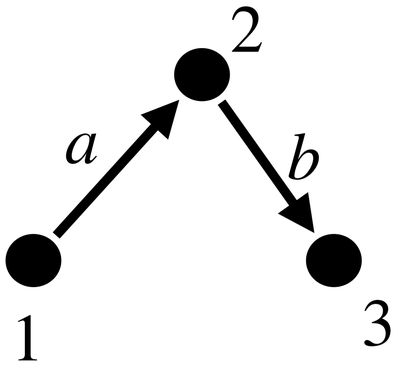


$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Walks, Paths, and Components



walk : 1,2,3,1,2
length : 4



path : 1,2,3
length : 2

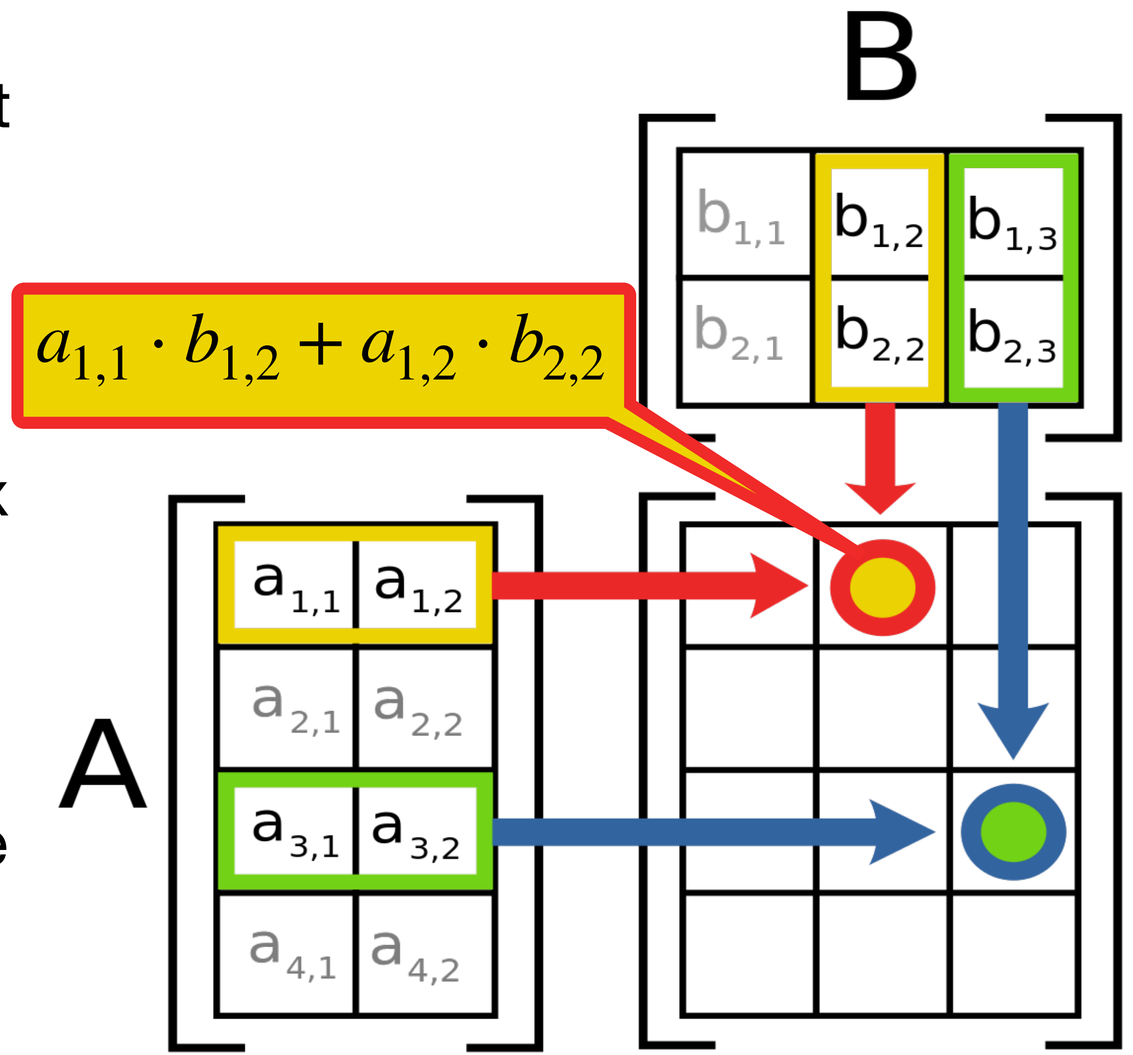
A ***walk*** in a network is a sequence of nodes such that every consecutive pair of nodes in the sequence is connected by an edge. The ***length*** of a walk in a network is the number of edges traversed along the walk (not the number of nodes), counted separately as they are traversed.

Walks that do not intersect themselves are called ***paths***.

Walks, Paths, and Components • Background

Matrix multiplication is an operation that produces a matrix from two matrices. The operation is written AB (shorthand for $A \cdot B$), where A is the "left" matrix and B the right one. To be applicable, the number of columns of the left matrix must be equal to the number of rows of the right matrix—in that case, they are called "conformable" to multiplication.

Let $C = AB$, then the element C_{ij} of the product is defined as $C_{ij} = \sum_{k=1}^{n} A_{ik} \cdot B_{kj}$



Walks, Paths, and Components • Background

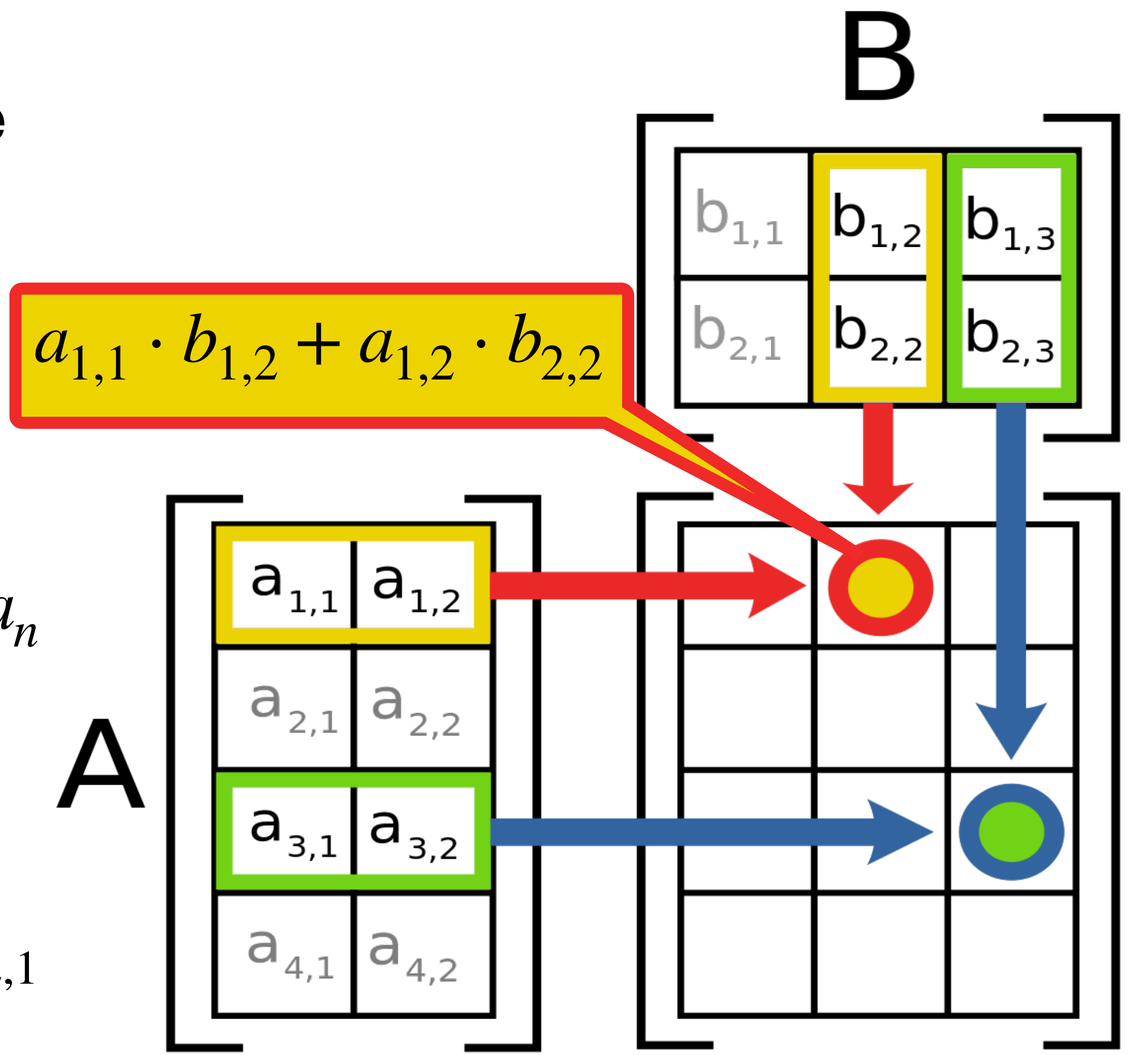
Let $C = AB$, then the element C_{ij} of the product is defined as $C_{ij} = \sum_{k=1}^{n} A_{ik} \cdot B_{kj}$

where

$$\sum_{i=m}^{n} a_i = a_m + a_{m+1} + a_{m+2} + \cdots + a_{n_1} + a_n$$

So, e.g.,

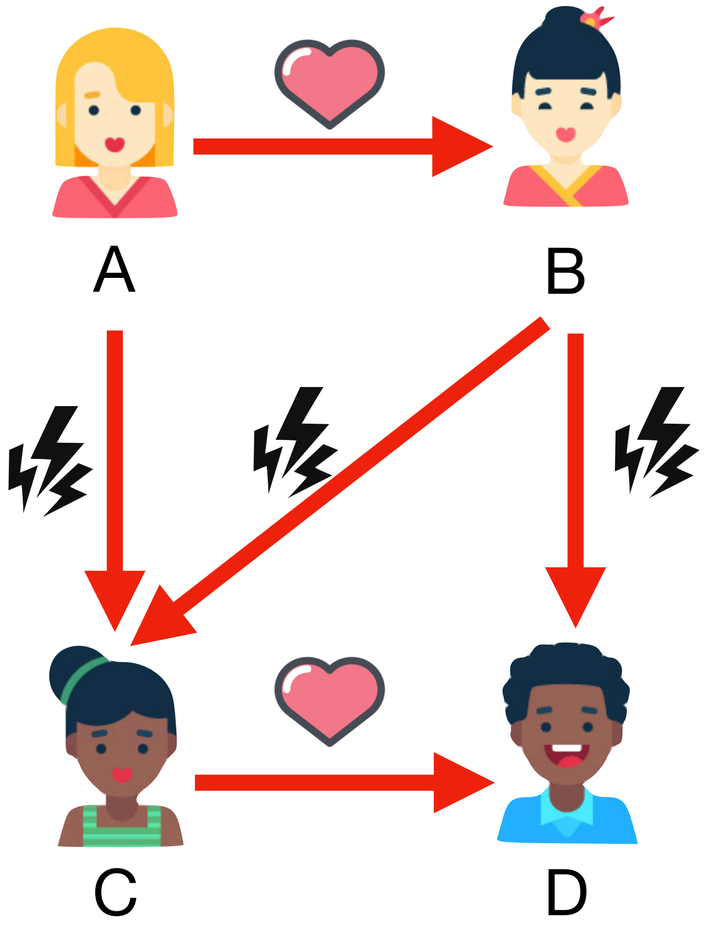
$$C_{1,1} = \sum_{k=1}^{2} A_{1,k} \cdot B_{k,1} = A_{1,1} \cdot B_{1,1} + A_{1,2} \cdot B_{2,1}$$



Walks, Paths, and Components • Background

Matrix multiplication is not commutative, so AB and BA are not necessarily the same (and possibly not even multiplication-conformable).

As an example, let us have two matrices L, representing love relations, and H representing hate ones. The product LH represents the "hated among the lovers", so that the ij^{th} cell of LH indicates the *lovers of i who are hated by j*. On the other hand, HL represents the "haters with lovers", so that the ij^{th} cell of HL indicates the *haters of i that are loved by j.*



Walks, Paths, and Components • Background

L	A	B	C	D
A	0	0	0	0
B	1	0	0	0
C	0	0	0	0
D	0	0	1	0

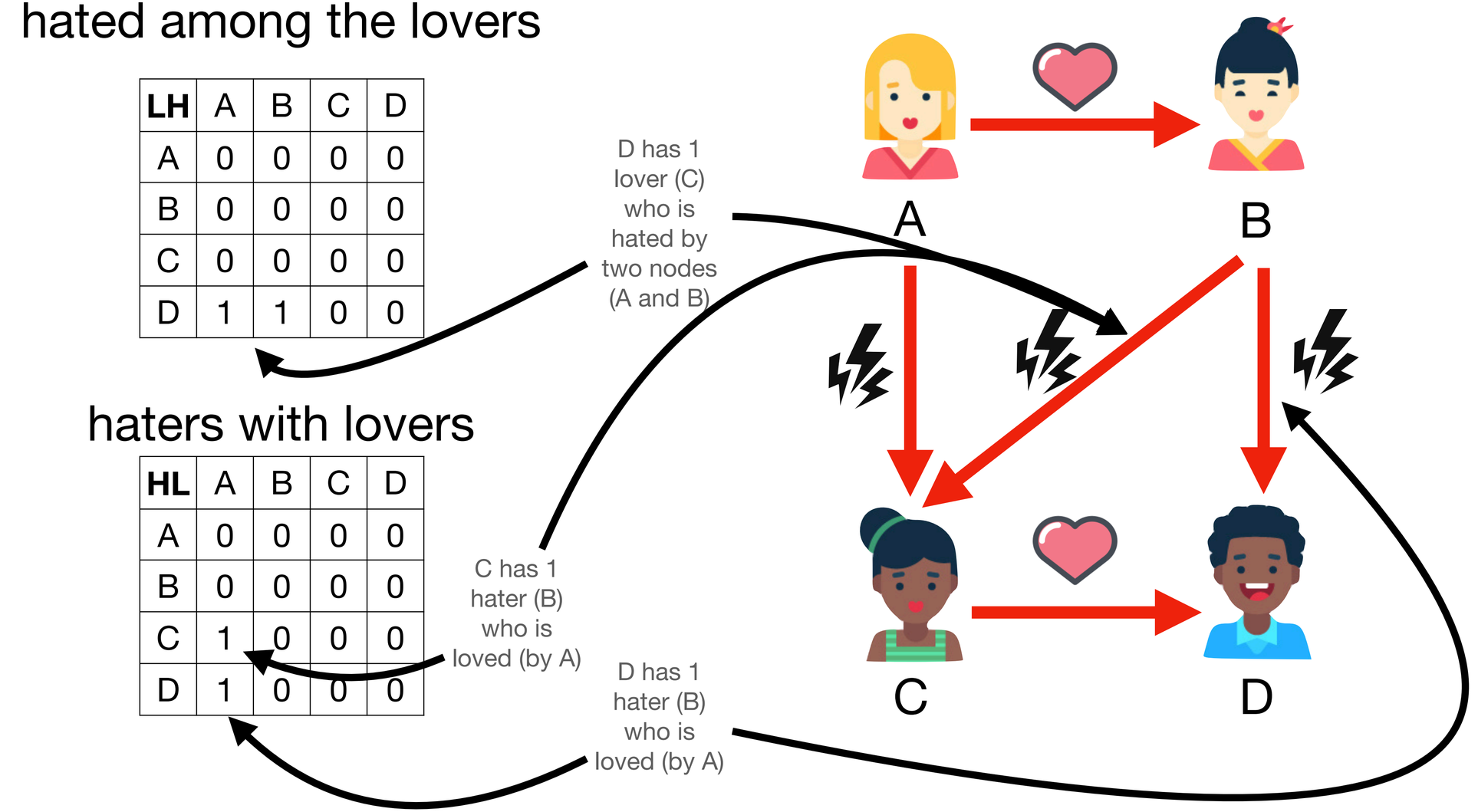
H	A	B	C	D
A	0	0	0	0
B	0	0	0	0
C	1	1	0	0
D	0	1	0	0

hated among the lovers

LH	A	B	C	D
A	0	0	0	0
B	0	0	0	0
C	0	0	0	0
D	1	1	0	0

haters with lovers

HL	A	B	C	D
A	0	0	0	0
B	0	0	0	0
C	1	0	0	0
D	1	0	0	0



Walks, Paths, and Components • Background

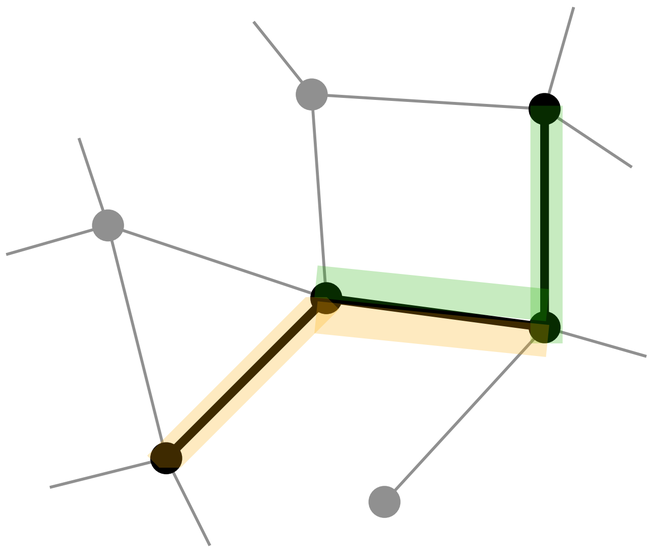
We can also compute products of matrices with themselves. E.g., if F is the friendship matrix, then FF is the "friend of friend" relation. When the ij^{th} cell of FF is greater than 0, it indicates the number of friends of i who have j as a friend.

A useful application of matrix products is to formalise social theories, e.g., let F be the matrix of friendship and E the matrix of enemies, we can hypothesise that:

- FF, the friends of my friends are my friends;
- EF, the friends of my enemies are my enemies;
- FE, the enemies of my friends are my enemies;
- EE, the enemies of my enemies are my enemies.

Then, if we have actual surveys of the relations, we can compare those (theoretical) measures and test our theories.

Walks, Paths, and Components



To calculate the number of walks of a given length r on a network, for either a directed or an undirected simple network, the element A_{ij} is 1 if there is an edge from node j to node i, and 0 otherwise (the condition for non-simple networks is slightly different). Then the product $A_{ik}A_{kj}$ is 1 if there is a walk from j to i via k, and 0 otherwise. In that case, we know the walk has length 2.

$$N_{ij}^{(2)} = \sum_{k=1}^{n} A_{ik}A_{kj} = [A^2]_{ij}$$

In general

$$N_{ij}^{(r)} = [\, A^r \,]_{ij}$$

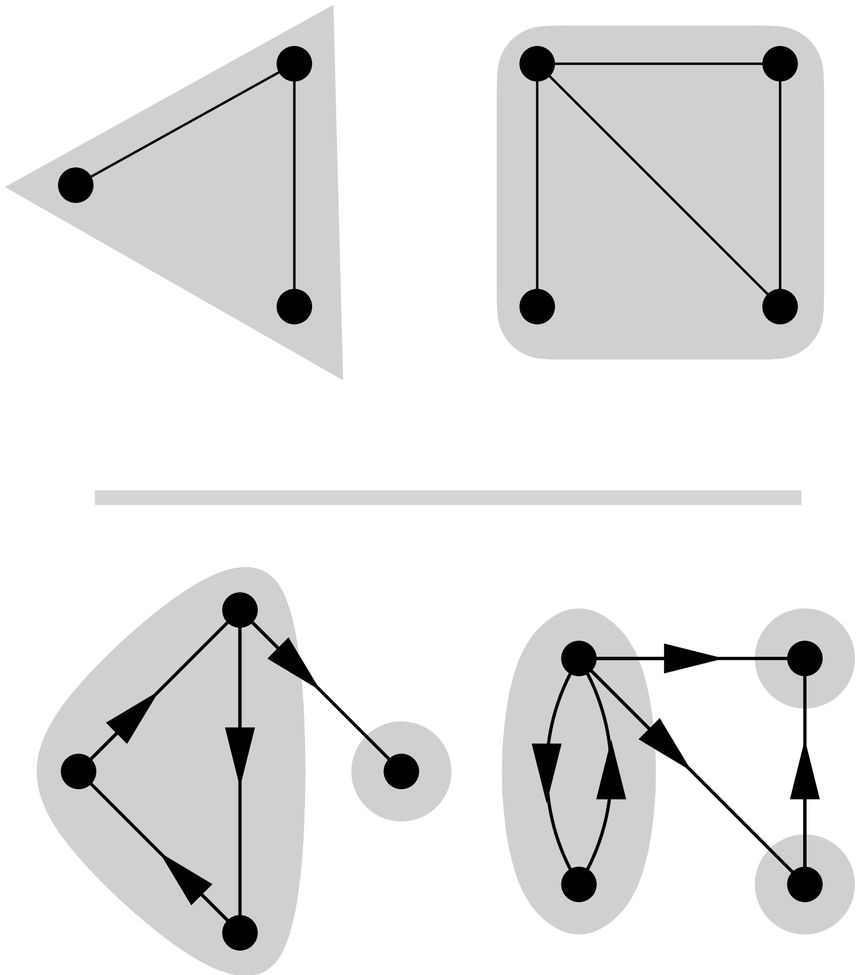
Walks, Paths, and Components

The ***shortest path*** between two nodes i and j, also called the ***geodesic path***, is the shortest walk (hence, self-avoiding) that connects i and j. The length of the shortest path (also called their **distance**) is the *minimum* r such that $[\ A^r\ ]_{ij} > 0$.

Shortest paths are important in many contexts, e.g., in communication and transportation networks, they affect how rapidly it is possible to get goods or data from one node to another, e.g., if, at each “leg”, there is some overhead (distances to be bridged, data manipulation).

The ***diameter*** of a network is the length of the longest among all existing shortest paths between every pair of nodes in the network. The diameter of the network is useful, e.g., to understand the connectedness of networks.

Walks, Paths, and Components

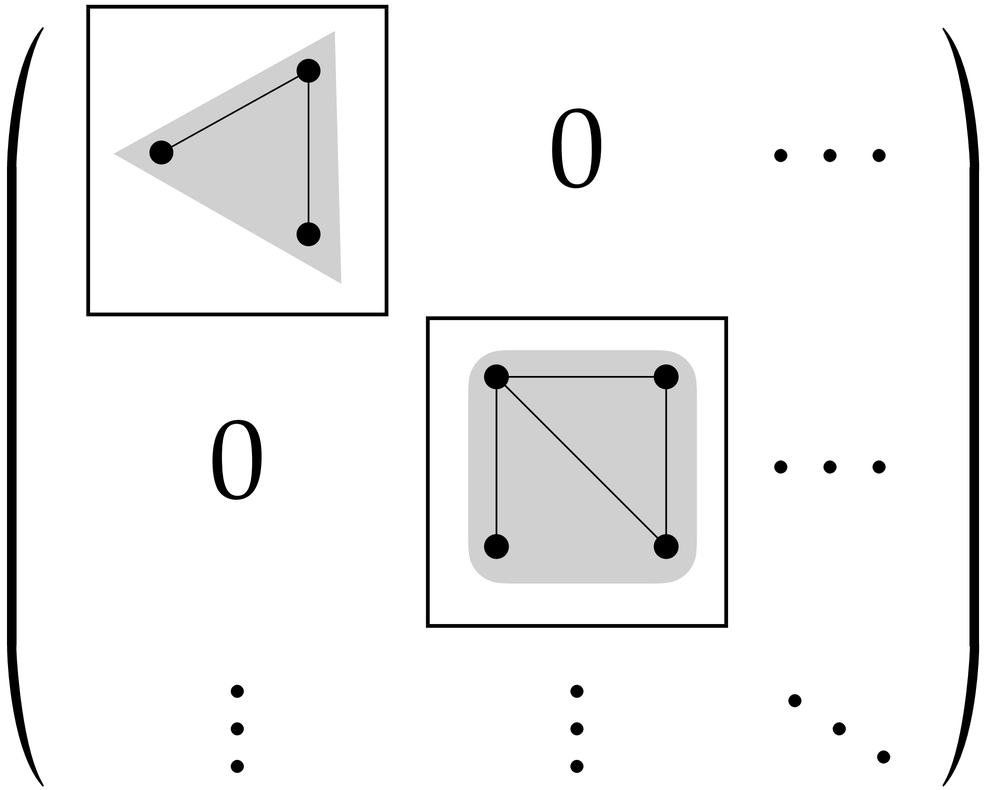


A network does not necessarily consist of a single connected set of nodes, indeed, frequently networks have *separate parts* that are disconnected from one another. Such parts are called ***components.***

Technically, a component is a subset of the nodes of a network with the properties**:**

- (connectedness) there exists **at least one path from each member to each other member of that subset**;
- (maximality) **no other node in the network can be added to the subset**.

Walks, Paths, and Components



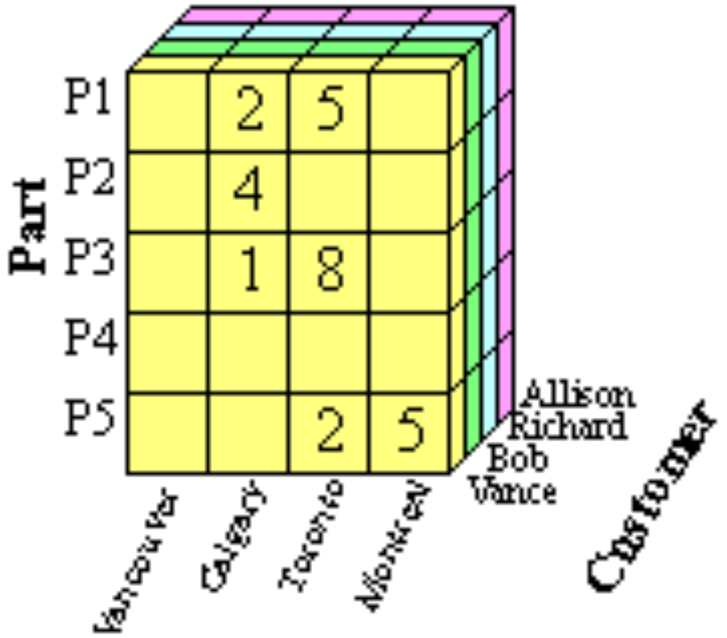
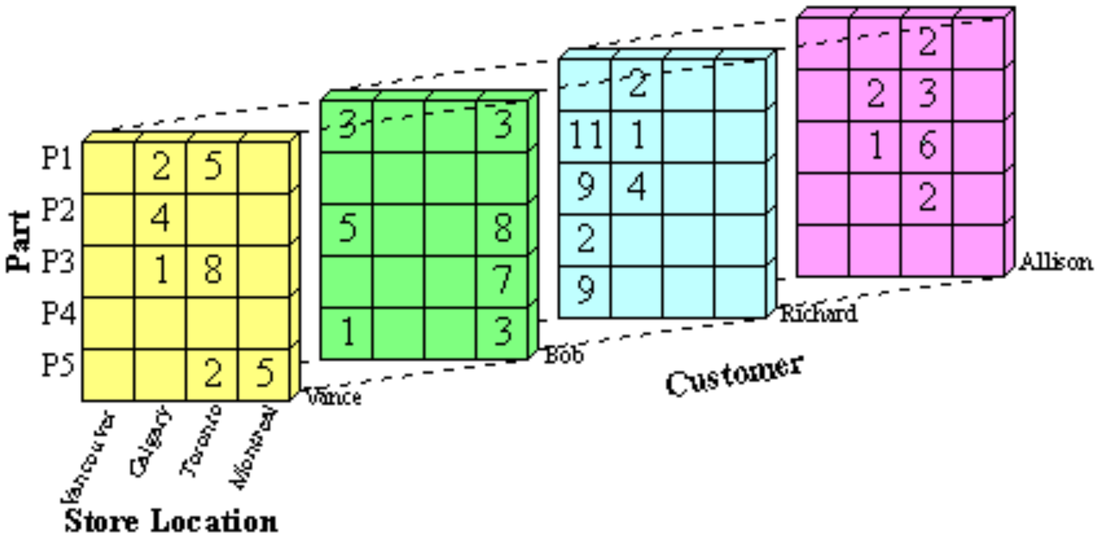
A network in which all nodes belong to the same single component is said to be ***connected***. Conversely, a network with more than one component is ***disconnected***.

The adjacency matrix of a network with more than one component can be written (after proper labelling) in block diagonal form, meaning that the non-zero elements of the matrix are “confined” to square blocks along the diagonal of the matrix, with all other elements being zero

Ways and Modes

The adjacency matrix of a graph **is always square**: it has the same number of rows as columns. It is also called a “one-mode matrix”, meaning that the rows and columns both **refer to the same single set of entities**.

Matrices have **ways** and **modes**: ways are the dimensions of the matrix—normally two—while modes are the kinds of entities represented.

A three-way matrix has rows, columns, and levels, as in a data-cube. For example, suppose we are retailers and we want to represent what parts which customer has at each of our stores.

Combining all of these into a single matrix we get a three-way, three-mode matrix.