Introduction to Graph Theory

Optimization Methods in Management Science
Master in Management
HEC Lausanne

Dr. Rodrigue Oeuvray

Fall 2019 Semester

Introduction to Graph Theory

Some basic concepts:

- Directed and undirected graphs
- Adjacency and incidence matrices
- Subgraph and partial graph
- Chain, path, cycle and circuit
- Connectivity and strong connectivity
- Tree and forest

Undirected Graph

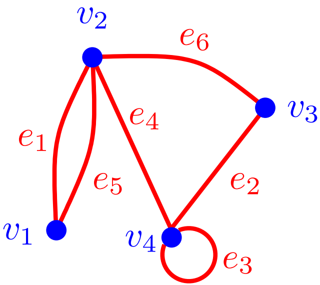
A finite undirected graph is a 3-tuple $G = (V, E, \Psi)$ where

- V is a finite set ($|V| = n$) whose elements are called vertices,
- E is a finite set ($|E| = m$) whose elements are called edges,
- $\Psi : E \to V \times V$ is called an incidence function which maps an edge $e \in E$ to an unordered pair $\{u(e), v(e)\}$ of vertices of V. The vertices $u(e)$ and $v(e)$ are called the endpoints of e.

If $u(e) = a$ and $v(e) = b$, we say that a and b are the endpoints of the edge e, that vertices a and b are adjacent or incident with e, or that the edge e is incident with a and b

Representation of an Undirected Graph in the Plane

We associate a point in the plane with each vertex and we represent each edge by a simple curve that connects its two endpoints



$$V = \{v_1, v_2, v_3, v_4\}$$

$$E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$

Ψ	e_1	e_2	e_3	e_4	e_5	e_6
$u(e)$	v_1	v_4	v_4	v_2	v_1	v_2
$v(e)$	v_2	v_3	v_4	v_4	v_2	v_3

Directed Graphs

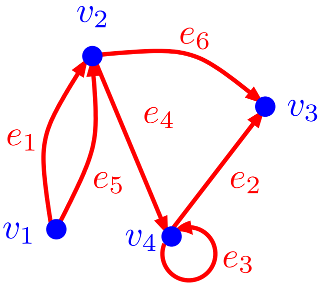
A finite directed graph is a 3-tuple $G = (V, E, \Psi)$ where

- V is a finite set ($|V| = n$) whose elements are called vertices
- E a finite set ($|E| = m$) whose elements are called arcs
- $\Psi : E \to V \times V$ is an incidence function which maps an arc to an ordered pair $(u(e), v(e))$ of vertices of V. The vertices $u(e)$ and $v(e)$ are called the endpoints of e.

If $u(e) = a$ and $v(e) = b$, we say that e is an incoming arc into b, an outgoing arc out of a, that the vertex a is the initial endpoint of e and b its terminal endpoint. A directed graph is also called a digraph

Representation of a Directed Graph in the Plane

It is similar to what we have seen before for undirected graphs:



$$V = \{v_1, v_2, v_3, v_4\}$$

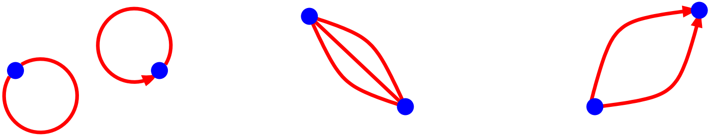
$$E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$

Ψ	e_1	e_2	e_3	e_4	e_5	e_6
$u(e)$	v_1	v_4	v_4	v_2	v_1	v_2
$v(e)$	v_2	v_3	v_4	v_4	v_2	v_3

Simple Graph and Multigraph

- An edge (an arc) whose endpoints are the same vertex is a loop
- A simple graph is a graph without loops and without multiples egdes (arcs)

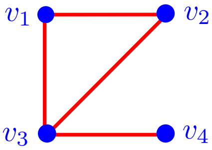
From now on and unless otherwise specified, the term graph will denote a finite simple graph. If loops and multiple edges (arcs) are allowed, we use the term multigraph



Important remark: two arcs in the opposite direction are not multiple arcs !

Simple Graph and Multigraph (Cont'd)

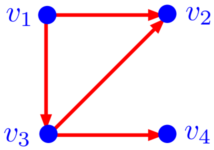
In a simple graph, one can identify unequivocally each edge (arc) with the pair (ordered or not) formed by its endpoints. The incidence function becomes useless et we denote $G = (V, E)$ such a graph.



$G = (V, E)$

$$V = \{v_1, v_2, v_3, v_4\}$$

$$E = \{\{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_3, v_4\}\}$$



$G = (V, E)$

$$V = \{v_1, v_2, v_3, v_4\}$$

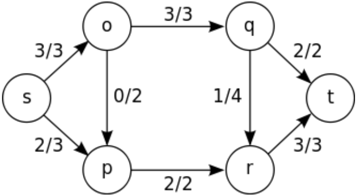
$$E = \{(v_1, v_2), (v_1, v_3), (v_3, v_2), (v_3, v_4)\}$$

Degree of a Vertex

- Let G be a multigraph, the degree of a vertex v, denoted by $deg(v)$, is the number of edges (arcs) incident to v
- If a vertex has one or several loops, each of them is counted twice to determine its degree
- If G is a directed multigraph, the external degree of vertex v, denoted by $deg_{+}(v)$, is equal to the number of outgoing arcs out of v. In a similar way, the internal degree of vertex v, denoted by $deg_{-}(v)$, is the number of incoming arcs into v

Degree of a Vertex (Cont'd)

- A leaf is a vertex with degree one
- In a directed graph, a source v is a vertex with $deg_{-}(v) = 0$
- In a directed graph, a sink is a vertex with $deg_{+}(v) = 0$
- In the graph below, s is a **source** and t is a **sink**



Matrices Associated with a Simple Undirected Graph

Adjacency matrix $\boldsymbol{B} : n \times n$

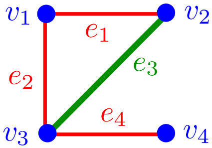
$$b_{ij} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

$$\boldsymbol{B} = \begin{array}{cc} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} & \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} & \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \end{array}$$

Incidence matrix $\boldsymbol{A} : n \times m$

$$a_{ik} = \begin{cases} 1 & \text{if } v_i \text{ is incident to } e_k \\ 0 & \text{otherwise} \end{cases}$$

$$\boldsymbol{A} = \begin{array}{cc} \begin{array}{cccc} e_1 & e_2 & e_3 & e_4 \end{array} & \\ \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \end{array}$$



Interpretation

- 1st row of $\boldsymbol{B}$: v_1 is adjacent to v_2 and v_3
- As the matrix is symmetric, we can also interpret the 1st column of $\boldsymbol{B}$ the same way as its first row
- 1st row of $\boldsymbol{A}$: v_1 is incident to e_1 and e_2
- 1st column of $\boldsymbol{A}$: e_1 is incident to v_1 and v_2
- $\boldsymbol{A}$ is not symmetric !

$$\boldsymbol{B} = \begin{array}{c c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} & \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} & \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \end{array}$$

$$\boldsymbol{A} = \begin{array}{c c} \begin{array}{cccc} e_1 & e_2 & e_3 & e_4 \end{array} & \\ \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \end{array}$$

Matrices Associated to a Simple Directed Graph

Adjacency matrix

$\boldsymbol{B} : n \times n$

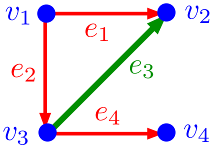
Incidence matrix $\boldsymbol{A} : n \times m$

$$b_{ij} = \begin{cases} 1 & \text{if } (v_i, v_j) \in E \\ 0 & \text{otherwise} \end{cases}$$

$$a_{ik} = \begin{cases} -1 & \text{if } v_i \text{ is the init. endpt of } e_k \\ 1 & \text{if } v_i \text{ is the term. endpt of } e_k \\ 0 & \text{otherwise} \end{cases}$$

$$\boldsymbol{B} = \begin{array}{cc} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} & \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} & \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \end{array}$$

$$\boldsymbol{A} = \begin{array}{cc} \begin{array}{cccc} e_1 & e_2 & e_3 & e_4 \end{array} & \\ \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 0 & \mathbf{1} & 0 \\ 0 & 1 & \mathbf{-1} & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \end{array}$$



Matrices Associated to a Simple Directed Graph

- 1st row of $\boldsymbol{B}$: v_1 is adjacent to v_2 and v_3
- 1st row of $\boldsymbol{A}$: v_1 is the initial endpoint for e_1 and e_2
- 1st column of $\boldsymbol{A}$: e_1 goes from v_1 to v_2
- Neither $\boldsymbol{A}$ nor $\boldsymbol{B}$ are symmetric !

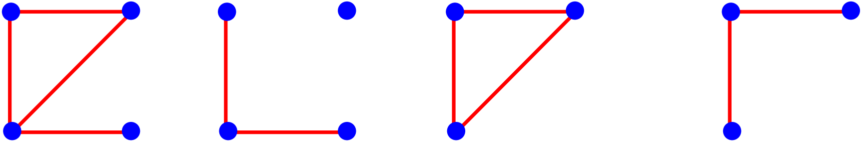
$$\boldsymbol{B} = \begin{array}{cc} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} & \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} & \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \end{array}$$

$$\boldsymbol{A} = \begin{array}{cc} \begin{array}{cccc} e_1 & e_2 & e_3 & e_4 \end{array} & \\ \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \end{array}$$

Subgraph and Partial Graph

Let $G = (V, E)$ be a graph

- G' is a partial graph of G if $G' = (V, E')$ with $E' \subseteq E$
- G' is a subgraph G induced by W if $G' = (W, E(W))$ where $W \subseteq V$ and $E(W)$ is the set of edges (arcs) having their endpoints in W
- G' is a partial subgraph of G if G' is a partial graph of a subgraph of G



Graph Partial-graph Subgraph Partial-subgraph

Chain and Cycle

- A chain C is an alternating sequence of vertices and edges:

$$C = (u_0, f_1, u_1, f_2, u_2, \ldots, u_{k-1}, f_k, u_k),$$

  where $u_i \in V \; \forall \, i$, $f_i \in E \; \forall \, i$ and $f_i = \{u_{i-1}, u_i\} \; \forall \, i$

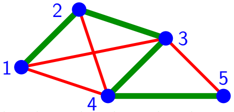
- A cycle is a chain whose endpoints are the same vertex

- Remarks:
  - A chain starts and ends with a vertex
  - By convention every chain must contain at least one edge
  - The sequence $C' = (u_k, f_k, u_{k-1}, \ldots, u_2, f_2, u_1, f_1, u_0)$ is the same chain as C
  - A loop is a cycle

Chain and Cycle (Cont'd)

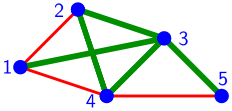
- A chain (a cycle) is elementary if each vertex is present at most once
- A chain (a cycle) is simple if each edge is present at most once
- The length of a chain (resp. of a cycle) is the number of edges of the chain (resp. of the cycle)
- An undirected graph is acyclic if it has no **simple** cycle
- A graph can have a **non-simple** cycle, for example $(v_i, \{v_i, v_j\}, v_j, \{v_i, v_j\}, v_i)$, and be acyclic !

Simple-and-elementary-chain



$C = (1, \{1,2\}, 2, \{2,3\}, 3, \{3,4\}, 4, \{4,5\}, 5)$

Non-elementary-chain



$C = (1, \{1,3\}, 3, \{2,3\}, 2, \{2,4\}, 4, \{3,4\}, 3, \{3,5\}, 5)$

Path and Circuit

- A path C is an alternating sequence of vertices and arcs

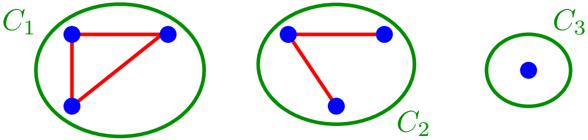
$$C = (u_0, f_1, u_1, f_2, u_2, \ldots, u_{k-1}, f_k, u_k),$$

where $u_i \in V \; \forall \, i$ and $f_i = (u_{i-1}, u_i) \in E \; \forall \, i$

- A circuit is a path whose endpoints are the same vertex
- The definitons of a simple and elementary path (circuit) as well as the concept of length are similar to the undirected case
- A directed graph is acyclic if it has no circuit

Connectivity

- Let $G = (V, E, \Psi)$ be an **undirected** multigraph. We define on V a relation defined as follows: 2 vertices v_i and v_j belong to the same connected component if and only if it exists a chain between v_i and v_j
- Concretely, two vertices belong to the same component if one can "move" from one vertex to the other one. If not, they belong to different components
- In the graph below, there are three connected components



Strong Connectivity

- Let $G = (V, E, \Psi)$ be a **directed** multigraph. We define on V a relation defined as follows: 2 vertices v_i and v_j belong to the same strongly connected component if and only if it exists a path between v_i and v_j **and** a path between v_j and v_i

- Concretely, two vertices belong to the same strongly component if one can "move" from one vertex to the other one **and vice versa**. If not, they belong to different components

- We say that a graph is strongly connected if it has only one strongly connected component

Marking Algorithm

Input: A directed multigraph $G = (V, E, \Psi)$.
Ouput: The number k of strongly connected components of G as well as the list $\{C_1, \ldots, C_k\}$ of its strongly connected components

(1) Initialization:

$k = 0,\ W = V$

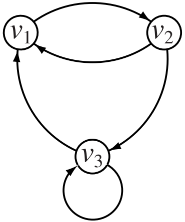
(2) Main loop:

As long as $W \neq \emptyset$, choose a vertex $v \in W$ and mark it with $+$ and $-$

(2.1) Mark all the direct and undirect successors of v with $+$
(2.2) Mark all the direct and undirect predecessors of v with $-$
(2.3) Update $k = k + 1$ and C_k equals to the set of vertices marked by $+$ and $-$
(2.4) Withdraw from W the vertices of C_k and remove the marks

(3) The number of strongly connected components of G is k. Every set $C_i,\ i = 1 \ldots, k$ corresponds to the vertices of the a strongly connected component

Marking Algorithm: Example

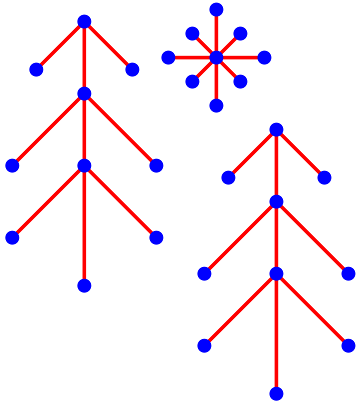


- $k = 0. W = \{v_1, v_2, v_3\}$
- We select v_1 and mark it with $+$ and $-$
- Direct and indirect sucessors of v_1: v_2 and v_3. They are marked with $+$
- Direct and indirect predecessors of v_1: v_3 and v_2. They are marked with $-$
- $k = 1,\ C_1 = \{v_1, v_2, v_3\}$
- $W = \emptyset$. STOP

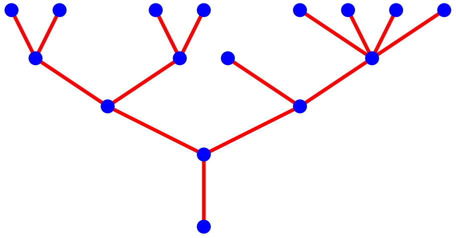
We conclude that there is only **one strongly connected** component. This graph is **strongly connected**

Tree and Forest

- A undirected multigraph without any cycle is a forest
- A undirected multigraph without any cycle and connected is a tree



A forest



A-tree

Remark: each connected component of a forest is a tree

Characterization of a Tree

Theorem

Let $G = (V, E, \Psi)$ be a multigraph with n vertices. The following statements are equivalent:

(a) *G is a tree*

(b) *G is connected and has no cycle*

(c) *G has no cycle and has $n - 1$ edges*

(d) *G is connected and has $n - 1$ edges*

(e) *G has no loop and each pair of distinct vertices is connected by a simple chain*