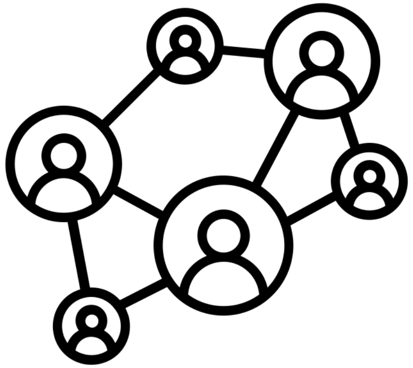
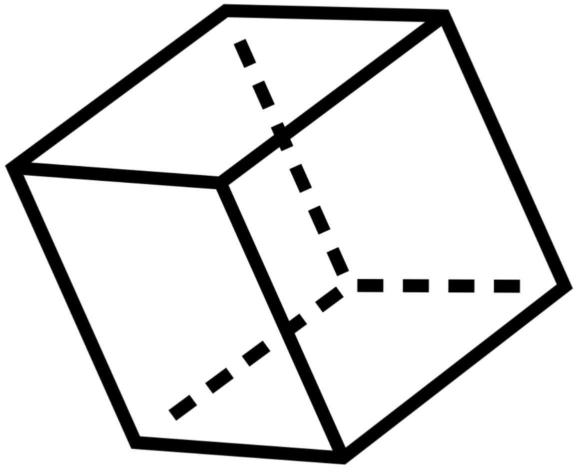


ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Fifth Week

The Network Measurement



1. The Network Measurement

5.1 Centrality: Degree

Definition 5.1.1 (Degree)

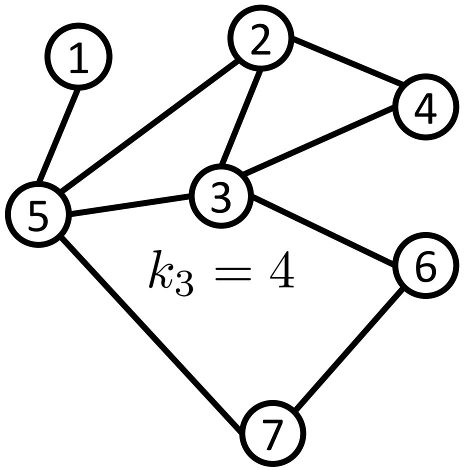
The **degree** of a vertex in a graph is the nubmer of edges connected to it.

⋆ We will denote the degree of vertex i by k_i.

⋆ For an undirected graph of n vertices the degree can be written in terms of the adjacency matrix as

$$k_i = \sum_{j=1}^{n} A_{ij}.$$

Example (Degree)



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

$$k_3 = \sum_{j=1}^{n} A_{3j} = 4.$$

5.1 Centrality: Degree

Note

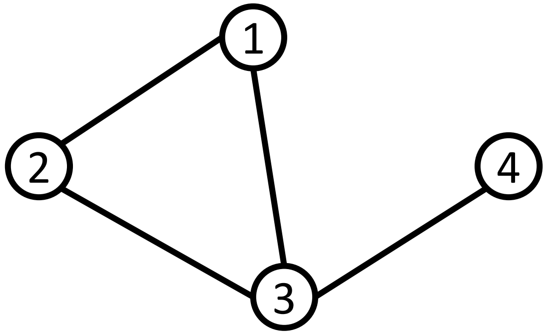
If there are m-edges in total then there are $2m$ ends of edges

$$\Rightarrow 2m = \sum_{i=1}^{n} k_i$$

$$\Rightarrow \quad m = \frac{1}{2}\sum_{i=1}^{n} k_i$$

$$= \frac{1}{2}\sum_{i=1}^{n}\left(\sum_{j=1}^{n} A_{ij}\right)$$

$$= \frac{1}{2}\sum_{i,j=1}^{n} A_{ij}$$



$\text{edges} = 4,\ 2m = 8$

$k_1 = 2,\ k_2 = 2,\ k_3 = 3,\ k_4 = 1$

5.1 Centrality: Degree

Note

The mean degree c of a vertex in an undirected graph is

$$c = \frac{1}{n}\sum_{i=1}^{n} k_i,$$
$$\Rightarrow cn = \sum_{i=1}^{n} k_i = 2m$$
$$\Rightarrow c = \tfrac{2m}{n}$$

Definition 5.1.2 (Density)

The maximum possible number of edges in a simple graph is $\binom{n}{2} = \frac{n(n-1)}{2}$.

The **connectance** or **density** ρ of a graph is the fraction of these edges that are actually present:

$$\rho = \tfrac{m}{\binom{n}{2}} = \tfrac{2m}{n(n-1)} = \tfrac{2c}{n(n-1)} = \tfrac{c}{n-1}$$

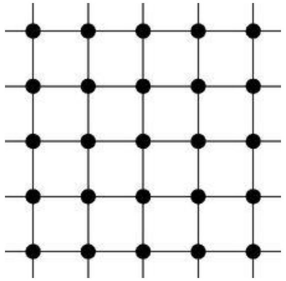
5.1 Centrality: Degree

Note

$0 < \rho < 1$

① As $n \to \infty$, $\rho \to$ constant $> 0 \Rightarrow$ A network is **dense**.

② As $n \to \infty$, $\rho \to 0 \Rightarrow$ A network is **sparse**.



$k_i = 4 \longrightarrow$ regular network (same degree)

5.1 Centrality: Degree

Definition 5.1.3 (in,out - degree)

In a directed network each vertex has two degrees. The **in-degree** is the number of ingoing edges connected to a vertex and the **out-degree** is the number of outgoing edges.

Note

The adjacency matrix of a directed network has element $A_{ij} = 1$ if there is an edge from j to i, in- and out-degrees can be written

$$k_i^{in} = \sum_{j=1}^{n} A_{ij}, \qquad k_j^{out} = \sum_{i=1}^{n} A_{ij}$$

(n is nubmer of vertices)

5.1 Centrality: Degree

Note

The number of edges m in a directed network is equal to the total number of ingoing ends of edges at all vertices, or equivalently to the total number of outgoing ends of edges, so

$$m = \sum_{i=1}^{n} k_i^{in} = \sum_{j=1}^{n} k_j^{out} = \sum_{ij} A_{ij}$$

Note

Thus the mean in-degree c_{in} and the mean out-degree c_{out} of every directed network are equal:

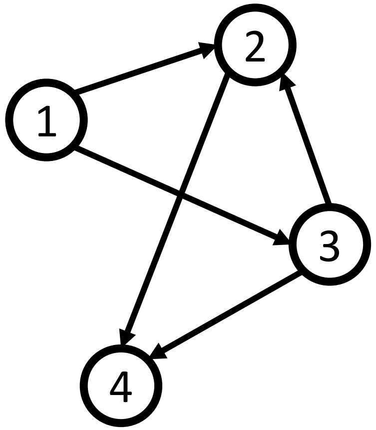
$$c_{in} = \frac{1}{n} \sum_{i=1}^{n} k_i^{in} = \frac{1}{n} \sum_{i=1}^{n} k_j^{out} = c_{out}$$

$$c = c_{in} = c_{out} = \frac{m}{n}$$

(n and m are number of vertices and edges in total)

5.1 Centrality: Degree

Example



$$n = 4,\ m = 5$$

$$k_1^{in} = 0,\ k_2^{in} = 2,\ k_3^{in} = 1,\ k_4^{in} = 2$$

$$k_1^{out} = 2,\ k_2^{out} = 1,\ k_3^{out} = 2,\ k_4^{out} = 0$$

$$c_{in} = \frac{0+2+1+2}{4} = \frac{5}{4} = \frac{2+1+2+0}{4} = c_{out}$$

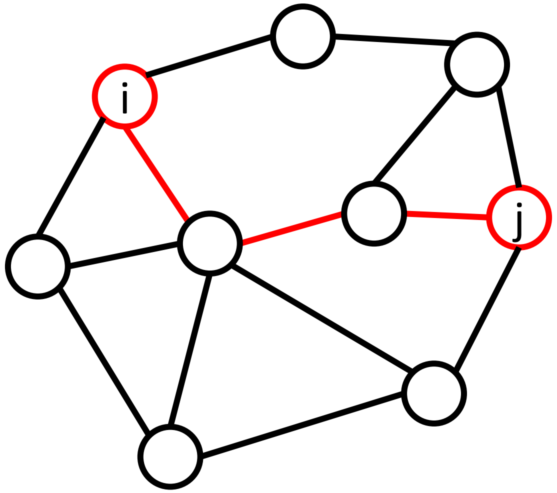
5.2 Centrality: Path

Definition 5.2.1 (Path)

The **path** in a network is any sequence of vertices such that every consecutive pair of vertices in the sequence is connected by an edge in the network.

The **length of a path** in a network is the number of edges traversed along th path (not the number of vertices).

Example (Path)

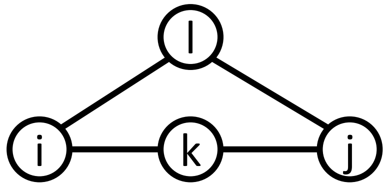


Red color vertices are path between i and j, and length of a path is 3.

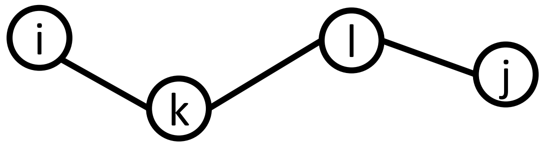
5.2 Centrality: Path

$\star\ N_{ij}^{(n)}$: The total number of paths of length n from j to i.

i) $N_{ij}^{(2)} = \sum_{k=0}^{n} A_{ik}A_{kj} = \left[A^2\right]_{ij}$



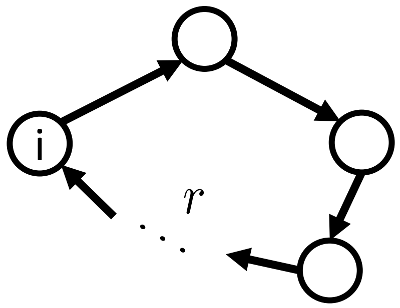
ii) $N_{ij}^{3} = \sum_{k,l=1}^{n} A_{ik}A_{kl}A_{lj} = \left[A^3\right]_{ij}$



$$\therefore \left[N^{(r)}\right]_{ij} = [A^r]_{ij}$$

A special case of this result is that the number of paths of length r that start and end at the same vertex i is

$$\left[N^{(r)}\right]_{ii} = [A^r]_{ii}$$



5.2 Centrality: Path

Note

The total number L_r of loops of length r anywhere in a network is the sum of this quantity over all possible starting points i:

$$L_r = \sum_{i=1}^{n} [A^r]_{ii} = \mathrm{Tr}(A^r) \qquad - \ (*)$$

This expression counts separately loops consisting of the same vertices in the same order but with different starting points.

i) $1 \to 2 \to 3 \to 1 \quad \neq \quad 2 \to 3 \to 1 \to 2$

ii) $1 \to 2 \to 3 \to 1 \quad \neq \quad 1 \to 3 \to 2 \to 1$

Equation $(*)$ can also be expressed in terms of the eigenvalues of the adjacency matrix.

5.2 Centrality: Path

Note

For an undirected network, A : symmetric and all real eigenvalues.

$$A = UKU^\top$$

$$U \text{ : orthogonal matrix } \left(UU^\top = I\right), \text{ Col}_i(U) \text{ is eigenvectors of } A$$

$$K \text{ : a diagonal matrix whose element is eignevlaues of } A$$

$$A^r = (UKU^\top)^r = (UKU^\top)(UKU^\top)\cdots(UKU^\top) = UK^rU^\top$$

$$\because L_r = \mathrm{Tr}(A^r),\ L_r = \mathrm{Tr}(UK^rU^\top)$$

Recall

$$\mathrm{Tr}(PQ) = \mathrm{Tr}(QP)$$

$$\because) \text{ PQ} = \sum_{i=1}^{n}(PQ)_{ii} = \sum_{i=1}^{n}\left[\sum_{j=1}^{n} P_{ij}Q_{ij}\right]$$

$$= \sum_{i=1}^{n}\left[Q_{ji}P_{ij}\right] = \sum_{j=1}^{n}\left[\sum_{i=1}^{n} Q_{ji}P_{ij}\right] = \sum_{j=1}^{n}(QP)_{jj} = \mathrm{Tr}(QP) \quad \square$$

5.2 Centrality: Path

Note

$$\therefore L_r = \mathrm{Tr}(UK^kU^\top) = \mathrm{Tr}(K^rUU^\top) = \mathrm{Tr}(K^r) = \sum_{i=1}^{n} \kappa_i^r$$

$$\text{where } \kappa_i \text{ is an eignevalues of } A,\ K = \begin{bmatrix} k_1 & 0 & \cdots & 0 \\ 0 & k_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k_n \end{bmatrix}$$

Theorem 5.1.1

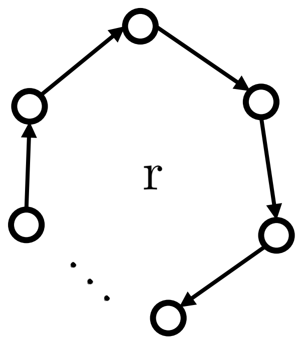
All of eigenvalues are zero $\Leftrightarrow$ a network is acycle.

Proof) $\Rightarrow$) zero $\Rightarrow$ acycle

$\Leftarrow$) A cycle $\Rightarrow$ there exists a nonzero eigenvalue of A

Suppose a network is cycle, $\exists r > 0$ such that $L_r > 0$

then, $L_r = \sum_{i=1}^{n} \kappa_i^r > 0 \Rightarrow \exists \kappa_j > 0$ for some j. $\square$



5.2 Centrality: Path

Note

For directed networks the situation is more complicated. in general, asymmetric adjacency matrices, and some asymmetric matrices cannot be diagonalized.

Every real matrix, whether diagonalizable or not, can be written in the form $A = QTQ^\top$, where Q is an orthogonal matrix ($QQ^\top = I$) and T is an upper triangular matrix.

Note (Schur decomposition)

Assume $A\mathbf{x} = \kappa\mathbf{x}$

$$A\mathbf{x} = QTQ^\top\mathbf{x} = \kappa\mathbf{x} \Rightarrow TQ^\top\mathbf{x} = \kappa Q^\top\mathbf{x}$$

$$\therefore\ \kappa \text{ is an eigenvalue of } A.$$

5.2 Centrality: Path

Note

$$T = \begin{bmatrix} t_{11} & & & \mathbf{X} \\ & t_{22} & & \\ & \mathbf{0} & \ddots & \\ & & & t_{nn} \end{bmatrix} \Longrightarrow \text{eigenvalue of } T \text{ is diagonal component}$$

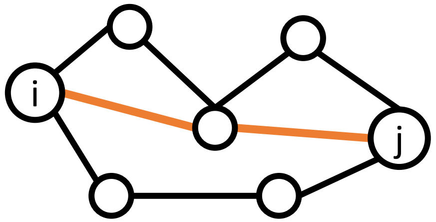
$$\Rightarrow \; T = \begin{bmatrix} \kappa_{11} & & & \mathbf{X} \\ & \kappa_{22} & & \\ & \mathbf{0} & \ddots & \\ & & & \kappa_{nn} \end{bmatrix} \Rightarrow \; T^r = \begin{bmatrix} \kappa_{11}^r & & & \mathbf{X} \\ & \kappa_{22}^r & & \\ & \mathbf{0} & \ddots & \\ & & & \kappa_{nn}^r \end{bmatrix}$$

$$\begin{aligned} \therefore \; L_r &= \mathrm{Tr}(A^r) = \mathrm{Tr}(QTQ^\top)^r \\ &= \mathrm{Tr}(QT^rQ^\top) = \mathrm{Tr}(T^rQ^\top Q) \\ &= \mathrm{Tr}(T^r) = \sum_{i=1}^{n} \kappa_i^r \end{aligned}$$

5.2 Centrality: Path

Definition 5.2.2 (Geodesic Paths)

A **geodesic path**, also called simply a shortest path, is a path between two vertices such that no shorter path exists:



Geodesic path between i and j is 2.

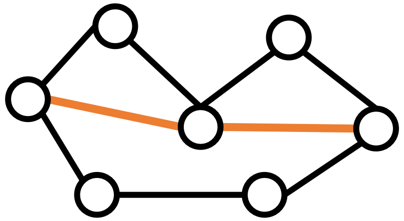
Definition 5.2.3 (Geodesic Distance)

The length of a geodesic path, often called the **geodesic distance** or **shortest distance**

The geodesic distance between vertices i and j is the smallest value of r such that $A^r_{ij} > 0$.

Definition 5.2.4 (Diameter)

The **diameter** of a graph is the length of the longest geodesic path between any pair of vertices in the network



Diameter is 4

5.2 Centrality: Path

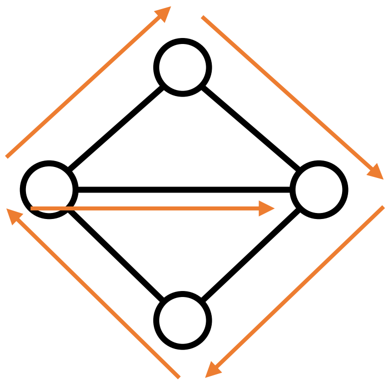
Definition5.2.5 (Eulerian path)

A **Hamiltonian path** is a path that visits each vertex exactly once.

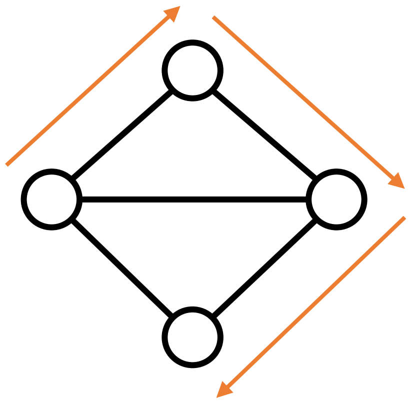
Definition5.2.6 (Hamiltonian path)

An **Eulerian path** is a path that traverses each edge in a network exactly once.

Example



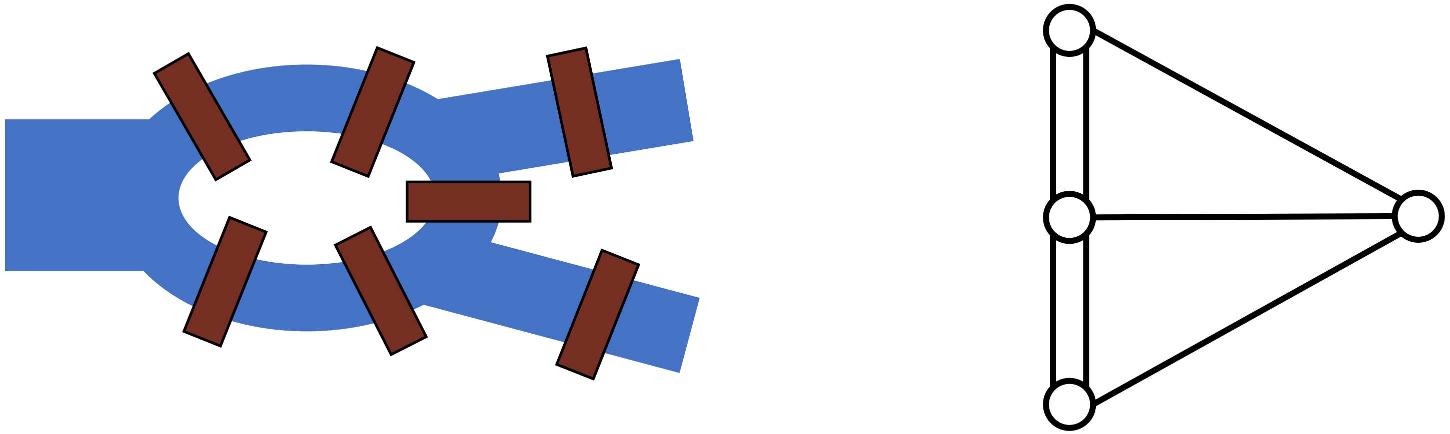
Eulerian path



Hamiltonian path

5.2 Centrality: Path

Example (The Königsberg bridges)



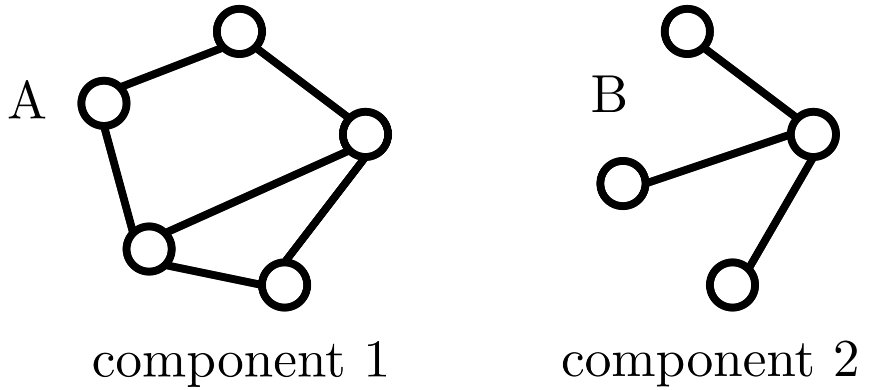
The Königsberg bridges.

Any Eulerian path must both enter and leave every vertex it passes through except the first and last, there can at most be two vertices in the network with odd degree if such a path is to exist. Since all four vertices in the Königsberg network have odd degree, the bridge problem necessarily has no solution.

5.2 Path: Component

Definition 5.2.7 (Component)

The subgroups in a network like that of bellow figure are called **components**.



Definition 5.2.8 (disconnected, connected)

There is no path from the vertex labeled A to the vertex labeled B. A network of this kind is said to be **disconnected**. Conversely, if there is a path from every vertex in a network to every other the network is **connected**.

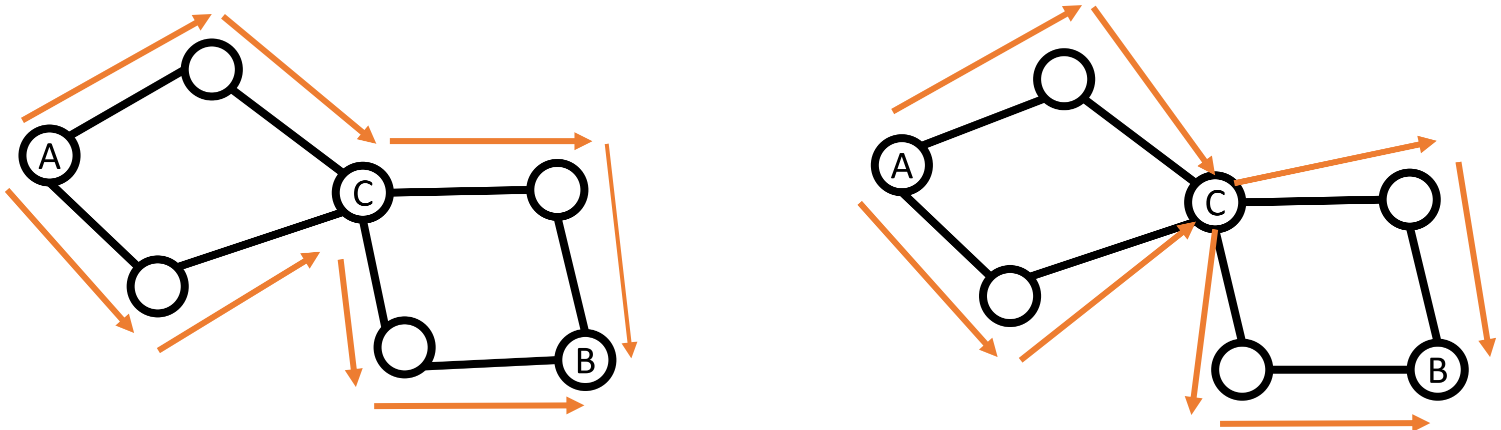
5.3 Path: Independent path

Definition 5.3.1 (Independent path)

Two paths connecting a given pair of vertices are **edge-independent** if they share no edges.

Two paths are **vertex-independent** (or **node-independent**) if they share no vertices other than the starting and ending vertices.

Example



(a) (b)

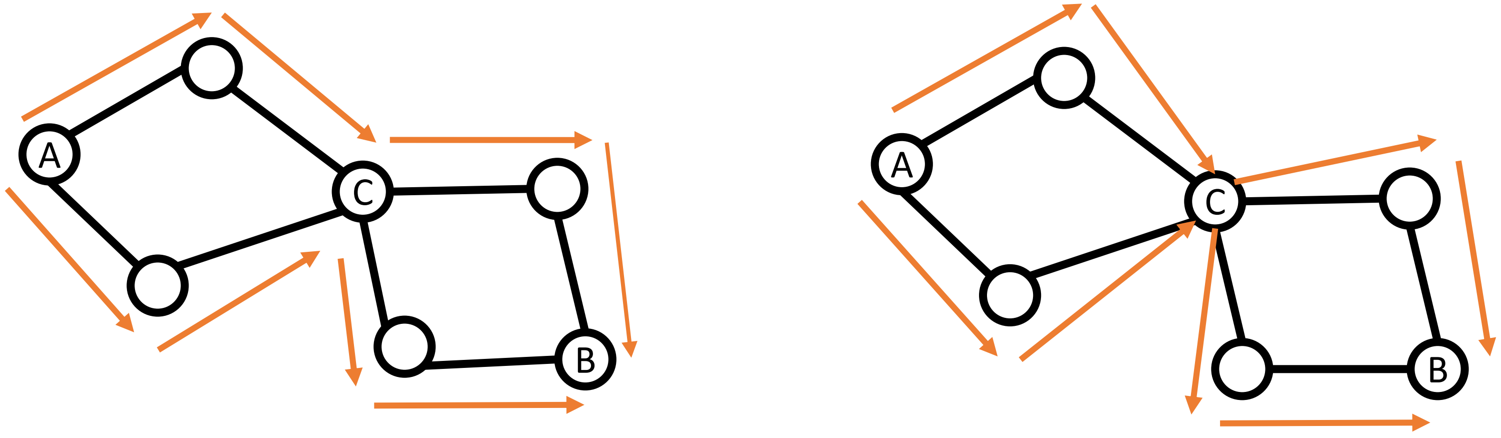
(a) There are two edge-independent paths from A to B in this figure, as denoted by the arrows, but there is only one vertex-independent path, because all paths must pass through the center vertex C. (b) The edge-independent paths are not unique; there are two different ways of choosing two independent paths from A to B in this case.

5.3 Path: Independent path

Definition 5.3.2 (Connectivity)

The number of independent paths between a pair of vertices is called the **connectivity** of the vertices. If we wish to be explicit about whether we are considering **edge-** or **vertex independence**, we refer to **edge** or **vertex connectivity**.

Example



Vertex-connectivity $= 1$ Edge-connectivity $= 2$

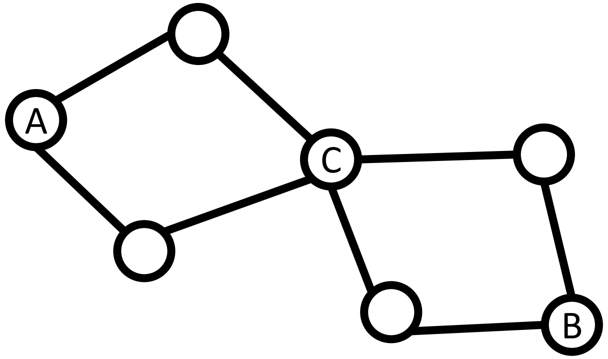
5.3 Path: Independent path

Definition 5.3.3 (Cut set)

A **cut set**, or more properly a **vertex cut set**, is a set of vertices whose removal will disconnect a specified pair of vertices.

An **edge cut set** is the equivalent construct for edges—it is a set of edges whose removal will disconnect a specified pair of vertices.

Example



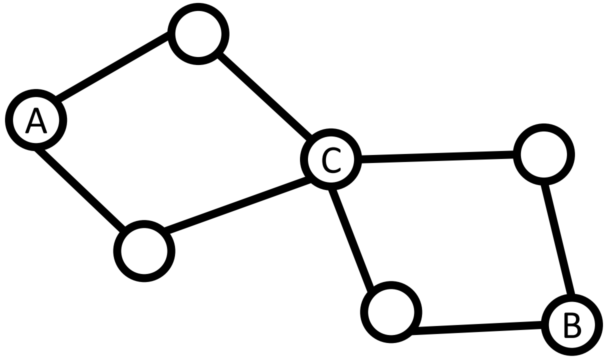
The central vertex C forms a cut set of size 1 for the vertices A and B. There are also other cut sets for A and B in this network, although all the others are larger than size 1.

5.3 Path: Independent path

Definition 5.3.4 (Minimum Cut set)

A **minimum cut set** is the smallest cut set that will disconnect a specified pair of vertices.

Example

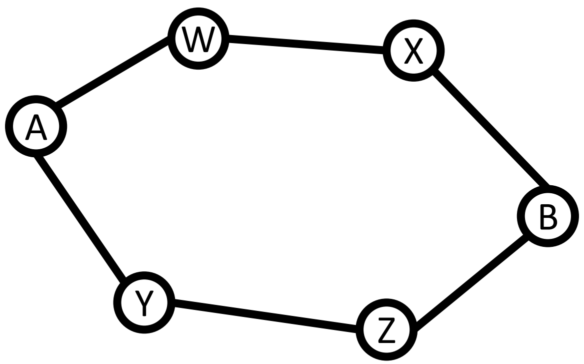


The single vertex C is a minimum vertex cut set for vertices A and B.

5.3 Path: Independent path

Note

A minimum cut set need not be unique. For instance, there is a variety of minimum vertex cut sets of size two between the vertices A and B in this network:



{W,Y}, {W,Z}, {X,Y}, and {X,Z} are all minimum cut sets for this network.

Theorem 5.3.1 (Menger's theorem)

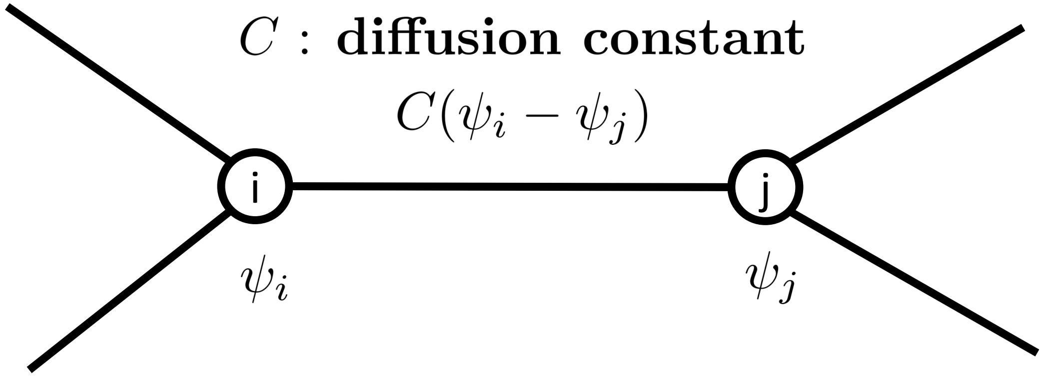
If there is no cut set of size less than n between a given pair of vertices, then there are at least n independent paths between the same vertices.

5.4 Graph Laplacian

Definition 5.4.1 (Diffusion)

Diffusion is, among other things, the process by which gas moves from regions of high density to regions of low, driven by the relative pressure (or partial pressure) of the different regions.

One can also consider diffusion processes on networks, and such processes are sometimes used as a simple model of spread across a network, such as the spread of an idea or the spread of a disease.



That is, in a small interval of time the amount of fluid flowing from j to i is $C(\psi_j - \psi_i)dt$.

$$\Rightarrow \frac{d\psi_i}{dt} = C \sum_j A_{ij}(\psi_j - \psi_i)$$

5.4 Graph Laplacian

$$\frac{d\psi_i}{dt} = C\sum_j A_{ij}\psi_j - C\psi_i \sum_j A_{ij} = C\sum_j A_{ij}\psi_j - C\psi_i k_i = C\sum_j \left(A_{ij} - \delta_{ij}k_i\right)\psi_j$$

where k_i is the degree of vertex i as usual and we have made use of the result $k_i = \sum_j A_{ij}$, and $\delta_i j$ is Kronecker delta, which is 1 if $i = j$ and 0 otherwise.

$$\therefore \frac{d\psi}{dt} = C(A - D)\psi$$

$$\text{where } D = \begin{bmatrix} k_1 & & & \mathbf{0} \\ & k_2 & & \\ & & \ddots & \\ \mathbf{0} & & & k_n \end{bmatrix}, \quad \psi = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_n \end{bmatrix}$$

$$\text{Let } L := D - A : \text{ Laplacian matrix}$$

$$\Rightarrow \frac{d\psi}{dt} + CL\psi = 0 \qquad - (*)$$

$$\left(\frac{d\psi}{dt} + C\Delta\psi = 0,\ \Delta \text{ is called Laplacian, and defined by } \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \Delta\right)$$

5.4 Graph Laplacian

$$L_{ij} = \begin{cases} k_i & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and there is an edge}, \\ 0 & \text{otherwise} \end{cases}$$

$$(L_{ij} = \delta_{ij} k_i - A_{ij})$$

Note

We can solve the diffusion equation $(*)$ by writing the vector ψ as a linear combination of the eigenvectors v_i of the Laplacian thus:

$$\psi(t) = \sum_i a_i(t) v_i$$

where $a_i(t)$ varying over time and v_i is an eigenvector of A.

$$\because \frac{d\psi}{dt} + CL\psi = 0, \Rightarrow \sum_i \frac{da_i(t)}{dt} v_i + CL \sum_i a_i(t) v_i \frac{da_i(t)}{dt} v_i + C \sum_i a_i(t) L v_i = 0$$

5.4 Graph Laplacian

Let $Lv_i = \lambda_i v_i$, where λ_i is the eigenvalue corresponding to the eigenvector v_i,

$$\Rightarrow \sum_i \frac{da_i(t)}{dt} v_i + C \sum_i a_i(t) \lambda_i v_i = 0$$

$$\Rightarrow \sum_i \left(\frac{da_i(t)}{dt} + C\lambda_i a_i(t) \right) v_i = 0$$

basis

$$\Rightarrow \frac{da_i(t)}{dt} + C\lambda_i a_i(t) = 0, \forall i.$$

$$\therefore a_i(t) = e^{-C\lambda_i t} a_i(0),\ a_i(0) : \text{initial conditino.}$$

Therefore $\psi(t) = \sum_i a_i(0) e^{-C\lambda_i t} v_i$

5.4 Graph Laplacian

Note

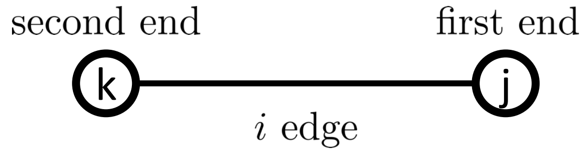
$$\text{Recall. at undirected network } L = D - A \Rightarrow L^\top = L : \text{ symmetric}$$

$$\Rightarrow \text{All eigenvalues are real}$$

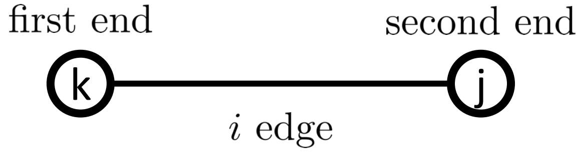
$$\text{Claim : All eigenvalues are non-negative}$$

Introducing $m \times n$ incidence matrix B, consider an undirected network with m vertices and m edges,

$$B_{ij} = \begin{cases} +1 & \text{if end 1 of edge } i \text{ is attached to vertex } j, \\ -1 & \text{if end 2 of edge } i \text{ is attached to vertex } j, \\ 0 & \text{otherwise} \end{cases}$$



$B_{ij} = 1$



$B_{ij} = -1$

5.4 Graph Laplacian

Note

$$\sum_k B_{ki} B_{kj} = \sum_k B^\top_{ik} B_{kj} = \begin{cases} -1 & \text{if } i \neq j \text{ and there is an edge from } i \text{ to } j \\ k_i & \text{if } i = j,\ (\sum_k B^2_{ki} = 1) \\ 0 & \text{otherwise} \end{cases}$$

$$= L_{ij}\ (L = D - A)$$

$$\therefore L = B^\top B$$

Note

Suppose $Lv_i = \lambda_i v_i$ with $\|v_i\| = 1$

$$\|Bv_i\|^2 = (Bv_i) \cdot (Bv_i) = (Bv_i)^\top (Bv_i) = v_i^\top B^\top B v_i$$

$$= v_i^\top \perp v_i = v_i^\top \lambda_i v_i = \lambda_i v_i^\top v_i$$

$$\therefore \lambda_i = \|Bv_i\|^2 \geq 0,\ (\because \|v_i\|^2 = 1)$$

$$\Rightarrow \psi(t) \longrightarrow 0 \text{ as } t \longrightarrow \infty$$

5.4 Graph Laplacian

Note

While the eigenvalues of the Laplacian cannot be negative, they can be zero, and in fact the Laplacian always has at least one zero eigenvalue.

$$\text{consider } \mathbf{1} = (1, 1, \cdots, 1)^\top$$

$$\sum_j L_{ij} \times 1 = \sum_j (\delta_{ij} k_i - A_{ij}) = k_i - \sum_j A_{ij} = k_i - k_i = 0$$

$$\left(\because \sum_j A_{ij} = k_i \right)$$

$$L_{i,:} \cdot \mathbf{1} = \begin{bmatrix} L_{i1} & L_{i2} & \cdots & L_{in} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\Rightarrow L \cdot \mathbf{1} = 0 = \underset{\text{eignevalue}}{0} \cdot \underset{\text{non-zero eigenvector}}{\mathbf{1}}$$