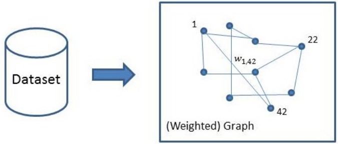
Lecture 1: From Data to Graphs, Weighted Graphs and Graph Laplacian

Radu Balan

University of Maryland

February 5, 2019

From Data to Graphs
Datasets Diversity



- Social Networks: Set of individuals ("agents", "actors") interacting with each other (e.g., Facebook, Twitter, joint paper authorship, etc.)
- Communication Networks: Devices (phones, laptops, cars) communicating with each other (emails, spectrum occupancy)
- Biological Networks: Macroscale: How animals interact with each other; Microscale: How proteins interact with each other.
- Databases of signals: speech, images, movies; graph relationship tends to reflect signal similarity: the higher the similarity, the larger the weight.
- Chemical networks: chemical compunds; each node is an atom.

Databases of Graphs

Public Datasets

Here are several public databases:

1. Duke: https://dnac.ssri.duke.edu/datasets.php
2. Stanford: https://snap.stanford.edu/data/
3. Uni. Koblenz: http://konect.uni-koblenz.de/
4. M. Newman (U. Michigan): http://www-personal.umich.edu/ mejn/netdata/
5. A.L. Barabasi (U. Notre Dame): http://www3.nd.edu/ networks/resources.htm
6. UCI: https://networkdata.ics.uci.edu/resources.php
7. Google/YouTube: https://research.google.com/youtube8m/
8. Chemical Compounds: http://quantum-machine.org/datasets/

Weighted Graphs
W

The main goal this lecture is to introduce basic concepts of weighted and undirected graphs, its associated graph Laplacian, and methods to build weight matrices.

Graphs (and weights) reflect either similarity between nodes, or functional dependency, or interaction strength.

- SIMILARITY: Distance, similarity between nodes $\Rightarrow$ weight $w_{i,j}$
- PREDICTIVE/DEPENDENCY: How node i is predicted by its nighbor node j $\Rightarrow$ weight $w_{i,j}$
- INTERACTION: Force, or interaction energy between atom i and atom j

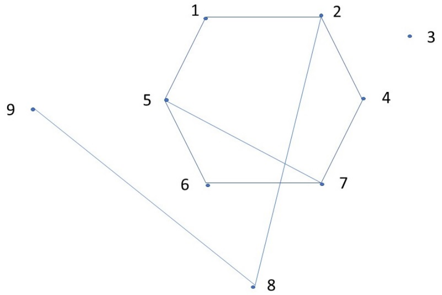
Definitions
$G = (\mathcal{V}, \mathcal{E})$ and $G = (\mathcal{V}, \mathcal{E}, w)$

An *undirected graph* G is given by two pieces of information: a set of *vertices* $\mathcal{V}$ and a set of *edges* $\mathcal{E}$, $G = (\mathcal{V}, \mathcal{E})$.
A *weighted graph* has three pieces of information: $G = (\mathcal{V}, \mathcal{E}, w)$, the set of vertices $\mathcal{V}$, the set of edges $\mathcal{E}$, and a *weight function* $w : \mathcal{E} \to \mathbb{R}$.

Definitions

$G = (\mathcal{V}, \mathcal{E})$ and $G = (\mathcal{V}, \mathcal{E}, w)$

An *undirected graph* G is given by two pieces of information: a set of *vertices* $\mathcal{V}$ and a set of *edges* $\mathcal{E}$, $G = (\mathcal{V}, \mathcal{E})$.
A *weighted graph* has three pieces of information: $G = (\mathcal{V}, \mathcal{E}, w)$, the set of vertices $\mathcal{V}$, the set of edges $\mathcal{E}$, and a *weight function* $w : \mathcal{E} \to \mathbb{R}$.



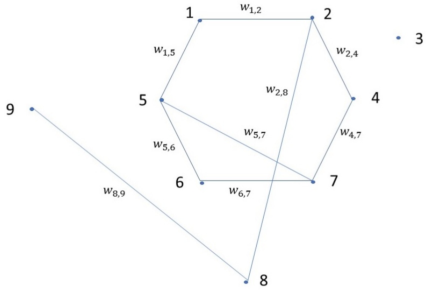
$\mathcal{V} = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$
$\mathcal{E} = \{(1, 2), (2, 4), (4, 7), (6, 7), (1, 5), (5, 6), (5, 7), (2, 8), (8, 9)\}$
9 *vertices* , 9 *edges*
Undirected graph, edges are not oriented. Thus $(1, 2) \sim (2, 1)$.

Definitions
$G = (\mathcal{V}, \mathcal{E})$

A weighted graph $G = (\mathcal{V}, \mathcal{E}, w)$ can be directed or undirected depending whether $w(i,j) = w(j,i)$.
Symmetric weights $==$ Undirected graphs

Definitions
$G = (\mathcal{V}, \mathcal{E})$

A weighted graph $G = (\mathcal{V}, \mathcal{E}, w)$ can be directed or undirected depending whether $w(i,j) = w(j,i)$.
Symmetric weights == Undirected graphs



Example 1 of a Weighted Graph
UCINET IV Datasets: Bernard & Killworth Office

Available online at:
http://vlado.fmf.uni-lj.si/pub/networks/data/ucinet/ucidata.htm
Content: Two 40×40 matrices: symmetric (B) and non-symmetric (C)
Bernard & Killworth, later with the help of Sailer, collected five sets of data on human interactions in bounded groups and on the actors' ability to recall those interactions. In each study they obtained measures of social interaction among all actors, and ranking data based on the subjects' memory of those interactions. The names of all cognitive (recall) matrices end in C, those of the behavioral measures in B.
These data concern interactions in a small business office, again recorded by an "unobtrusive" observer. Observations were made as the observer patrolled a fixed route through the office every fifteen minutes during two four-day periods. BKOFFB contains the observed frequency of interactions; BKOFFC contains rankings of interaction frequency as recalled by the employees over the two-week period.

Example 1 of a Weighted Graph

UCINET IV Datasets: Bernard & Killworth Office

bkoff.dat

DL

N=40 NM=2

FORMAT = FULLMATRIX DIAGONAL PRESENT

LEVEL LABELS:

BKOFFB

BKOFFC

DATA:

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
0 0 0 0 0 1 0 0 0 0 0 2 1 0
0 0 4 8 0 3 3 0 1 1 0 0 3 0 0 0 2 1 1 0 0 2 2 0 0 0
0 1 0 0 2 9 0 1 1 0 0 3 0 0
0 4 0 2 1 0 14 0 0 0 1 0 0 1 1 1 0 0 2 0 0 10 0 1 1 0
0 0 0 4 0 1 0 1 8 1 0 0 2 1
...

Example 1 of a Weighted Graph

UCINET IV Datasets: Bernard & Killworth Office

. . .

0 0 1 3 0 0 2 0 0 0 0 1 0 5 0 0 0 0 0 0 0 5 0 0 0 0

0 1 1 0 0 0 2 4 5 0 0 0 0 0

0 27 3 36 23 34 14 19 13 9 3 26 21 25 1 8 22 12 11 4 2 37 35 17 5 20

7 33 32 39 38 16 28 30 29 24 6 10 18 31

. . .

29 38 17 4 31 37 6 35 36 22 17 24 39 20 19 26 12 30 32 28 25 1 18 14 33 34

27 8 9 21 11 10 5 3 2 15 23 16 13 0

Example 2 of a Weighted Graph
QM9 Dataset: Molecular Compunds

QM9 Dataset available online at:
http://quantum-machine.org/datasets/
It contains 133,885 stable small molecules made up of $\{C, H, O, N, F\}$ each of about 10-30 atoms. For instance, for $C_7H_{10}O_2$ there are 6,095 isomers.
For instance file dsC7O2H10nsd_0300.xyz contains one such isomer. It has 24 (=2+19+3) lines. The mid 19 lines are important for the second solo homework.

Example 2 of a Weighted Graph

dsC7O2H10nsd_0300.xyz

19
gdb 300 2.6468919 2.0498681 1.8207938 1.876 72.89 -0.23607 0.0807 0.31677 866.986 0.161363 -422.551272 -422.544945 -422.544001 -422.581568 26.731
C 1.8119167747 -2.9989969945 3.3314921125 -0.266598
C 0.8677645122 -3.1880895657 2.1142035421 0.109841
O 1.5196288392 -2.9011626304 0.886880295 -0.272853
C 2.4559686885 -1.8275703466 1.0088791998 -0.097637
C 1.9689535703 -0.7816147977 2.0205224486 0.084531
O 0.8672780033 -0.0464843805 1.4906303854 -0.271398
C -0.3086335295 -0.8572021813 1.6209899755 -0.100804
C 0.0602954945 -1.9404530581 2.6286631808 -0.051904
C 1.4076218717 -1.5045220158 3.2719948006 -0.060027
H 1.4026278139 -3.4826781968 4.2219020256 0.104162
H 2.8612441374 -3.2835066136 3.2227123719 0.102564
H 0.3456726136 -4.1362959194 1.9757045623 0.077988

Example 2 of a Weighted Graph

dsC7O2H10nsd_0300.xyz – cont.

H 3.4414164369 -2.2113424376 1.3119407265 0.091892
H 2.5536336613 -1.3784780795 0.0173053432 0.11705
H 2.7646419377 -0.062340385 2.2342927387 0.081125
H -1.1244173217 -0.2177796796 1.9759113534 0.100289
H -0.5891314998 -1.2804120314 0.6481165581 0.107001
H -0.7324214782 -2.2127499243 3.3267153723 0.075012
H 1.425737714 -0.8881989422 4.1719569977 0.069768
195.7759 269.7508 375.0772 408.3447 441.6589 527.52 588.116 683.8729
730.6984 769.8084 830.7567 834.9175 876.9447 908.7187 934.7508 949.6675
957.0372 1003.4101 1021.008 1039.3247 1060.6831 1072.7141 1098.4612
1109.0223 1127.5117 1185.0693 1195.1973 1237.3076 1244.2752 1261.0481
1266.5805 1300.5391 1305.287 1325.1493 1342.8624 1363.584 1387.4244
1402.4656 1486.7371 1512.4786 1513.4557 2994.0025 3021.5295 3063.1291
3063.8915 3075.5124 3094.7683 3097.3621 3098.317 3109.794 3116.4017
C1C2OCC3OCC2C13 C1[C@@H]2OC[C@H]3OC[C@@H]2[C@@H]13
InChI=1S/C7H10O2/c1-4-5-2-8-7(4)3-9-6(1)5/h4-7H,1-3H2
InChI=1S/C7H10O2/c1-4-5-2-8-7(4)3-9-6(1)5/h4-7H,1-3H2/t4-,5-,6+,7-/m1/s1

Example 2 of a Weighted Graph

dsC7O2H10nsd_0300.xyz – cont.

Format:

Line 1: Number of Atoms

Line 2: Various properties

Lines 3-21: ElementType X Y Z [Angstrom] Q (=Mulliken charge) [e]

Line 22: Frequencies

Line 23: SMILES (Simplified Molecular-Input Line-Entry System)

Line 24: InChI (International Chemical Identifier)

Example 2 of a Weighted Graph

dsC7O2H10nsd_0300.xyz – cont.

Format:
Line 1: Number of Atoms
Line 2: Various properties
Lines 3-21: ElementType X Y Z [Angstrom] Q (=Mulliken charge) [e]
Line 22: Frequencies
Line 23: SMILES (Simplified Molecular-Input Line-Entry System)
Line 24: InChI (International Chemical Identifier)

How to create a weighted graph?

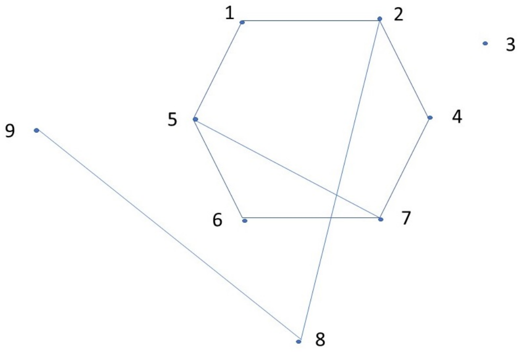
Use the electric interaction strength as weight:
Interaction Energy: $V_{ij} = \frac{Q_i Q_j}{R_{ij}}$, where
$R_{i,j} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2 + (Z_i - Z_j)^2}$.
To avoid signature problems, set $W_{i,j} = |V_{i,j}|$.

Definitions

Paths

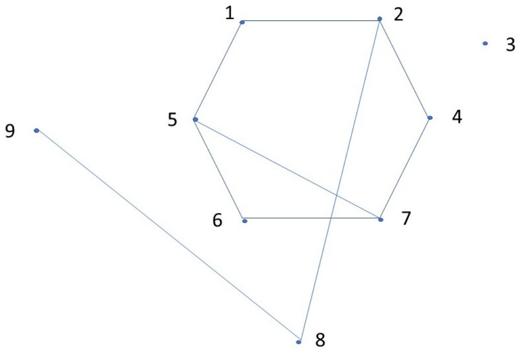
Concept: A path is a sequence of edges where the right vertex of one edge coincides with the left vertex of the following edge.

Example:



Definitions
Paths

Concept: A path is a sequence of edges where the right vertex of one edge coincides with the left vertex of the following edge.
Example:



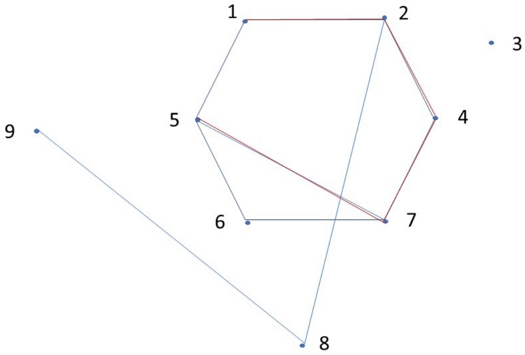
$$\{(1,2),(2,4),(4,7),(7,5)\} =$$
$$= \{(1,2),(2,4),(4,7),(5,7)\}$$

Definitions

Paths

Concept: A path is a sequence of edges where the right vertex of one edge coincides with the left vertex of the following edge.

Example:



$$\{(1,2),(2,4),(4,7),(7,5)\} =$$

$$= \{(1,2),(2,4),(4,7),(5,7)\}$$

Definitions
Graph Attributes

Graph Attributes (Properties):

- Connected Graphs: Graphs where any two distinct vertices can be connected through a path.

Definitions

Graph Attributes

Graph Attributes (Properties):

- Connected Graphs: Graphs where any two distinct vertices can be connected through a path.
- Complete (or Totally Connected) Graphs: Graphs where any two distinct vertices are connected by an edge.

Definitions
Graph Attributes

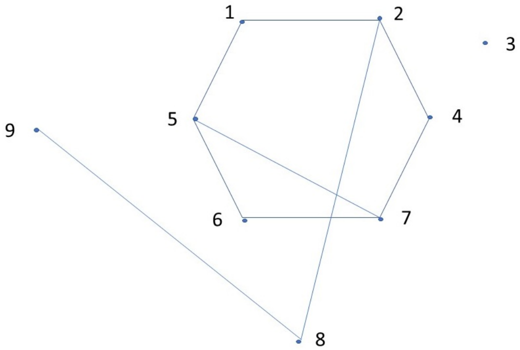
Graph Attributes (Properties):

- Connected Graphs: Graphs where any two distinct vertices can be connected through a path.
- Complete (or Totally Connected) Graphs: Graphs where any two distinct vertices are connected by an edge.

A complete graph with n vertices has $m = \binom{n}{2} = \frac{n(n-1)}{2}$ edges.

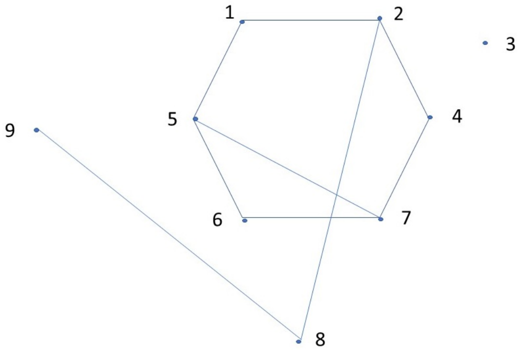
Definitions
Graph Attributes

Example:



Definitions
Graph Attributes

Example:



- This graph is not connected.
- It is not complete.
- It is the union of two connected graphs.

Definitions
Metric

Distance between vertices: For two vertices x, y, the distance $d(x, y)$ is the length of the shortest path connecting x and y. If $x = y$ then $d(x, x) = 0$.

Definitions
Metric

Distance between vertices: For two vertices x, y, the distance $d(x, y)$ is the length of the shortest path connecting x and y. If $x = y$ then $d(x, x) = 0$.
In a connected graph the distance between any two vertices is finite.
In a complete graph the distance between any two distinct vertices is 1.

Definitions
Metric

Distance between vertices: For two vertices x, y, the distance $d(x, y)$ is the length of the shortest path connecting x and y. If $x = y$ then $d(x, x) = 0$.
In a connected graph the distance between any two vertices is finite.
In a complete graph the distance between any two distinct vertices is 1.
The converses are also true:

1. If $\forall x, y \in \mathcal{E}$, $d(x, y) < \infty$ then $(\mathcal{V}, \mathcal{E})$ is connected.
2. If $\forall x \neq y \in \mathcal{E}$, $d(x, y) = 1$ then $(\mathcal{V}, \mathcal{E})$ is complete.

Definitions
Metric

Graph diameter: The diameter of a graph $G = (\mathcal{V}, \mathcal{E})$ is the largest distance between two vertices of the graph:

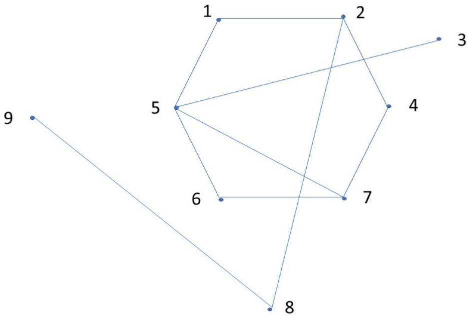
$$D(G) = \max_{x,y\in\mathcal{V}} d(x,y)$$

Definitions
Metric

Graph diameter: The diameter of a graph $G = (\mathcal{V}, \mathcal{E})$ is the largest distance between two vertices of the graph:

$$D(G) = \max_{x,y \in \mathcal{V}} d(x, y)$$

Example:

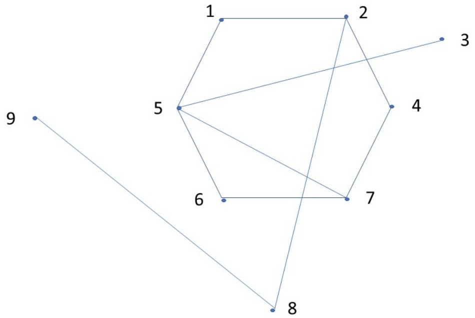


Definitions
Metric

Graph diameter: The diameter of a graph $G = (\mathcal{V}, \mathcal{E})$ is the largest distance between two vertices of the graph:

$$D(G) = \max_{x,y \in \mathcal{V}} d(x,y)$$

Example:



$$D = 5 = d(6,9) = d(3,9)$$

Definitions

The Adjacency Matrix

For a graph $G = (\mathcal{V}, \mathcal{E})$ the adjacency matrix is the $n \times n$ matrix A defined by:

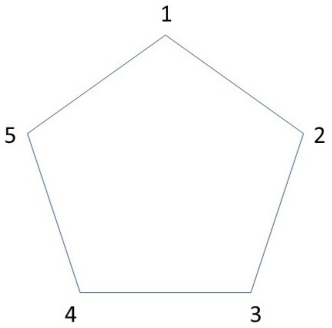
$$A_{i,j} = \begin{cases} 1 & \textit{if} \quad (i,j) \in \mathcal{E} \\ 0 & \quad \textit{otherwise} \end{cases}$$

Definitions
The Adjacency Matrix

For a graph $G = (\mathcal{V}, \mathcal{E})$ the adjacency matrix is the $n \times n$ matrix A defined by:

$$A_{i,j} = \begin{cases} 1 & if \quad (i,j) \in \mathcal{E} \\ 0 & \quad otherwise \end{cases}$$

Example:



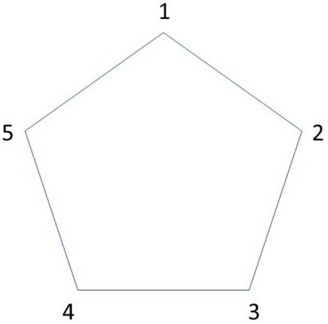
Definitions

The Adjacency Matrix

For a graph $G = (\mathcal{V}, \mathcal{E})$ the adjacency matrix is the $n \times n$ matrix A defined by:

$$A_{i,j} = \begin{cases} 1 & if \quad (i,j) \in \mathcal{E} \\ 0 & \quad otherwise \end{cases}$$

Example:



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

Definitions
The Adjacency Matrix

For undirected graphs the adjacency matrix is always symmetric:

$$A^T = A$$

For directed graphs the adjacency matrix may not be symmetric.

Definitions
The Adjacency Matrix

For undirected graphs the adjacency matrix is always symmetric:

$$A^T = A$$

For directed graphs the adjacency matrix may not be symmetric.
For weighted graphs $G = (\mathcal{V}, \mathcal{E}, W)$, the weight matrix W is simply given by

$$W_{i,j} = \begin{cases} w_{i,j} & \text{if} \quad (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$$

Degree Matrix

$d(v)$ and D

For an undirected graph $G = (\mathcal{V}, \mathcal{E})$, let $d_v = d(v)$ denote the number of edges at vertex $v \in \mathcal{V}$. The number $d(v)$ is called the degree (or valency) of vertex v. The n-vector $d = (d_1, \cdots, d_n)^T$ can be computed by

$$d = A \cdot 1$$

where A denotes the adjacency matrix, and 1 is the vector of 1's, ones(n,1).
Let D denote the diagonal matrix formed from the degree vector d: $D_{k,k} = d_k$, $k = 1, 2, \cdots, n$. D is called the *degree matrix*.

Degree Matrix
$d(v)$ and D

For an undirected graph $G = (\mathcal{V}, \mathcal{E})$, let $d_v = d(v)$ denote the number of edges at vertex $v \in \mathcal{V}$. The number $d(v)$ is called the degree (or valency) of vertex v. The n-vector $d = (d_1, \cdots, d_n)^T$ can be computed by

$$d = A \cdot 1$$

where A denotes the adjacency matrix, and 1 is the vector of 1's, ones(n,1).
Let D denote the diagonal matrix formed from the degree vector d: $D_{k,k} = d_k$, $k = 1, 2, \cdots, n$. D is called the *degree matrix*.
Key obervation: $(D - A) \cdot 1 = 0$ always holds. This means the matrix $D - A$ has a non-zero null-space (kernel), hence it is rank deficient.

Degree Matrix
$d(v)$ and D

For an undirected graph $G = (\mathcal{V}, \mathcal{E})$, let $d_v = d(v)$ denote the number of edges at vertex $v \in \mathcal{V}$. The number $d(v)$ is called the degree (or valency) of vertex v. The n-vector $d = (d_1, \cdots, d_n)^T$ can be computed by

$$d = A \cdot 1$$

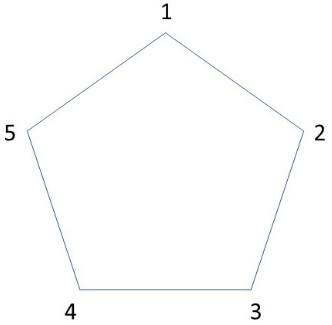
where A denotes the adjacency matrix, and 1 is the vector of 1's, ones(n,1).
Let D denote the diagonal matrix formed from the degree vector d: $D_{k,k} = d_k$, $k = 1, 2, \cdots, n$. D is called the *degree matrix*.
Key obervation: $(D - A) \cdot 1 = 0$ always holds. This means the matrix $D - A$ has a non-zero null-space (kernel), hence it is rank deficient.
Second observation: The dimension of the null-space of $D - A$ equals the number of connected components in the graph.

Vertex Degree
Matrix D

For an undirected graph $G = (\mathcal{V}, \mathcal{E})$ of n vertices, we denote by D the $n \times n$ diagonal matrix of degrees: $D_{i,i} = d(i)$.

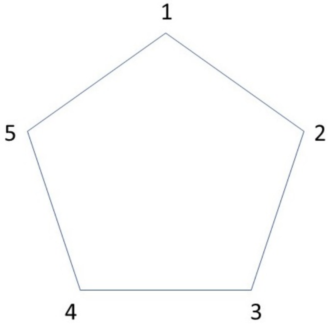
Vertex Degree
Matrix D

For an undirected graph $G = (\mathcal{V}, \mathcal{E})$ of n vertices, we denote by D the $n \times n$ diagonal matrix of degrees: $D_{i,i} = d(i)$. Example:



Vertex Degree
Matrix D

For an undirected graph $G = (\mathcal{V}, \mathcal{E})$ of n vertices, we denote by D the $n \times n$ diagonal matrix of degrees: $D_{i,i} = d(i)$. Example:



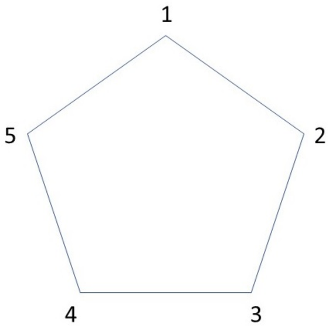
$$D = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

Graph Laplacian
Δ

For a graph $G = (\mathcal{V}, \mathcal{E})$ the graph Laplacian is the $n \times n$ symmetric matrix Δ defined by:

$$\Delta = D - A$$

Example:

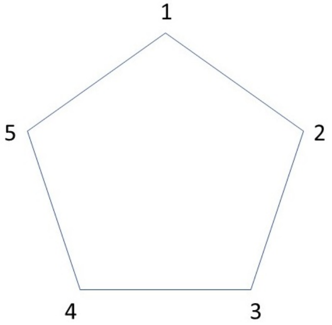


Graph Laplacian
Δ

For a graph $G = (\mathcal{V}, \mathcal{E})$ the graph Laplacian is the $n \times n$ symmetric matrix Δ defined by:

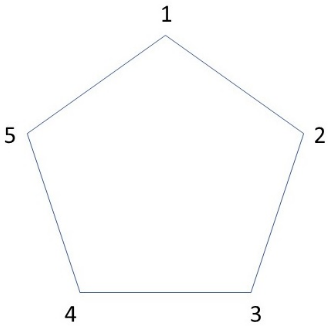
$$\Delta = D - A$$

Example:



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

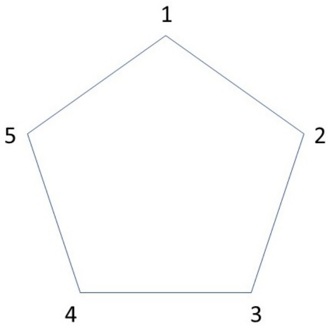
Graph Laplacian
Intuition



Assume $x = [x_1, x_2, x_3, x_4, x_5]^T$ is a signal of five components defined over the graph. The *Dirichlet* energy E, is defined as

$$E = \sum_{(i,j)\in\mathcal{E}} (x_i - x_j)^2 =$$

Graph Laplacian
Intuition

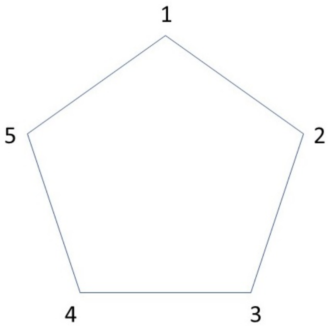


Assume $x = [x_1, x_2, x_3, x_4, x_5]^T$ is a signal of five components defined over the graph. The *Dirichlet* energy E, is defined as

$$E = \sum_{(i,j)\in\mathcal{E}} (x_i - x_j)^2 = (x_2 - x_1)^2 + (x_3 - x_2)^2 +$$

$$+(x_4 - x_3)^2 + (x_5 - x_4)^2 + (x_1 - x_5)^2.$$

Graph Laplacian
Intuition



Assume $x = [x_1, x_2, x_3, x_4, x_5]^T$ is a signal of five components defined over the graph. The *Dirichlet* energy E, is defined as

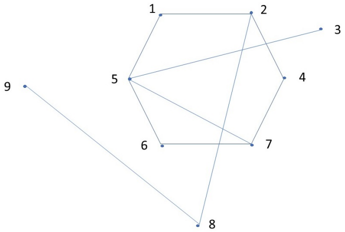
$$E = \sum_{(i,j)\in\mathcal{E}} (x_i - x_j)^2 = (x_2 - x_1)^2 + (x_3 - x_2)^2 +$$

$$+(x_4 - x_3)^2 + (x_5 - x_4)^2 + (x_1 - x_5)^2.$$

By regrouping the terms we obtain:

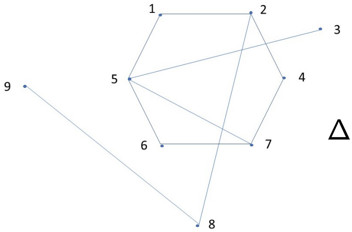
$$E = \langle \Delta x, x\rangle = x^T \Delta x = x^T (D - A)x$$

Graph Laplacian
Example



Graph Laplacian

Example



$$\Delta = \begin{bmatrix} 2 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 3 & 0 & -1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 2 & 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 & 4 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & -1 & 3 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

Normalized Laplacians
$\tilde{\Delta}$

Normalized Laplacian: (using pseudo-inverses)

$$\tilde{\Delta} = D^{\dagger/2} \Delta D^{\dagger/2} = \tilde{I} - D^{\dagger/2} A D^{\dagger/2}$$

$$\tilde{\Delta}_{i,j} = \left\{ \begin{array}{ccc} 1 & if & i=j \text{ and } d_i > 0 \text{ (non} - \text{isolated vertex)} \\ -\frac{1}{\sqrt{d(i)d(j)}} & if & (i,j) \in \mathcal{E} \\ 0 & & otherwise \end{array} \right.$$

where $\tilde{I}$ denotes the diagonal matrix: $\tilde{I}_{k,k} = 1$ if $d(k) > 0$ and $\tilde{I}_{k,k} = 0$ otherwise. Thus $\tilde{I}$ is the identity matrix if and only if the graph has no isolated vertices.

$D^{\dagger}$ denotes the pseudo-inverse, which is the diagonal matrix with elements:

$$D^{\dagger}_{k,k} = \left\{ \begin{array}{ccc} \frac{1}{d(k)} & if & d(k) > 0 \\ 0 & if & d(k) = 0 \end{array} \right.$$

Normalized Laplacians

L

Normalized Asymmetric Laplacian:

$$L = D^{\dagger}\Delta = \tilde{I} - D^{\dagger}A$$

$$L_{i,j} = \begin{cases} 1 & if \quad i = j \text{ and } d_i > 0 \text{ (non} - \text{isolated vertex)} \\ -\frac{1}{d(i)} & if \quad (i,j) \in \mathcal{E} \\ 0 & otherwise \end{cases}$$

Normalized Laplacians
L

Normalized Asymmetric Laplacian:

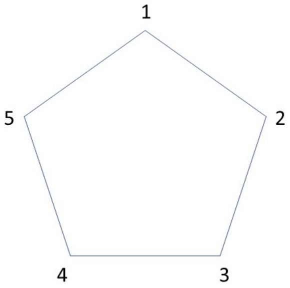
$$L = D^{\dagger}\Delta = \tilde{I} - D^{\dagger}A$$

$$L_{i,j} = \begin{cases} 1 & \textit{if} \quad i = j \text{ and } d_i > 0 \text{ (non} - \text{isolated vertex)} \\ -\frac{1}{d(i)} & \textit{if} \quad (i,j) \in \mathcal{E} \\ 0 & \textit{otherwise} \end{cases}$$

$$\Delta D^{\dagger} = \tilde{I} - AD^{\dagger} = L^T$$

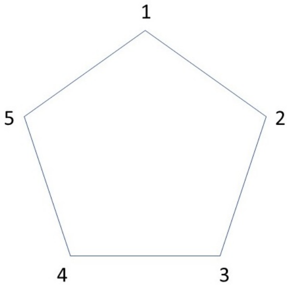
Normalized Laplacians
Example

Example:



Normalized Laplacians
Example

Example:



$$\tilde{\Delta} = \begin{bmatrix} 1 & -0.5 & 0 & 0 & -0.5 \\ -0.5 & 1 & -0.5 & 0 & 0 \\ 0 & -0.5 & 1 & -0.5 & 0 \\ 0 & 0 & -0.5 & 1 & -0.5 \\ -0.5 & 0 & 0 & -0.5 & 1 \end{bmatrix}$$

Laplacian and Normalized Laplacian for Weighted Graphs

In the case of a weighted graph, $G = (\mathcal{V}, \mathcal{E}, w)$, the weight matrix W replaces the adjacency matrix A.

Laplacian and Normalized Laplacian for Weighted Graphs

In the case of a weighted graph, $G = (\mathcal{V}, \mathcal{E}, w)$, the weight matrix W replaces the adjacency matrix A.
The other matrices:

$$D = diag(W \cdot 1) \quad , \quad D_{k,k} = \sum_{j \in \mathcal{V}} W_{k,j}$$

$$\Delta = D - W \quad , \quad \dim\ker(D - W) = \text{number connected components}$$

$$\tilde{\Delta} = D^{\dagger/2} \Delta D^{\dagger/2}$$

$$L = D^{\dagger} \Delta$$

where $D^{\dagger/2}$ and $D^{\dagger}$ denote the diagonal matrices:

$$(D^{\dagger/2})_{k,k} = \begin{cases} \frac{1}{\sqrt{D_{k,k}}} & if \quad D_{k,k} > 0 \\ 0 & if \quad D_{k,k} = 0 \end{cases} \quad , \quad (D^{\dagger})_{k,k} = \begin{cases} \frac{1}{D_{k,k}} & if \quad D_{k,k} > 0 \\ 0 & if \quad D_{k,k} = 0 \end{cases} .$$

Laplacian and Normalized Laplacian for Weighted Graphs
Dirichlet Energy

For symmetric (i.e., undirected) weighted graphs, the Dirichlet energy is defined as (note edges contribute two terms in the sum)

$$E = \frac{1}{2} \sum_{i,j \in \mathcal{V}} w_{i,j} |x_i - x_j|^2$$

Expanding the square and grouping the terms together, the expression simplifies to

$$\sum_{i \in \mathcal{V}} |x_i|^2 \sum_j w_{ij} - \sum_{i,j \in \mathcal{V}} w_{i,j} x_i x_j = \langle Dx, x \rangle - \langle Wx, x \rangle = x^T (D - W) x.$$

Hence:

$$E = \frac{1}{2} \sum_{i,j \in \mathcal{V}} w_{i,j} |x_i - x_j|^2 = x^T \Delta x$$

where $\Delta = D - W$ is the weighted graph Laplacian.

Spectral Analysis
Eigenvalues and Eigenvectors

Recall the eigenvalues of a matrix T are the zeros of the characteristic polynomial:

$$p_T(z) = det(zI - T) = 0.$$

There are exactly n eigenvalues (including multiplicities) for a $n \times n$ matrix T. The set of eigenvalues is calles its *spectrum*.

Spectral Analysis
Eigenvalues and Eigenvectors

Recall the eigenvalues of a matrix T are the zeros of the characteristic polynomial:

$$p_T(z) = det(zI - T) = 0.$$

There are exactly n eigenvalues (including multiplicities) for a $n \times n$ matrix T. The set of eigenvalues is calles its *spectrum*.
If λ is an eigenvalue of T, then an associated eigenvector is the non-zero n-vector x such that $Tx = \lambda x$.

Spectral Analysis
Eigenvalues and Eigenvectors

Recall the eigenvalues of a matrix T are the zeros of the characteristic polynomial:

$$p_T(z) = det(zI - T) = 0.$$

There are exactly n eigenvalues (including multiplicities) for a $n \times n$ matrix T. The set of eigenvalues is calles its *spectrum*.
If λ is an eigenvalue of T, then an associated eigenvector is the non-zero n-vector x such that $Tx = \lambda x$.
Recall: If $T = T^T$ then T is called a *symmetric matrix*. Furthermore:

- Every eigenvalue of T is real.
- There is a set of n eigenvectors $\{e_1, \cdots, e_n\}$ normalized so that the matrix $U = [e_1 | \cdots | e_n]$ is orthogonal ($UU^T = U^T U = I_n$) and $T = U \Lambda U^T$, where Λ is the diagonal matrix of eigenvalues.

Spectral Analysis

Eigenvalues and Eigenvectors

Recall the eigenvalues of a matrix T are the zeros of the characteristic polynomial:

$$p_T(z) = det(zI - T) = 0.$$

There are exactly n eigenvalues (including multiplicities) for a $n \times n$ matrix T. The set of eigenvalues is calles its *spectrum*.

If λ is an eigenvalue of T, then an associated eigenvector is the non-zero n-vector x such that $Tx = \lambda x$.

Recall: If $T = T^T$ then T is called a *symmetric matrix*. Furthermore:

- Every eigenvalue of T is real.
- There is a set of n eigenvectors $\{e_1, \cdots, e_n\}$ normalized so that the matrix $U = [e_1|\cdots|e_n]$ is orthogonal ($UU^T = U^T U = I_n$) and $T = U\Lambda U^T$, where Λ is the diagonal matrix of eigenvalues.

Remark. Since $det(A_1 A_2) = det(A_1)det(A_2)$ and $L = D^{-1/2}\tilde{\Delta}D^{1/2}$ it follows that $eigs(\tilde{\Delta}) = eigs(L) = eigs(L^T)$.

Spectral Analysis
UCINET IV Database: Bernard & Killworth Office Dataset

For the Bernard & Killworth Office dataset (bkoff.dat) dataset we obtained the following results:
The graph is connected. $rank(\Delta) = rank(\tilde{\Delta}) = rank(L) = 39$.

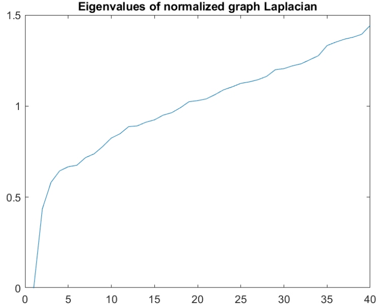


Figure: Adjacency Matrix based Graph Laplacian

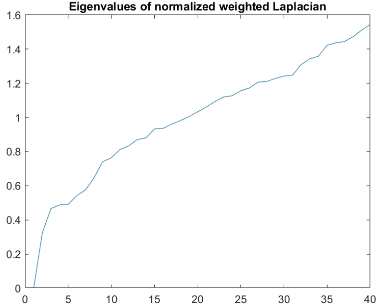


Figure: Weight Matrix based Graph Laplacian

Spectral Analysis

Symmetric Matrices

Recall the following result:

Theorem

Assume T is a real symmetric $n \times n$ matrix. Then:

1. *All eigenvalues of T are real numbers.*
2. *There are n eigenvectors that can be normalized to form an orthonormal basis for $\mathbb{R}^n$.*
3. *The largest eigenvalue λ_{max} and the smallest eigenvalue λ_{min} satisfy*

$$\lambda_{max} = \max_{x \neq 0} \frac{\langle Tx, x \rangle}{\langle x, x \rangle} \quad , \quad \lambda_{min} = \min_{x \neq 0} \frac{\langle Tx, x \rangle}{\langle x, x \rangle}$$

Spectral Analysis
Rayleigh Quotient

For two symmetric matrices T, S we say $T \leq S$ if $\langle Tx, x\rangle \leq \langle Sx, x\rangle$ for all $x \in \mathbb{R}^n$.

Spectral Analysis
Rayleigh Quotient

For two symmetric matrices T, S we say $T \leq S$ if $\langle Tx, x \rangle \leq \langle Sx, x \rangle$ for all $x \in \mathbb{R}^n$.
Consequence 3 can be rewritten:

$$\lambda_{min} I \leq T \leq \lambda_{max} I$$

Spectral Analysis

Rayleigh Quotient

For two symmetric matrices T, S we say $T \leq S$ if $\langle Tx, x\rangle \leq \langle Sx, x\rangle$ for all $x \in \mathbb{R}^n$.
Consequence 3 can be rewritten:

$$\lambda_{min} I \leq T \leq \lambda_{max} I$$

In particular we say T is positive semidefinite $T \geq 0$ if $\langle Tx, x\rangle \geq 0$ for every x.
It follows that T is positive semidefinite if and only if every eigenvalue of T is positive semidefinite (i.e. non-negative).

References

- B. Bollobás, **Graph Theory. An Introductory Course**, Springer-Verlag 1979. **99**(25), 15879–15882 (2002).
- F. Chung, L. Lu, The average distances in random graphs with given expected degrees, Proc. Nat.Acad.Sci.
- R. Diestel, **Graph Theory**, 3rd Edition, Springer-Verlag 2005.
- P. Erdös, A. Rényi, On The Evolution of Random Graphs
- G. Grimmett, **Probability on Graphs. Random Processes on Graphs and Lattices**, Cambridge Press 2010.
- J. Leskovec, J. Kleinberg, C. Faloutsos, Graph Evolution: Densification and Shrinking Diameters, ACM Trans. on Knowl.Disc.Data, **1**(1) 2007.