1 Graphs

Graphs play a key role in data analysis. A graph $G = (V, E)$ contains a set of nodes $V = \{v_1, \ldots, v_n\}$ and edges $E \subseteq \binom{V}{2}$. An edge $(i, j) \in E$ if v_i and v_j are connected. Here is one of the graph theorists favorite examples, the Petersen graph[1]:

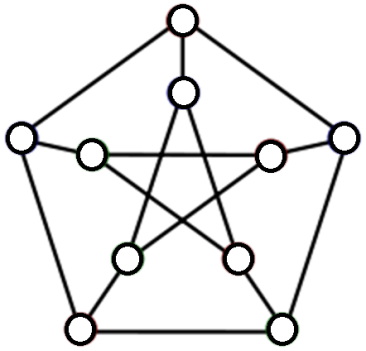


Figure 1: The Petersen graph

Graphs are crucial tools in many fields, the intuitive reason being that many phenomena, while complex, can often be thought about through pairwise interactions between objects (or data points), which can be nicely modeled with the help of a graph.

Let us recall some concepts about graphs that we will need.

- A graph is complete if, for all pairs of vertices, there is a path between these vertices on the graph. The number of connected components is simply the size of the smallest partition of the nodes into connected subgraphs. The Petersen graph is connected (and thus it has only 1 connected component).
- A clique of a graph G is a subset S of its nodes such that the subgraph corresponding to it is complete. In other words S is a clique if all pairs of vertices in S share an edge. The clique number $c(G)$ of G is the size of the largest clique of G. The Petersen graph has a clique number of 2.
- An independence set of a graph G is a subset S of its nodes such that no two nodes in S share an edge. Equivalently it is a clique of the complement graph $G^c := (V, E^c)$. The independence number of G is simply the clique number of S^c. The Petersen graph has an independence number of 4.

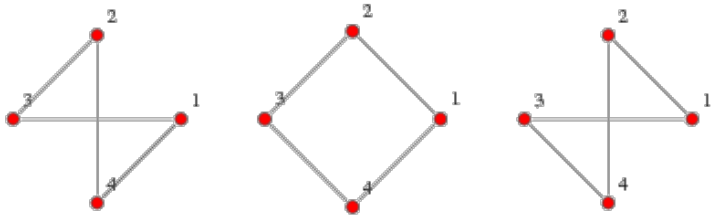
[1]The Peterson graph is often used as a counter-example in graph theory.

2 Graph Laplacian

A particularly useful way to represent a graph is through its adjacency matrix. Given a graph $G = (V, E)$ on n nodes ($|V| = n$), we define its adjacency matrix $A \in \mathbb{R}^{n \times n}$ as the symmetric matrix with entries

$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{if } i,j \text{ are not connected} \\ 0 & \text{if } i = j \end{cases}$$

Figure 2 shows some examples of graphs and their adjacency matrices.



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

Figure 2: Some graphs and their adjacency matrices.

Sometimes, we will consider weighted graphs $G = (V, E, W)$, where edges may have weights w_{ij}, we think of the weights as non-negative $w_{ij} \geq 0$ and symmetric $w_{ij} = w_{ji}$, $w_{ij} = 0$ if two vertices are not connected and $w_{ii} = 0$. For undirected graph, $w_{ij} = w_{ji}$. *Unweighted graphs* are weighted graphs with $w_{ij} = 0$ if not connected and $w_{ij} = 1$ if connected.

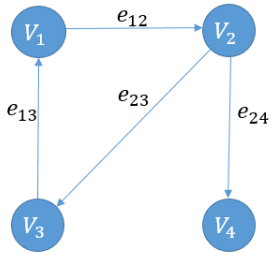
The *degree of a vertex* $v_i \in V$ is $d_i = \sum_{j=1}^{n} w_{ij}$.
The *degree matrix* D is the diagonal matrix with entries d_i.
The *(unweighted) adjacancy matrix* A is defined as:

The *incidence matrix* ∇ is a matrix of size $|E| \times |V|$

$$\nabla := \begin{cases} \nabla = -1 & \text{if } v \text{ is the initial vertex of the edge} \\ \nabla = 1 & \text{if } v \text{ is the terminal vertex of the edge} \\ \nabla = 0 & \text{if } v \text{ is not in the edge,} \end{cases}$$

where we assume each edge has an arbitrary, but fixed orientation. If we have an undirected graph, we can obtain the incidence matrix by simply choosing a (fixed) orientation of the edges.

Here is an example:



$$\nabla = \begin{array}{c|cccc} & v_1 & v_2 & v_3 & v_4 \\ \hline e_{12} & -1 & 1 & 0 & 0 \\ e_{13} & 1 & 0 & -1 & 0 \\ e_{23} & 0 & -1 & 1 & 0 \\ e_{24} & 0 & -1 & 0 & 1 \end{array}$$

Let f be a function acting on the vertices, then $\nabla f = f(v_j) - f(v_i)_{i,j}$. Hence ∇ is a kind of difference operator.

The *unnormalized graph Laplacian* L is defined as $L := \nabla^T \nabla$. It is easy to see that there holds $L = D - A$. For the same example above,

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow L = D - A = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} = \nabla^T \nabla.$$

Sometimes L is defined directly via $L = D - A$. For weighted graphs, we define $L = D - W$.

Theorem 1. *The graph Laplacian L satisfies the following properties:*

1. *For every vector $f \in \mathbb{R}^n$ there holds*

$$f^* L f = \frac{1}{2} \sum_{i,j=1}^{n} w_{ij} (f_i - f_j)^2. \tag{1}$$

2. *L is symmetric and positive semidefinite.*
3. *0 is an eigenvalue of L and the associated eigenvector is $\frac{1}{\sqrt{n}} [1, 1, \cdots, 1]^T$.*

Proof. 1. We have

$$f^*Lf = f^*Df - f^*Wf = \sum_{i=1}^{n} d_i f_i^2 - \sum_{i,j} w_{ij} f_i f_j$$
$$= \frac{1}{2}\Big(\sum_{i=1}^{n} d_i f_i^2 - 2\sum_{i,j} w_{ij} f_i f_j + \sum_{j=1}^{n} d_j f_j^2\Big) = \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2.$$

2. The symmetry of L follows from the symmetry of D and W and the positive-semidefiniteness follows e.g. from (5) which implies that $f^*Lf \geq 0$ for all $f \in \mathbb{R}^n$.
3. Follows immediately from (5). □

The unnormalized graph Laplacian and its eigenvalues and eigenvectors can be used to describe many properties of graphs [1]. In connection with clustering it leads to an approach called *spectral clustering*. The following theorem illustrated one such connection to clustering [2].

Theorem 2. *[Number of connected components and the spectrum of L] Let G be an undirected graph with non-negative weights. Then the multiplicity k of the eigenvalue 0 of L equals the number of connected components $A_1, \ldots, A_k$ in the graph. The eigenspace associated with the eigenvalue 0 is spanned by the indicator vectors $\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_k}$ of those components.*

Proof. We start with the case $k = 1$, that is, the graph is connected. Assume that f is an eigenvector with eigenvalue 0. Then we know from Theorem 1 that

$$0 = f^*Lf = \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2. \tag{2}$$

Since $w_{i,j} \geq 0$, the sum in (2) can only be zero if all terms $w_{ij}(f_i - f_j)^2$ are zero. Thus, if two vertices v_i and v_j are connected (i.e., $w_{ij} > 0$), then f_i must be equal to f_j . Hence, f must be constant for all vertices which can be connected by a path in the graph, which in turn implies that f is constant on the whole connected component. In a graph consisting of only one connected component we thus only have the constant-one vector as eigenvector with eigenvalue 0, which obviously is the indicator vector of the connected component.

Now consider the case of k connected components. Without loss of generality we assume that the vertices are ordered according to the connected components they belong to. In this case, the adjacency matrix W is block diagonal, and so is the matrix L:

$$L = \begin{bmatrix} L_1 & & & \\ & L_2 & & \\ & & \ldots & \\ & & & L_k \end{bmatrix}.$$

Note that each of the blocks L_i can be considered as a graph Laplacian on its own, namely the Laplacian corresponding to the subgraph of the i-th connected component. Since L is a

block diagonal matrix, its spectrum is the union of the spectra of L_i and the corresponding eigenvectors of L are the eigenvectors of L_i, padded with zeroes at the positions of the other blocks. As each L_i is a graph Laplacian of a connected graph, we know that each L_i has eigenvalue 0 with multiplicity 1, and the corresponding eigenvector is the constant-one vector on the i-th connected component. Thus, the matrix L has as many eigenvalues 0 as there are connected components, and the corresponding eigenvectors are the indicator vectors of the connected components. □

There are two normalized versions of the graph Laplacian, a symmetric one and a non-symmetric one, given by

$$L_S = D^{-\frac{1}{2}} L D^{-\frac{1}{2}} = I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}} \tag{3}$$

$$L_N = D^{-1} L = I - D^{-1} W. \tag{4}$$

We will later see that L_N has a nice and natural interpretation of a random walk on a graph.

Theorem 3. *The normalized graph Laplacians satisfy the following properties:*

1. *For every vector $f \in \mathbb{R}^n$ there holds*

$$f^* L_S f = \frac{1}{2} \sum_{i,j=1}^{n} w_{ij} \Big(\frac{f_i}{\sqrt{d_i}} - \frac{f_j}{\sqrt{d_j}} \Big)^2. \tag{5}$$

2. *λ is an eigenvalue of L_N with eigenvector u if and only if λ is an eigenvalue of L_S with eigenvector $w = D^{\frac{1}{2}} u$.*
3. *λ is an eigenvalue of L_N with eigenvector u if and only if λ and u solve the generalized eigenproblem $Lu = \lambda Du$.*
4. *0 is an eigenvalue of L_N and the associated eigenvector is $\mathbf{1}$. 0 is an eigenvalue of L_S and the associated eigenvector is $D^{\frac{1}{2}} \mathbf{1}$.*

Theorem 2 can be easily adapted to L_S and L_N.

References

[1] Fan RK Chung. *Spectral graph theory*, volume 92. American Mathematical Soc., 1997.

[2] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007.