CISC 271 Class 2

Graphs: Adjacency Matrix and Laplacian Matrix

Text Correspondence, Goodaire & Parmentier: §9.1–9.2, §10.3

Main Concepts:

- *Graph $\mathcal{G}$: structure in discrete mathematics*
- *Vertex set $\mathcal{V}$: finite non-empty set*
- *Edge set $\mathcal{E}$: set of distinct pairs of vertices*
- *Adjacency matrix: symmetric nonnegative matrix of a graph*
- *Degree matrix D of graph $\mathcal{G}$: diagonal matrix, each entry is sum of rows/columns*
- *Laplacian matrix of graph $\mathcal{G}$: $L = D - A$*
- *Laplacian matrix is symmetric positive semidefinite*
- *Fiedler vector: eigenvector of second-smallest eigenvalue*
- *Binary cluster: vertices classified by sign of Fiedler vector*

Sample Problem, Data Analysis: How can we cluster a graph from a matrix?

A *graph* is a term for the abstract structure that represents many common drawings and ideas. For example, when we look at a "map" of part of the London Underground subway system, company, we see subway stops as circles or ellipses that are connected by straight or curved lines; a sample of such a "map" is shown in Figure 2.1.

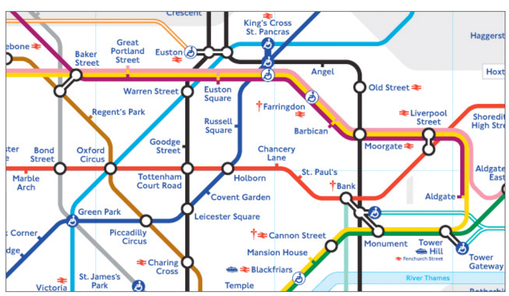

Figure 2.1: A "map" of part of the underground rail system of London, Great Britain. Subway stops are shown as circles, ellipses, or bars. Stops for which a train is directly available are shown as various straight and curved lines. (This image is sourced from `https://www.visitlondon.com` and is presented as "fair use".)

A graph is used to represent many familiar complex relationships. Examples in computer science include the network connectivity between computing nodes and the layout of transistors on a silicon chip. In this class, we are less interested in the applications of graphs or in the theory of computation for graphs. Our interest is in the representation of a graph as a matrix and some uses of linear algebra for a graph matrix.

In the intellectual field of discrete mathematics, a *graph* $\mathcal{G}$ is a structure that has two sets. We must be careful when defining and using graphs because there is no effective standardization of notation or meanings. When we read each text or other reference work, we need to read carefully to ensure that we understand the meaning of the author(s).

The first set of a graph $\mathcal{G}$, the *vertices*, is non-empty and finite. Each member of this set is a *vertex*. We will write the vertex set as $\mathcal{V}$ and each vertex as v_j.

The second set of a graph $\mathcal{G}$, the *edges*, is finite. Each member of the set is a pair of vertices. We will write the edge set as $\mathcal{E}$ and each edge as $e_k = \{v_i \;\, v_j\}$ or as $e_k = (v_i \;\, v_j)$, with the understanding that the second notation is not an ordered set. We will generally write a graph as $\mathcal{G}(\mathcal{V}, \mathcal{E})$. We may use subscripts to distinguish graphs from each other.

We will make two assumptions about a graph that are important and that vary considerably between authors. The first assumption is that a graph is *undirected*. In the subway example, this implies that if it is possible to go from station v_i to station v_j without visiting any other station, then it is likewise possible to go from station v_j to station v_i without visiting any other station. If the graph represents city streets, then every street is "two-way" and there are no "one-way" streets.

The second important assumption that we will make is that there is no *loop* in a graph. By this, we intend that for each vertex v_j there is no edge of the form $\{v_j \;\, v_j\}$.

For example, suppose that we construct a graph $\mathcal{G}_1(\mathcal{V}, \mathcal{E})$ that has five vertices and four edges:

$$\mathcal{G}_1 \stackrel{\text{def}}{=} \left\{ \begin{array}{lcl} \mathcal{V}_1 & = & \{1\;2\;3\;4\;5\} \\ \mathcal{E}_1 & = & \{(1\;2)\;\;(1\;3)\;\;(2\;3)\;\;(4\;5)\} \end{array} \right. \tag{2.1}$$

We can produce a diagram of the graph $\mathcal{G}_1$ in a similar way that the map in Figure 2.1 is drawn. We can use a circle for each vertex and a line for each edge. The graph defined in Equation 2.1 can be diagrammed as shown in Figure 2.2.

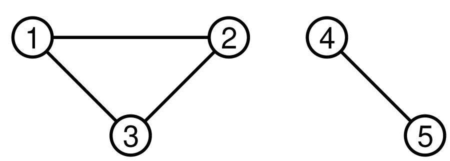

Figure 2.2: An illustration of the graph $\mathcal{G}_1$, having five vertices and four edges. The graph has two distinct sub-graphs that have no shared edge.

2.1 Relevant Definitions

In this course, we will use very limited graph theory. For our graphs, which are undirected and have no loops, some relevant definitions are:

incident vertex and edge: a vertex v_i and an edge e for which $e = (v_i\ v_j) = (v_j\ v_i)$

adjacent vertices: vertices v_i and v_j for which there exists an edge $e = (v_i\ v_j)$

edge weight: $w_{ij} > 0$ that represents "importance" or a related concept

degree: for a vertex v_j, the weighted sum of edges e_i that are incident to v_j

multiple edge: distinct edges e_p and e_q that are equal, so that $e_p = e_q$

loop: an edge $e = (v_j\ v_j)$

pseudograph: a set $\{\mathcal{V}\ \mathcal{E}\}$ that has at least one multiple edge and/or at least one loop

subgraph: for a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, a graph $\mathcal{G}_S(\mathcal{V}_S, \mathcal{E}_S)$ is a *subgraph* of $\mathcal{G}$ is defined as: $\mathcal{V}_S \subseteq \mathcal{V}$ and $\mathcal{E}_S \subseteq \mathcal{E}$

bipartite graph: a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ such that: (a) $\mathcal{V}$ can be partitioned into two disjoint non-empty sets $\mathcal{V}_L$ and $\mathcal{V}_R$, and (b) each edge e_k is incident with a vertex $v_i \in \mathcal{V}_L$ and is incident with a vertex $v_j \in \mathcal{V}_R$

From these definitions, we can make a few observations on the graph $\mathcal{G}_1$, such as:

- edge $(1\ 2)$ is incident with vertex 1
- vertex 3 is not incident with $(1\ 2)$
- the degree of vertex 3 is 2 and the degree of vertex 4 is 1
- vertex 2 is adjacent to vertex 3
- vertex 2 is not adjacent to vertex 4
- no edge is adjacent to vertex 4 and to any of vertices 1–3
- the graph $\mathcal{G}_1$ is not bipartite

The observation that graph $\mathcal{G}_1$ is not bipartite follows from the "triangle" formed by the first three vertices.

An example of a bipartite graph is the graph $\mathcal{G}_2$ that has six vertices and five edges:

$$\mathcal{G}_2 \stackrel{\text{def}}{=} \left\{ \begin{array}{lcl} \mathcal{V}_2 & = & \{1\ 2\ 3\ 4\ 5\ 6\} \\ \mathcal{E}_2 & = & \{(1\ 2)\ \ (2\ 3)\ \ (2\ 4)\ \ (4\ 5)\ \ (4\ 6)\} \end{array} \right. \tag{2.2}$$

The graph defined in Equation 2.2 can be drawn as shown in Figure 2.3. One way to partition the vertices of graph $\mathcal{G}_2$ is as $\mathcal{V}_L = \{2\ 5\ 6\}$ and $\mathcal{V}_R = \{1\ 3\ 4\}$.

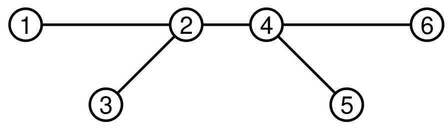

Figure 2.3: An illustration of the bipartite graph $\mathcal{G}_2$, having six vertices and five edges.

An observation on the graph $\mathcal{G}_2$ is that every nontrivial sub-graph is also a bipartite graph; this is a fact about all bipartite graphs, which we will not prove. The interested reader is encouraged to explore graph theory for more on these topics.

2.2 Adjacency Matrix of a Graph

Although the intellectual field of discrete mathematics uses a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ directly, we are more interested in using a matrix representation of a graph. There are two fundamental ways of representing a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ in linear algebra: as a square *adjacency* matrix that has a size that varies according to the number of vertices in $\mathcal{V}$; and as an *incidence* matrix that has a size that varies according to the number of vertices in $\mathcal{V}$ and the number of edges in $\mathcal{E}$. We will work with the adjacency matrix and recognize that the incidence matrix is also useful.

Definition: adjacency matrix $A(\mathcal{V}, \mathcal{E})$ of a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$

> For any graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ that has n vertices in $\mathcal{V}$ labeled $v_1, v_2, \ldots, v_n$ and that has m edges labeled $(v_i\ v_j)$, the *adjacency matrix* $A(\mathcal{G})$ is the $n \times n$ matrix for which each entry is defined as
>
> $$a_{ij} = \left\{ \begin{array}{ll} 1 & \text{if } (v_i\ v_j) \in \mathcal{E} \\ 0 & \text{if } (v_i\ v_j) \notin \mathcal{E} \end{array} \right. \tag{2.3}$$

According to Definition 2.3, any adjacency matrix A of a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ must have these properties:

- A is real, because either $a_{ij} = 0$ or $a_{ij} = 1$
- A is symmetric, because if $(v_i \;\; v_j) \in \mathcal{E}$ then $a_{ij} = a_{ji} = 1$; this is true of all undirected graphs, which is the kind of graph that we are using
- A has diagonal entries $a_{jj} = 0$, because the kind of graph that we are using has no loop
- The eigenvalues of A are real, because A is a real symmetric matrix
- The eigenvalues of A add up to zero, because $\mathrm{tr}(A) = 0$

We can easily create the adjacency matrices for the two graphs that we have defined so far in this class. The matrix for graph $\mathcal{G}_1(\mathcal{V}_1, \mathcal{E}_1)$ is

$$A_1 = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{2.4}$$

The adjacency matrix for graph $\mathcal{G}_2(\mathcal{V}_2, \mathcal{E}_2)$ is

$$A_2 = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \tag{2.5}$$

A simple calculation that we can perform is to multiply an adjacency matrix and the "ones" vector. The "ones" vector, which has n entries and each entry is 1, is written as

$$\vec{1} \stackrel{\text{def}}{=} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

If we compute $[A_1]\vec{1}$, we find that the result is the vector

$$\vec{d_1} = [A_1]\vec{1} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 1 \\ 1 \end{bmatrix} \tag{2.6}$$

Each entry of the vector $\vec{d_1}$ in Equation 2.6 is the *degree* of the vertex that has the corresponding row number. This makes sense because multiplication of a matrix and the "ones" vector takes the sum of the rows of the matrix; the sum of a row of an adjacency matrix is the number of edges that are incident with the vertex, which is the definition of the degree of the vertex.

We can likewise compute $[A_2]\vec{1}$ for graph $\mathcal{G}_2(\mathcal{V}_2, \mathcal{E}_2)$, to get

$$\vec{d_2} = [A_2]\vec{1} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ 3 \\ 1 \\ 1 \end{bmatrix} \tag{2.7}$$

A new graph can be derived from graph $\mathcal{G}_1(\mathcal{V}_1, \mathcal{E}_1)$ by adding one vertex and two edges. This graph $\mathcal{G}_3(\mathcal{V}_3, \mathcal{E}_3)$ can be defined as

$$\mathcal{G}_3 \stackrel{\text{def}}{=} \begin{cases} \mathcal{V}_3 & = & \{1\ 2\ 3\ 4\ 5\ 6\} \\ \mathcal{E}_3 & = & \{(1\ 2)\ \ (1\ 3)\ \ (2\ 3)\ \ (4\ 5)\ \ (5\ 6)\} \end{cases} \tag{2.8}$$

The graph $\mathcal{G}_3(\mathcal{V}_3, \mathcal{E}_3)$ defined in Equation 2.8 can be diagrammed as shown in Figure 2.4.

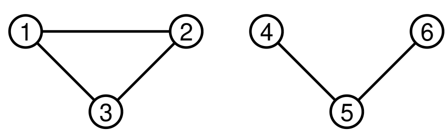

Figure 2.4: An illustration of the graph $\mathcal{G}_3$, having six vertices and five edges. The graph has two distinct sub-graphs that have no shared edge.

The graph $\mathcal{G}_1(\mathcal{V}_1, \mathcal{E}_1)$ and the $\mathcal{G}_3(\mathcal{V}_3, \mathcal{E}_3)$ have a common property: each graph has two distinct *components*. In graph theory, a component is a subgraph of a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with two special properties:

- For any two vertices of the component, there is a "walk" along incident edges between the vertices; and
- Each vertex of the component is not adjacent to any non-component vertex of $\mathcal{G}$

The adjacency matrix for graph $\mathcal{G}_3(\mathcal{V}_3, \mathcal{E}_3)$ is

$$A_3 = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{2.9}$$

The graph property of having two distinct components has a correspondence in the adjacency matrix. In general, a graph that has distinct components can be partitioned with non-zero diagonal blocks and zero off-diagonal blocks.

We can modify graph $\mathcal{G}_3(\mathcal{V}_3, \mathcal{E}_3)$, creating a new graph $\mathcal{G}_4(\mathcal{V}_4, \mathcal{E}_4)$, by adding two edges. This new graph can be defined as

$$\mathcal{G}_4 \stackrel{\text{def}}{=} \begin{cases} \mathcal{V}_4 & = & \{1\;2\;3\;4\;5\;6\} \\ \mathcal{E}_4 & = & \{(1\;2)\;\;(1\;3)\;\;(2\;3)\;\;(2\;4)\;\;(4\;5)\;\;(4\;6)\;\;(5\;6)\} \end{cases} \tag{2.10}$$

The graph $\mathcal{G}_4(\mathcal{V}_4, \mathcal{E}_4)$ defined in Equation 2.10 can be diagrammed as shown in Figure 2.5. The

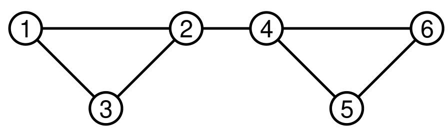

Figure 2.5: An illustration of the graph $\mathcal{G}_4$, having six vertices and seven edges. The graph has no distinct sub-graphs, but there are two "clusters" of vertices in the graph.

adjacency matrix for graph $\mathcal{G}_4(\mathcal{V}_4, \mathcal{E}_4)$ is

$$A_4 = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \tag{2.11}$$

The graph $\mathcal{G}_4(\mathcal{V}_4, \mathcal{E}_4)$ is not bipartite, because there are two "triangles" that prevent the vertices from being appropriately partitioned. Graph $\mathcal{G}_4$ also cannot be partitioned into components, because there is always a "path" from any subgraph to a vertex that is not in the subgraph.

The graph $\mathcal{G}_4(\mathcal{V}_4, \mathcal{E}_4)$ is interesting to us because it suggests that there are "clusters" of vertices; in this instance the clusters appear, visually, to be $\{1 \quad 2 \quad 3\}$ and $\{4 \quad 5 \quad 6\}$. Although they are not components, the vertices of each "cluster" have a higher connectivity within the cluster than to outside the cluster.

We will use a different matrix, which is derived from the adjacency matrix, to explore clustering of vertices in graphs.

2.3 Weighted Adjacency Matrix of a Graph

For some graphs, one edge may have a greater "importance", or *weight*, than another edge. This is represented by a positive value $w_{ij} > 0$. The weights are in a set $\mathbb{W}$ where there is a 1:1 correspondence between the weights and the edges. The definition of a weighted adjacency matrix is an extension of the above adjacency matrix.

Definition: weighted adjacency matrix $A(\mathcal{V}, \mathcal{E}, \mathbb{W})$ of a graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathbb{W})$

For any graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathbb{W})$ that has n vertices in $\mathcal{V}$ labeled $v_1, v_2, \ldots, v_n$, m edges labeled $(v_i\ v_j)$, and exactly one weight $w_{ij} > 0$ for each edge, the *weighted adjacency matrix* $A(\mathcal{G})$ is the $n \times n$ matrix for which each entry is defined as

$$a_{ij} = \begin{cases} w_{ij} & \text{if } (v_i\ v_j) \in \mathcal{E} \\ 0 & \text{if } (v_i\ v_j) \notin \mathcal{E} \end{cases} \tag{2.12}$$

We can see, from Definition 2.12, that a simple adjacency matrix is a weighted adjacency matrix where $w_{ij} = 1$.

2.4 Laplacian Matrix of a Graph

In graph theory, a commonly used matrix is derived from the adjacency matrix. The *Laplacian matrix* is usually written as L, despite the fact that it is a symmetric matrix and is not a lower-triangular matrix. We can define the Laplacian matrix of a graph by using some of the extensions that are defined in the extra notes for Class 1. First, we will define the commonly used *degree matrix* by using the $\operatorname{diag}(\cdot)$ operator of Definition 1.13.

Definition: degree matrix of a graph

For any graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ that has an adjacency matrix $A_{n\times n}(\mathcal{G})$, the *degree matrix* $D(\mathcal{G})$ is the $n \times n$ diagonal matrix that is defined as

$$D(\mathcal{G}) \stackrel{\text{def}}{=} \text{diag}([A]\vec{1}) \tag{2.13}$$

The Laplacian matrix is most often defined as the difference between the degree matrix of Equation 2.13 and the adjacency matrix of Equation 2.3.

Definition: Laplacian matrix of a graph

For any graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ that has an adjacency matrix $A_{n\times n}(\mathcal{G})$ and a degree matrix $D_{n\times n}(\mathcal{G})$, the *Laplacian matrix* $L(\mathcal{G})$ is the $n \times n$ matrix that is defined as

$$L(\mathcal{G}) \stackrel{\text{def}}{=} D(\mathcal{G}) - A(\mathcal{G}) \tag{2.14}$$

Observation: The graph that is used to define the Laplacian matrix is usually understood by context, so it is common to abbreviate Equation 2.14 as

$$L = D - A$$

Let us re-visit the graph $\mathcal{G}_1$ that is specified in Equation 2.1 and that has the adjacency matrix A_1 that is specified in Equation 2.4. The degree matrix D_1 for the graph $\mathcal{G}_1$ is

$$D_1 \stackrel{\text{def}}{=} \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.15}$$

The Laplacian matrix of the graph $\mathcal{G}_1$ is the difference between the degree matrix and the adjacency matrix, so

$$L_1 \stackrel{\text{def}}{=} D_1 - A_1 = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \tag{2.16}$$

We previously observed that the graph $\mathcal{G}_1$ has two components, which are distinct subgraphs. We can easily verify that the Laplacian matrix for the graph $\mathcal{G}_1$ can be partitioned as diagonal blocks that are non-zero matrices and off-diagonal blocks that are zero matrices, so

$$L_1 = \begin{bmatrix} L_{1S1} & \mathbf{0} \\ \mathbf{0} & L_{1S2} \end{bmatrix} \tag{2.17}$$

Each diagonal block of the Laplacian matrix L_1, described in Equation 2.17, is a Laplacian matrix of a subgraph of $\mathcal{G}_1$.

We can also compute the Laplacian matrix for the graph $\mathcal{G}_2$ that is specified in Equation 2.2 and that has the adjacency matrix A_2 that is specified in Equation 2.5. Omitting the presentation of the degree matrix, the Laplacian matrix of the graph $\mathcal{G}_2$ is

$$L_2 \stackrel{\text{def}}{=} D_2 - A_2 = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 3 & -1 & -1 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix} \tag{2.18}$$

Consider: re-ordering the vertices of the graph $\mathcal{G}_2$. From the diagram in Figure 2.3, we can see that the integer that is assigned to each vertex must be distinct but that there is no required order. An alternative graph $\mathcal{G}_5$ can be defined as

$$\mathcal{G}_5 \stackrel{\text{def}}{=} \left\{ \begin{array}{lcl} \mathcal{V}_5 & = & \{1\ 2\ 3\ 4\ 5\ 6\} \\ \mathcal{E}_5 & = & \{(1\ 4)\ \ (1\ 5)\ \ (1\ 6)\ \ (6\ 2)\ \ (6\ 3)\} \end{array} \right. \tag{2.19}$$

The graph defined in Equation 2.19 can be drawn as shown in Figure 2.6. One way to partition the vertices of graph $\mathcal{G}_5$ is as $\mathcal{V}_L = \{1\ 2\ 3\}$ and $\mathcal{V}_R = \{4\ 5\ 6\}$.

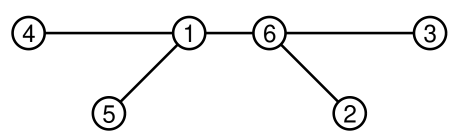

Figure 2.6: An illustration of the bipartite graph $\mathcal{G}_5$, having six vertices and five edges.

The adjacency matrix for graph $\mathcal{G}_5(\mathcal{V}_5, \mathcal{E}_5)$ is

$$A_5 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{2.20}$$

It is easy to see that the adjacency matrix A_5 of graph $\mathcal{G}_5$ can be partitioned into diagonal blocks that are zero matrices and off-diagonal blocks that are non-zero matrices, so that

$$A_5 = \begin{bmatrix} \mathbf{0} & A_{5S} \\ A_{5S}^T & \mathbf{0} \end{bmatrix} \tag{2.21}$$

2.5 Eigenvalues of a Laplacian Matrix

By definition, a Laplacian matrix is symmetric, real, and diagonally dominant. Such a matrix is guaranteed to be positive semidefinite, so its eigenvalues are non-negative. Let us explore these properties with some computational experiments.

Consider: multiplying a Laplacian matrix and a "ones" vector $\vec{1}$. As we discovered in the previous class, the resulting vector is the sum of the rows of the matrix. By definition, each diagonal entry of a Laplacian matrix is the sum of the corresponding row of the adjacency matrix, which is the degree of the graph. The remaining entries of each row of the Laplacian matrix are the negation of the corresponding row of the adjacency matrix. Consequently, the product of each row of a Laplacian matrix and the "ones" vector must be zero!

This can be verified for the above Laplacian matrices, for which we can compute

$$\begin{array}{rcl} [L_1]\vec{1} & = & \vec{0} \\ [L_2]\vec{1} & = & \vec{0} \end{array} \tag{2.22}$$

We can see, from Equation 2.22, that the "ones" vector $\vec{1}$ must be an eigenvector of the Laplacian matrix, with the corresponding eigenvalue being $\lambda_1 = 0$.

Let us try a numerical experiment on the graph $\mathcal{G}_1$. If we use `MATLAB` to find the eigenvalues of the Laplacian matrix L_1, we find that these are approximately

$$\begin{array}{rcl} \lambda_1 & = & 0 \\ \lambda_2 & = & 0 \\ \lambda_3 & = & 2 \\ \lambda_4 & = & 3 \\ \lambda_5 & = & 3 \end{array}$$

Observation: The graph $\mathcal{G}_1$ has two components and there are two zero eigenvalues of the Laplacian matrix. In this case, the dimension of the nullspace of the Laplacian matrix is equal to the number of components.

If we are curious, we might want to see whether this is also true of graph $\mathcal{G}_3$, because it also has two components. The Laplacian matrix for graph $\mathcal{G}_3$ is

$$L_2 \stackrel{\text{def}}{=} D_2 - A_2 = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \tag{2.23}$$

By now we are unsurprised by the block structure of the Laplacian matrix L_3 in Equation 2.23 and we are keen to find its eigenvalues. Again using `MATLAB`, we get

$$\begin{aligned} \lambda_1 &= 0 \\ \lambda_2 &= 0 \\ \lambda_3 &= 2 \\ \lambda_4 &= 3 \\ \lambda_5 &= 3 \\ \lambda_6 &= 3 \end{aligned}$$

The pattern persists and we can consult external reference material to determine that this is a fact about the Laplacian matrix of a graph:

> *The dimension of the nullspace of the Laplacian matrix is equal to the number of components of the graph*

What can we discover by numerical experiments on our other graphs? We can try graph $\mathcal{G}_2$, which has eigenvalues that can be numerically approximated as

$$\begin{aligned} \lambda_1 &= 0 \\ \lambda_2 &\approx 0.4384 \\ \lambda_3 &= 2 \\ \lambda_4 &= 3 \\ \lambda_5 &= 3 \\ \lambda_6 &\approx 4.5616 \end{aligned}$$

This is not especially fruitful, so let us examine one of the eigenvectors. The eigenvector for the second-smallest eigenvalue, λ_2, is approximately

$$\vec{v}_2 \approx \begin{bmatrix} -0.4647 \\ -0.2610 \\ -0.4647 \\ 0.2610 \\ 0.4647 \\ 0.4647 \end{bmatrix} \tag{2.24}$$

For the eigenvector $\vec{v}_2$ of Equation 2.24, the absolute values of the entries do not seem to be informative. However, the sign $(\pm)$ of the entries correspond to the visual clusters that we previously observed, which were $\{1\ \ 2\ \ 3\}$ and $\{4\ \ 5\ \ 6\}$. Is this true for the other clustered graph, $\mathcal{G}_4$?

Another graph that we have written is $\mathcal{G}_5$, which is a re-ordering of the graph $\mathcal{G}_2$. Doing the same computations, we find that eigenvalues and the eigenvector $\vec{v}_2$ are

$$\begin{array}{lcl} \lambda_0 & & \\ \lambda_2 & \approx & 0.4384 \\ \lambda_3 & = & 1 \\ \lambda_4 & = & 1 \\ \lambda_5 & = & 3 \\ \lambda_6 & \approx & 4.5616 \end{array} \qquad \vec{v}_2 \approx \begin{bmatrix} 0.2610 \\ -0.4647 \\ -0.4647 \\ 0.4647 \\ 0.4647 \\ -0.2610 \end{bmatrix} \tag{2.25}$$

The sign $(\pm)$ of the entries of $\vec{v}_2$ for graph $\mathcal{G}_5$ correspond to the visual clusters, which were $\{1\ \ 4\ \ 5\}$ and $\{2\ \ 3\ \ 6\}$.

2.6 The Fiedler Vector of a Graph

The eigenvector of the second-smallest eigenvalue is called the Fielder vector, named after Miroslav Fiedler who presented its use in 1989. His term for the property was “algebraic connectivity”.

Currently, the signs of the entries of the Fiedler vector are used as a *binary clustering* of a graph $\mathcal{G}$. The process for graph clustering is the same as we used in our numerical experiments above:

1. Compute the Laplacian matrix $L(\mathcal{G})$ of the graph $\mathcal{G}$
2. Compute the Fiedler vector as the eigenvector of the second-smallest eigenvalue
3. Determine the positive-negative sign of each entry

4. Assign negative entries to Set -1 and non-negative entries to Set +1

Now that we have many examples of how eigenvectors of a matrix can help us to analyze data, we will find it useful to explore the kinds of matrices to which we can apply these methods.

Extra Notes

2.7 Extra Notes: Incidence Matrix of a Graph

Another commonly used matrix of a graph is the *incidence matrix*. There are two ways to write this matrix – the ways are transposes of each other – so we must be careful when we read a source to determine which way the matrix is written. Here, we will use the `MATLAB` convention for the incidence matrix.

Suppose that a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ has n vertices and m edges. The incidence matrix is an $n \times m$ matrix that we will write as $C_{n \times m}(\mathcal{G})$, or in context simply as C. Each row of the incidence matrix C corresponds to an edge $e_i \in \mathcal{E}$ of the graph $\mathcal{G}$ and each column the incidence matrix C corresponds to a vertex $v_i \in \mathcal{V}$ of the graph $\mathcal{G}$.

<u>Definition:</u> incidence matrix $C(\mathcal{V}, \mathcal{E})$ of a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$

> For any graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ that has n vertices in $\mathcal{V}$ labeled $v_1, v_2, \ldots, v_n$ and that has m edges labeled $(v_i\ v_j)$, the *incidence matrix* $C(\mathcal{G})$ is the $n \times m$ matrix for which each entry is defined as
>
> $$c_{ij} = \begin{cases} 1 & \text{if } v_i \in e_j \\ 0 & \text{if } v_i \notin e_j \end{cases} \tag{2.26}$$

It is often easier to understand Definition 2.26 from examples. For the graph $\mathcal{G}_1$ that is specified in Equation 2.1, the incidence matrix is

$$C_1 \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.27}$$

From the first row of C_1 in Equation 2.27, we can say that vertex v_1 is incident with edge e_1, and edge e_1 also is incident with edge e_2. From the second column of C_1 in Equation 2.27, we can say that edge e_2 is incident with vertex v_1, and edge e_2 also is incident with vertex v_3.

For the bipartite graph $\mathcal{G}_2$ that is defined in Equation 2.2, the incidence matrix is

$$C_2 \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.28}$$

The columns of C_2 in Equation 2.28 correspond with the edges in graph $Gset_2$, and each non-zero entry of a column corresponds with a vertex to which the edge is incident.

In this course, we will not make practical use of the incidence matrix. There are many powerful applications of the incidence matrix, particularly for directed graphs, that the interested student is encouraged to explore.

2.8 Extra Notes: Properties of a Laplacian Matrix

A Laplacian matrix $L(\mathcal{G})$ of a graph $\mathcal{G}$ has, among others, these properties:

- L is real and symmetric
- L is diagonally dominant
- From the above, L is positive semidefinite; written as $L \succeq 0$
- Both the row sum $[L]\vec{1}$ and the column sum $\vec{1}^T[L]$ are zero
- From the above, $\vec{1}$ is an eigenvector of the eigenvalue $\lambda_1 = 0$
- The number of connected components in the graph $\mathcal{G}$ is: (a) the dimension of the nullspace of the Laplacian matrix, which is the geometric multiplicity of the zero eigenvalue, and (b) is the algebraic multiplicity of the zero eigenvalue
- The eigenvector of the second-smallest eigenvalue, called the Fiedler vector, is a binary clustering of the vertices of the graph $\mathcal{G}$
- If the graph $\mathcal{G}$ has k components, then the Laplacian matrix of $\mathcal{G}$ is similar to a Laplacian matrix that has k non-zero diagonal blocks and zero off-diagonal blocks
- If the graph $\mathcal{G}$ is bipartite, then then the Laplacian matrix of $\mathcal{G}$ is similar to a Laplacian matrix that has two zero diagonal blocks and two non-zero off-diagonal blocks; from the above, the off-diagonal blocks are transposes of each other
- The trace of the Laplacian matrix, $\mathrm{tr}(L)$, is two times the number of edges in the graph $\mathcal{G}$

The Laplacian matrix is also closely related to the incidence matrix of a graph $\mathcal{G}$. For a graph $\mathcal{G}$ that has an adjacency matrix $A(\mathcal{G})$ as specified in Definition 2.3, a degree matrix $D(\mathcal{G})$ as specified in Definition 2.13, and an incidence matrix $C(\mathcal{G})$ as specified in Definition 2.26, a matrix identity is

$$L = 2D - CC^T \tag{2.29}$$

From the definition of the Laplacian matrix as $L = D - A$, the degree matrix can be deduced from the incidence as

$$D = \mathrm{diag}(\mathrm{diag}(CC^T)) \tag{2.30}$$

Combining Equation 2.29 and Equation 2.30, we can compute the adjacency matrix from the incidence matrix as

$$A = CC^T - D \tag{2.31}$$

End of Extra Notes

