Linear Algebra - Project

Meet Gera   Yashas S. B.   Rohan

Course Instructor: Dr. Chittraranjan Hens
IIIT Hyderabad

Abstract

Eigenvalues offer a window into the intrinsic properties of graphs, providing a means to comprehend their connectivity, partitioning, symmetry, and dynamics. The exploration of the spectral realm has paved the way for advancements in diverse fields, ranging from computer science and mathematics to social sciences and biology. By leveraging the power of eigenvalues, researchers can unravel the hidden intricacies of graphs and harness their insights to tackle real-world challenges.

1 Introduction

Eigenvalues serve as a key tool in understanding graph connectivity and community structure. By examining the eigenvalues of a graph's adjacency matrix, which captures essential structural information about the graph's connectedness, we can gain valuable insights. The number of zero eigenvalues, known as the algebraic connectivity, provides a measure of the robustness and resilience of the graph. Moreover, the eigenvectors corresponding to the smallest nonzero eigenvalue aid in graph partitioning. Beyond that, eigenvalues unlock further insights into the behavior of graphs. The spectrum of eigenvalues reveals information about the graph's symmetry, regularity, and expansion properties. Large eigenvalues indicate the presence of well-connected clusters or highly influential nodes, while small eigenvalues suggest sparseness or potential bottlenecks. Furthermore, the distribution of eigenvalues allows for the identification of graph families and the comparison of graph similarity. Eigenvalues offer a window into the intrinsic properties of graphs, providing a means to comprehend their connectivity, partitioning, symmetry, and dynamics. The exploration of the spectral realm has paved the way for advancements in diverse fields, ranging from computer science and mathematics to social sciences and biology. By leveraging the power of eigenvalues, researchers can unravel the hidden intricacies of graphs and harness their insights to tackle real-world challenges.

2 Bipartite Graphs

2.1 Mathematical Representation

A graph is a collection of nodes and edges, denoted as $G = (V, E)$. In the case of bipartite graphs, there are two sets of nodes such that there are no edges between any two nodes of the same set. Mathematically, we can represent this as $\exists V_1, V_2$ such that $V_1 \cup V_2 = V$ and $V_1 \cap V_2 = \emptyset$. The representation of bipartite graphs is given by $G = (V_1, V_2, E)$.
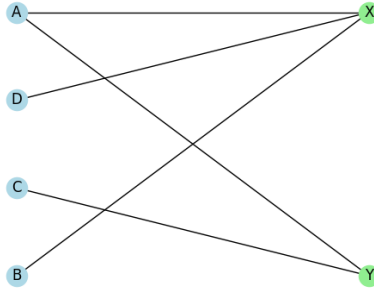

Figure 1: Example of a Bipartite Graph

2.2 Adjacency Matrix

The adjacency matrix $\mathbf{A}$ is a square matrix of size $n \times n$, where n is the number of nodes in the graph. For a bipartite graph, the adjacency matrix is partitioned into four blocks, capturing the connections between nodes in V_1, nodes in V_2, and edges between V_1 and V_2. The adjacency matrix of a bipartite graph takes the form:

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{bmatrix}$$

where $\mathbf{0}$ represents the zero matrix and $\mathbf{B}$ captures the edges connecting V_1 and V_2. The adjacency matrix provides a concise representation of the graph's structure and facilitates various graph analysis techniques.

2.3 Eigenvalues of Bipartite Graphs

The eigenvalues of a bipartite graph's adjacency matrix exhibit an interesting property. Let $\mathbf{A}$ be the adjacency matrix of a bipartite graph. By definition, a bipartite graph has its vertices divided into two disjoint sets such that all edges

connect vertices from different sets. The eigenvalues of **A** appear in pairs of $+a$ and $-a$, where a is a real number. This property holds true due to the specific structure of bipartite graphs and has important implications for understanding their spectral properties.
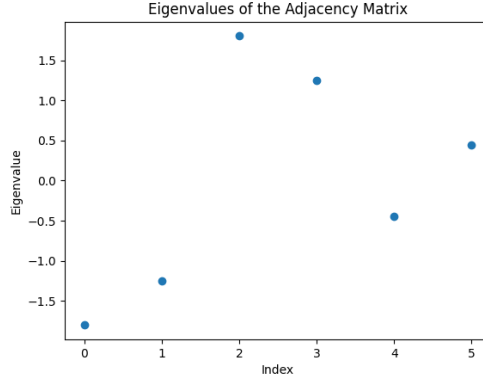

Figure 2: Example of a eigenvalues in Complex Plane

3 Relation with Eigenvalue and Proof

Theorem 1 (Eigenvalues of Bipartite Graphs). *Let* **A** *be the adjacency matrix of a bipartite graph. By definition, a bipartite graph has its vertices divided into two disjoint sets such that all edges connect vertices from different sets. The eigenvalues of* **A** *appear in pairs of* $+a$ *and* $-a$, *where a is a real number.*

Proof. We will prove the statement by considering the properties of the adjacency matrix of a bipartite graph.

Step 1: Matrix A^2

Consider the matrix $\mathbf{A}^2$. The (i,j)th entry of $\mathbf{A}^2$ represents the number of paths of length 2 between vertex i and vertex j. In a bipartite graph, there are no direct edges between vertices within the same set, so all entries on the diagonal of $\mathbf{A}^2$ are zero.

Step 2: Equivalence of A^2v and Av^2

For any vector **v**, we have $(\mathbf{A}^2)\mathbf{v} = \mathbf{A}(\mathbf{v}^2)$. This equation can be proven using matrix multiplication.

Step 3: Eigenvalue λ and eigenvector v

Let λ be an eigenvalue of **A** with eigenvector **v**. By definition, $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

Step 4: From $Av = \lambda v$ to $Av^2 = \lambda(Av)$

Multiply both sides of the equation by $\mathbf{A}$: $\mathbf{A}(\mathbf{Av}) = \mathbf{A}(\lambda\mathbf{v})$. This yields $\mathbf{A}^2\mathbf{v} = \lambda(\mathbf{Av})$.

Step 5: Substituting Step 2

By substituting Step 2, we have $\mathbf{Av}^2 = \lambda(\mathbf{Av})$. This implies that λ is an eigenvalue of $\mathbf{A}^2$ with eigenvector $\mathbf{v}^2$.

Step 6: Non-negativity of A^2

The eigenvalues of $\mathbf{A}^2$ are non-negative. This is because $\mathbf{A}^2$ is a non-negative matrix representing the number of paths of length 2 in the bipartite graph.

Step 7: Eigenvalue λ is non-negative

Since λ is an eigenvalue of $\mathbf{A}^2$, we have $\lambda \geq 0$.

Step 8: Assumption: $\lambda \neq 0$

Let's assume $\lambda \neq 0$. If $\lambda \neq 0$, then the eigenvector $\mathbf{v}^2$ is nonzero.

Step 9: From $Av = \lambda v$ to $\lambda = \frac{v^T A^T v}{v^T v}$

By the definition of the eigenvalue, we have $\mathbf{Av} = \lambda\mathbf{v}$. Multiplying both sides of the equation by $\mathbf{v}^T$, we get $(\mathbf{v}^T\mathbf{Av}) = \lambda(\mathbf{v}^T\mathbf{v})$.

Step 10: Symmetric property

Since $\mathbf{A}^T = \mathbf{A}$ for a bipartite graph, we have $\mathbf{A}^T\mathbf{v} = \mathbf{Av}$. This implies that $(\mathbf{v}^T\mathbf{A}^T\mathbf{v}) = (\mathbf{v}^T\mathbf{Av})$.

Step 11: Combining Step 9 and Step 10

Combining steps 9 and 10, we get $\lambda = \frac{v^T A^T v}{v^T v}$.

Step 12: Eigenvalue λ is real

Since $\lambda = \frac{v^T A^T v}{v^T v}$, it follows that λ is a real number.

Step 13: Conclusion: Eigenvalues appear in pairs

Therefore, we conclude that the eigenvalues of $\mathbf{A}$ appear in pairs of $+a$ and $-a$, where a is a real number.
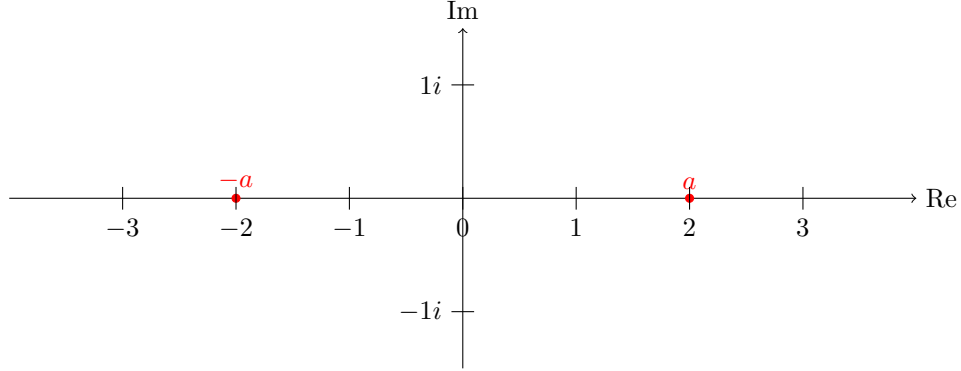
$\square$

Figure 3: Eigenvalues of a Bipartite Graph

4 Conclusion for Bipartite Graphs

In conclusion, the study of bipartite graphs and their eigenvalues holds great promise for understanding graph connectivity, community structure, and various real-world applications. The exploration of eigenvalues provides valuable insights into the robustness, resilience, and behavior of bipartite graphs.

The application of bipartite graphs extends to various domains, including matching and assignment problems, biological networks, and social network analysis. By leveraging the power of eigenvalues, researchers can unravel the hidden intricacies of bipartite graphs and harness their insights to tackle real-world challenges.

Looking to the future, the utilization of bipartite graphs and their eigenvalues is expected to continue advancing. As data becomes increasingly complex and interconnected, bipartite graphs offer a powerful framework for modeling and analyzing large-scale networks. From personalized recommendations to understanding information diffusion, bipartite graphs have the potential to revolutionize various fields and drive innovative solutions.

In conclusion, the study of bipartite graphs and their eigenvalues opens up new avenues for research and practical applications. By delving into the spectral realm of eigenvalues, we can unlock the hidden patterns and structures within bipartite graphs, paving the way for further advancements in graph theory and its applications.

So, let us embrace the power of eigenvalues and embark on a journey to unravel the mysteries of bipartite graphs. With each step, we deepen our understanding of their connectivity, partitioning, symmetry, and dynamics, ultimately enabling us to tackle real-world challenges with confidence and precision.

5 Graph Connectivity and Eigenvalues

5.1 Definition

In graph theory, connectivity refers to a measure of how connected or interconnected a graph is. It quantifies the ability to travel or navigate between vertices (nodes) in a graph. The connectivity of a graph is important in understanding its structural properties and determining how easily information or resources can flow through the network represented by the graph.

Strong Connectivity: A directed graph is strongly connected if there is a directed path between any two vertices in the graph. In other words, it is possible to reach any vertex from any other vertex in the graph.

Weak Connectivity: A directed graph is weakly connected if replacing all the directed edges with undirected edges produces a connected undirected graph. In other words, it is possible to reach any vertex from any other vertex when considering the graph as undirected.

5.2 Graph Connectivity

The connectivity of a graph can be determined using the Laplacian matrix. By calculating the eigenvalues of the Laplacian matrix and examining the second smallest eigenvalue, you can determine whether the graph is disconnected or has multiple connected components.

Theorem: The second smallest eigenvalue λ_2 tells you about the connectivity of the graph. If the graph has two disconnected components, $\lambda_2 = 0$. And if λ_2 is small, this suggests the graph is nearly disconnected, with two components that are not very connected to each other.

Proof:

1. Let $v_1, v_2, \ldots, v_k$ be k eigenvectors corresponding to the eigenvalue 0, where k is the multiplicity of 0 as an eigenvalue.

2. Consider the vector $x = v_1 + v_2 + \ldots + v_k$. Since L is a linear operator, we have $Lx = Lv_1 + Lv_2 + \ldots + Lv_k = 0 + 0 + \ldots + 0 = 0$.

3. Now, suppose the graph G has m connected components, where $m \leq k$.

4. Let $C_1, C_2, \ldots, C_m$ be the m connected components of G. Each connected component C_i has its corresponding eigenvector v_i, where $1 \leq i \leq m$. Note that each v_i has entries corresponding to vertices only in the connected component C_i, and all other entries are 0.

5. Therefore, when we sum up all the eigenvectors $v_1, v_2, \ldots, v_k$, we get $x = v_1 + v_2 + \ldots + v_k$, which has entries corresponding to vertices in all connected components C_i, and all other entries are 0.

6. Since $Lx = 0$, we have $Lx = L(v_1 + v_2 + \ldots + v_k) = Lv_1 + Lv_2 + \ldots + Lv_k = 0$. But since each $Lv_1, Lv_2, \ldots, Lv_k$ corresponds to a connected

6

component C_i and all other entries are 0, we can rewrite the equation as $Lx = L(v_1 + v_2 + \ldots + v_k) = L(v_1) + L(v_2) + \ldots + L(v_k) = 0$.

7. This implies that for each connected component C_i, we have $L(v_i) = 0$. In other words, each connected component has an eigenvector corresponding to the eigenvalue 0.

8. Since m is the number of connected components and $m \leq k$, we conclude that the multiplicity of 0 as an eigenvalue of L is equal to the number of connected components in the graph.

Hence, if the second smallest eigenvalue of the Laplacian matrix is 0 (i.e., $\lambda_2 = 0$), it implies that the graph is disconnected.

5.3 Example

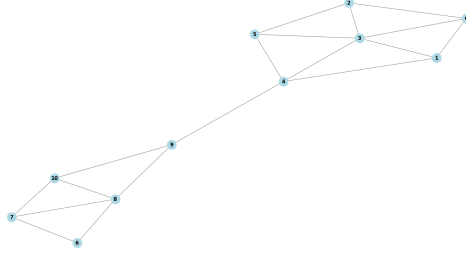### 5.3.1 Let us see this graph:

Figure 4: A Connected Graph

The eigenvalues for the Laplacian matrix from the above graph are: 0, 0.214, 1.81, 2.382, 2.655, 3.368, 4.441, 4.618, 4.94, 5.461, 6.11.

5.3.2 But here in this graph:

The eigenvalues for the Laplacian matrix from the above graph are: 0, 0, 1.586, 2.382, 2.382, 3, 4.414, 4.618, 4.618, 5, 6.

From the above two examples, we can tell that a graph with the second lowest eigenvalue as 0 can be a disconnected graph.

6 Applications of Laplacian Matrix Eigenvalues

6.1 Graph Partitioning

The eigenvalues of the Laplacian matrix can be used in graph partitioning algorithms. These algorithms aim to divide a graph into multiple components or

Figure 5: A Disconnected Graph

clusters based on the connectivity of the graph. The second smallest eigenvalue, being related to graph connectivity, can be utilized as a criterion for determining the quality of graph partitioning methods.

6.2   Spectral Graph Theory

The Laplacian matrix and its eigenvalues play a significant role in spectral graph theory. Spectral graph theory studies the properties of graphs using linear algebraic techniques. The eigenvalues of the Laplacian matrix provide insights into the structure, connectivity, and properties of graphs. The theorem you mentioned is a fundamental result in spectral graph theory.

6.3   Network Analysis

The theorem can be applied in the analysis of various networks, such as social networks, transportation networks, and communication networks. By examining the eigenvalues of the Laplacian matrix, particularly the second smallest eigenvalue, researchers can gain insights into the connectivity and structural properties of these networks.

7   Laplacian Eigen Values

Laplacian eigenvalues, also known as Laplacian spectrum or Laplacian eigenfrequencies, are a set of mathematical properties associated with the Laplacian operator.

In the context of graph theory, the Laplacian eigenvalues are defined for a graph, which is a collection of vertices (nodes) connected by edges. The Laplacian matrix of a graph is a square matrix that encodes the graph's connectivity information. Each entry of the Laplacian matrix represents the connection between two vertices, with positive values indicating the presence of an edge and negative values representing the weights or conductance of the edges.

The eigenvalues of a Laplacian matrix provide valuable insights into the structure and properties of a graph. Here, I will elaborate on the characteristics associated with each eigenvalue of a Laplacian matrix.

7.1 Zero Eigenvalue

One eigenvalue must equal to 0 for a graph:

Proof:

When we subtract the adjacency matrix from the degree matrix to form the Laplacian matrix, we can observe that the row sums of the resulting matrix are all zero. This property implies that the vector of all ones (a constant vector) is an eigenvector associated with the eigenvalue zero.

This zero eigenvalue and its corresponding eigenvector have important implications. It reflects that there is always a trivial solution to the Laplacian matrix equation, where the sum of values at each vertex is zero. This property corresponds to the graph being connected, as there is no isolated component or disconnected subgraph.

7.2 Classification of Laplacian Eigenvalues

Zero Eigenvalue: As mentioned earlier, a Laplacian matrix always has a zero eigenvalue. It corresponds to a constant eigenvector, where all elements of the vector are the same. This eigenvalue signifies the connectivity of the graph and indicates that the graph is connected without any isolated components.

Positive Eigenvalues: The positive eigenvalues of the Laplacian matrix provide information about the connectivity and expansion properties of the graph. The magnitude of these eigenvalues indicates how fast information can spread or diffuse across the graph. Larger positive eigenvalues imply faster diffusion and better connectivity within the graph.

The sum of all positive eigenvalues corresponds to the sum of all degrees of the graph. This property is known as the "degree-sum theorem" and provides a useful relationship between the Laplacian eigenvalues and the graph's vertex degrees.

The first positive eigenvalue, often referred to as the "Fiedler eigenvalue," plays a crucial role in spectral graph theory. It divides the graph into two or more clusters based on the associated eigenvector. Spectral clustering algorithms utilize this eigenvalue to perform graph partitioning and clustering tasks.

Negative Eigenvalues: In some cases, the Laplacian matrix may have negative eigenvalues. The presence of negative eigenvalues indicates that the graph has a non-trivial bipartite structure or a sign of imbalance in the graph's connectivity.

The magnitude of negative eigenvalues provides information about the extent of the imbalance or bipartite structure within the graph. Larger negative eigenvalues suggest a stronger bipartite structure or imbalance.

It's important to note that the Laplacian matrix of a graph is always a symmetric and positive semi-definite matrix. Therefore, all the eigenvalues are real numbers, and they can be ordered from smallest to largest.

Spectral Gap: The difference between consecutive eigenvalues is referred to as the "spectral gap." The spectral gap provides insights into the connectivity and conductance properties of the graph. A larger spectral gap indicates a graph with stronger connectivity and better conductance between its components.

These are some additional aspects related to Laplacian eigenvalues. They offer a rich set of information about graph properties, connectivity, conductance, community structure, and even applications in graph drawing and random walks. By examining the Laplacian eigenvalues, we can gain insights into the underlying structure and behavior of graphs.

8 Uses:

Graph Partitioning and Clustering: The Laplacian eigenvalues and eigenvectors are extensively employed in spectral graph partitioning and clustering algorithms. By analyzing the eigenvalues, particularly the second smallest eigenvalue (algebraic connectivity), and the corresponding eigenvector, the graph can be divided into clusters or partitions, revealing underlying community structures.

Graph Drawing and Visualization: Laplacian eigenvalues and eigenvectors are utilized in graph drawing and visualization techniques. They help determine the optimal layout of vertices in a graph, with connected vertices positioned close to each other and disconnected vertices positioned far apart. This aids in visualizing the graph's structure and relationships.

Graph-based Image Segmentation: Laplacian eigenvalues are utilized in image processing tasks, particularly image segmentation. The Laplacian matrix is constructed based on the image's graph representation, and the eigenvalues and eigenvectors of the Laplacian matrix assist in identifying distinct regions or objects within the image.

Community Detection and Social Network Analysis: Laplacian eigenvalues are employed in community detection algorithms in social network analysis. By analyzing the eigenvalues and eigenvectors of the Laplacian matrix, communities or groups of densely connected nodes can be identified in the network, revealing patterns and structures in the social system.

Spectral Graph Theory: Laplacian eigenvalues play a fundamental role in spectral graph theory, a branch of mathematics that investigates graph properties using the eigenvalues and eigenvectors of matrices associated with the graph. The Laplacian eigenvalues provide insights into graph connectivity, expansion, and symmetries.

Random Walks on Graphs: Laplacian eigenvalues are connected to random walks on graphs. The second smallest eigenvalue, known as the algebraic connectivity, and its associated eigenvector describe the behavior and convergence properties of random walks on the graph. This information is useful in studying

diffusion processes and information propagation on networks.

Machine Learning and Data Analysis: Laplacian eigenvalues have found applications in machine learning and data analysis tasks. They are used in dimensionality reduction techniques like Laplacian eigenmaps and spectral embedding, where the eigenvalues and eigenvectors of the Laplacian matrix are employed to transform high-dimensional data into a lower-dimensional representation while preserving the underlying structure.

Network Resilience and Robustness: The Laplacian eigenvalues provide insights into the resilience and robustness of networks. Large eigenvalues indicate strong connectivity and faster information diffusion, while the spectral gap between eigenvalues reveals the network's vulnerability to perturbations and structural changes.

9 Clustering in Graphs using Eigenvalues

9.1 What is Clustering?

Clustering is the task of grouping objects in a way such that objects belonging to a certain cluster are more similar to each other than to those objects not in the cluster. It is one of the more important techniques in several data analysis fields like pattern recognition, image compression, bioinformatics, machine learning and several others. Clustering in graphs using eigenvalues and eigenvectors of the adjacency matrix is a widely used technique known as spectral clustering.

9.2 Spectral Clustering

Spectral clustering is a technique with roots in graph theory, where the approach is used to identify communities of nodes in a graph based on the edges connecting them. Spectral clustering leverages the spectral properties of the graph Laplacian, which is derived from the adjacency matrix, to partition the graph nodes into cohesive clusters. In statistics, spectral clustering techniques make use of the eigenvalues of the similarity matrix of the data to perform dimensionality reduction before clustering in fewer dimensions. The similarity matrix which is provided as the input consists of quantitative assessment of the relative similarity of different pairs of points in the dataset.

9.3 Basic Terminologies used in Spectral Clustering

We will try to explain the basic terminologies in Spectral Clustering using a graph.

9.3.1 Adjacency and Affinity Matrix (A)

The graph (or set of data points) can be represented as an Adjacency Matrix, where the row and column indices represent the nodes, and the entries represent the absence or presence of an edge between the nodes (i.e. if the entry in row
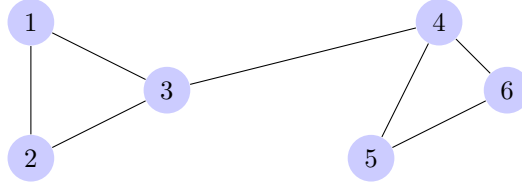
Figure 6: A simple graph with 6 nodes

0 and column 1 is 1, it would indicate that node 0 is connected to node 1). An Affinity Matrix is like an Adjacency Matrix, except the value for a pair of points expresses how similar those points are to each other. If pairs of points are very dissimilar then the affinity should be 0. If the points are identical, then the affinity might be 1. In this way, the affinity acts like the weights for the edges on our graph.

Below is the Adjacency/Affinity matrix of the graph above:

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

9.3.2 Degree Matrix (D)

A Degree Matrix is a diagonal matrix, where the degree of a node (i.e. values) of the diagonal is given by the number of edges connected to it. We can also obtain the degree of the nodes by taking the sum of each row in the adjacency matrix.

Below is the Degree matrix of the adjacency matrix we defined above:

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

9.3.3 Laplacian Matrix (L)

This is another representation of the graph/data points, which attributes to the beautiful properties leveraged by Spectral Clustering. One such representation is obtained by subtracting the Adjacency Matrix from the Degree Matrix (i.e.

$L = D{-}A$).

Below is the Laplacian matrix of the graph we defined above:

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

1. **Spectral Gap**: The first non-zero eigenvalue of L is called the Spectral Gap. The Spectral Gap gives us some notion of the density of the graph.

2. **Fiedler Value**: The second eigenvalue of L is called the Fiedler Value, and the corresponding vector is the Fiedler vector. Each value in the Fiedler vector gives us information as to which side of the decision boundary a particular node belongs to.

3. Using L, we find the first large gap between eigenvalues which generally indicates that the number of eigenvalues before this gap is equal to the number of clusters.

9.4 Algorithm for spectral clustering in graphs

The three major steps involved in spectral clustering are: constructing a similarity graph, projecting data onto a lower-dimensional space, and clustering the data.

Step 1: Form a distance matrix.

Step 2: Transform the distance matrix into an affinity matrix A.

Step 3: Compute the degree matrix D and the Laplacian matrix $L = D{-}A$.

Step 4: Find eigenvalues and eigenvectors of L.

Step 5: With the eigenvectors of k largest eigenvalues computed from the previous step form a matrix.

Step 6: Normalize the vectors.

Step 7: Cluster the data points in k-dimensional space.

9.5 Why use spectral clustering over other forms of clustering?

1. **Flexibility in Handling Complex Data:** Spectral clustering is capable of identifying clusters with complex shapes and structures. It is not limited to spherical or convex clusters like some other algorithms such as k-means.

Spectral clustering can handle clusters that are non-linear, disconnected, or overlapping, making it more suitable for a wide range of datasets.

2. **Connectivity-Based Clustering:** Spectral clustering takes into account the connectivity or pairwise relationships between data points. Points that are close in proximity but not directly connected can still be assigned to the same cluster. This allows spectral clustering to capture meaningful relationships between data points and identify clusters based on connectivity patterns.

3. **Scalability with Sparse Graphs:** Spectral clustering can be particularly efficient when working with sparse graphs. The use of sparse matrix representations and efficient algorithms for eigenvalue computation can help reduce computational costs, making it feasible for large-scale datasets.
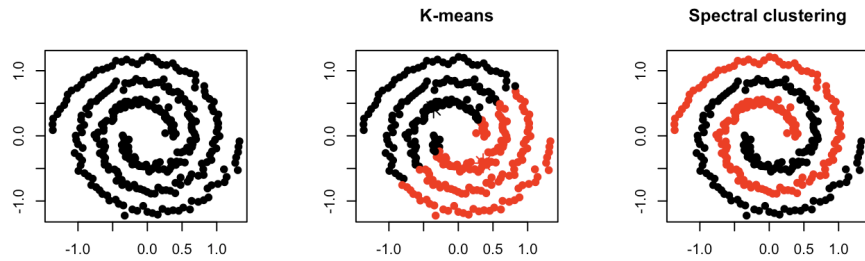
.

Figure 7: Clustering with different clustering algorithms

9.6 Some cons of Spectral Clustering

1. **Computational Complexity:** Spectral clustering involves the computation of eigenvalues and eigenvectors of the similarity matrix or Laplacian matrix, which can be computationally expensive for large datasets. The time and memory requirements of spectral clustering can be higher compared to simpler clustering algorithms like k-means.

2. **Parameter Selection:** Spectral clustering requires the selection of various parameters, such as the number of clusters (k) and the type of affinity matrix or similarity measure to use. Choosing the optimal values for these parameters can be challenging and may require domain knowledge or trial-and-error experimentation.

3. **Lack of Robustness to Noise and Outliers:** Spectral clustering is not inherently robust to noise or outliers in the data. Outliers can influence the computation of the affinity matrix and affect the clustering results. Preprocessing steps or outlier detection techniques may be needed to handle noisy data effectively.

9.7 Uses of Spectral Clustering

1. **Recommendation Systems:** Spectral clustering can be employed in recommendation systems to group users or items based on their preferences or similarities. By clustering users or items with similar preferences, personalized recommendations can be made, enhancing the accuracy and relevance of the recommendations.

2. **Social network analysis:** Spectral clustering is often used in analyzing social networks to identify communities or groups of nodes with dense connections. By representing the network as a graph and applying spectral clustering, it becomes possible to uncover hidden structures, detect communities, and understand the relationships within the network.

3. **Image segmentation:** Spectral clustering can be applied to segment images by grouping pixels or regions with similar characteristics, such as color, texture, or intensity. It has proven to be effective in various computer vision tasks, including object recognition, image compression, and image retrieval.

10 Eigenvector Centrality

10.1 What is Eigenvector Centrality in a Graph?

In graph theory, eigenvector centrality is a measure of the influence of a node in a network. If a graph node is connected to a node with a high eigenvector centrality score, the graph node in question gets a higher eigenvector centrality score as well. A high eigenvector score means that a node is connected to many nodes who themselves have high scores.

10.2 How to compute Eigenvector Centrality?

Mathematically, eigenvector centrality is computed using the eigenvector corresponding to the largest eigenvalue of the adjacency matrix of the network. The adjacency matrix represents the connections or relationships between nodes in the network. The eigenvector centrality of a node is given by the corresponding entry in the eigenvector.

10.3 Using and Adjacency matrix to compute eigenvector centrality

For a graph, $G = (V, E)$, with V vertices let $A = (a_{i,j})$ be the adjacency matrix. Here, $a_{i,j} = 1$ if vertex i is connected to vertex j. Otherwise, $a_{i,j} = 0$.

For this adjacency matrix, the relative centrality score x_i of a vertex i can be expressed as:

$$x_i = \frac{1}{\lambda} \sum_{j \in V} a_{i,j} \cdot x_j$$

The above equation can also be represented as:

$$Ax = \lambda x$$

Here λ represents the eigenvalue of the adjacency matrix.

As, we are measuring the influence of a node in a graph, where the influence cannot be lesser than zero, the eigenvector of the graph (which contains the centrality score of each corresponding node) cannot have entries which are lesser than zero. So, from the Perron Frobenius' theorem, we can say that only the greatest eigenvalue results in the desired centrality scores.

10.4   Considerations to make while using Eigenvector Centrality

1. Centrality scores for nodes with no incoming relationships will converge to 0.

2. Due to missing degree normalization, high-degree nodes have a very strong influence on their neighbors' score.

10.5   Example usage to find Eigenvector Centrality of a graph

We can achieve the task of finding the centrality of a graph using Python. A simple Python code fulfils all our need and executes as required to find the centrality score of each node.
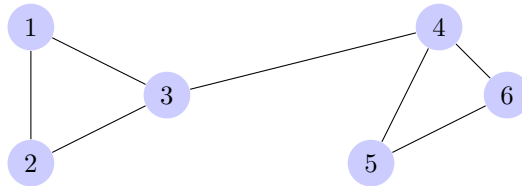
Here is a graph:

Figure 8: A graph

The centrality score of the graph nodes are as follows:
Node 1: 0.35355343614372065
Node 2: 0.35355343614372065
Node 3: 0.49999993558193184

Node 4: 0.49999993558193184
Node 5: 0.35355343614372065
Node 6: 0.35355343614372065

As we can clearly see, nodes with more edges have higher centrality scores.

10.6   Uses of Eigenvector centrality

Eigenvector centrality has various applications and has been used in different fields. Here are some notable uses:

1. In a study using data from the Philippines, researchers showed how political candidates' families had disproportionately high eigenvector centrality in local intermarriage networks.

2. Eigenvector centrality has been extensively applied to study economic outcomes, including cooperation in social networks. In economic public goods problems, a person's eigenvector centrality can be interpreted as how much that person's preferences influence an efficient social outcome.

3. Eigenvector centrality is commonly employed in social network analysis to identify influential individuals within a social network. Nodes with high eigenvector centrality are considered to have a greater impact and influence on the overall network dynamics. It helps in identifying key individuals who can spread information or influence behavior effectively.

4. In academic research, eigenvector centrality is used in citation networks to identify influential papers or authors. Nodes with high eigenvector centrality represent papers that are frequently cited by other influential papers, indicating their impact within the research community.