Journal of Discrete Mathematics and Its Applications 10 (4) (2025) 359–374

Journal of Discrete Mathematics and Its Applications

Available Online at: http://jdma.sru.ac.ir

Shahid Rajaee Teacher Training University

Research Paper

Detection of communities by modularity matrix

Seyedeh Akram Hosseini[1], **Gholam Hossein Fath-Tabar**[1,*], **Hossein Shabani**[2]

[1]Department of Pure Mathematics, Faculty of Mathematical Sciences, University of Kashan, Kashan, I. R. Iran

[2]Faculty of Computer, Network and Communication , Imam Hossein Comprehensive University, Tehran, I. R. Iran

Academic Editor: Majid Arezoomand

Abstract. One of the most discussed topics in social networks is community detection. As these networks become more complex, spectral graph properties and graph-related structures are increasingly used for community detection. In this paper, we examine the properties of the modularity matrix, including the eigenvalues of the modularity matrix, modularity energy, and the Estrada modularity index. Additionally, we investigate the bounds for the energy and Estrada indices. Furthermore, considering the significant issue of estimating the number of communities in some community detection algorithms in networks, we focus on the modularity eigenvalues.

Keywords. Estrada index, eigenvalues, communities estimate, graph energy, modularity matrix.
Mathematics Subject Classification (2020): 05C50, 05C85, 05C82.

1 Introduction

Community detection, the process of identifying cohesive subgroups within networks, has become a fundamental approach to analyzing complex systems. These subgroups, often referred to as communities, represent clusters of nodes that share stronger connections among themselves than with nodes outside the group. By uncovering these communities,

*Corresponding author (Email address: fathtabar@kashanu.ac.ir).
Received 26 January 2025; Revised 13 March 2025; Accepted 21 April 2025
First Publish Date: 01 December 2025

researchers can gain valuable insights into the structure and dynamics of networks. For instance, in social networks, communities might represent groups of friends, colleagues, or people with shared interests. In biological networks, they could correspond to protein complexes or functional modules [10, 25]. Understanding the underlying principles of community detection is essential for researchers across various disciplines, from sociology and biology to computer science and engineering. As network data continues to grow in volume and complexity, the development of advanced community detection techniques remains a vibrant area of research [31, 32]. A multitude of algorithms and approaches have been developed to tackle the challenge of community detection, each with its own strengths and limitations. These methods range from simple clustering techniques to sophisticated optimization-based approaches. Understanding the underlying principles of these algorithms is crucial for effectively applying them to different types of networks and extracting meaningful information [2, 6]. As networks grow increasingly intricate, spectral graph theory has emerged as a powerful tool for unraveling their underlying structure [23]. The graph is a mathematical structure consisting of vertices (nodes) and edges that represent relationships or connections between these vertices. It serves as an effective model for networks by illustrating relationships between nodes at an abstract level. In spectral graph theory, graphs are commonly represented using matrices, such as the adjacency matrix or the Laplacian matrix, to facilitate analysis and computation. These matrices encapsulate the structure of the graph, with entries indicating the presence or absence of edges between vertices. By analyzing the eigenvalues and eigenvectors of these matrices, we can extract valuable information about the graph's properties, such as connectivity, clustering, and spectral partitioning [9, 11, 33]. In this paper, we focus on leveraging the spectral properties of the modularity matrix to advance community detection.

Let $G = (V, E)$ be a simple graph with vertex set $V(G) = \{v_1, \cdots, v_n\}$ and edge set $E(G)$. The adjacency matrix $A(G)$ of the graph G is a symmetric matrix of order n with entries a_{ij}, such that $a_{ij} = 1$ if $ij \in E(G)$ and $a_{ij} = 0$ otherwise. The Laplacian matrix is defined by $L = D - A$, where D is the degree matrix, a diagonal matrix where each entry represents the degree of a vertex, and A is the adjacency matrix [33]. The energy of the graph G is defined as $E(G) = \sum_{i=1}^{n} |\lambda_i|$, where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of the adjacency matrix [20]. The Estrada index is defined as $EE(G) = \sum_{i=1}^{n} e^{\lambda_i}$, where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of the adjacency matrix [11]. References [1, 14, 26] are provided for further study.
Modularity, introduced by Newman in 2004 [27], measures how effectively a network can be partitioned into distinct groups or communities. It compares the actual number of connections within these groups to the expected number in a random network with the same number of nodes and edges. A higher modularity value indicates a stronger community structure, making it a key metric for analyzing the organization and relationships within complex networks.

As a fundamental metric, modularity quantifies the strength of a network's division into communities, providing valuable insights into the underlying structure and connectivity patterns. Given a graph $G = (V, E)$, where V is the set of vertices and E is the set of edges, and a division of the vertices into communities $C_1, C_2, \cdots, C_k$, the modularity Q is defined as follows:

$$Q = \frac{1}{2m} \sum_{i,j} (A_{i,j} - \frac{k_i k_j}{2m}) \delta(C_i, C_j),$$

where

- A_{ij} is the ij-th entry of the adjacency matrix of the network (1 if there is an edge between nodes i and j, otherwise 0).
- k_i and k_j are the degrees of nodes i and j, respectively.
- m is the total number of edges in the network.
- C_i and C_j are the community assignments of nodes i and j, respectively.
- $\delta(C_i, C_j)$ is the Kronecker delta function, which equals 1 if $C_i = C_j$ (i.e., if nodes i and j belong to the same community) and 0 otherwise [29].

The modularity matrix serves as a cornerstone in community detection, offering a structured approach to identifying cohesive groups within a network. Its spectral properties, particularly its eigenvalues and eigenvectors, provide deep insights into the underlying community structure [29]. The modularity matrix $M = [M_{ij}]$ is defined as follows:

$$M_{ij} = A_{ij} - \frac{k_i k_j}{2m}. \tag{1}$$

The modularity matrix essentially captures the difference between the observed number of edges between two nodes and the expected number of edges based on a random graph model. By examining the spectral characteristics of the modularity matrix, including its eigenvalues and eigenvectors, we aim to uncover novel insights into community structure [28]. Furthermore, we investigate the relationship between modularity and other graph-theoretic concepts, such as graph energy and the Estrada index [11].

In graph theory, the Estrada index and graph energy are valuable tools with applications in diverse fields. The Estrada index is particularly useful in analyzing chain molecular structures, protein degrees, and complex networks, while graph energy is closely connected to molecular energy in chemistry [15, 16]. A critical challenge in network analysis is accurately estimating the number of communities. This is essential for many community detection algorithms, which often require a predetermined number of clusters. Consequently, determining the optimal number of communities remains an open research question and is frequently addressed through heuristic methods [12].

This paper provides a comprehensive analysis of the modularity matrix and its spectral properties, offering new insights into community detection and paving the way for advanced network analysis techniques.

2 Modularity matrix

The modularity matrix is a fundamental tool in network analysis, offering a quantitative measure of the quality of a community division, facilitating the comparison of partitions, and providing insights into network structure. By capturing the difference between the observed number of edges within and between communities, the modularity matrix serves as a valuable tool for identifying cohesive subgroups within a network. The eigenvalues and eigenvectors of the modularity matrix provide deeper insights into the underlying community structure. These spectral properties enable the development of effective algorithms for community detection and network analysis, improving the understanding of how nodes are organized into meaningful clusters. As mentioned in equation (2), the modularity matrix of a graph is defined as $M = A - P$ where:

- A is the adjacency matrix of the graph.
- P is the edge average matrix, with the entries $p_{ij} = \frac{k_i k_j}{2m}$, where k_i and k_j are the degrees of nodes i and j, respectively, and m is the total number of edges in the graph [29].

If G is a connected, simple, and k-regular graph, then it is obvious that $M(G) = A - \frac{k}{n}J$ where J is the unit matrix.
For example if G is a complete and regular bipartite graph $K_{n,n}$ then $M(G) = A - \frac{1}{2}J$.

Example 2.1. *Let G be a complete bipartite graph $K_{n,m}$. Then*

$$M = \begin{bmatrix} -\frac{m}{2n}J_{n\times n} & \frac{1}{2}J_{n\times m} \\ \frac{1}{2}J_{m\times n} & -\frac{n}{2m}J_{m\times m} \end{bmatrix}.$$

For example, we will have for a star graph $S_n = K_{1,n}$, the referenced matrix is written as follows:

$$M = \frac{1}{2}\begin{bmatrix} -n & 1 & \cdots & 1 \\ 1 & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ 1 & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & -\frac{1}{n} & \cdots & -\frac{1}{n} \end{bmatrix}.$$

Proposition 2.2. *Let G be the complete multipartite graph $G = K_{n_1,n_2,\cdots,n_k}$. The modularity matrix is a block-matrix. The entries in the block of size $n_i \times n_j$ are $m_{ij} = (1-\delta_{ij}) - \frac{k_j k_i}{2m}$, where δ_{ij} stands for the Kronecker delta-symbol.*

Proof. According to the adjacency matrix A of a complete multipartite graph $G = K_{n_1,n_2,\cdots,n_k}$, where V is the vertex set partitioned into k disjoint clusters $V_1, \cdots, V_k$, the entries a_{ij} are $1 \; if \; C(i) \neq C(j) \; and \; 0 \; if \; C(i) = C(j)$, where $C(i)$ refers to the cluster membership of vertex i.

Let $|V_i| = n_i$ for each $i = 1, \cdots, k$, meaning the number of vertices in each cluster V_i is n_i, $k_j = n - n_i$ where $j \in V_i$, and $2m = \sum_{j=1}^{n} k_j$.

If $n_1 \neq n_2 \neq \cdots \neq n_k$, then the entries in the block equals

$$m_{ij} = (1 - \delta_{ij}) - \frac{k_j k_i}{2m}.$$

If $n_1 = n_2 = \cdots = n_k$, then the entries in the block equals

$$m_{ij} = (1 - \delta_{ij}) - \frac{k_i^2}{2m}.$$

□

Graph partitioning is the process of dividing a graph into smaller, more manageable components. These components, often referred to as clusters or communities, can be characterized by specific properties. Numerous algorithms have been developed to address graph partitioning problems. Fiedler [17] pioneered the spectral clustering method, while Newman and Girvan [30] introduced the modularity clustering approach. Newman [28] later provided a detailed explanation of modularity clustering. It has been shown [22] that using normalized modularity matrices and normalized adjacency matrices yields identical clustering results. The allocation of positive modularity eigenvalues plays a crucial role in community detection. Bolla [4] demonstrated that having zero as the largest eigenvalue of the modularity matrix is a sufficient but not necessary condition for the network to be indivisible. In the following we present relevant theorems related to these issues.

As an alternative to the standard modularity matrix, Bolla [3] proposed the normalized modularity matrix, defined as:

$$M_D = D^{-1/2} M D^{-1/2},$$

where D is the diagonal degree-matrix. The spectrum of the normalized modularity matrix lies within the interval $[-1, 1]$. A key property of the normalized modularity matrix is that it always has an eigenvalue equal to 0, with the corresponding eigenvector given by $\sqrt{d}$, where $d = (d_1, d_2, \cdots, d_n)$ represents the degree vector of the graph (where each d_i is the degree of vertex i). The normalized adjacency matrix is defined as follows:

$$A_D = D^{-1/2} A D^{-1/2},$$

where D is the diagonal degree-matrix [22].

Sylvester's inertia theorem states that if a symmetric matrix is congruent to a diagonal matrix (i.e., there exists an invertible matrix that transforms the symmetric matrix into a diagonal form), the number of positive, negative, and zero entries on the diagonal (the inertia of the matrix) is invariant under congruence transformations. In simpler terms, the law tells us that the number of positive, negative, and zero eigenvalues of a symmetric matrix does not change if the matrix is transformed by a congruent matrix. These counts are called the inertia of the matrix. The proof of this theorem is presented in reference [21].

Lemma 2.3. *(Sylvester's inertia theorem) If A is a symmetric matrix of order n and B is a nonsingular matrix of order n, then $B^T AB$ and A have the same number of positive (pos), negative (neg), and zero (zero) eigenvalues. This means they share the same inertia, denoted as*

inertia(pos,neg,zero),

where ***pos+neg+zero=n***.

Theorem 2.4. *[22] Let 0 be a simple eigenvalue of M_D, and 1 a simple eigenvalue of A_D. If $\lambda \neq 0,1$ then (λ, u) is an eigenpair of A_D if and only if (λ, u) is an eigenpair of M_D.*

Theorem 2.5. *Let G be a graph with the adjacency matrix A with*

$$inertia(pos, neg, zero),$$

$pos \geq 1$. Then the normalized modularity matrix $M_D = D^{-1/2}MD^{-1/2}$ and the modularity matrix M have the following inertia:

inertia(pos-1,neg,zero+1).

Proof. According to Lemma 2.3, A and its normalized adjacency matrix $A_D = D^{-1/2}AD^{-1/2}$ have the same inertia, so it holds for M and M_D as well. On the other hand, the corresponding eigenvector of eigenvalues 1 of A_D is $\sqrt{d}$ and also M_D has an eigenvalue of 0 with an eigenvector $\sqrt{d}$. According to theorem 2.4, if $\lambda \neq 0,1$ then (λ, u) is a eigenpair of A_D if and only if (λ, u) is a eigenpair of M_D.

□

Theorem 2.6. *Let G be a k-regular graph of order n, then the edge average matrix P contain eigenvalue 0, k with multiplicity $n-1$ and 1, respectively.*

Proof. The sum of the entries in each row of the matrix P is equal to k, so k and 0 are eigenvalues. □

The independence number of a graph G, denoted by $\alpha(G)$, is a valuable concept in community detection. It represents the maximum number of vertices that can be selected from G such that no pair of vertices is directly connected [5]. The upper bound for the independence number of a graph, using the maximum eigenvalue λ_{max} of the normalized Laplacian, is given by

$$\alpha(G) \leq n(1 - \frac{1}{\lambda_{max}})\frac{\Delta}{\delta}. \tag{2}$$

Specifically, Δ and δ are maximum and minimum degree of G respectively. This bound is derived from the fact that the normalized Laplacian matrix is positive semidefinite, and its eigenvalues are real and non-negative. Moreover, the normalized Laplacian matrix has an eigenvalue of 0 with multiplicity equal to the number of connected components in the graph [8]. In the following we employ eigenvalue-based approaches to estimate the independence number, utilizing both the modularity matrix and its normalized counterpart.

Proposition 2.7. *Let λ_1 be the smallest eigenvalue of M_D then*

$$\alpha(G) \leq n(1 - \frac{1}{1-\lambda_1(M_D)})\frac{\Delta}{\delta}. \tag{3}$$

Proof. According to [13], the relationship between the largest eigenvalue of the normalized Laplacian matrix L_D and the smallest eigenvalue of the normalized modularity matrix M_D is derived as $\lambda_{max} = 1 - \lambda_1(M_D)$. Therefore, according to relation (3), the theorem is proved. □

3 Modularity energy and modularity Estrada

Modularity energy is a concept analogous to graph energy, but it is specifically designed to measure the strength of connections within and between communities in a network. It is calculated using the eigenvalues of the modularity matrix. If G is a graph and M is its modularity matrix. The modularity energy of G is defined as $E_{mod}(G) = E(M) = \sum\limits_{i=1}^{n} |\lambda_i(M)|$, where $\lambda_i(M)$ are the eigenvalues of M.

Example 3.1. *If $G = K_n$ and M and M_D denote the modularity matrix and the normalized modularity matrix, respectively. The modularity energy equals:*

$$E(M) = \sum_{i=1}^{n} \|\lambda_i(M)\| = n - 1.$$
$$E(M_D) = \sum_{i=1}^{n} \|\lambda_i(M_D)\| = 1.$$

In this section, we derive the upper and lower bounds for the modularity energy. The first Zagreb index, denoted as $M_1 = \sum_{i=1}^{n} k_i^2$, is a topological index used to characterize the degree-based properties of a graph G [19].

Theorem 3.2. *Let G be a connected and simple graph and M it's modularity matrix, then*

$$E(M) \geq \delta + \Delta - \frac{n-1}{n}\sqrt{\frac{nM_1 - 4m^2}{n-1}} - \frac{2m}{n}, \tag{4}$$

where the minimum and maximum degrees are indicated by δ and Δ respectively.

Proof.

$$\begin{aligned} M_1 = \sum_{i=1}^{n} k_i^2 = k_1^2 + k_2^2 + \cdots + k_n^2 &\geq \Delta^2 + \frac{(k_2 + \cdots + k_n)^2}{n-1} \\ &= \Delta^2 + \frac{(2m - \Delta)^2}{n-1}, \end{aligned}$$

on the other hand, by multiplying both sides by $n - 1$, we obtain

$$(n-1)M_1 \geq \Delta^2(n-1) + 4m^2 - 4m\Delta + \Delta^2.$$

Now consider function $f(\Delta) = n\Delta^2 - 4m\Delta + 4m^2 - (n-1)M_1 \leq 0$. To compute the discriminant of $f(\Delta)$, we obtian:

$$\Delta_{1,2} = \frac{2m}{n} \pm \frac{n-1}{n}\sqrt{\frac{nM_1-4m^2}{n-1}}.$$

Therefore, from the relation $\delta \leq \frac{2m}{n}$, the result is $\delta + \Delta \leq \frac{2m}{n} + \Delta$. The inequality (4) follows easily from $E(M) \geq 2m/n$. □

Theorem 3.3. *Let G be a simple, connected and k-regular graph with n vertices and m edges. Then*

$$E(M) \leq \sqrt{n(2m-k^2)}, \tag{5}$$

equality holds in (7) if and only if G is isomorphic to $\bar{k_n}$.

Proof. Suppose that $\lambda_i(M)$ be the eigenvalues of the modularity matrix $M(G)$, then

$$E(M) = \sum_{i=1}^{n} |\lambda_i(M)| \leq \sqrt{\sum_{i=1}^{n} 1^2} \times \sqrt{\sum_{i=1}^{n} \lambda_i(M)^2} = \sqrt{n} \times \sqrt{2m-k^2}.$$

□

To quantify the overall strength of community structure in a graph, we introduce the modularity Estrada index, a topological invariant derived from the spectral properties of the modularity matrix. If G is a graph and M is it's modularity matrix the modularity Estrada index is defined as $EE_{mod}(G) = EE(M) = \sum_{i=1}^{n} e^{\lambda_i(M)}$ where $\lambda_i(M)$ are the eigenvalues of M.

This index, leverages the concept of the Estrada index, a well-established graph invariant, to characterize the community structure in terms of the exponential sum of the eigenvalues of the modularity matrix. It seems that, a higher value of $EE_{mod}(G)$ indicates a stronger and more cohesive community structure within the graph. The spectral moment of order k of a graph is defined as $M(G,k) = \sum_{i=1}^{n} \lambda_i^k$.

Example 3.4. *Let G be a complete graph of order n with the modularity matrix M. Also, let $\lambda_i(M)$ be the eigenvalues of the modularity matrix and $M(G,k) = \sum_{i=1}^{n} (\lambda_i(M))^k$ is the spectral moment of graph. Then*

$$\begin{aligned} EE(M) &= \sum_{i=1}^{n} e^{\lambda_i(M)} = \sum_{i=1}^{n}\sum_{k=0}^{\infty} \frac{(\lambda_i(M))^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{M(G,k)}{k!} \\ &= n + \frac{-(n-1)}{1!} + \frac{n-1}{2!} + \cdots = n + (n-1)(\frac{1}{e} - 1) = 1 + \frac{n-1}{e}. \end{aligned}$$

Example 3.5. *Let G be a complete graph of order n with the normalized modularity matrix M_D. Also, let $\lambda_i(M_D)$ be the eigenvalues of the normalized modularity matrix. Then*

$$
\begin{aligned}
EE(M_D) &= \sum_{i=1}^{n} e^{\lambda_i(M_D)} = \sum_{i=1}^{n}\sum_{k=0}^{\infty} \frac{(\lambda_i(M_D))^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{M(G,k)}{k!} \\
&= n + \frac{-(n-1)}{n-1} + \cdots = 1 + \frac{n-1}{e^{1/n-1}}.
\end{aligned}
$$

Theorem 3.6. *Let G be a simple, connected, and k-regular graph of order n with the modularity matrix M. Then*

$$
EE(M) \le n - 1 + e^{\sqrt{2m-k^2}}, \tag{6}
$$

the equality holds if and only if $G = \bar{K}_n$.

Proof. Using the definition Estrada index of M, we have:

$$
\begin{aligned}
EE(M) &= \sum_{i=1}^{n} e^{\lambda_i(M)} = \sum_{i=1}^{n}\sum_{k\ge 0} \frac{(\lambda_i(M))^k}{k!} \\
&= n + \sum_{i=1}^{n}\sum_{k\ge 1} \frac{(\lambda_i(M))^k}{k!} \\
&= n + \sum_{k\ge 1} \frac{1}{k!} (\sum_{i=1}^{n} (\lambda_i(M))^k \\
&= n + \sum_{k\ge 1} \frac{1}{k!} (\sum_{i=1}^{n} ((\lambda_i(M))^2)^{k/2}) \\
&\le n + \sum_{k\ge 1} \frac{1}{k!} (\sum_{i=1}^{n} (\lambda_i(M))^2)^{k/2} \\
&= n - 1 + \sum_{k\ge 0} \frac{1}{k!} (2m - k^2)^{k/2}.
\end{aligned}
$$

□

In the following section, we show the trace of the matrix M, defined as $Tr(M) = \sum_{i=1}^{n} m_{ii}$, where m_{ii} represents the diagonal elements of the matrix.

Theorem 3.7. *If M is the modularity matrix of the simple graph G, and M_1 is the first Zagreb index, then $Tr(M) = \frac{-M_1}{2m}$.*

Proof. If A is the adjacency matrix and P is the average edge matrix of the graph G then $M = A - P$ is the matrix modularity. Therefore

$$\begin{aligned} Tr(M) &= Tr(A-P) \\ &= Tr(A) - Tr(P) \\ &= 0 - Tr(P) \\ &= -\sum_{i=1}^{n} \frac{k_i k_i}{2m} = -\sum_{i=1}^{n} \frac{k_i^2}{2m} \\ &= -\frac{\sum_{i=1}^{n} k_i^2}{2m} = \frac{-M_1}{2m}. \end{aligned}$$

□

4 Community detection and modularity maximization

Community detection in networks is a fundamental task aimed at identifying groups of nodes that are densely connected within themselves but sparsely connected to nodes in other groups. A key metric in community detection is modularity Q, which quantifies the quality of a network partition into communities. Maximizing modularity is a well-known NP-hard problem, requiring efficient algorithms to find optimal or near-optimal solutions [24]. Various approaches, including greedy algorithms, spectral methods, and optimization techniques, have been proposed to address this challenge. Another critical aspect of community detection is estimating the number of communities. Many algorithms require the number of clusters as input, and the choice of this parameter significantly impacts the results. Although several methods for estimating the number of clusters, such as spectral methods and statistical techniques, exist, these approaches can be computationally expensive [7]. In this work, we propose a simpler approach based on the eigenvalues of the modularity matrix. By analyzing the distribution of eigenvalues, particularly the largest eigenvalue, we can obtain a reasonable estimate of the number of communities in a network. This approach offers a practical and computationally efficient solution for estimating the number of clusters without the need complex statistical methods. In this section, relationships between the eigenvalues of the modularity matrix M and the average degree matrix P are presented.

Theorem 4.1. *Let G be a simple, connected, k-regular graph with adjacency matrix A and average degree matrix P. Then $\lambda_i(M) = \lambda_i(A) - \lambda_i(P)$.*

Proof. Since A and P are real symmetric matrices, A and P are diagonalizable. Moreover, $AP = PA$ which implies A and P are simultaneously diagonalizable. Therefore, the i-th eigenvalue of M is given by $\lambda_i(M) = \lambda_i(A) - \lambda_i(P)$.

□

Lemma 4.2. *[33] Let M be the modularity matrix of a graph G with eigenvalues $\lambda_1(M) \geq \lambda_2(M) \geq \cdots \geq \lambda_n(M)$. The maximum modularity value Q_{max} is given by:*

$$Q_{max} = \tfrac{n}{4m}\lambda_1(M).$$

Theorem 4.3. *Let A be the adjacency matrix of a graph G with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. The maximum modularity value Q_{max} has the following bounds:*

$$\frac{n}{4m}\lambda_2 \leq Q_{max} \leq \frac{n}{4m}\lambda_1. \tag{7}$$

Proof. According to Lemma 4.2, the maximize Q_{max}, is corresponding with

$$Q_{max} = \tfrac{n}{4m}\lambda_1(M),$$

where $\lambda_1(M)$ is the largest eigenvalue of the modularity matrix. Using the [22], we have

$$\lambda_1(A) \geq \lambda_1(M) \geq \lambda_2(A) \geq \lambda_2(M) \geq \cdots \geq \lambda_n(A) \geq \lambda_n(M).$$

Therefore, the result is obtained. □

Theorem 4.4. *Let M be the modularity matrix of a graph G of order n with c communities, and let M_i be the matrix corresponding to the i-th community on the main diagonal of the modularity matrix in the block form as follows:*

$$M = \begin{bmatrix} M_1 & M_{12} \dots & M_{1c} \\ M_{12}^T & M_2 \dots & M_{2c} \\ \vdots & \vdots \ddots & \vdots \\ M_{1c}^T & M_{2c}^T \dots & M_c \end{bmatrix}.$$

Then

$$Q = \frac{\sum\limits_{i=1}^{c}(\sum\limits_{ij} m_{ij})}{2m}. \tag{8}$$

Proof. Consider the adjacency matrix A of a graph G with c clusters. The nodes of the graph can always be reordered so that they are arranged according to their cluster memberships. The $n_i \times n_i$ adjacency matrices A_i of these clusters are located on the diagonal of the adjacency matrix A and contain only intracluster links:

$$A = \begin{bmatrix} A_1 & A_{12} \dots & A_{1c} \\ A_{12}^T & A_2 \dots & A_{2c} \\ \vdots & \vdots \ddots & \vdots \\ A_{1c}^T & A_{2c}^T \dots & A_c \end{bmatrix}.$$

Therefore, the modularity matrix corresponding to it is as follows:

$$M = A - P = \begin{bmatrix} M_1 & M_{12} \dots & M_{1c} \\ M_{12}^T & M_2 \dots & M_{2c} \\ \vdots & \vdots \ddots & \vdots \\ M_{1c}^T & M_{2c}^T \dots & M_c \end{bmatrix}.$$

On the other hand, according to [18],

$$Q = \sum_{c \in C} \left(\frac{|E(c)|}{|E|} - \left(\frac{\sum_{v \in c} k_v}{2|E|} \right)^2 \right). \tag{9}$$

The modularity can be decomposed into the sum of the modularity of its communities. Since the modularity matrix is block-diagonal, with each block representing a community, the overall modularity value can be calculated by summing the diagonal elements of these blocks and dividing by twice the number of edges, $2m$. □

5 Application of the eigenvalue of modularity

Let M be the modularity matrix of the graph G. Let

$$\lambda_1(M) \geq \lambda_2(M) \geq \cdots \geq \lambda_n(M),$$

be the eigenvalues of M, where $\lambda_i(M)$ represents the i-th eigenvalue of the modularity matrix M. The estimate (c^*) is the number of eigenvalues greater than the square root of the largest eigenvalue:

$$c^* = \sum_{i=1}^{n} 1_{\{\lambda_i(M) \geq \sqrt{\lambda_1(M)}\}}. \tag{10}$$

We investigated the relationship between the number of communities in a network and the eigenvalues of its modularity matrix. Through Python-based experiments on various types of graphs (complete, cyclic, Petersen, path, star, multipartite complete, and ladder graphs), we observed a strong correlation between the number of communities and the number of eigenvalues greater than the square root of the maximum eigenvalue. This finding suggests that the eigenvalue spectrum of the modularity matrix can provide valuable insights into the underlying community structure of a network. We further validated this approach by applying it to several well-known real-world networks. The estimated number of communities closely matched the results obtained from traditional modularity-based algorithms.

In the following sections, we a brief overview of some of these networks and analyze them, including social, technological, and biological networks, to identify community structures. Additionally, in Table 1, we present the calculated eigenvalues of the modularity matrix and the corresponding estimated number of communities based on these eigenvalues.

The **Les Misérables network** is an undirected graph representing the co-occurrences of characters in Victor Hugo's novel. It consists of 77 nodes, each representing a character, and 254 edges, indicating co-appearances in the same chapter. The weight of each edge reflects the frequency of co-occurrence. This network exhibits a community structure, with various algorithms suggesting a division into 5 to 6 communities. Our proposed method, based on the eigenvalues of the modularity matrix, estimates the number of communities to be 6, aligning with these previous findings. By analyzing the eigenvalues that exceed the

Table 1. N_c= Number of detectioned communities
$N(\lambda_i(M) \geq \sqrt{\lambda_1(M)})$=Number of eigenvalues greater than the square root of the largest eigenvalue.

Graph	Vertex	Edge	$\sqrt{\lambda_1(M)}$	$N(\lambda_i(M) \geq \sqrt{\lambda_1(M)})$	c^*	N_c
Complete	n	n(n-1)/2	0	0	1	1
P_{25}	25	24	1.39	6	6	6
P_{14}	14	13	1.34	3	3	3
Petersen	10	15	1	1	2	1
C_5	5	5	0.78	1	2	2
C_{40}	40	40	1.40	10	10	8
$K_{2,2}$	4	4	1.49	1	1	1
$K_{6,6}$	12	36	2.98	1	1	1
$K_{2,5,6,55}$	68	767	2.53	1	1	1
W_{10}	10	18	1.23	3	3	4
W_{55}	55	108	1.40	10	10	7
CL_{23}	46	69	1.71	8	8	6
L_{43}	8	10	1.27	2	2	3
S_4	5	4	1.05	1	1	1
S_{26}	21	20	1.82	1	1	1
Davis-southern-women	32	69	2.09	4	4	3
Florentine-families	15	20	2.25	3	3	4
Les-miserables	77	254	3-05	6	6	6
Karate-club	34	78	2.23	3	3	3
Footbal	115	616	3.04	12	12	13
Facebook	4039	88234	12.21	62	62	60

square root of the maximum eigenvalue, we arrive at this estimate, offering a reliable and computationally efficient approach for community detection.

Zachary's karate club, a well-known benchmark dataset, represents social ties among 34 individuals in a karate club. This network is frequently used to evaluate community detection algorithms, which typically identify between 2 and 4 communities. In this study, we leverage the modularity eigenvalue method to estimate the number of communities in this network, and our results indicate the presence of 3 distinct communities.

The **American football network**, compiled by Girvan and Newman, represents games between Division IA colleges during the 2000 regular season. Comprising 115 nodes and 616 edges, this network is organized into 12 teams. Various community detection algorithms have identified between 10 and 13 communities within this network. Our Python-based analysis, utilizing an eigenvalue-based modularity method, has yielded 13 communities, aligning with the upper end of previously reported findings.

The **Facebook network**, a dataset comprising 4,039 nodes and 888,234 edges, includes node features, circles, and ego networks. It has been previously partitioned into 60 communities using various community detection algorithms. Our investigation using Python, employing an eigenvalue-based modularity method, identified 62 communities in this network.

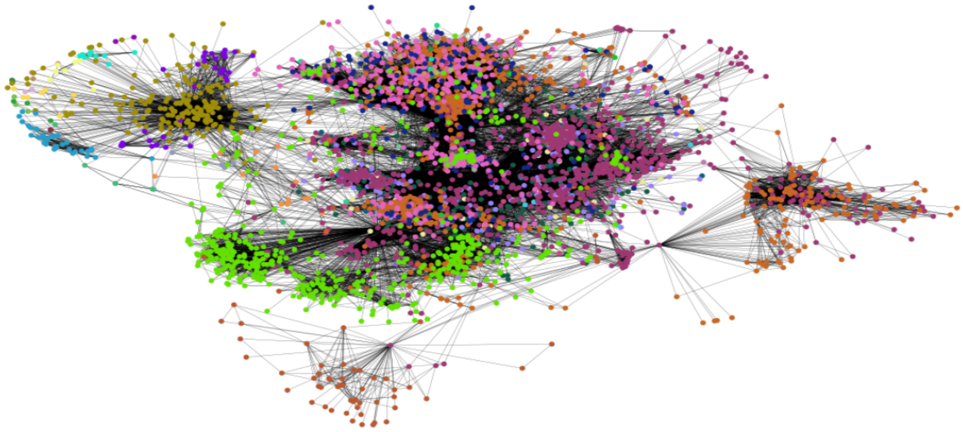

Figure 1. Communities of the Facebook network.

The **Florentine families network**, a dataset comprising 15 nodes and 20 edges, depicts marriage relationships among 16 Italian families in 15th-century Florence. This network has been studied extensively, with various community detection algorithms identifying 3 to 4 communities. Our investigation, using a Python-based approach relying on the eigenvalue of the modularity matrix, also converges to a 3 community structure.

The **Davis Southern women network** is a social network dataset consisting of 32 nodes representing 18 Southern women and 14 social events, collected by Davis and colleagues in the 1930s. This network has been extensively studied in the context of community detection. While previous analyses have identified 3 communities, our investigation using Python and an eigenvalue-based modularity method suggests the presence of 4 distinct communities.

6 Conclusion

In this study, the role of the Estrada index and graph energy in analyzing the modularity spectrum and estimating the number of communities in complex networks was investigated. By integrating these spectral indices with the modularity matrix, a novel method for evaluating the community structure of networks was introduced, which not only enhances the accuracy of estimation but also offers deeper insights into the stability and organization of communities. This matrix modularity spectral approach establishes a robust framework for analyzing large-scale networks without requiring computationally intensive full clustering procedures. Furthermore, it lays the groundwork for developing efficient algorithms to identify dynamic and hierarchical community structures.

Funding

This research received no external funding.

Data Availability Statement

Data is contained within the article.

Conflicts of Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

[1] A. R. Ashrafi, G. H. Fath-Tabar, Bounds on the Estrada index of ISR (4, 6)-fullerenes, Appl. Math. Lett. 24 (2011) 337–339. https://doi.org/10.1016/j.aml.2010.10.018
[2] R. Bhattacharya, N. K. Nagwani, S. Tripathi, A community detection model using node embedding approach and graph convolutional network with clustering technique, Decis. Anal. J. 9 (2023) 100362. https://doi.org/10.1016/j.dajour.2023.100362
[3] M. Bolla, Penalized versions of the Newman-Girvan modularity and their relation to normalized cuts and k-means clustering, Phys. Rev. E 84 (2011) 016108. https://doi.org/10.1103/PhysRevE.84.016108
[4] M. Bolla, B. Bullins, S. Chaturapruek, S. Chen, K. Friedl, Spectral properties of modularity matrices, Linear Algebra Appl. 473 (2015) 359–376. https://doi.org/10.1016/j.laa.2014.10.039
[5] J. A. Bondy, U. S. R. Murty, Graph Theory, Springer, New York, 2008. https://doi.org/10.1007/978-1-84628-970-5
[6] S. P. Borgatti, M. G. Everett, J. C. Johnson, Analyzing Social Networks, SAGE Publications, London, 2022. https://doi.org/10.1080/0022250x.2015.1053371
[7] G. Budel, P. Van Mieghem, Detecting the number of clusters in a network, J. Complex Netw. 8 (2020) cnaa047. https://doi.org/10.1093/comnet/cnaa047
[8] F. Chung, Spectral Graph Theory, American Mathematical Society, Providence, 1997. https://doi.org/10.1090/cbms/092
[9] D. Cvetković, S. Simić, Graph spectra in computer science, Linear Algebra Appl. 434 (2011) 1545–1562. https://doi.org/10.1016/j.laa.2010.11.035
[10] A. K. Dey, Y. Tian, Y. R. Gel, Community detection in complex networks: From statistical foundations to data science applications, WIREs Comput. Stat. 14 (2022) e1566. https://doi.org/10.1002/wics.1566
[11] E. Estrada, P. A. Knight, A First Course in Network Theory, Oxford University Press, Oxford, 2015. https://doi.org/10.1007/s00362-017-0961-1
[12] D. Fasino, F. Tudisco, An algebraic analysis of the graph modularity, SIAM J. Matrix Anal. Appl. 35 (2014) 997–1018. https://doi.org/10.1137/130943455
[13] D. Fasino, F. Tudisco, Modularity bounds for clusters located by leading eigenvectors of the normalized modularity matrix, J. Math. Ineq. 11(2016) 701–714. https://doi.org/10.7153/jmi-11-56
[14] G. H. Fath-Tabar, A. R. Ashrafi, New upper bounds for Estrada index of bipartite graphs, Linear Algebra Appl. 435 (2011) 2607–2611. https://doi.org/10.1016/j.laa.2011.01.034
[15] G. H. Fath-Tabar, A. R. Ashrafi, I. Gutman, Note on Laplacian energy of graphs, Bull. Acad. Serbe Sci. Arts (Cl. Sci. Math. Natur.) 33 (2008) 1–10. http://eudml.org/doc/258685
[16] G. H. Fath-Tabar, A. R. Ashrafi, I. Gutman, Note on Estrada and L-Estrada indices of graphs, Bull.

Acad. Serbe Sci. Arts (Cl. Sci. Math. Natur.) 139 (2009) 1–16. http://eudml.org/doc/253352

[17] M. Fiedler, Algebraic connectivity of graphs, Czechoslov. Math. J. 23 (1973) 298–305. https://doi.org/10.21136/CMJ.1973.101168

[18] I. Gutman, The energy of a graph, Ber. Math.-Statist. Sekt. Forsch. Graz 103 (1978) 1–22.

[19] I. Gutman, N. Trinajstić, Graph theory and molecular orbitals, in: A. Graovac (Ed.), Topics in Molecular Organization and Engineering, Vol. 2, Springer, Dordrecht, 2005, pp. 49–93. https://doi.org/10.1007/3-540-06399-4_5

[20] I. Gutman, X. Li, J. Zhang, Graph energy, in: M. Dehmer, F. Emmert-Streib (Eds.), Analysis of Complex Networks: From Biology to Linguistics, Wiley-VCH, Weinheim, 2009, pp. 145–174. https://doi.org/10.1002/9783527627981.ch7

[21] R. A. Horn, C. R. Johnson, Matrix Analysis, 2nd ed., Cambridge University Press, Cambridge, 2012. https://doi.org/10.1017/CBO9781139020411

[22] H. Jiang, C. Meyer, Relations between adjacency and modularity graph partitioning, in: H. Kashima, T. Ide, W. Peng (Eds.), Advances in Knowledge Discovery and Data Mining, PAKDD 2023, Lecture Notes in Computer Science, vol. 13936, Springer, Cham, 2023, pp. 547–559. https://doi.org/10.1007/978-3-031-33377-4_15

[23] D. Jin, B. Yang, C. Liu, D. He, J. Zhang, W. Zhang, A survey of community detection approaches: From statistical modeling to deep learning, IEEE Trans. Knowl. Data Eng. 35 (2021) 1149–1170. https://doi.org/10.1109/TKDE.2021.3104155

[24] D. Koutra, M. van Leeuwen, C. Plant, M. M. Gaber, E. Ntoutsi, A. Schubert, A. Zimek (Eds.), Machine Learning and Knowledge Discovery in Databases: Research Track, ECML PKDD 2023, Proceedings, Part V, Lecture Notes in Computer Science, vol. 14173, Springer, Cham, 2023. https://doi.org/10.1007/978-3-031-43418-1

[25] J. Li, Z. Wang, K. Yu, J. Cao, Y. Wang, A comprehensive review of community detection in graphs, Neurocomputing 576 (2024) 128169. https://doi.org/10.1016/j.neucom.2024.128169

[26] R. Nasiri, G. H. Fath-Tabar, A. Gholami, Z. Yarahmadi, Resolvent Estrada and signless Laplacian Estrada indices of graphs, MATCH Commun. Math. Comput. Chem. 77 (2017) 157–176. https://match.pmf.kg.ac.rs/electronic_versions/Match77/n1/match77n1_157-176.pdf

[27] M. E. J. Newman, Detecting community structure in networks, Eur. Phys. J. B 38 (2004) 321–330. https://doi.org/10.1140/epjb/e2004-00124-y

[28] M. E. J. Newman, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2004) 026113. https://doi.org/10.1103/PhysRevE.69.026113

[29] M. E. J. Newman, Networks: An Introduction, Oxford University Press, Oxford, 2010. https://doi.org/10.1093/acprof:oso/9780199206650.001.0001

[30] M. E. J. Newman, Community detection in networks: Modularity optimization and maximum likelihood are equivalent, arXiv:1606.02319 (2016). https://doi.org/10.48550/arXiv.1606.02319

[31] M. Rosvall, A. V. Esquivel, A. Lancichinetti, J. D. West, R. Lambiotte, Different approaches to community detection, in: P. Doreian, V. Batagelj, A. Ferligoj (Eds.), Advances in Network Clustering and Blockmodeling, John Wiley & Sons, Hoboken, 2019, pp. 105–119. https://doi.org/10.1002/9781119483298.ch4

[32] L. Tang, H. Liu, Community Detection and Mining in Social Media, Springer, Cham, 2022. https://doi.org/10.1007/978-3-031-01900-5

[33] P. Van Mieghem, Graph Spectra for Complex Networks, Cambridge University Press, Cambridge, 2023. https://doi.org/10.1017/CBO9780511921681

Citation: S. A. Hosseini, G. H. Fath-Tabar, H. Shabani, Detection of communities by modularity matrix, J. Disc. Math. Appl. 10(4) (2025) 359–374.

https://doi.org/10.22061/jdma.2025.11684.1112