Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 611-618, August 2014

ELA

A NOTE ON GRAPHS WHOSE LARGEST EIGENVALUE OF THE MODULARITY MATRIX EQUALS ZERO*

SNJEŽANA MAJSTOROVIĆ† AND DRAGAN STEVANOVIĆ‡

Abstract. Informally, a community within a graph is a subgraph whose vertices are more connected to one another than to the vertices outside the community. One of the most popular community detection methods is the Newman's spectral modularity maximization algorithm, which divides a graph into two communities based on the signs of the principal eigenvector of its modularity matrix in the case that the modularity matrix has positive largest eigenvalue. Newman defined a graph to be indivisible if its modularity matrix has no positive eigenvalues. It is shown here that a graph is indivisible if and only if it is a complete multipartite graph.

Key words. Modularity matrix, Community structure, Largest eigenvalue, Complete multipartite graph.

AMS subject classifications. 05C50, 05C82, 05C90, 92C42.

1. Introduction.

In a real-world network originating from a particular application, communities represent vertices that are usually bonded together by a defining property, and as a consequence, tend to be more connected to one another than to other vertices in the network. For example, communities in social networks may arise based on common location, interests or occupation, while citation networks may form communities based on research topic. Usually, the information on the network's community structure is not available, and being able to effectively identify communities within a network may provide further insight into its topology and node properties. This problem of community detection has been thoroughly surveyed in [6].

Let $G = (V, E)$ be a graph with $n = |V|$ vertices, $m = |E|$ edges, adjacency matrix A and let $d \in \mathbb{R}^n$ be the vector of the degrees of its vertices. For a particular partition of V into the communities $C_1, \ldots, C_k$, $k \geq 2$, a widely used measure de-

*Received by the editors on November 17, 2013. Accepted for publication on August 11, 2014. Handling Editor: Leslie Hogben.

†University of Osijek, Department of Mathematics, Trg Ljudevita Gaja 6, 31000 Osijek, Croatia (smajstor@mathos.hr). The work of this author was partially supported by GReGAS project within the EuroGIGA Collaborative Research Network.

‡University of Primorska, Institute Andrej Marušič, Muzejski trg 2, 6000 Koper, Slovenia, and Mathematical Institute, Serbian Academy of Sciences and Arts, Knez Mihajlova 36, 11000 Belgrade, Serbia (dragan.stevanovic@upr.si). The work of this author was supported by Research Programme P1-0285 and Research Project J1-4021 of the Slovenian Research Agency, and Research Grant ON174033 of the Ministry of Education and Science of Serbia.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 611-618, August 2014

ELA

scribing its "quality" is the *modularity*, introduced by Newman and Girvan [10] and defined as

$$q = \frac{1}{2m} \sum_{p=1}^{k} \sum_{u \in C_p} \sum_{v \in C_p} M_{uv},$$

where

$$M = A - \frac{1}{2m} dd^T$$

is the *modularity matrix* of G. All the row sums of M are zeros, due to the sum of vertex degrees being equal to $2m$, so the modularity matrix has an eigenvalue 0 with the all-one vector $j \in \mathbb{R}^n$ as its eigenvector.

Newman proposed in [9] the spectral partitioning algorithm which divides a graph into two communities, in the case that its modularity matrix has positive largest eigenvalue, in such a way that all vertices with positive components of the principal eigenvector of modularity matrix are placed in one community and all the rest in the other community. The discussion in [9] implicitly implies that the graphs with the largest eigenvalue of the modularity matrix equal to zero are totally lacking the community structure, and Newman called such graphs *indivisible*.

On the other hand, apparent examples of graphs with pure anticommunity structure are the complete multipartite graphs, in which any two vertices from the same part are not adjacent to each other, while any two vertices from different parts are adjacent to each other. We show that these two observations are related through the following

THEOREM 1.1. *A connected graph has the largest eigenvalue of the modularity matrix equal to zero if and only if it is a complete multipartite graph.*

The paper is organized as follows. In the next section, we describe the modularity spectrum of complete multipartite graphs, proving one direction of Theorem 1.1. In Section 3, we combine the facts that the modularity matrix is a rank-one modification of the adjacency matrix and that the graphs with exactly one positive eigenvalue of the adjacency matrix are complete multipartite graphs, to prove the other direction of Theorem 1.1. Using Sylvester's law of inertia, an analogue of Theorem 1.1 is proved for the normalized modularity matrix as well. Finally, in Section 4, we discuss differences between our approach and the approach taken in [3], where these results have been obtained independently.

2. The modularity spectrum of complete multipartite graphs. Recall that G is a *complete multipartite graph* if there exists a partition $V = V_1 \cup \cdots \cup V_k$ of its vertices such that any two vertices from different parts are adjacent to each

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 611-618, August 2014

ELA

other, while any two vertices from the same part are not adjacent to each other. The complete multipartite graph with parts of size $n_p = |V_p|$, $p = 1, \ldots, k$, is denoted by $K_{n_1,\ldots,n_k}$. Denote the vertices of the part V_p, $p = 1, \ldots, k$, by $v_{p,1}, \ldots, v_{p,n_p}$. Note that the degree of any vertex from V_p is equal to $d(v_{p,q}) = n - n_p$, $p = 1, \ldots, k$, $q = 1, \ldots, n_p$.

For each $p = 1, \ldots, k$, if $n_p \geq 2$, let $x_{p,q} \in \mathbb{R}^n$, $q = 2, \ldots, n_p$, denote the vector with

$$x_{p,q}(v_{p,1}) = 1, \quad x_{p,q}(v_{p,q}) = -1, \quad x_{p,q}(v) = 0 \text{ for } v \in V \setminus \{v_{p,1}, v_{p,q}\}.$$

Then

$$Mx_{p,q} = Ax_{p,q} - \frac{1}{2m} dd^T x_{p,q} = [d(v_{p,1}) - d(v_{p,q})] - \frac{1}{2m} d\,[d(v_{p,1}) - d(v_{p,q})] = 0,$$

showing that $x_{p,q}$ is an eigenvector of M corresponding to eigenvalue 0. Since j and the vectors $x_{p,q}$, $p = 1, \ldots, k$, $q = 2, \ldots, n_p$, are linearly independent (j is orthogonal to each $x_{p,q}$, while for each vertex $u \in V \setminus \{v_{1,1}, \ldots, v_{k,1}\}$ exactly one vector among all $x_{p,q}$ has a nonzero u-component), this shows that the multiplicity of 0 as an eigenvalue of M is at least $1 + \sum_{p=1}^{k}(n_p - 1) = n - k + 1$.

The remaining $k - 1$ eigenvectors of M can be chosen in such a way that they are orthogonal to j and all of $x_{p,q}$. Such an eigenvector y has to be constant within any part V_p, $p = 1, \ldots, k$: if $n_p \geq 2$, then y is orthogonal to all of $x_{p,q}$ which implies that $y(v_{p,1}) = y(v_{p,q})$ for each $q = 2, \ldots, n_p$ (and if $n_p = 1$, then y is trivially constant in the singleton V_p).

Let $j_p \in \mathbb{R}^n$, $p = 1, \ldots, k$, be the vector being equal to 1 for vertices in V_p and 0 for vertices in $V \setminus V_p$. Then there exist coefficients $\alpha_p \in \mathbb{R}$, $p = 1, \ldots, k$, such that

$$y = \sum_{p=1}^{k} \alpha_p j_p.$$

From the orthogonality of y and j we get

$$y^T j = \sum_{p=1}^{k} \alpha_p n_p = 0. \tag{2.1}$$

Let λ be an eigenvalue of M corresponding to the eigenvector y. Noting that the column of A corresponding to a vertex in V_p is equal to $j - j_p$, $p = 1, \ldots, k$, and that

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 611-618, August 2014

ELA

$d = \sum_{p'=1}^{k} j_{p'}(n - n_{p'})$, we get

$$\begin{aligned}
0 = My - \lambda y &= \sum_{p=1}^{k} \alpha_p (Mj_p - \lambda j_p) \\
&= \sum_{p=1}^{k} \alpha_p \left[Aj_p - \frac{1}{2m} dd^T j_p - \lambda j_p \right] \\
&= \sum_{p=1}^{k} \alpha_p \left[(j - j_p)n_p - d\frac{(n - n_p)n_p}{2m} - \lambda j_p \right] \\
&= \sum_{p=1}^{k} \alpha_p \left[\sum_{p' \neq p} j_{p'} n_p - \sum_{p'=1}^{k} j_{p'} \frac{(n - n_{p'})(n - n_p)n_p}{2m} - \lambda j_p \right] \\
&= \sum_{p'=1}^{k} j_{p'} \left[\sum_{p \neq p'} \alpha_p n_p - \sum_{p=1}^{k} \alpha_p \frac{(n - n_{p'})(n - n_p)n_p}{2m} - \alpha_{p'} \lambda \right] \\
&= \sum_{p'=1}^{k} j_{p'} \left[-\alpha_{p'} n_{p'} - \sum_{p=1}^{k} \alpha_p \frac{(n - n_{p'})(n - n_p)n_p}{2m} - \alpha_{p'} \lambda \right].
\end{aligned}$$

The vectors $j_{p'}$, $p' = 1, \ldots, k$, are linearly independent, which implies that the coefficient of each $j_{p'}$ above is equal to 0. After dividing it by $n - n_{p'}$ and adding $\frac{n}{2m} \sum_{p=1}^{k} \alpha_p n_p = 0$, we get

$$\alpha_{p'} \frac{\lambda + n_{p'}}{n - n_{p'}} = \sum_{p=1}^{k} \alpha_p \frac{n_p^2}{2m}. \tag{2.2}$$

Since the right-hand side of Eq. (2.2) does not depend on p', the expression $\alpha_{p'} \frac{\lambda + n_{p'}}{n - n_{p'}}$ has a constant value c for each p'. In particular,

$$\alpha_p \frac{\lambda + n_p}{n - n_p} = \alpha_{p'} \frac{\lambda + n_{p'}}{n - n_{p'}} = c.$$

If $\lambda = -n_p$ for some $p = 1, \ldots, k$, then $c = 0$, and consequently, $\alpha_{p'} = 0$ holds for all p' such that $n_{p'} \neq n_p$. Let P be the set of all indices i such that $n_i = n_p$. Then both Eq. (2.1) and (2.2) reduce to

$$\sum_{i \in P} \alpha_i = 0,$$

which has $|P| - 1$ linearly independent solutions, yielding that $-n_p$ is an eigenvalue of M with multiplicity $|P| - 1$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 611-618, August 2014

ELA

If $\lambda + n_p \neq 0$ for each p, then

$$\alpha_p = \alpha_{p'} \frac{\lambda + n_{p'}}{n - n_{p'}} \frac{n - n_p}{\lambda + n_p}.$$

Replacing the values α_p back in Eq. (2.2) and dividing it by $\alpha_{p'} \frac{\lambda + n_{p'}}{n - n_{p'}}$, we obtain

$$f(\lambda) := \sum_{p=1}^{k} \frac{n_p^2(n - n_p)}{2m(\lambda + n_p)} = 1. \tag{2.3}$$

Let $n'_1 < \cdots < n'_{k'}$ be the sequence of all distinct values among $n_1, \ldots, n_k$. The function $f(\lambda)$ is defined on the union of intervals $(-\infty, -n'_{k'}) \cup (-n'_{k'}, -n'_{k'-1}) \cup \cdots \cup (-n'_2, -n'_1) \cup (-n'_1, +\infty)$. The derivative $f'(\lambda) = \sum_{p=1}^{k} -\frac{n_p^2(n-n_p)}{2m(\lambda+n_p)^2}$ is negative everywhere, so that $f(\lambda)$ is strictly decreasing on each interval. The function $f(\lambda)$ is negative on $(-\infty, -n'_{k'})$, so that Eq. (2.3) has no solution on this interval. Since

$$\lim_{\lambda \to -n'_i -} f(\lambda) = -\infty, \lim_{\lambda \to -n'_i +} f(\lambda) = +\infty, \lim_{\lambda \to +\infty} f(\lambda) = 0,$$

Eq. (2.3) has a unique solution on each of the intervals $(-n'_i, -n'_{i-1})$, $i = 2, \ldots, k'$, and $(-n'_1, +\infty)$. Its solution on the interval $(-n'_1, +\infty)$, however, is equal to 0, which has been dealt with already.

To conclude, we have:

THEOREM 2.1. *For a complete multipartite graph $K_{n_1,\ldots,n_k}$ with $n = n_1 + \cdots + n_k$ vertices, let $n'_1 < \cdots < n'_{k'}$ be the sequence of all distinct values among $n_1, \ldots, n_k$, and let s_i, $i = 1, \ldots, k'$ be the number of occurrences of n'_i among $n_1, \ldots, n_k$. The spectrum of the modularity matrix of $K_{n_1,\ldots,n_k}$ consists of:*

- *(i) an eigenvalue 0 of multiplicity $n - k + 1$,*
- *(ii) an eigenvalue $-n'_i$ of multiplicity $s_i - 1$, whenever $s_i \geq 2$, and*
- *(iii) $k' - 1$ eigenvalues λ, one from each of the intervals $(-n'_i, -n'_{i-1})$, $i = 2, \ldots, k'$, satisfying*

$$\sum_{p=1}^{k} \frac{n_p^2(n - n_p)}{2m(\lambda + n_p)} = 1.$$

Hence, a complete multipartite graph has the largest eigenvalue of the modularity matrix equal to zero, proving one direction of Theorem 1.1.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 611-618, August 2014

ELA

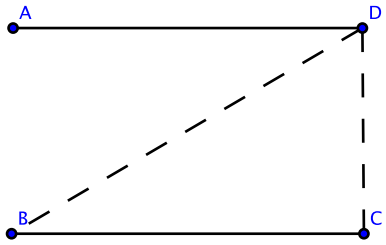

FIG. 3.1. *The subgraph induced by vertices A, B, C and D.*

3. Noncomplete multipartite graphs have a positive modularity eigenvalue. Let us first recall an 1982 result of Miroslav Petrović, that was originally stated in terms of infinite graphs.

THEOREM 3.1 ([11]). *A connected graph has exactly one positive eigenvalue of its adjacency matrix if and only if it is a complete multipartite graph.*

In the process of proving this result, Petrović first determined the adjacency spectrum of complete multipartite graphs. It is interesting that the same spectrum has been independently determined 30 years later by Charles Delorme in [5]. For the proof of the other direction of Theorem 3.1, Petrović relied on the simple use of the Interlacing theorem [4]: if a graph is not a complete multipartite graph, then it contains two nonadjacent vertices A and B, together with a vertex C that is, say, adjacent to B, but not adjacent to A. Since A is not an isolated vertex, there exists a vertex D adjacent to A (see Fig. 3.1). Then, regardless of the existence of edges BD and CD, the subgraph induced by vertices A, B, C and D has two positive adjacency eigenvalues, implying that a graph which is not a complete multipartite graph has at least two positive adjacency eigenvalues.

We now need the following classical result that may be found, for example, in [7, Theorem 8.1.8] or [12, pp. 94–97].

LEMMA 3.2. *Let A be a real symmetric $n \times n$ matrix with eigenvalues $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$. Let $B = A + zz^T$, where $z \in \mathbb{R}^n$, have eigenvalues $\lambda_1(B) \geq \cdots \geq \lambda_n(B)$. Then*

$$\lambda_{i-1}(A) \geq \lambda_i(B) \geq \lambda_i(A), \qquad i = 1, \ldots, n$$

under the convention that $\lambda_0(A) = +\infty$.

To finish the proof of the other direction of Theorem 1.1, we observe that $M = A - \frac{1}{2m}dd^T$ is a rank-one modification of A with $z = d/\sqrt{2m}$, i.e., that $A = M + zz^T$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 611-618, August 2014

ELA

Applying Lemma 3.2 to A and M, we get that $\lambda_1(M) \geq \lambda_2(A)$. Now, if G is not a complete multipartite graph, then it has at least two positive eigenvalues of its adjacency matrix by Theorem 3.1, which then implies that the largest eigenvalue of its modularity matrix is positive.

REMARK 3.3. As an alternative to modularity matrix, Bolla [2] studied the *normalized modularity matrix* defined as

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

with $D = diag(d)$. The spectrum of M_D is contained within $[-1, 1]$ and always contains an eigenvalue 0 with $\sqrt{d}$ as its eigenvector. The analogue of Theorem 1 holds for the normalized modularity matrix as well: *A connected graph has the largest eigenvalue of the normalized modularity matrix equal to zero if and only if it is a complete multipartite graph.*

The proof stems immediately from the proof of Theorem 1: since M_D and M are congruent matrices, by Sylvester's law of inertia [8] the matrix M_D has no positive eigenvalues (and has eigenvalue 0) if and only if the matrix M has no positive eigenvalues (and has eigenvalue 0), which happens if and only if G is a complete multipartite graph.

4. Concluding remarks. Theorem 1.1 resolves the open problem that the second author posed at CRM Conference on Applications of Graph Spectra in Barcelona in July, 2012. In the meantime, one of the participants of that conference and her group solved this open problem independently in [3] (see also addendum to [1] at `http://media.wiley.com/product_ancillary/28/11183449/DOWNLOAD/addendum.pdf`). We would like to shed some light on the differences between the approaches taken here and in [3].

Bolla et al. [3] first derive the modularity and the normalized modularity spectra of complete graphs (not the multipartite ones), determine the multiplicity of modularity eigenvalue zero of a complete multipartite graph, and then show that the remaining modularity eigenvalues are negative, without determining them. In contrast, we completely describe the modularity eigenstructure of complete multipartite graphs by determining their eigenvectors and providing an expression that implicitly defines non-zero modularity eigenvalues (Theorem 2.1).

To prove the other part of Theorem 1.1, Bolla et al. [3] rely on characterization of complete multipartite graphs in terms of 3-vertex induced subgraphs and use the characterization of negative semidefinite matrices in terms of signs of their principal minors (specifically, the minors of order three). In contrast, we use a very insightful observation of the first author that the modularity matrix is a rank-one perturbation of the adjacency matrix, which enabled us to complete the proof by using a classical

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 611-618, August 2014

ELA

interlacing result (Lemma 3.2) and a slightly less classical characterization of the adjacency spectra of complete multipartite graphs (first obtained by Petrović [11] in 1982, then independently reproved by Delorme [5] in 2012).

REFERENCES

[1] M. Bolla. *Spectral Clustering and Biclustering: Learning Large Graphs and Contingency Tables.* John Wiley & Sons, Ltd., West Sussex, 2013.
[2] M. Bolla. Penalized versions of the Newman-Girvan modularity and their relation to normalized cuts and k-means clustering. *Phys. Rev. E (3)*, 84:016108, 2011.
[3] M. Bolla, B. Bullins, S. Chaturapruek, S. Chen, and K. Friedl. When the largest eigenvalue of the modularity and the normalized modularity matrix is zero. Preprint arXiv:1305.2147 [math.SP], 2013. Available at `http://arxiv.org/abs/1305.2147`.
[4] D. Cvetković, M. Doob, and H. Sachs. *Spectra of Graphs—Theory and Application*, third edition. Johann Ambrosius Barth Verlag, Heidelberg-Leipzig, 1995.
[5] C. Delorme. Eigenvalues of complete multipartite graphs. *Discrete Math.*, 312:2532–2535, 2012.
[6] S. Fortunato. Community detection in graphs. *Phys. Rep.*, 486:75–174, 2010.
[7] G.H. Golub and C.F. Van Loan. *Matrix Computations*, third edition. Johns Hopkins University Press, Baltimore, 1996.
[8] R.A. Horn and C.R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.
[9] M.E.J. Newmann. Modularity and community structure in networks, *Proc. Natl. Acad. Sci. USA*, 103:8577–8582, 2006.
[10] M.E.J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E (3)*, 69:026113, 2004.
[11] M.M. Petrović. The Spectrum of Infinite Complete Multipartite Graphs. *Publ. Inst. Math. (Belgrade) (N.S.)*, 31:169–176, 1982.
[12] J.H. Wilkinson. *The Algebraic Eigenvalue Problem.* Clarendon Press, Oxford, 1965.