Bounds on the Largest Eigenvalue of the Distance Matrix of Connected Graphs

Natalie Denny

Advisor: John Caughman
2^{nd} Reader: Derek Garton

Math 501 Project
in partial fulfillment of the requirements for
the Master of Science in Mathematics

Portland State University
2020

Abstract

In this MTH 501 - Mathematical Literature project, we present the results of the article “On the Largest Eigenvalue of the Distance Matrix of a Connected Graph” by Bo Zhou and Nenad Trinajstic [Chemical Physics Letters 447 (2007) p.384-387]. The main results provide bounds on the largest eigenvalue of the distance matrix of a connected graph. These bounds are of interest as a molecular descriptor of chemical compounds.

1 Introduction

1.1 Overview

In this MTH 501 - Mathematical Literature project, we present the results by Zhou and Trinajstic found in the article "On the Largest Eigenvalue of the Distance Matrix of a Connected Graph" [6]. The eigenvalues of distance matrices, as well as other graph invariants, are of use in structure-property-activity modeling such as QSAR and QSPR models [3]. These models incorporate knowledge and methods from chemistry, physics, biology, mathematics, and statistics to predict the activity (e.g. anti-malarial, anti-cancer, etc), properties (e.g. water solubility, melting point, etc.), or toxicological data (e.g. organ toxicity, genotoxicity, etc.) of new chemicals [4]. The main results in the Zhou and Trinajstic paper provide bounds on the largest eigenvalue of the distance matrix of a connected graph.

To achieve those results, first we need to explore properties of connected graphs and their distance matrices. This will lead to the main bounds on the eigenvalues of distances matrices. Using these bounds, we can prove a Nordhaus-Gaddum type result for the upper and lower bounds of the largest eigenvalue of a connected graph and its complement.

1.2 Ethane and C_{20}

I have chosen to illustrate the properties and results throughout this paper with two molecules–ethane (C_2H_6) and the fullerene C_{20}, depicted below.

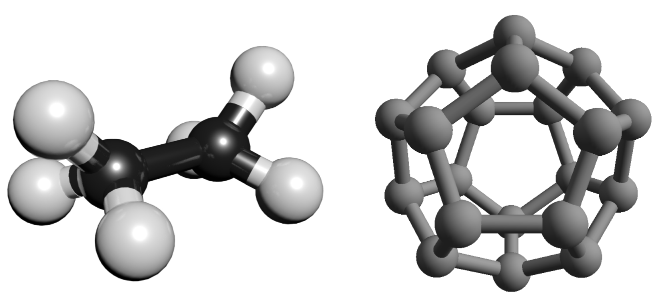


Figure 1: Ethane and C_{20} molecules

As these two examples illustrate, molecules can be modeled using graphs, where vertices represent atoms and the edges represent bonds between those atoms. In Figure 2, we see the graphs of ethane (Graph **E**) and the C_{20} fullerene (Graph **C**). The solid black vertices represent carbon atoms and the open vertices in Graph **E** represent hydrogen atoms.

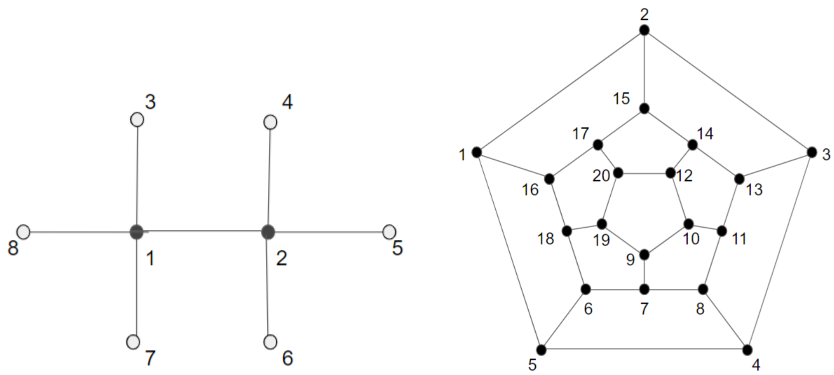


Figure 2: Graph **E** and Graph **C**

2 Definitions

In this section, we will collect a few definitions needed for our study of these topics. We begin by defining the distance matrix.

Definition 2.1. *Let G be a connected graph with n vertices. The* ***distance matrix*** $\mathbf{D}$ *is the $n \times n$ matrix whose ij-entry is*

$$\mathbf{D}_{ij} = dist(i,j)$$

for all $1 \leq i,j \leq n$, where $dist(i,j)$ denotes the path-length distance function for G.

Let $\mathbf{D_E}$ and $\mathbf{D_C}$ be the distance matrices of the the graphs **E** and **C**. The entry $[\mathbf{D_E}]_{ij}$ encodes the distance between vertices i and j. For example, $[\mathbf{D_E}]_{14}$ means that $dist(1,4) = 2$. The distance matrix $\mathbf{D_E}$ also encodes the edge set of **E** because any $[\mathbf{D_E}]_{ij} = 1$ implies $i \sim j$. Therefore the graph of **E** is uniquely determined up to isomorphism by the matrix $\mathbf{D_E}$. (The matrix $\mathbf{D_C}$ has been included in the appendix.)

$$\mathbf{D_E} = \begin{bmatrix} 0 & 1 & 1 & 2 & 2 & 2 & 1 & 1 \\ 1 & 0 & 2 & 1 & 1 & 1 & 2 & 2 \\ 1 & 2 & 0 & 3 & 3 & 3 & 2 & 2 \\ 2 & 1 & 3 & 0 & 2 & 2 & 3 & 3 \\ 2 & 1 & 3 & 2 & 0 & 2 & 3 & 3 \\ 2 & 1 & 3 & 2 & 2 & 0 & 3 & 3 \\ 1 & 2 & 2 & 3 & 3 & 3 & 0 & 2 \\ 1 & 2 & 2 & 3 & 3 & 3 & 2 & 0 \end{bmatrix}$$

Figure 3: The distance matrix of **E**.

As this example demonstrates, distance matrices of graphs are real, non-negative, and symmetric. Since they are symmetric, the sum of the entries above the diagonal and below the diagonal are equal. This sum is called the Wiener index of a chemical graph G, as defined below.

Definition 2.2. *Let G be a connected graph with n vertices. Define*

$$W(G) = \sum_{1\le i<j\le n} \boldsymbol{D}_{ij}.$$

*In other words, $W(G)$ is the sum of the distances between all unordered pairs of vertices. We refer to $W(G)$ as the **Wiener index** of G.*

The Wiener index is a topological index of a molecule that is used in chemical graph theory. For our examples, we find that

$$W(\mathbf{E}) = 58 \quad \text{and} \quad W(\mathbf{C}) = 500.$$

Another metric of interest in chemical graph theory is the sum of the squares of the distances of all unordered pairs, denoted $S(G)$ and defined below.

Definition 2.3. *Let G be a connected graph with n vertices. We define $S(G)$ to be the sum:*

$$S(G) = \sum_{u,v\in V(G)} dist(u,v)^2.$$

Definition 2.4. *For any symmetric $n \times n$ matrix with (real) eigenvalues λ_i $(1 \le i \le n)$, we use Λ to denote the maximum eigenvalue, so that*

$$\Lambda \ge \lambda_i \quad (1 \le i \le n).$$

Definition 2.5. *Let $\mathbb{G}$ be the class of connected graphs for which the distance matrix has exactly one positive eigenvalue.*

Familar graphs that are members of $\mathbb{G}$ include the dodecahedral and icosahedral graphs, the Petersen graph, the Gosset graph, the Schlafli graph, and the three Chang graphs (see [1] for definitions). There are also many infinite families of graphs that belong to $\mathbb{G}$. These include the cycle graphs, the Johnson graphs, the Hamming Graphs, the Cocktail Party graphs, the Doob graphs, the Doubled Odd Graphs, and the Halved Cubes. These graphs (and more) have been extensively studied in [1]. Since the C_{20} molecule is viewed as a dodecahedral graph, which has distance spectra

$$\{50, 0^{(9)}, (-7+3\sqrt{5})^{(3)}, -2^{(4)}, (-7-3\sqrt{5})^{(3)}\},$$

it follows that $\mathbf{C} \in \mathbb{G}$.

Interestingly, there are many additional examples of graphs in this class $\mathbb{G}$ found in chemistry. The graphs that correspond to methane, ethane, acetic acid, and alcohol all belong to $\mathbb{G}$. This is because the graphs of these molecules are trees. In Zhou's work on the largest eigenvalue of trees [5], it was shown that the spectrum of a tree has exactly one positive eigenvalue, and so it follows that

$\mathbf{E} \in \mathbb{G}$. Additionally, some compounds that are realized as cyclic graphs, such as octasulfur S_8, also belong to $\mathbb{G}$.

Some graphs of chemical molecules that do not belong to $\mathbb{G}$ are the "bucky ball" (buckminsterfullerene C_{60}) and the fullerene C_{70}.

3 Connected Graphs

Let us begin with a few lemmas about connected graphs that will be used later in the results about their distance matrices.

Lemma 3.1. *If G is a connected graph and $diam(G) \geq 3$, then $diam(\overline{G}) \leq 3$.*

Proof. Let $diam(G) \geq 3$. So there exist vertices u, v with $dist_G(u, v) = 3$, and in $\overline{G}$, we have $u \sim v$.

Consider any vertices z, w. In $\overline{G}$, either $z \sim u$ or $z \sim v$, because if $z \nsim u$ and $z \nsim v$ then $dist_G(u, v) \leq 2$, which is a contradiction. Similarly, $w \sim u$ or $w \sim v$ in $\overline{G}$.

Because $u \sim v$ in $\overline{G}$, this means $dist_{\overline{G}}(z, w) \leq 3$. Since z, w were chosen arbitrarily, this implies $\overline{G}$ has diameter at most 3. □

Lemma 3.2. *Let G be a connected graph with n vertices and m edges. Then*

$$W(G) \geq n(n-1) - m,$$

with equality if and only if $G = K_n$ or G has diameter two.

Proof. Since G is connected, we have

$$W(G) = \sum_{i<j} \mathbf{D}_{ij} \geq m.$$

Suppose $G = K_n$, then $m = \binom{n}{2} = \frac{n(n-1)}{2}$. Therefore $W(G) = m$. So,

$$W(G) = m = n(n-1) - \frac{n(n-1)}{2} = n(n-1) - m,$$

as desired.

Now suppose $G \neq K_n$, then $W(G) > m$. Let $i = \frac{n(n-1)}{2} - m$, so that i counts the number of pairs of vertices that are not adjacent in G. The distance between any pair of non-adjacent vertices is at least two. Therefore, if $G \neq K_n$, then

$$W(G) \geq m + 2i = m + 2\Big[\frac{n(n-1)}{2} - m\Big] = n(n-1) - m$$

with equality if the diameter of G is at most two. □

Since **E** has diameter greater than two, equality does not hold in the above lemma. Indeed,

$$W(\mathbf{E}) = 58 > 49 = 8(7) - 7.$$

Since the dodecahedron graph **C** is distance regular, the row sums of the distance matrix for **C** are constant ($[D_{\mathbf{C}}]_i = 50$ where $[D_{\mathbf{C}}]_i$ is the i-th row sum of the distance matrix). So

$$W(\mathbf{C}) = \frac{[20D_{\mathbf{C}}]_i}{2} = 500 > 350 = 20(19) - 30.$$

Lemma 3.3 ([6], Eqn (3)). *Let G be a connected graph with n vertices. Then*

$$\frac{n(n-1)}{2} \leq S(G) \leq \frac{(n+1)n^2(n-1)}{12}$$

and $S(G)$ is minimum for K_n and maximum for P_n, in which cases equality is attained.

Proof. Since G is a connected graph, $dist(u, v) \geq 1$ for all $u, v \in V(G)$.

Suppose $G = K_n$, so that $dist(u, v) = 1$ for all $u, v \in V(G)$. Then $S(G) = \binom{n}{2} = \frac{n(n-1)}{2}$, which counts the number of unordered pairs of vertices.

Suppose $G = P_n$, so that $dist(u, v) = n - 1$ for some $u, v \in V(G)$. Then there

$$S(G) = \sum_{i=1}^{n-1} i(n-i)^2 = \sum_{i=1}^{n-1} i^2(n-i) = \sum_{i=1}^{n-1} ni^2 - \sum_{i=1}^{n-1} i^3.$$

Evaluating the sums gives us

$$S(G) = n\left[\frac{(n-1)n(2n-1)}{6}\right] - \frac{n^2(n-1)^2}{4} = \frac{(n-1)n^2(n+1)}{12}.$$

If $G \neq K_n$ and $G \neq P_n$, then $dist(u, v) < n - 1$ for all u, v in $V(G)$ and $dist(u, v) > 1$ for some u, v in $V(G)$. Therefore $S(G) < \sum_{i=1}^{n-1} i(n-i)^2$ and $S(G) > \binom{n}{2}$. And thus,

$$\frac{n(n-1)}{2} < S(G) < \frac{(n+1)n^2(n-1)}{12},$$

as desired. □

Applying this result to our graphs **E** on 8 vertices and **C** on 20 vertices, we find that

$$28 \leq S(\mathbf{E}) \leq 336 \quad \text{and} \quad 190 \leq S(\mathbf{C}) \leq 13300.$$

Since the graphs for ethane and C_{20} are neither complete graphs nor paths, they are not examples of the minimum or maximum $S(G)$. By calculation, $S(\mathbf{E}) = 136$ and $S(\mathbf{C}) = 1540$.

4 Distance Matrices of Connected Graphs

Distance matrices of connected graphs are real, symmetric, non-negative, and irreducible. Row sums of these matrices contribute to the upperbounds on the eigenvalues. This first lemma implies that the spectral radius is at most the largest row sum.

Lemma 4.1. *Let λ be an eigenvalue of a non-negative matrix A. Then λ is bounded above by the maximum row sum of A. Furthermore, if the row sums of A are equal, then the maximum eigenvalue is the maximum row sum.*

Proof. Let $\vec{x}$ be an eigenvector in the λ-eigenspace for any eigenvalue λ of A. Then, $|A\vec{x}| = |\lambda\vec{x}|$. Let x_i be a component of $\vec{x}$ such that $|x_i| \geq |x_j|$ for all j. Then

$$|\lambda||x_i| = \left|\sum_{j=1}^{n} a_{ij}x_j\right| \leq \sum_{j=1}^{n} |a_{ij}||x_j| = \sum_{j=1}^{n} a_{ij}|x_j| \leq |x_i|\sum_{j=1}^{n} a_{ij}.$$

Note that $|x_i| > 0$ since there must be a non-zero entry of $\vec{x}$. So

$$|\lambda| \leq \sum_{j=1}^{n} a_{ij} \leq \max_{1\leq i\leq n} \sum_{j=1}^{n} a_{ij}.$$

Therefore the spectral radius of a non-negative matrix is bounded above by the maximum row sum. Furthermore, if the row sums are equal, then Λ equals the maximum row sum. □

In relation to ethane and C_{20} (graphs **E** and **C**), the above result says that the eigenvalues have an upper bound of 16 and 50 respectively. As noted earlier, the dodecahedron is distance-regular, so the rows sums of the distance matrix for C_{20} are constant. Therefore the largest eigenvalue for that matrix is 50.

In the results that follow, we will refine the bounds by constraining the maximum row sum, basing our constraints upon the number of vertices and diameter.

Definition 4.1. *Let G be a connected graph on $n \geq 4$ vertices. We say G is a **Broom graph**, if $diam(G) = 3$ and there exists a vertex v such that*

$$\big|\{u \mid dist(u, v) = 3\}\big| = n - 3.$$

Note that such a graph G contains a star $(K_{1,n-1})$ as a subgraph. We denote the class of all broom graphs by $\mathcal{B}$.

See Figure 3 for an illustration. Notice that the definition allows for there to exist edges between vertices that are at distance three from the vertex v.

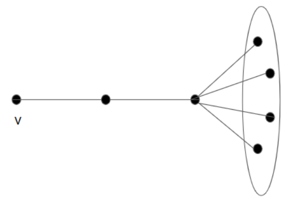
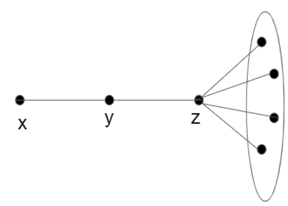


Figure 3: A broom graph

The next result gives a sufficient condition for a graph to be a broom.

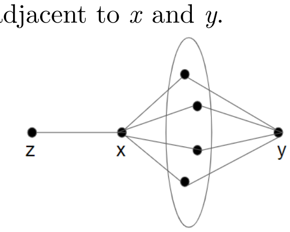
Lemma 4.2. *Suppose G is connected with $n \geq 5$ vertices and diameter 3. If any row of the distance matrix $\mathbf{D}$ has a sum equal to $1+2+3(n-3)$, then that row sum is maximal in $\mathbf{D}$ and $G \in \mathcal{B}$. Moreover, $\overline{G}$ is connected with diameter 3 and a maximum row sum of $1+2(n-3)+3$.*

Proof. Let x be a vertex that corresponds to a row with the specified sum in the distance matrix. Then there exists exactly one vertex at distance 1 (call it y), exactly one vertex at distance 2 (call it z), and the remaining $n-3$ vertices are at distance 3. The graph $G \backslash x$ contains a star ($K_{1,n-1}$) subgraph and thus G $\in \mathcal{B}$.



Claim: The diameter of $\overline{G}$ is 3.

In $\overline{G}$, we have $deg(x) = n-2$ (adjacent to all but y), $deg(y) = n-3$ (adjacent to all but x and z), and $deg(z) = 1$ (adjacent to x). The remaining vertices have degree at least 2 and are adjacent to x and y.



Consider the vertices partitioned by their distance in $\overline{G}$ from z.

Let $D_i = \{u \,|\, i = dist(z,u)\}$. Then $D_1 = \{x\}$, $D_2 = V(G) \backslash \{x,\ y,\ z\}$, and $D_3 = \{y\}$. It follows that $\overline{G}$ has diameter 3 and z has the maximum row sum. $\square$

Lemma 4.3. *Let D_M be the maximum row sum of the distance matrix $\boldsymbol{D}$. Let G be a connected graph with n vertices and diameter d. Then*

$$D_M \leq \sum_{i=1}^{d-1} i + (n-d)(d) \leq \frac{n(n-1)}{2}$$

and equality holds if and only if G is a path of length $n-1$.

Proof. Since G is a connected graph, $diam(G) \leq n-1$. Pick u, v such that $dist(u,v) = d$. Then there are a maximum of $n-d$ vertices at distance d from u. Let D_M be the maximum row sum of $\mathbf{D}$. Then

$$D_M \leq \sum_{i=1}^{d-1} i + (n-d)(d) \leq \frac{n(n-1)}{2}$$

with equality if and only if $d = n-1$.

Let G be a path of length $n-1$. Then $d = n-1$, so $D_M = \frac{n(n-1)}{2}$. Conversely, suppose $D_M = \frac{n(n-1)}{2}$. Then $d = n-1$ and G is a path of length $n-1$. $\square$

Looking at the distance matrices for $\mathbf{E}$ and $\mathbf{C}$, you can see the diameter in each of the matrices. Since the diameter is the maximum distance over all pairs of vertices, the diameter of $\mathbf{E}$ is three. By Lemma 4.3, the maximum row sum must satisfy

$$D_M \leq \sum_{i=1}^{2} i + (8-3)(3) = 18.$$

As exhibited in the distance matrix below, the maximum row sum is 16.

$$\mathbf{D}_E = \begin{bmatrix} 0 & 1 & 1 & 2 & 2 & 2 & 1 & 1 \\ 1 & 0 & 2 & 1 & 1 & 1 & 2 & 2 \\ 1 & 2 & 0 & 3 & 3 & 3 & 2 & 2 \\ 2 & 1 & 3 & 0 & 2 & 2 & 3 & 3 \\ 2 & 1 & 3 & 2 & 0 & 2 & 3 & 3 \\ 2 & 1 & 3 & 2 & 2 & 0 & 3 & 3 \\ 1 & 2 & 2 & 3 & 3 & 3 & 0 & 2 \\ 1 & 2 & 2 & 3 & 3 & 3 & 2 & 0 \end{bmatrix}$$

Corollary 4.3.1. *Let G be a path with n vertices. Then the row sums of the distance matrix are not equal.*

Proof. Let D_w be the row sum of $\mathbf{D}$ corresponding to a vertex, w, where $deg(w) = 2$. Then

$$D_w \leq 1 + \sum_{i=1}^{n-2} i = 1 + \frac{(n-2)(n-1)}{2} < \frac{n(n-1)}{2} = D_M.$$

$\square$

5 Eigenvalues of Distance Matrices

The eigenvalues of distance matrices have many properties that are not seen in the eigenvalues of other real and symmetric matrices. Since the diagonal of a distance matrix is all zeros, the sum of the eigenvalues is zero.

Lemma 5.1. *[6, Eqn (1)], Let G be a connected graph with $n \geq 2$ vertices and let λ_i $(1 \leq i \leq n)$ be the eigenvalues of the distance matrix $\boldsymbol{D}$ of G. Then*

$$\sum_{i=1}^{n} \lambda_i = 0.$$

Proof. Since $\mathbf{D}$ is real and symmetric, $\mathbf{D}$ is diagonalizable by the spectral theorem. So $\mathbf{D} = \mathbf{P}\mathbb{D}\mathbf{P}^{-1}$ where $\mathbb{D}$ is the diagonal matrix of eigenvalues of $\mathbf{D}$. Now

$$Tr(\mathbf{D}) = Tr(\mathbf{P}\mathbb{D}\mathbf{P}^{-1}) = Tr(\mathbb{D}\mathbf{P}^{-1}\mathbf{P}) = Tr(\mathbb{D}).$$

But $Tr(\mathbf{D}) = 0$ because $\mathbf{D}_{ii} = 0$ for all i. It follows that $Tr(\mathbb{D}) = 0$. □

Additionally, since the distance matrices are symmetric, the sum of the squares of the eigenvalues are related to the sum of the squares of the distances between all pairs of vertices. This relationship is detailed in the following lemma.

Lemma 5.2. *[6, Eqn (2)] Let G be a connected graph with $n \geq 2$ vertices and let λ_i $(1 \leq i \leq n)$ be the eigenvalues of the distance matrix $\boldsymbol{D}$ of G. Then*

$$\sum_{i=1}^{n} \lambda_i^2 = 2S(G)$$

Proof. Since $\mathbf{D}$ is real and symmetric, $\mathbf{D}$ is diagonalizable by the spectral theorem. So $\mathbf{D} = \mathbf{P}\mathbb{D}\mathbf{P}^{-1}$ where $\mathbb{D}$ is the diagonal matrix of eigenvalues of $\mathbf{D}$ and

$$\mathbf{D}^2 = \mathbf{P}\mathbb{D}^2\mathbf{P}^{-1}.$$

Now we compute as follows:

$$\sum_{i=1}^{n} \lambda_i^2 = Tr(\mathbb{D}^2) = Tr(\mathbb{D}^2\mathbf{P}^{-1}\mathbf{P}) = Tr(\mathbf{P}\mathbb{D}^2\mathbf{P}^{-1}) = Tr(\mathbf{D}^2).$$

Since $\mathbf{D}$ is symmetric, $\mathbf{D}^2 = \mathbf{D}\mathbf{D}^T$. Thus $\mathbf{D}^2_{ii} = \sum_{j=1}^{n} dist(i,j)^2$. Since $\mathrm{tr}(\mathbf{D}^2)$ counts the sum of the squares of the distances of ordered pairs, it double-counts the unordered pairs. So, $Tr(\mathbf{D}^2) = 2S(G)$. □

From the above results, the largest eigenvalue of a distance matrix is positive because we know that the sum of the squares must be non-zero ($S(G) > 0$) and the sum of the eigenvalues is zero. The next lemma provides a lower bound on the largest eigenvalue for all connected graphs in terms of the sum of the distances between all unordered pairs of vertices, $W(G)$.

Lemma 5.3. *[6, Cor.7] Let Λ be the largest eigenvalue of the distance matrix $\boldsymbol{D}$. Then*

$$\Lambda \geq \frac{2}{n}W(G)$$

with equality if and only if the row sums of $\boldsymbol{D}$ are all equal.

Proof. By the Rayleigh quotient for $\mathbf{D}$,

$$\Lambda = max\{\frac{\vec{x}^T\mathbf{D}\vec{x}}{\vec{x}^T\vec{x}} \mid \vec{x} \neq \vec{0}\}.$$

So it follows that

$$\Lambda \geq \frac{1^T\mathbf{D}1}{1^T1} = \frac{1^T[D_1D_2...D_n]^T}{n} = \frac{\sum_{i=1}^n D_i}{n} = \frac{2}{n}W(G).$$

Now suppose $\Lambda = \frac{2}{n}W(G)$. Then

$$\Lambda = \frac{1^T\mathbf{D}1}{1^T1}.$$

Moreover, $\Lambda = \frac{\vec{x}^T\mathbf{D}\vec{x}}{\vec{x}^T\vec{x}}$ if and only if $\vec{x}$ is an eigenvector for Λ. Therefore, 1 is an eigenvector for Λ, and

$$\Lambda 1 = \mathbf{D}1 = \begin{bmatrix} D_1 \\ D_2 \\ ... \\ D_n \end{bmatrix}$$

so therefore $D_1 = D_2 = ... = D_n$.

Conversely, suppose $D_1 = D_2 = ... = D_n$. Then $D_M = D_i$ for all i. So,

$$\mathbf{D}1 = \begin{bmatrix} D_1 \\ D_2 \\ ... \\ D_n \end{bmatrix} = \begin{bmatrix} D_M \\ D_M \\ ... \\ D_M \end{bmatrix} = D_M 1$$

Therefore, D_M is an eigenvalue. By Lemma 4.1, this implies $D_M = \Lambda$, and

$$\frac{2}{n}W(G) = \frac{2}{n}\sum_{i<j} D_{ij} = \frac{2}{n}\left[\frac{1}{2}\sum_{i=1}^{n} D_i\right] = \frac{1}{n}[nD_M] = D_M = \Lambda,$$

as desired. □

For the ethane and C_{20} graphs, the above bound gives us that $\Lambda(\mathbf{E}) > 14.5$ and $\Lambda(\mathbf{C}) \geq 50$. We realize equality in graph $\mathbf{C}$ because the row sums are equal. For graph $\mathbf{E}$, we now know that $14.5 < \Lambda < 16$.

The next result is another way to find a lower bound. As seen in our examples, the bound is less than optimal for many graphs. For example, it says that

$\Lambda(\mathbf{E}) > 12.25$ and $\Lambda(\mathbf{C}) > 34.84$. However, its usefulness will be seen when we seek bounds on the sum of the eigenvalues for a graph and its complement in a Nordhaus-Gaddum type result (Theorem 7.2).

Lemma 5.4. *[6, Cor.8] Let G be a connected graph with $n \geq 2$ vertices and m edges. Then*

$$\Lambda \geq 2(n-1) - \frac{2m}{n}$$

with equality if and only if $G = K_n$ or G is a regular graph of diameter two.

Proof. By Lemma 3.2,

$$W(G) \geq n(n-1) - m$$

with equality if and only if the diameter of G is at most two. By Lemma 5.3,

$$\Lambda \geq \frac{2}{n} W(G) \geq 2(n-1) - \frac{2m}{n}.$$

Suppose equality holds throughout. Then $G = K_n$ or G is of diameter two and by Lemma 5.3 the row sums of $\mathbf{D}$ are equal. This implies that $\mathbf{D} = 2\mathbf{J} - A - 2\mathbf{I}$. So, for all i, $D_M = D_i = k \cdot 1 + j \cdot 2$ where k is the number of vertices adjacent to vertex i and j is the number of vertices at distance two from vertex i. Since all of the row sums are the same, and since $1 + j + k = n$, it follows that k and j must be constant for each vertex. Therefore, G is k-regular. So, either $G = K_n$ or G is regular with diameter 2.

Conversely, suppose $G = K_n$ or G is a regular graph of diameter two. Then the row sums of the distance matrix are equal. So by Lemma 5.3 and Lemma 3.2,

$$\Lambda = \frac{2}{n} W(G) = \frac{2}{n}[n(n-1) - m] = 2(n-1) - \frac{2m}{n},$$

as desired. □

6 Distance Matrices of Graphs in $\mathbb{G}$

Given the abundance of chemical molecules whose graphs belong to the $\mathbb{G}$ class, including ethane and C_{20}, looking at this special class leads to further refinement of the bounds on the largest eigenvalue. The next lemma does not provide a bound but will be needed for the proof of Theorem 6.2.

Lemma 6.1. *[2] Let A be a non-negative, irreducible, symmetric matrix with exactly two distinct eigenvalues. Then $A = uu^T + r\boldsymbol{I}$ for some positive column vector u and some $r \in \mathbb{R}$.*

Theorem 6.2. *[6, Eqn(4)] Let $G \in \mathbb{G}$ with $n \geq 2$ vertices. Then*

$$\Lambda \leq \sqrt{\frac{2(n-1)}{n}S(G)},$$

with equality if and only if $G = K_n$.

Proof. By Lemma 5.1, $\Lambda = -\sum_{i=2}^{n} \lambda_i$, so $\Lambda^2 = \left[\sum_{i=2}^{n} \lambda_i\right]^2$. By the Cauchy-Schwartz inequality,

$$\left[\sum_{i=2}^{n} \lambda_i\right]^2 \leq (n-1)\sum_{i=2}^{n} \lambda_i^2.$$

But by Lemma 5.2,

$$\sum_{i=2}^{n} \lambda_i^2 = 2S(G) - \Lambda^2.$$

Therefore,

$$\Lambda^2 \leq (n-1)[2S(G) - \Lambda^2],$$

with equality if and only if $\lambda_2 = ... = \lambda_n$. Adding $(n-1)\Lambda^2$ to both sides, we obtain

$$n\Lambda^2 \leq 2(n-1)S(G),$$

which yields the desired result:

$$\Lambda \leq \sqrt{\frac{2(n-1)}{n}S(G)}.$$

Claim: $\Lambda = \sqrt{\frac{2(n-1)}{n}S(G)}$ if and only if $G = K_n$

Suppose $\Lambda = \sqrt{\frac{2(n-1)}{n}S(G)}$. Then $\lambda_2 = ... = \lambda_n$. This implies $\mathbf{D}$ has two distinct eigenvalues. By Lemma 6.1, $\mathbf{D} = \vec{u}\vec{u}^T + r\mathbf{I}$ for some positive column vector $\vec{u}$. Since $\mathbf{D}_{ii} = 0$ for all i and $[\vec{u}\vec{u}^T]_{ii} = u_i^2$, we have $0 = \mathbf{D}_{ii} = u_i^2 + r$. So, $u_i = \sqrt{-r}$, and $\vec{u} = \sqrt{-r}1$. Then, for all $i \neq j$, $\mathbf{D}_{ij} = u_i u_j + r\mathbf{I}_{ij} = u_i u_j = -r$. Since G is a connected graph, then for all $i \neq j$, $\mathbf{D}_{ij} > 0$ which implies that $r < 0$. This implies $r = -1$ (because each vertex must have a neighbor in a connected graph and $dist(v, u) = 1$ for $u \in N(v)$.) Therefore, for all $i \neq j$, $\mathbf{D}_{ij} = 1$, thus $G = K_n$.

Conversely, suppose $G = K_n$. Then $\mathbf{D} = \mathbf{J} - \mathbf{I}$ and $D_1 = ... = D_n$. So by Lemma 4.1, $\Lambda = D_1 = ... = D_n$. Also, $S(G) = \binom{n}{2} = \frac{n(n-1)}{2}$. So,

$$\sqrt{\frac{2(n-1)}{n}S(G)} = \sqrt{\frac{2(n-1)}{n}\frac{n(n-1)}{2}} = n - 1 = D_i = \Lambda,$$

as desired. ∎

Returning to ethane and C_{20}, Theorem 6.2 posits $\Lambda(\mathbf{E}) < 15.427$ and $\Lambda(\mathbf{C}) < 54.093$. Even though this is not an improvement for $\mathbf{C}$, it does provide us with an improvement on the bounds of the largest eigenvalue of the distance matrix for **E**. So, the search has narrowed to $14.5 < \Lambda < 15.427$. In fact, by calculation,

$$\Lambda(\mathbf{E}) \approx 14.937.$$

Corollary 6.2.1. *[6] Let $G \in \mathbb{G}$ with $n \geq 3$ vertices. Then*

$$\Lambda < \sqrt{\frac{(n+1)n(n-1)^2}{6}}.$$

Proof. By Theorem 6.2,

$$\Lambda \leq \sqrt{\frac{2(n-1)}{n} S(G)},$$

with equality if and only if $G = K_n$. By Lemma 3.3,

$$S(G) \leq \frac{(n+1)n^2(n-1)}{12}$$

with equality if and only if $G = P_n$. For $n \geq 3$, we have $P_n \neq K_n$, therefore

$$\Lambda < \sqrt{\frac{2(n-1)}{n} \frac{(n+1)n^2(n-1)}{12}} = \sqrt{\frac{(n+1)n(n-1)^2}{6}},$$

as desired. □

Unfortunately the corollary above does not provide a better bound on the largest eigenvalue for either example – ethane or C_{20}. Indeed, Corollary 6.2.1 only tells us that $\Lambda(\mathbf{E}) < 24.248$ and $\Lambda(\mathbf{C}) < 158.965$. The large discrepancy in the bound for graph **C** highlights that this upper bound on the largest eigenvalue was a result on $S(G)$ attaining a maximum value when the graph is a path.

7 Nordhaus-Gaddum Type Result

The final two theorems give us Nordhaus-Gaddum type results for the bounds on the largest eigenvalue of the distance matrix of a graph G and its complement, $\overline{G}$. The second theorem refines the first in the case that G or $\overline{G}$ is in the graph class $\mathbb{G}$.

In order to have eigenvalues for the distance matrix, a graph must be connected. So for a Nordhaus-Gaddum result, G and $\overline{G}$ must be connected. So as we will see in our next lemma, there needs to be at least 4 vertices in G.

Lemma 7.1. *Suppose that G and $\overline{G}$ are connected and nonempty graphs. Then G has $n \geq 4$ vertices.*

Proof. Since G and $\overline{G}$ are connected, $|E(G)| \geq n-1$ and $|E(\overline{G})| \geq n-1$. But $|E(G)| + |E(\overline{G})| = \binom{n}{2}$, so

$$2(n-1) \leq \frac{n(n-1)}{2},$$

which implies $n \geq 4$. $\square$

Theorem 7.2. *[6, Eqn (11)] Let G be a connected graph on $n \geq 4$ vertices with a connected complement $\overline{G}$. Then*

$$3(n-1) \leq \Lambda(G) + \Lambda(\overline{G}) < \frac{n(n+3)}{2} - 3,$$

with left equality if and only if G and $\overline{G}$ are both regular graphs of diameter two.

Proof. Let m_G and $m_{\overline{G}}$ be the number of edges of G and $\overline{G}$. Then we have $m_G + m_{\overline{G}} = \binom{n}{2}$, so, $2(m_G + m_{\overline{G}}) = n(n-1)$. By Lemma 5.4,

$$\begin{aligned}
\Lambda(G) + \Lambda(\overline{G}) &\geq \left[2(n-1) - \frac{2m_G}{n}\right] + \left[2(n-1) - \frac{2m_{\overline{G}}}{n}\right] \\
&= 4(n-1) - \frac{2(m_G + m_{\overline{G}})}{n} \\
&= 4(n-1) - \frac{n(n-1)}{n} \\
&= 3(n-1),
\end{aligned}$$

and equality holds if and only if G and $\overline{G}$ are regular graphs of diameter two.

Let $f(n) = \frac{n(n+3)}{2} - 3$. Let D_M be the maximum row sum of the distance matrix for G. By Lemma 4.1, we have $\Lambda \leq D_M$ with equality if and only if the row sums are equal. By Corollary 4.3.1, if the row sums of **D** are equal, then G is not a path and thus $\Lambda < \frac{n(n-1)}{2}$.

Consider the possible diameters for G and $\overline{G}$. Note that $diam(G) \neq 1$ since $\overline{G}$ is connected.

Suppose G has diameter two. First we claim that $\Lambda < 2n-3$. To see why, note that G cannot be a path. This is because if G is a path of diameter two, then $n = 3$, which is impossible. Now, by Lemma 4.3, $D_M < 1 + 2(n-2) = 2n-3$, and so

$$\Lambda < 2n - 3. \tag{1}$$

Now, by Lemma 3.1, since G has diameter 2, it must be the case that $\overline{G} > 3$. This implies

$$\Lambda(G) + \Lambda(\overline{G}) < \frac{n(n-1)}{2} + 2n - 3 = \frac{n(n-1) + 2(2n)}{2} - 3 = f(n).$$

Suppose G has diameter three. Then by Lemma 3.1 $\overline{G}$ has diameter three.

Case $n = 4$, then $G = \overline{G} = P_4$. So $\Lambda(G) = \Lambda(\overline{G}) < \frac{n(n-1)}{2} = 6$. Then, $\Lambda(G) + \Lambda(\overline{G}) < 12 \nleq f(4) = 11$. However, $\Lambda(G) = \Lambda(\overline{G}) = \sqrt{10} + 2$, so

$$\Lambda(G) + \Lambda(\overline{G}) = 4 + 2\sqrt{10} < 11 = f(4).$$

Case $n \geq 5$ and either G or $\overline{G}$ has a maximum row sum in the distance matrix of $1+2+3(n-3)$. Then by Lemma 4.2, the other graph has a maximum row sum of $1 + 2(n-3) + 3$. Therefore,

$$\Lambda(G) + \Lambda(\overline{G}) < [1 + 2 + 3(n-3)] + [1 + 2(n-3) + 3] = 5n - 8$$

Case $n \geq 5$ and neither G or $\overline{G}$ has a maximum row sum in the distance matrix of $1 + 2 + 3(n-3)$. Then,

$$\Lambda(G) + \Lambda(\overline{G}) < 2[1 + 2 + 3(n-3) - 1] = 6n - 14. \tag{2}$$

For $n = 5$, $6n - 14 < 5n - 8 = 17$. So, $\Lambda(G) + \Lambda(\overline{G}) < 17 = f(5)$.
For $n \geq 6$, $6n-14 \geq 5n-8$ and thus $\Lambda(G)+\Lambda(\overline{G}) < 6n-14$. Also, $6n-17 < f(n)$ for all n. Therefore, $\Lambda(G) + \Lambda(\overline{G}) < f(n)$ as desired. □

Coming back to the distance matrix for graph $\mathbf{E}$, this result suggests

$$21 < \Lambda(\mathbf{E}) + \Lambda(\overline{\mathbf{E}}) < 41.$$

Together with Corollary 6.2.1, we can deduce that $5.573 < \Lambda(\overline{\mathbf{E}}) < 26.5$. But since the maximum row sum of the distance matrix for graph $\overline{\mathbf{E}} = 12$ and using Lemma 4.1, $\Lambda(\overline{\mathbf{E}}) < 12$. Graph $\mathbf{E} \in \mathbb{G}$, so the next result will give improved bounds on the sum of the largest eigenvalues.

Applying Theorem 7.2 to graph $\mathbf{C}$ yields

$$57 < \Lambda(\mathbf{C}) + \Lambda(\overline{\mathbf{C}}) < 227.$$

Theorem 7.3. *[6, Eqn (12)] Let G be a connected graph on $n \geq 4$ vertices with a connected complement $\overline{G}$. If $G \in \mathbb{G}$ or $\overline{G} \in \mathbb{G}$, then*

$$\Lambda(G) + \Lambda(\overline{G}) < \sqrt{\frac{(n+1)n(n-1)^2}{6}} + 2n - 3.$$

Proof. Suppose $G \in \mathbb{G}$ or $\overline{G} \in \mathbb{G}$. Suppose WLOG, $G \in \mathbb{G}$ then from Corollary 6.2.1, $\Lambda(G) < \sqrt{\frac{(n+1)n(n-1)^2}{6}}$.
Consider the possible diameters for G and $\overline{G}$.

Suppose $\overline{G}$ has diameter two (or both G and $\overline{G}$ have diameter two) . Then by Theorem 7.2,

$$\Lambda(G) + \Lambda(\overline{G}) < \sqrt{\frac{(n+1)n(n-1)^2}{6}} + 2n - 3.$$

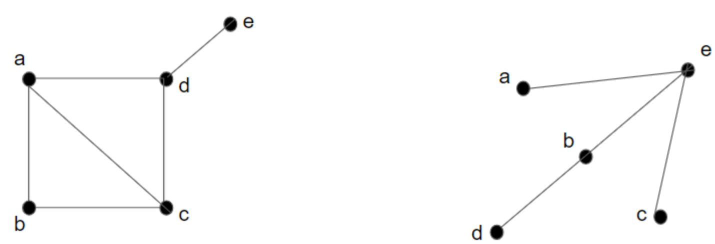
Else suppose $G \in \mathbb{G}$, $\overline{G} \notin \mathbb{G}$ and G has diameter 2. The case that was not made clear in the Zhou paper [6] is whether $G \in \mathbb{G}$ and G has diameter 2 determines that either $\overline{G}$ must also have a diameter 2 or that $\overline{G}$ must be an element of $\mathbb{G}$. So this case has not been proven to meet the improved bound.

Else suppose both G and $\overline{G}$ have diameter three. Then by (2) in the proof of Theorem 7.2, $\Lambda(G) + \Lambda(\overline{G}) < 6n - 14$. For $n = 4$ and $n \geq 6$, $6n - 14 < \sqrt{\frac{(n+1)n(n-1)^2}{6}} + 2n - 3$.

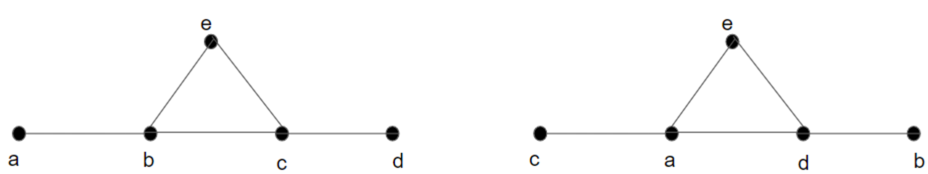
Suppose $n = 5$: If both G and $\overline{G}$ have diameter three with n = 5, then G and $\overline{G}$ both have at least 4 edges (n-1) and $|E(G)| + |E(\overline{G})| = \binom{5}{2} = 10$. But suppose G has 4 edges, then $G = P_5$, but $diam(P_5) = 4$ which is impossible. This implies $|E(G)| = |E(\overline{G})| = 5$.

There are three pairs graphs G and $\overline{G}$ on 5 vertices of diameter 3 up to isomorphism.

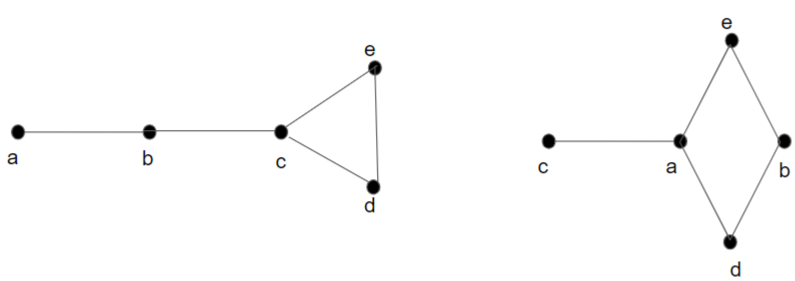
Pair 1 (G_1 and $\overline{G_1}$): $\Lambda(G_1) + \Lambda(\overline{G_1}) \approx 13.6754$



Pair 2 (G_2 and $\overline{G_2}$): $\Lambda(G2) + \Lambda(\overline{G_2}) \approx 13.2750$



Pair 3 (G_3 and $\overline{G_3}$): $\Lambda(G_3) + \Lambda(\overline{G_3}) \approx 13.5467$



And thus for $n = 5$, $\Lambda(G) + \Lambda(\overline{G}) < \sqrt{\frac{(n+1)n(n-1)^2}{6}} + 2n - 3 \approx 15.94427$. $\square$

Additionally, in all of the above 5-vertex graphs with diameter 3, both G_i and $\overline{G}_i$ have only one positive eigenvalue so they all belong to $\mathbb{G}$.

This theorem applies to both graph $\mathbf{E}$ and $\mathbf{C}$ since they belong to the graph class $\mathbb{G}$. This new result to graph $\mathbf{E}$ gives $\Lambda(\mathbf{E}) + \Lambda(\overline{\mathbf{E}}) < 34.39$. This new upper bound narrows further what the largest eigenvalue of graph $\overline{\mathbf{E}}$ to be at most 19.885. However, the previous lemmas could be applied for better bounds on the largest eigenvalue of $\overline{\mathbf{E}}$. We know from Lemma 4.1, $\Lambda\overline{E} < 12$ and from Lemma 6.2 $\Lambda(\mathbf{E}) < 15.427$. So $\Lambda(\mathbf{E}) + \Lambda(\overline{\mathbf{E}}) < 27.427$.

For graph C, Theorem 7.3 gives $\Lambda(\mathbf{C}) + \Lambda(\overline{\mathbf{C}}) < 195.97$. However, Lemma 3.1 that $diam(\overline{\mathbf{C}}) \leq 3$. Therefore by Lemma 4.2, the maximum row sum is $1+2+3(17) = 54$. Therefore, from Lemma 4.1, that $\Lambda(\mathbf{C}) = 50$ and $\Lambda(\overline{\mathbf{C}}) \leq 54$ and therefore the sum of the eigenvalues is at most 104. This is significantly less than what Theorem 7.3 gives.

8 Conclusion

While the Nordhaus-Gaddum type theorem can be appreciated, it seems like there can be improvements to this result. It also has yet to be proven that Theorem 7.3 applies to the case when G has diameter two, $G \in \mathbb{G}$, but $\overline{G} \notin \mathbb{G}$. What is known is that $\overline{G}$ cannot belong to any of the families of graphs like a Johnson, Hamming, Doob, Cocktail party, cyclic graphs and many more. Also, the diameter of $\overline{G}$ is not bounded.

Using Zhou's results [6], this paper has outlined upper and lower bounds on the largest eigenvalue of a distance matrix based upon the Wiener index, $S(G)$, as well as the largest row sum. A few of the results are also applicable to other symmetric, non-negative matrices.

9 Appendix

$$
\mathbf{D}_C = \begin{bmatrix}
0 & 1 & 2 & 2 & 1 & 2 & 3 & 3 & 4 & 5 & 4 & 4 & 3 & 3 & 2 & 1 & 2 & 2 & 3 & 3 \\
1 & 0 & 1 & 2 & 2 & 3 & 4 & 3 & 5 & 4 & 3 & 3 & 2 & 2 & 1 & 2 & 2 & 3 & 4 & 3 \\
2 & 1 & 0 & 1 & 2 & 3 & 3 & 2 & 4 & 3 & 2 & 3 & 1 & 2 & 2 & 3 & 3 & 4 & 5 & 4 \\
2 & 2 & 1 & 0 & 1 & 2 & 2 & 1 & 3 & 3 & 2 & 4 & 2 & 3 & 3 & 3 & 4 & 3 & 4 & 5 \\
1 & 2 & 2 & 1 & 0 & 1 & 2 & 2 & 3 & 4 & 3 & 5 & 3 & 4 & 3 & 2 & 3 & 2 & 3 & 4 \\
2 & 3 & 3 & 2 & 1 & 0 & 1 & 2 & 2 & 3 & 3 & 4 & 4 & 5 & 4 & 2 & 3 & 1 & 2 & 3 \\
3 & 4 & 3 & 2 & 2 & 1 & 0 & 1 & 1 & 2 & 2 & 3 & 3 & 4 & 5 & 3 & 4 & 2 & 2 & 3 \\
3 & 3 & 2 & 1 & 2 & 2 & 1 & 0 & 2 & 2 & 1 & 3 & 2 & 3 & 4 & 4 & 5 & 3 & 3 & 4 \\
4 & 5 & 4 & 3 & 3 & 2 & 1 & 2 & 0 & 1 & 2 & 2 & 3 & 3 & 4 & 3 & 3 & 2 & 1 & 2 \\
5 & 4 & 3 & 3 & 4 & 3 & 2 & 2 & 1 & 0 & 1 & 1 & 2 & 2 & 3 & 4 & 3 & 3 & 2 & 2 \\
4 & 3 & 2 & 2 & 3 & 3 & 2 & 1 & 2 & 1 & 0 & 2 & 1 & 2 & 3 & 5 & 4 & 4 & 3 & 3 \\
4 & 3 & 3 & 4 & 5 & 4 & 3 & 3 & 2 & 1 & 2 & 0 & 2 & 1 & 2 & 3 & 2 & 3 & 2 & 1 \\
3 & 2 & 1 & 2 & 3 & 4 & 3 & 2 & 4 & 2 & 1 & 2 & 0 & 1 & 2 & 4 & 3 & 5 & 4 & 3 \\
3 & 2 & 2 & 3 & 4 & 5 & 4 & 3 & 3 & 2 & 2 & 1 & 1 & 0 & 1 & 3 & 2 & 4 & 3 & 2 \\
2 & 1 & 2 & 3 & 3 & 4 & 5 & 4 & 4 & 3 & 3 & 2 & 2 & 1 & 0 & 2 & 1 & 3 & 3 & 2 \\
1 & 2 & 3 & 3 & 2 & 2 & 3 & 4 & 3 & 4 & 5 & 3 & 4 & 3 & 2 & 0 & 1 & 1 & 2 & 2 \\
2 & 2 & 3 & 4 & 3 & 3 & 4 & 5 & 3 & 3 & 4 & 2 & 3 & 2 & 1 & 1 & 0 & 2 & 2 & 1 \\
2 & 3 & 4 & 3 & 2 & 1 & 2 & 3 & 2 & 3 & 2 & 3 & 5 & 4 & 3 & 1 & 2 & 0 & 1 & 2 \\
3 & 4 & 5 & 4 & 3 & 2 & 2 & 3 & 1 & 2 & 3 & 2 & 4 & 3 & 3 & 2 & 2 & 1 & 0 & 1 \\
3 & 3 & 4 & 5 & 4 & 3 & 3 & 4 & 2 & 2 & 3 & 1 & 3 & 2 & 2 & 2 & 1 & 2 & 1 & 0
\end{bmatrix}
$$

References

[1] Ghodratollah Aalipour, Aida Abiad, Zhanar Berikkyzy, Jay Cummings, Jessica De Silva, Wei Gao, Kristin Heysse, Leslie Hogbend, Franklin H.J. Kenter, Jephian C.-H.Lin, and Michael Taite. On the distance spectra of graphs. *Linear Algebra and its Applications*, 497:66–87, 2016.

[2] D. Cao, V. Chvatal, A.J. Hoffman, and A. Vince. Variations on a theorem of ryser. *Linear Algebra and its Applications*, 260:215–222, 1997.

[3] James Devillers and Alexandru T Balaban. *Topological Indices and Related Descriptors in QSAR and QSPR*. CRC Press, 2000.

[4] K Roy, S Kar, and RN Das. *A primer on QSAR/QSPR modeling: Fundamental Concepts*. Springer-Verlag Inc., 2015.

[5] Bo Zhou. On the largest eigenvalue of the distance matrix of a tree. *MATCH Communications in Mathematical and Computer Chemistry*, 58:657–662, 2007.

[6] Bo Zhou and Nenad Trinajstic. On the largest eigenvalue of the distance matrix of a connected graph. *Chemical Physical Letters*, 447:384–387, 2007.