



A TEORIA ESPECTRAL DE GRAFOS E ALGUNS PROBLEMAS INTERESSANTES APLICADOS NA QUÍMICA E BIOLOGIA

Autor: Jeferson Dias da Silva

Orientador: Prof. Dr. Rodrigo Orsini Braga

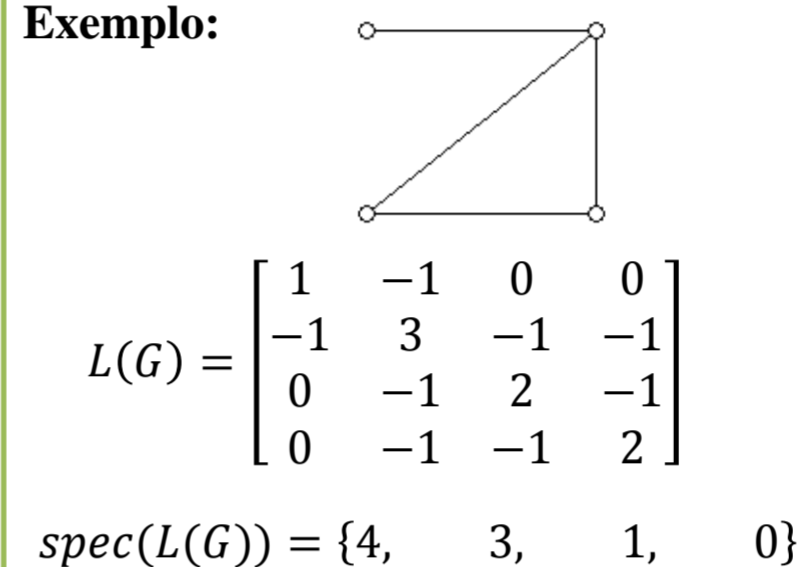
Um grafo é uma estrutura algébrica $G = G(V, E)$ onde V é um conjunto finito e não vazio de elementos denominados **vértices** e E é um subconjunto de pares não ordenados de V denominados **arestas**. A **Teoria Espectral de Grafos** procura estudar propriedades estruturais de grafos a partir de matrizes e de seus espectros.

- Um grafo é **conexo** se existe um caminho ligando quaisquer pares de vértices.
- A **distância** entre dois vértices é o número de arestas do menor caminho entre eles. O **diâmetro** de um grafo G , denotado por $diam(G)$, é o máximo das distâncias entre dois vértices de G .
- Uma **árvore** é um grafo conexo e sem ciclos.
- O **grau** de um vértice v , denotado por $d(v)$, é o número de arestas incidentes em v . Denotamos por $\Delta = \Delta(G)$ o máximo dos graus dos vértices de G , isto é, $\Delta = \max\{d(v) \mid v \in V\}$.

- Se $G = G(V, E)$ é um grafo com n vértices, a **Matriz de Adjacência** $A(G)$ é a matriz $n \times n$ cujas entradas são dadas por

$$a_{ij} = \begin{cases} 1, & \text{se } \{v_i, v_j\} \in E, \text{ para } v_i, v_j \in V; \\ 0, & \text{caso contrário.} \end{cases}$$
- Seja D a matriz diagonal dos graus dos vértices de G , isto é, $d_{ii} = d(v_i)$. A **Matriz Laplaciana** $L(G)$ é definida por $L(G) = D - A(G)$. Denotamos por $spec(L(G))$ o espectro de $L(G)$.

Exemplo:



- $L(G)$ é positiva semidefinida, ou seja, todos os seus autovalores são reais e não negativos. O menor autovalor é sempre zero e o segundo menor é positivo se e somente se G é conexo [1]. Tal autovalor, denotado por $a(G)$, é chamado **conectividade algébrica** de G .
- O maior autovalor de $L(G)$, denotado por $\mu_1(G)$, é simples (**Teorema de Perron-Frobenius** [1]) e é chamado **índice laplaciano**.

Árvores Químicas

Podemos relacionar o grau máximo de um grafo com o índice laplaciano através da desigualdade

$$\Delta + 1 \leq \mu_1 < \Delta + 2\sqrt{\Delta - 1}. \quad [5].$$

Árvores com $\Delta \leq 4$ (**árvores químicas**) modelam isômeros de **alcanos**, que são estruturas moleculares compostas por carbono e hidrogênio de maneira que os carbonos formem apenas ligações simples entre si. Se $\Delta = 2$, o grafo é o caminho P_n e sabe-se que $\mu_1(P_n) = 2 + 2\cos(\pi/n) < 4$. Se $4 \leq \mu_1 < 5$, $\Delta = 3$ (árvores terciárias). Logo, teremos árvores quaternárias ($\Delta = 4$) somente se $\mu_1 \geq 5$.

Através do software *Sagemath* [4] determinamos o espectro de todas as árvores (não isomorfas) entre 10 e 15 vértices com $\Delta = 3$ e $\Delta = 4$. A tabela abaixo apresenta as árvores químicas com os maiores índices laplacianos de acordo com o número de vértices.

As árvores químicas com $\Delta = 3$ e $\Delta = 4$ que possuem os menores índices laplacianos são da forma

$$\Delta = 3$$



$$\Delta = 4$$



Conclusões

Como vimos, a determinação do índice laplaciano permite detectar a presença de carbonos terciários e quaternários na estrutura molecular dos isômeros de alcanos. Ademais, podemos obter outras informações dessa estrutura a partir das tabelas acima e do gráfico abaixo.

As tabelas indicam que as árvores mais ramificadas são as que apresentam os maiores índices laplacianos.

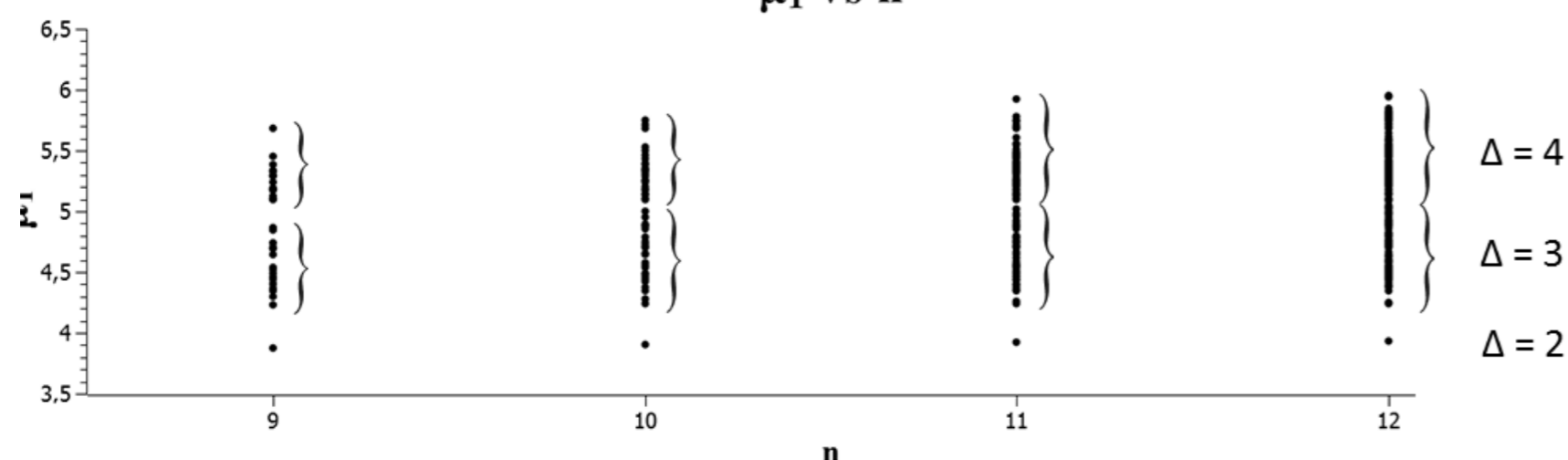
A partir do gráfico abaixo poderíamos erroneamente inferir que o maior índice das árvores terciárias é sempre menor do que o menor índice das árvores quaternárias. Esse resultado vale para $n \leq 12$, mas não é verdade para $n = 13$, visto que $\max(\mu_1 \mid \Delta = 3) = 5,108 > 5,098 = \min(\mu_1 \mid \Delta = 4)$. O mesmo ocorre para $n = 14$ e $n = 15$.

Árvores com os maiores índices laplacianos para $\Delta = 3$ e $\Delta = 4$

Número de vértices	$\mu_1 \mid \Delta = 3$	$\mu_1 \mid \Delta = 4$
10	5,0	5,74751962
11	5,02103472	5,92542292
12	5,08249675	5,95296997
13	5,10842955	6,00564283
14	5,17740968	6,13263749
15	5,18781133	6,15690387

O gráfico abaixo mostra a distribuição dos índices laplacianos das árvores químicas em função do número de vértices.

μ_1 vs n



Árvores e o RNA secundário

A conectividade algébrica $a(G)$ está relacionada com o diâmetro de G através da desigualdade

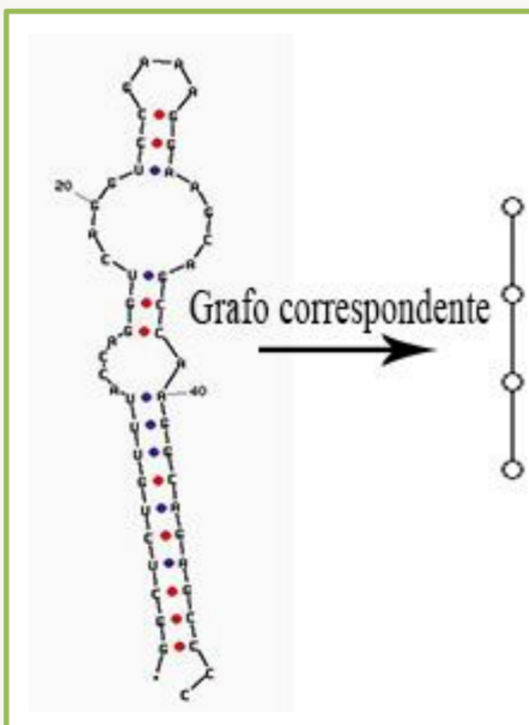
$$a(G) \geq \frac{4}{n \cdot diam(G)}, \quad [2]$$

onde n é o número de vértices de G . Além disso, para toda árvore T , $a(T) \leq 1$ e $a(T) = 1$ se e somente se T é a estrela S_n . [1]

O dobramento da molécula de RNA forma estruturas secundárias e terciárias. Uma estrutura secundária é a combinação de pares de bases nitrogenadas de maneira a formar laços, hastes e outras estruturas, enquanto que uma estrutura terciária é um arranjo tridimensional da interação entre os elementos da estrutura secundária.

Podemos modelar as estruturas secundárias através de grafos árvores com o objetivo de facilitar a compreensão da molécula de RNA. Diferentemente das árvores químicas, na estrutura secundária do RNA os vértices e as arestas de um grafo podem representar diferentes estruturas da molécula. A figura abaixo à esquerda mostra um exemplo de como representar uma estrutura secundária por uma árvore [3].

A determinação de $a(G)$ nos dá uma medida da compacidade do grafo G , isto é, uma conectividade algébrica grande indica um grafo mais compacto, enquanto uma conectividade algébrica pequena indica um grafo alongado. Com o auxílio do *Sagemath*, determinamos as conectividades algébricas das árvores com até 15 vértices. Na figura abaixo à direita apresentamos todas as árvores (não isomorfas) com 7 vértices e suas respectivas conectividades algébricas.



$a(G) = 0,19806226$	$a(G) = 0,2253771$	$a(G) = 0,26032269$
$a(G) = 0,26794919$	$a(G) = 0,29553219$	$a(G) = 0,32171965$
$a(G) = 0,38196601$	$a(G) = 0,38196601$	$a(G) = 0,39832087$
$a(G) = 0,4659298$	$a(G) = 1,0$	

Conclusões

Como as moléculas de RNA biologicamente correlacionadas tendem a ter estruturas secundárias semelhantes, determinar o espectro dos grafos associados a estas estruturas pode nos ajudar a identificar tais correlações. Em particular, a determinação de $a(G)$ permite relacionar as estruturas secundárias de acordo com a compacidade.

Referências

- [1] N. Abreu, R. Del-Vecchio, V. Trevisan, C. Vinagre, *Teoria Espectral de Grafos - Uma Introdução*, III^o Colóquio de Matemática da Região Sul, Florianópolis, SC, 2014.
- [2] B. Mohar, *Eigenvalues, Diameter, and Mean Distance in Graphs*, Graphs and Combinatorics 7 (1991) 53-64.
- [3] RNA as Graphs, Biomath, The New York University. Disponível em: <<http://www.biomath.nyu.edu/?q=rag/home>>. Acesso em 09 set. 2019.
- [4] Sagemath, The Sage Development Team. Disponível em: <<http://www.sagemath.org/pt/>>. Acesso em 29 ago. 2019.
- [5] D. Stevanović, *Bounding the largest eigenvalue of trees in terms of the largest vertex degree*, Linear Algebra and its Applications 360 (2003) 35-42.