

Topologia Molecular: uma interessante aplicação da Topologia

Heitor Souza Assis

Orientador: Diogo da Silva Machado

Departamento de Matemática - Centro de Ciências Exatas e Tecnológicas - UFV

heitor.assis@ufv.br - diogo.machado@ufv.br

1 Introdução

O uso da matemática em outros ramos do conhecimento como ferramenta ímpar para a solução de problemas é flagrante, não só nas páginas da história da ciência, como também nos mais recentes e atuais projetos, descobertas, aplicações e desenvolvimento da pesquisa científica. Física, Engenharia e até Biologia são alguns setores em que a Matemática é usada em suas órbitas de atividades, existindo na literatura diversos textos e livros que ressaltam e reforçam tal fato.

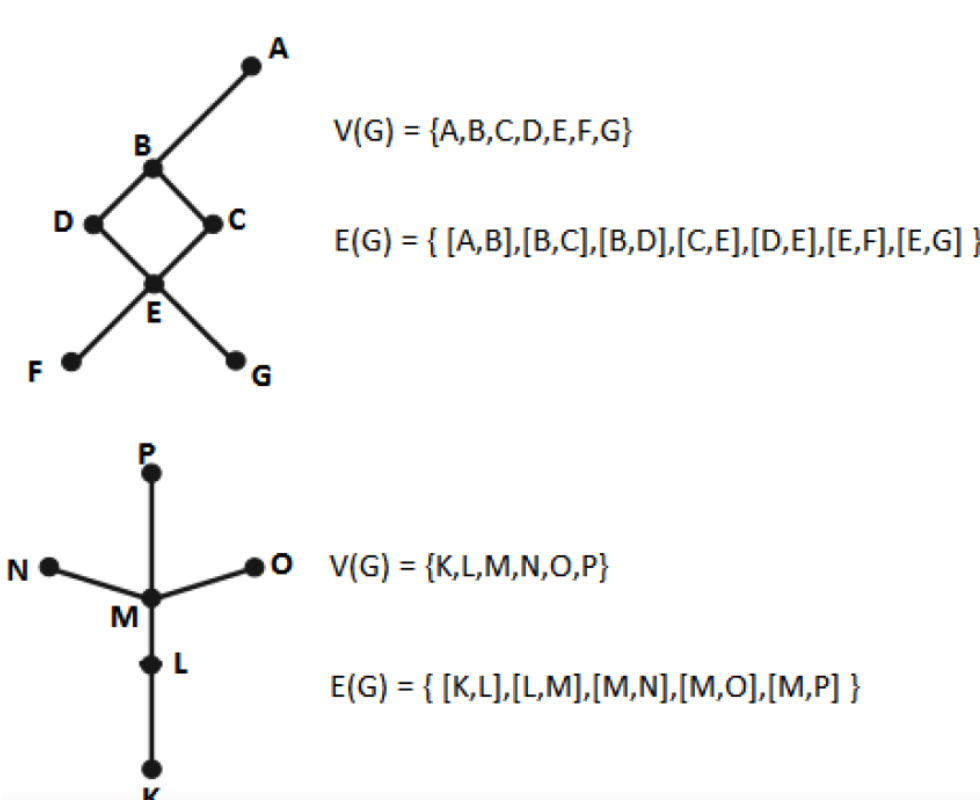
Mais recentemente, uma técnica com forte apelo à Topologia tem se mostrado muito eficiente na solução do problema de determinar as relações entre estruturas moleculares e propriedades químicas e, também, físicas e biológicas. Esta se chama *Topologia Molecular*, a qual é empregada, por exemplo, no estudos de propriedades físico-químicas de certos compostos orgânicos utilizando *índices* associados às propriedades de conectividade de suas moléculas.

Na indústria química e farmacêutica, onde se busca novos compostos com determinadas propriedades, a *Topologia Molecular* tem sido usada como instrumento para indicar de maneira prévia as potencialidades de uma molécula em manifestar as propriedades requeridas. Assim, o que é feito na prática é a predição das propriedades de uma molécula antes mesmo de sintetizá-la o que, de fato, constitui grande economia de recursos financeiros e agilidade quanto ao tempo gasto em todo o processo.

Neste trabalho, veremos como a Matemática atua junto à Topologia Molecular.

2 Teoria de Grafos

I. Grafo: um grafo G consiste de um conjunto finito $V(G)$ de *vértices* acrescido de um conjunto $E(G)$ de pares (não-ordenados) de vértices, chamados de *arestas*. Dada uma aresta $[u, v] \in E(G)$, os vértices u e v são ditos *adjacentes*.

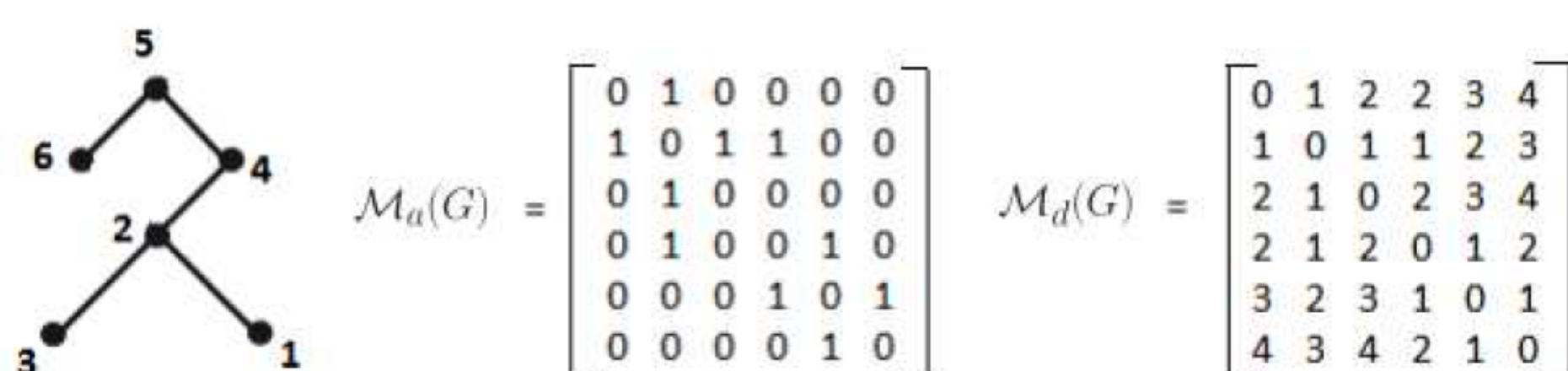


II. Grau: o grau $\eta(v)$ de um vértice v é o número de arestas que incidem nele.

III. Caminho: um *caminho* de um grafo G é uma seqüência alternada de vértices e arestas, começando e terminando com vértices. O número de arestas em um caminho é dito *comprimento*.

IV. Distância: a *distância* $d(u, v)$ entre os vértices u e v é o comprimento do menor caminho que os une.

V. Matrizes de Distâncias e Adjacências: seja G um grafo. Atribuindo-se a cada um dos seus vértices um número natural entre 1 e n (onde n é o número de vértices de G), sem repetição, a *Matriz de Distâncias* de G é a matriz $M_d(G) = [d(i, j)]_{n \times n}$. Já a *Matriz de Adjacências* de G , $M_a(G) = [a_{ij}]_{n \times n}$, é definida pela regra: $a_{ij} = 1$, se i e j são adjacentes e $a_{ij} = 0$, caso contrário.



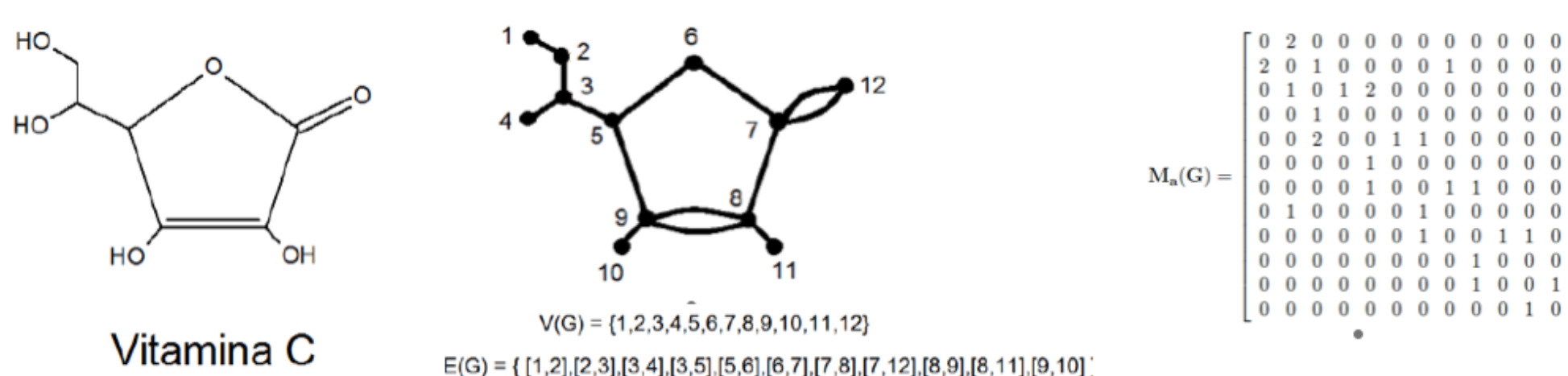
V. Automorfismo de Grafos: um automorfismo de um grafo G é uma bijeção $\varphi: V(G) \rightarrow V(G)$ que preserva adjacências, ou seja,

$$[\varphi(u), \varphi(v)] \in E(G) \iff [u, v] \in E(G).$$

O conjunto de todos os automorfismos de um grafo G , munido da operação composição de funções, constitui um grupo, o qual é chamado de *Grupo de Automorfismos* do grafo G .

3 Topologia Molecular

O grafo que nos mostra como os átomos estão ligados em um composto é chamado de *Grafo Molecular*.



Os Índices Topológicos são utilizados para analisar correlações entre a conectividade da molécula e suas propriedades físicas, químicas e biológicas. Dado um grafo G , alguns desses índices são:

I. Índice de Randić:
$$R_\chi(G) = \sum_{[u,v] \in E(G)} (\eta(u)\eta(v))^{-\frac{1}{2}}$$

II. Índices de Carga: A cada elemento químico Q , definimos a sua eletronegatividade relativa por

$$Re^{(-)}(Q) = 2, 2 \left(e^{(-)}(Q) - e^{(-)}(C) \right),$$

onde $e^{(-)}(Q)$ e $e^{(-)}(C)$ são, respectivamente, a eletronegatividade do elemento Q e do átomo de Carbono (C).

Dado um grafo molecular G , a sua *Matriz de Adjacências Modificada*, denotada por $M_a^V(G) = [\tilde{a}_{i,j}]$ é obtida substituindo-se os elementos da diagonal principal de $M_a(G)$ pela correspondente eletronegatividade do átomo, ou seja: $\tilde{a}_{i,i} = Re^{(-)}(i)$. A *Matriz Intermediária*, $M_{int}(G)$, é obtida pelo produto $M_{int}(G) = M_a^V(G) \times M_d^*(G)$, onde $M_d^*(G) = [d(i, j)^2]_{n \times n}$ é a matriz cujos elementos são os quadrados das distâncias. Finalmente, a *Matriz de Transferência de Carga* é definida como $M_t(G) = M_{int}(G) - (M_{int}(G))^T$.

III. Índice de Gálvez:
$$G_k^V = \sum_{u=1}^{n-1} \sum_{v=u+1}^n |C_{u,v}| d_k(u, v),$$

onde $C_{u,v}$ corresponde a um elemento arbitrário de $M_t(G)$ e o número $d_k(u, v)$ é definido pela seguinte regra: $d_k(u, v) = 0$ se $d(u, v) \neq k$ e $d_k(u, v) = 1$ se $d(u, v) = k$. Outro índice de carga, J_k^V , é definido por: $J_k^V = \frac{G_k^V}{n-1}$.

3.1 Exemplo e cálculos

Dado o grafo G do ácido ascórbico (vitamina C) e sua respectiva *matriz de adjacências*, conforme supracitado. Vamos calcular a seguir o índice de Randić:

Como temos a equação (II.), calculemos:

$$\begin{aligned} \bullet f(e_{1,2}) &= (deg_1 \cdot deg_2)^{-\frac{1}{2}} = (2 \cdot 4)^{-\frac{1}{2}} = \frac{1}{\sqrt{8}} = \frac{\sqrt{8}}{8} = \frac{2\sqrt{2}}{8} = \frac{\sqrt{2}}{4} \\ \bullet f(e_{2,3}) &= (deg_2 \cdot deg_3)^{-\frac{1}{2}} = (4 \cdot 4)^{-\frac{1}{2}} = \frac{1}{\sqrt{16}} = \frac{1}{4} \\ \bullet f(e_{2,8}) &= (deg_2 \cdot deg_8)^{-\frac{1}{2}} = (4 \cdot 2)^{-\frac{1}{2}} = \frac{1}{\sqrt{8}} = \frac{\sqrt{8}}{8} = \frac{2\sqrt{2}}{8} = \frac{\sqrt{2}}{4} \\ \bullet f(e_{3,4}) &= (deg_3 \cdot deg_4)^{-\frac{1}{2}} = (4 \cdot 1)^{-\frac{1}{2}} = \frac{1}{\sqrt{4}} = \frac{1}{2} \\ \bullet f(e_{3,5}) &= (deg_3 \cdot deg_5)^{-\frac{1}{2}} = (4 \cdot 4)^{-\frac{1}{2}} = \frac{1}{\sqrt{16}} = \frac{1}{4} \\ \bullet f(e_{5,6}) &= (deg_5 \cdot deg_6)^{-\frac{1}{2}} = (4 \cdot 1)^{-\frac{1}{2}} = \frac{1}{\sqrt{4}} = \frac{1}{2} \\ \bullet f(e_{5,7}) &= (deg_5 \cdot deg_7)^{-\frac{1}{2}} = (4 \cdot 3)^{-\frac{1}{2}} = \frac{1}{\sqrt{12}} = \frac{\sqrt{12}}{12} = \frac{4\sqrt{3}}{12} = \frac{\sqrt{3}}{3} \\ \bullet f(e_{7,8}) &= (deg_7 \cdot deg_8)^{-\frac{1}{2}} = (3 \cdot 2)^{-\frac{1}{2}} = \frac{1}{\sqrt{6}} = \frac{\sqrt{6}}{6} \\ \bullet f(e_{7,9}) &= (deg_7 \cdot deg_9)^{-\frac{1}{2}} = (3 \cdot 2)^{-\frac{1}{2}} = \frac{1}{\sqrt{6}} = \frac{\sqrt{6}}{6} \\ \bullet f(e_{9,10}) &= (deg_9 \cdot deg_{10})^{-\frac{1}{2}} = (3 \cdot 1)^{-\frac{1}{2}} = \frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{3} \\ \bullet f(e_{9,11}) &= (deg_9 \cdot deg_{11})^{-\frac{1}{2}} = (3 \cdot 2)^{-\frac{1}{2}} = \frac{1}{\sqrt{6}} = \frac{\sqrt{6}}{6} \\ \bullet f(e_{11,12}) &= (deg_{11} \cdot deg_{12})^{-\frac{1}{2}} = (2 \cdot 1)^{-\frac{1}{2}} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2} \end{aligned}$$

Como o índice de Randić é dado pela equação (I.), calculemos:

$$\chi_R(G) = 2 \cdot \frac{\sqrt{2}}{4} + 2 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} + 2 \cdot \frac{\sqrt{3}}{3} + 3 \cdot \frac{\sqrt{6}}{6} =$$

$$= 0.707106781 + 0.5 + 1 + 1.154700538 + 1.224744871 \Rightarrow \chi_R(G) = 4.58655219$$

4 Aplicações

Nesta seção, como aplicação da Topologia Molecular vamos apresentar os resultados do trabalho (Gálvez, [1]), onde os índices topológicos foram utilizados na determinação de uma equação que correlaciona o momento dipolo molecular de um conjunto de hidrocarbonetos.

Num primeiro momento, foram calculados os índices topológicos de um banco de moléculas do tipo hidrocarboneto. Os índices de cada molécula aparecem descritos na tabela seguinte (ver Gálvez,[1]).

Molecule	N	G ₁	G ₂	G ₃	G ₄	G ₅	J ₁	J ₂	J ₃	J ₄	J ₅
3	3	0.5000	0.0000	0.0000	0.0000	0.0000	0.2500	0.0000	0.0000	0.0000	0.0000
4	4	0.5000	0.2222	0.0000	0.0000	0.0000	0.1667	0.0714	0.0000	0.0000	0.0000
2M3	4	1.5000	0.0000	0.0000	0.0000	0.0000	0.5000	0.0000	0.0000	0.0000	0.0000
5	5	0.5000	0.2222	0.1250	0.0000	0.0000	0.1250	0.0556	0.0313	0.0000	0.0000
2M4	5	1.5000	0.4444	0.0000	0.0000	0.0000	0.3750	0.1111	0.0000	0.0000	0.0000
22MM3	5	3.0000	0.0000	0.0000	0.0000	0.0000	0.7500	0.0000	0.0000	0.0000	0.0000
6	6	0.5000	0.2222	0.1250	0.0800	0.0000	0.1000	0.0444	0.0250	0.0160	0.0000
3M5	6	1.5000	0.4444	0.2500	0.0000	0.0000	0.3000	0.0889	0.0500	0.0000	0.0000
3M5	6	1.5000	0.6667	0.1250	0.0000	0.0000	0.3000	0.1333	0.0250	0.0000	0.0000
22MM4	6	3.0000	0.667	0.0000	0.0000	0.0000	0.6000	0.1333	0.0000	0.0000	0.0000
23MM4	6	2.0000	0.8888	0.0000	0.0000	0.0000	0.4000	0.1778	0.0000	0.0000	0.0000
7	7	0.5000	0.2222	0.1250	0.0800	0.0556	0.0833	0.0370	0.0208	0.0133	0.0093
2M6	7	1.5000	0.4444	0.2500	0.1600	0.0000	0.2500	0.0714	0.0417	0.0267	0.0000
22MM4	7	3.0000	0.6667	0.0000	0.0000	0.0000	0.6000	0.1333	0.0000	0.0000	0.0000
23MM4	7	2.5000	0.4444	0.5000	0.0000	0.0000	0.4167	0.0714	0.0833	0.0000	0.0000
3E5	7	1.5000	0.6667	0.3750	0.0000	0.0000	0.2500	0.1111	0.0625	0.0000	0.0000
22MM5	7	3.0000	0.6667	0.3750	0.0000	0.0000	0.5000	0.1111	0.0625	0.0000	0.0000
23MM5	7	2.0000	1.1111	0.2500	0.0000	0.0000	0.3333	0.1852	0.0417	0.0000	0.0000
24MM5	7	2.5000	0.4444	0.5000	0.0000	0.0000	0.4167	0.0714	0.0833	0.0000	0.0000
33MM5	7	3.0000	1.1111	0.1250	0.0000	0.0000	0.5000	0.1852	0.0208	0.0000	0.0000
233MM4	7	3.5000	1.3333	0.0000	0.0000	0.0000	0.5833	0.2222	0.0000	0.0000	0.0000
8	8	0.5000	0.2222	0.1250	0.0800	0.0556	0.0714	0.0317	0.0179	0.0114	0.0079
3M7	8	1.5000	0.4444	0.2500	1.0000	0.1111	0.2143	0.0635	0.0357	0.0229	0.0159
4M7	8	1.5000	0.6667	0.3750	0.0800	0.0556	0.2143	0.0952	0.0536	0.0114	0.0079
3E8	8	1.5000	0.6667	0.3750	0.1600	0.0000	0.2143	0.0952	0.0536	0.0229	0.0100
22MM6	8	3.0000	0.6667	0.3750	0.3000	0.0000	0.4286	0.0952	0.0536	0.0343	0.0000
23MM6	8	2.0000	1.1111	0.2500	0.1600	0.0000	0.2857	0.1587	0.0893	0.0000	0.0000
24MM6	8	2.5000	0.6667	0.2500	0.2000	0.0000	0.3571	0.0952	0.0357	0.0457	0.0000
24E5	8	2.0000	1.1111	0.2500	0.0000	0.0000	0.2857	0.1587	0.0893	0.0000	0.0000
223MM5	8	3.5000	1.5556	0.3750	0.0000	0.0000	0.5000	0.2222	0.0536	0.0000	0.0000
234MM5	8	2.5000	1.3333	0.5000	0.0000	0.0000	0.3571	0.1905	0.0714	0.0000	0.0000
9	9	0.5000	0.2222	0.1250	0.0800	0.0556	0.0625	0.0278	0.0156	0.0100	0.0069
4M5	9	1.5000	0.6667	0.3750	0.1600	0.0556	0.1875	0.0823	0.0469	0.0250	0.0069
235MM6	9	3.0000	1.1111	0.5000	0.2000	0.0000	0.3750	0.1389	0.0625	0.0400	0.0000
24MM5E5	9	2.5000	1.3333	1.0000	0.0000	0.0000	0.3125	0.1667	0.1250	0.0000	0.0000
33E5	9	3.0000	1.3333	0.7500	0.0000	0.0000	0.3750	0.1667	0.0938	0.0000	0.0000
223MM5	9	4.0000	1.7778	0.7500	0.0000	0.0000	0.5000	0.2222	0.0938	0.0000	0.0000
224MM6	9	4.0000	0.8889	0.7500	0.2000	0.0000	0.5000	0.1111	0.0938	0.0030	0.0000
10	10	0.5000	0.2222	0.1250	0.0800	0.0556	0.0556	0.0247	0.0139	0.0089	0.0062

De posse dos índices topológicos, os autores aplicaram regressão linear e obtiveram equações que correlacionam o momento dipolo molecular μ , conforme apresentado na tabela a seguir (ver Gálvez,[1]):

Var.	Equation
1	$\mu = -0.200 + 2.535 \cdot J_2$
2	$\mu = -0.116 + 0.203 \cdot J_2 + 2.149 \cdot J_3$
3	$\mu = -0.078 - 0.129 \cdot G_2 + 0.569 \cdot G_3 + 2.791 \cdot J_2$
4	$\mu = -0.088 - 0.131 \cdot G_2 + 1.389 \cdot G_4 + 0.958 \cdot J_2 + 0.103 \cdot J_3$
5	$\mu = -0.053 - 0.139 \cdot G_2 + 0.315 \cdot J_1 + 0.858 \cdot J_2 + 1.840 \cdot J_3 + 7.212 \cdot J_4$
6	$\mu = -0.037 - 0.179 \cdot G_1 - 0.134 \cdot G_2 + 1.365 \cdot G_4 + 0.789 \cdot J_1 + 0.820 \cdot J_2 + 2.421 \cdot J_3$
7	$\mu = -0.185 - 0.043 \cdot N - 0.230 \cdot G_1 - 0.163 \cdot G_2 + 1.380 \cdot G_4 + 0.917 \cdot J_1 + 0.875 \cdot J_2 + 2.392 \cdot J_3$

Em seguida, elegendo uma das equações obtidas, foi possível aplicá-las a um grupo de hidrocarbonetos para a obtenção do momento dipolo de cada uma delas. A equação escolhida foi a equação (6) da tabela acima, considerando: $n = 30$; $r = 0.941$; $SD = 0.107$; $F = 28$; $\rho < 0.001$.

Na tabela a seguir (ver Gálvez,[1]) estão apresentados os valores obtidos pela equação. Também é possível observar a comparação dos valores obtidos com o valor conhecido na literatura. Observa-se que em geral que o valor residual é bem pequeno o que indica a precisão da equação obtida.

Compound	Observed	Calculated	Residual
1,1-dimethylbutyne	0,66	0,71	0,05
propene	0,36	0,4	-0,03
acetylene	0	-0,04	0,04
ethylbenzene	0,59	0,6	-0,02
toluene	0,36	0,42	-0,06
propane	0,08	0,07	0,01
butane	0,04	0,04	0
2,2-dimethylpropane	0	0,02	-0,02
cyclopropene	0,45	0,39	0,06
Propyne	0,78	0,73	0,05
propadiene	0	0,18	-0,18
cyclopropane	0	-0,04	0,04
cyclobutene	0,13	0,28	-0,15
1-butene	0,8	0,76	0,04
1,2-butadiene	0,4	0,45	-0,05
1,3-butadiene	0	0,22	-0,22
1-butene	0,34	0,38	-0,04
2-methylpropene	0,5	0,41	0,09
2-methylpropane	0,13	0,09	0,04
1,3-cyclopentadiene	0,42	0,22	0,2
cyclopentene	0,2	0,15	0,05
1-pentene	0,81	0,82	-0,01
2-methyl-1,3-butadiene	0,25	0,33	-0,08
2-methylbutane	0,13	0,02	0,11
benzene	0	0,12	-0,12
1,3-cyclohexadiene	0,44	0,31	0,13
hexene	0,83	0,72	0,11
1,3,5-cycloheptatriene	0,25	0,15	0,1
o-xylene	0,62	0,64	-0,02

5 Conclusões

Concluimos, então, que a *Topologia Molecular* se faz importante quando se pretende analisar propriedades moleculares e, à elas, relacionar outras propriedades, equações e índices topológicos que a descrevem de maneira a otimizar