应用拓扑指数预测化学性质

王化云 江元生

(吉林大学理论化学研究所,长春 130023)

摘要 拓扑指数能够提炼出分子中很有价值的结构信息,它已成功地用于模拟一系列的化学现象。 拓扑指数 的 理论和应用的研究已引起国际学术界的广泛关注。在国内,有关的研究刚刚起步。本文综述了常见拓扑指数的计算过程及其在结构/性质相关研究中的应用状况。

关键词 拓扑指数,图论,结构与性质相关

有机分子的拓扑指数应用研究是现代计算 化学,结构化学与量子化学相互交叉和结合的 产物,已成为当前信息化学的重要分支。

实验表明,分子的许多性质首先受制于分子的拓扑结构-原子的连接性^山。这些性质包括分子的宏观热力学性质,如沸点、分子体积、折光率等,也包括分子的微观特性¹²⁻⁴¹。除此之外,生物活性也可以从分子拓扑加以解释¹⁵¹。 这就启发化学家从一个侧面去总结结构与性质的定量相关关系。应用拓扑指数提炼分子信息要经过分子结构的图形化,矩阵化和数值化三个步骤。

一、分子结构的图形化

1. 图论与化学的早期渗透

数学家 Cayley 第一个将图论中"树"的概念用于饱和链烃的异构体计数中。后来Polya提出有名的计数定理,为异构体研究和寻找提供了理论工具。图用于表征分子,称为分子图^[6]。它的出现推动了分子结构表征的发展进程。

2. 图的基本概念和术语

图是描述给定分子拓扑性的一个数学结构,它与数学上用于记录和显示数据的图不同。图由一组点和一组连结点的边组成。图中顶点 表示原子,边表示分子中的化学键,每个分子都有相应的图。图1给出了丙烷和苯的分子图。利用图反映分子拓扑结构时,重点在于强调原子间的互连方式,它决定分子的最终构造¹⁰⁰。在

分子图中,分子的实际三维形状,连结原子的化学键的类别、长度、键角等都不重要。 关键是分子中有多少个原子,每个原子连接多少个其它原子,原子是连接成单一的直链还是带有支链,是否成环等。 化学图中通常省略氢原子,因为它对分析问题及最后结果不起原则作用。

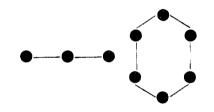


图 1 丙烷和苯的分子图

分子图是分子中原子键合的抽象,它提供了用图论来研究具体化学问题的基本数学模型。图论中所用到的许多术语与化学词汇可以做到——对应(见表 1)。

表 1 用于描述图论概念的相当的数学名称和化学名称

数学名称	化学名称	数学名称	化学名称
顶点	原子	圏数	环数
边	化学键	本征多 项式	特征多项式
化学图	结 构式	邻接矩阵	拓扑矩阵
树图	无环分子	本征值	能级
二分图	交替分子	零本征值	非键能级
顶点度数	原子的化合 价	正本征值	成键能级
# 顶点链	n-多烯	负本征值	反键能级
# 頭点回路	n-轮烯	谱理论	Hückel 理论

二、拓扑结构的矩阵描述

数学家 Sylvester 将分子图转化为矩阵的 形式来表示分子的拓扑结构。矩阵是以数字的 形式提供分子拓扑性的科学表达。这里仅讨论 在表征化学结构时最有用的两种矩阵。

1. 邻接矩阵



图 2 苯的分子图及其邻接矩阵

矩阵元为:

苯的分子图、原子编号及邻接矩阵。

2. 距离矩阵

Wiener^{t2} 在研究分子的加和性质时,以隐含的形式使用距离矩阵。对任一有n个顶点的图构成一个 $n \times n$ 阶矩阵D,矩阵元 d_{ij} 为:

 $d_{ij} = \begin{cases} d & d$ 为连接顶点i 和j 的最小边数 ∞ 当顶点i 和j 不连通 苯的距离矩阵见图 3, D 也是对称的。

	1	0	2 1 0 1 2 3	3 2	4	5 2	6 1	
	2	1	0	1	2	3	2	
D =	3	2	1	0	1	2	3	
	4	3	2	1	0	1	2	
	5	2	3	2	1	0	1	
	6	1	2	3	2	1	0]	

图 3 苯的距离矩阵

三、分子结构的数值化

化学结构本身是抽象的,难以定量抽述,而它们的各种物理化学性质则表现为一定的数值。抽象的结构与用数值表达的性质间无法进行定量的关联。因此,对结构的数值表征的研究十分必要。通过对分子图的矩阵实施某种数

字运算而获得的拓扑指数,建立了结构和一个 无量纲数据间的——对应,实现了结构的数字 形式的表达。

1936年,邻接矩阵 4开始被引入化学¹⁷。任

ſ1 当顶点;与顶点;相邻

(0 当顶点;与顶点;相隔

一有 n 个顶点的图,可构成一个 n × n 阶矩阵。

即当图中有一边相连,则在矩阵的相应位置记

自 1947 年第一个能表征分子的"支链性"的拓扑指数 Wiener 指数 W^[2] 提出后,又出现了 120 多个拓扑指数,但只有很小一部分与分子的性质有较好的相关性能^[8]。本文就较为常用的几种拓扑指数的计算及应用进行总结。象大多数拓扑指数一样,它们都是建立在距离矩阵或邻接矩阵及其不变量的基础上。

1. 距离矩阵指数

(1) Wiener 指数 W Wiener^[2] 首先将分子结构绘成隐氢图(化学图),由图得到表示分子拓扑结构的距离矩阵,加和此矩阵的上三角矩阵元便得到 Wiener 指数 W。利用图 3, 苯的 Wiener 指数为 27。

$$W = 1 + 2 + 3 + 2 + 1 + 1 + 2 + 3 + 2 + 1 + 2 + 3 + 1 + 2 + 1 = 27$$

Wiener 指数有较高的简并度。 对于有较多碳原子的分子来说,W通常较大。Mekengan^[9] 提出了一个改进方法,使得W对无穷大体系也能给出一个有限值,只要这类系统是由许多相同的有限基本单位构成。另外,采用 Hückel 加

权图¹¹⁰³, Wiener 指数还可用于含杂原子体系。

Wiener 提出W指数后,研究人员作了大量应用工作,发现该指数与某些类型烃分子的性质如沸点^[2]、临界常数、沾滞性、表面张力^[11]、色谱保留时间^[12]具有良好的相关性。 Bonchev^[10]利用 Wiener 指数与单取代和双取代苯系列的色谱保留时间相关,得到的回归方程为;

$$RI = (244 \pm 4) W^{(0.297 \pm 0.003)}$$

相关系数超过 0.999。 近几年来,Mekenyan^{tul} 将此指数用于聚合物的熔点、沸点的预测,这些 聚合物包括聚四氟乙烯、聚己酰胺、聚乙烯对苯 二酸等,也取得满意的结果。

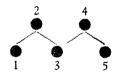
(2) Balaban 建立的指数 1983 年 Balaban 建立了两个以距离矩阵为 基 础 的 拓 扑 指数 如 5 ,平均距离矩阵指数 D 与平均距离总和连通性指数 J,分别定义为:

$$D = \left[\sum_{d_i=1}^{d_i (\max x)} g_i d_i^2 / \sum_{d_i=1}^{d_i (\max x)} g_i\right]^{\frac{1}{4}}$$
 (1)

$$J = q/u + 1 \sum_{s_i, i = j} (s_i s_j)^{-\frac{1}{2}}$$
 (2)

其中 d_i 为边长, q_i 为边长是 d_i 的边出现的次数, q 为邻接边的数目, u = q = n + 1 是环数, q_i 为顶点 i 到所有顶点拓扑距离的和。

对正戊烷,分子图为



距离矩阵

$$D = \sqrt{\frac{4 \times 1^2 + 3 \times 2^2 + 2 \times 3^2 + 1 \times 4^2}{4 + 3 + 2 + 1}}$$
= 2.2361

$$J = 4 \times [(10 \times 7)^{-\frac{1}{2}} + (7 \times 6)^{-\frac{1}{2}} + (6 \times 7)^{-\frac{1}{2}} + (7 \times 10)^{-\frac{1}{2}}]$$

$$= 2.1908$$

D具有较高的简并度。J是目前提出的拓扑指数中简并度最低的一个指数。对庚烷和辛烷异构系列的马达法辛烷值进行单变量相关,采用D时,相关系数分别为 -0.949 和-0.951.利用J时,相关系数分别为 0.914 和 0.932^[13]。Hansen^[14] 研究了J和分子图中独立环数之积与 93 种烃分子的煤烟量间的线 性 关系,相关系数 0.974。

2. 邻接矩阵指数

建立在邻接矩阵基础上的最常用的拓扑指数是由Randic^{fd}提出的,经量子化学家Kier^{fsf}扩展的由一组指数组成的分子联接性指数。此指数可用于处理含环,多重键和含杂原子的分子。自 1975 年以来,Kier 和 Hall 单独或合作共发表了近 40 篇关于分子联接性指数在定量结构与活性相关方面的研究论文^{fl6l},并有两本专著问世^{fl5l}。此指数是应用最广泛的一个拓扑指数。

Randic 指数 X 是分子联接性指数中 最 简单的指数,定义为:

$$\chi = \sum (P_i P_i)^{-\frac{1}{2}} \tag{3}$$

 P_i , P_i 为相邻点 i 和 i 的度数。这里仅就 χ 在物化性质方面具有代表性的应用简单介绍于表 1 中。

实验表明^[8], Randic 指数能较好地反映取 决于分子体积的性质,而与分子形状决定的性 质相关较差。分子联接性指数在生物学,环境 科学等领域均有很好的应用^[15,16]。

3. 多项式指数

(1) Hosoya 指数 第一个建立在本征 多项式基础上的拓扑指数是 1971 年 Hosoya^[3]提出的 Z 指数。

性 质	回归方程	回归点数	体 系	相关系数
沸点	$t_B = 57.85 \chi - 97.90$	51	烷烃	0.985
デース 辛醇/水分配系数	$\lg P = 1.48 - 0.950\%$	138	不同类有机物	0.986
水溶性	lgS = 6.702 - 2.666X	51	链醇	0.987
色谱保留时间	$T_{\rm R} = 482.12 \chi - 559.60$	51	粗砂	0.991
土壤吸附	$lgk = 0.550\chi - 0.450$	37	不同类有机物	0.973

表 2 Randie 指数与不同物化性质相关结果举例[8]

$$Z = \sum_{K=0}^{m} P(G,K) \tag{4}$$

其中 P(G,K) 为化学图 G中,K个边以不相邻状态出现的次数。当化学图为树形图时, Z可定义为本征多项式系数绝对值之 和。 Hosoya 指数象其它指数一样可用于多种物性数据的研究^[8],包括碳氢化合物和取代烷烃的沸点,烷烃熵的绝对值,此外,它还可以用于不饱和碳氢化合物 π 电子结构的研究。实验还表明,Hosoya 指数随其取代位置呈明显的交替变化(见图4),因此,此指数的显著特点是能反映取代烷烃沸点随取代位置变化的规律。

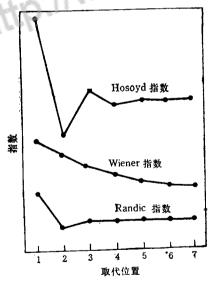


图 4 单取代十三烷的 Randic 指数, Wiener 指数, Hosowa 指数随其取代位置的变化

(2) a_N 指数 a_N 是由分子图所确定的本征多项式的常数项。在 HMO 理论中, a_N 构成一组环共轭异构体的相对稳定性的一种 判据。但是,对链状体系 a_N 只取两种数值,0 或 1。

为使 an 值能够反映链状图的支化状况¹¹⁷,针对饱和链烃相邻碳原子上的 sp³ 轨道,并考虑了它们间的轴向定域成键作用(b)及一种非轴向次级微扰作用(c),提出了饱和烃碳骨架衍



生图模型:每个 sp3 轨道对应一个顶点,紧邻碳

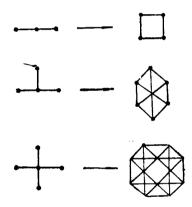


图 5 基本碳骨架和衍生图

原子的 sp³ 轨道的两种定域轨道作用对应两种不同权的边。由此每一个二,三,四度顶点分别衍生一个四、六、八元环。在图 5 中,粗线代表轴向定域作用,细线代表非轴向次级作用。复杂的树图是由这些基本片断构成的,它们的衍生图则由上述四、六、八元环组成。图 4 显示了2,4,4-三甲基己烷的衍生图。

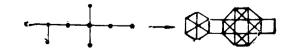
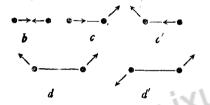


图 6 2,4,4-三甲基己烷的衍生图

定义衍生图本征多项式的常数 项 为 6、 相 数。此指数对 C,以下饱和烃的沸点、密度、生成焓、原子生成热、折光率和色谱指数等相关,得到了形式统一的计算式,

 $R = A \lg a_N + B n_3 + C n_4 + D N + E (5)$ R 为某一热力学量,A , B , C , D , E 相对每个 R 为一组常数, n_3 , n_4 代表顶点支化数,N 为碳原子数,对各热力学量的计算值和实验值相关系数均在 0.98 以上。

(3) 广义 a_N 指数 GAI 在 a_N 指数的基础上^[18],提出了适合一般有机化合物的广义 a_N 指数,在计算时考虑了轨道间轴向定域作用(b) 和轨道的两种非轴向次级微扰作用(c,d)。



这样,对每个分子可以得到相邻原子轨道作用 图 GOILA 和相邻原子间轨道作用矩阵 MOI LA、以丁烷为例,图 7显示了丁烷 GOILA 和 MOILA 的获得。

对含有n个轨道的 GOILA,将有一个 $n \times n$ 阶的 MOILA 相对应。在 MOILA 中,忽略非相邻原子的轨道作用,矩阵的相应矩阵元位

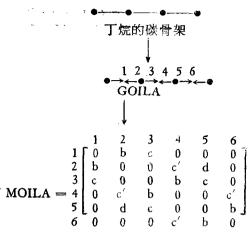


图7 丁烷的碳骨架,GOILA 和 MOILA

置为零元素;对相邻原子轨道,根据轨道作用形式不同,利用重迭积分的相对值估算轨道相互作用能的大小,得到相应矩阵元b, c, c', d 或 d'。 显然此矩阵是关于对角线对称的,对角线上的元素为相应轨道的相对能量。定义 MOILA本征多项式常数项的绝对值为广义 a_N 指数GAI。

GAI 的提出不仅解决了以往拓扑指数不能区分顺反异构体的困难,而且能定量地反映顺反异构体的密度、折光率、沸点和分子折射度(分别见表 2 和表 3)。将 GAI 应用于庚烷、辛烷系列马达法辛烷值,有机磷萃取剂的密度、折光率、分子折射度、纸上层析比移值、泳动值、色

化合物	GAI	ВР	D D	q10	n ž	d25	M R
顺-5-癸烯	1.50028	169.5	1.4252	0.7445	1.4230	0.7406	48.22
反-5-癸烯	1.47495	170.2	1.4235	.7401	1.4213	.7363	48.34
顺-4-辛烯	1.68656	121.7	1.4136	.7205	1.4113	.7163	38.92
反-4-辛烯	1.65613	121.4	1.4116	.7147	1.4091	.7104	39.05
顺-3-辛烯	1.68611	122.3	1.4125	.7189	1-4101	-7147	38.91
反-3-辛烯	1.65627	122.4	1.4124	.7163	1.4100	.7121	39.04
顺-2-辛烯	1.69367	124.6	1.4139	-	_	_	
反-2-辛烯	1.66414	123.4	1.4128	_	_		
顺-3-己烯	1.89278	66.85	1.3934	.6796	1.3908	.6749	29.61
反-3-己烯	1.85864	67.5	1.3938	6779	1.3912	.6730	29.72
顺-2-戊烯	2.01327		1.3828	.6554	_	_	_
反-2-戊烯	1.97818	_	1.3792	.6475	_		-
颠-2- j 烯	2.14145	3.7	_	-	 -	_	
反-2-丁烯	2.10539	0.88		_	_	_	_

表 3 几种烯烃顺反异构体的 GAI 和物理性质[20]

表 4 烯烃的 GAI 与物理性质的相关分析结果[20]

物理性质	п	а	b	R	s
B.P.	12	-264.21	563.98	-0.997	4.671
n 20	12	-0.09	1.56	0.981	0.003
d20	10	-0.18	1.01	-0.981	0.007
n 2.5	8	-0.08	1.54	-0.981	0.003
d.29	8	-0.16	0.98	0.975	0.006
MR_{D}	8	-47.46	118.56	-0.994	0.827

谱保留指数、萃取分配比等多种性质的研究,发现它们之间具有良好的单变量相关关系[18-20]。

拓扑指数在化学上应用的价值只是在最近 几年才得以广泛承认,在其它方面的应用还有 待进一步研究。如从分子拓扑结构出发寻找有 效的描述分子体系的参量,改善物质的性能,以 及进行可靠的常规预测都具有重要的研究价 值。

参 考 文 献

- [2] H. Wiener, J. Am. Chem. Soc., 69, 17(1947).
- [3] H. Hosoya, K. Kawasaki, Bull. Chem. Soc. Jap., 45, 3415(1972).
- [4] M. Randic, J. Am. Chem Soc., 97, 6609(1975).
- [5] L. B. Kier, J Med. Chem., 18, 1272(1975).
- [6] D. H. Rouvray, J. Chem. Educ., 768(1975).
- [7] A. T. Balaban, "Chem. Appl. Graph Theory". 1976.
- [8] D. H. Rouvray, J. Comput. Chem., 8, 470(1987).
- [9] D. Mekenyan, Eur. Polym. J., 19, 1135(1983).
- [10] R. S. Lall, V. K. Srivastava, Math. Chem., 106, 602 (1982).
- [11] D. H. Rouvray, South Afr. J. Sci., 72, 74(1976).
- [12] D. Papazova, J. Chromatogr., 188, 297(1980).
- [13] A. T. Balaban, Pure Appl. Chem., 55, 199(1983).
- [14] M. P. Hanson, D. H. Rouvray, J. Phys. Chem., 91, 2981(1987).
- [15] L. B. Kier, L. H. Hall, "Molecular Connectivity in Chemistry and Drug Research", New York. Academic Press, 1976: "Molecular Connectivity Structure Activity Analysis", New York, Wiley, 1986.
- [16] P. J. Hansen, P. C. Jurs, J. Chem. Educ., 574(1988).
- [17] 杨家安、江元生,化学学报,41,884(1983)。
- [18] 王化云、吕天雄、许禄等,化学学报,48,1159(1990)。
- [19] 王化云、许禄、苏铧,化学学报,49,424(1991)。
- [20] 王化云、许禄、苏锦,化学学报,50,22(1992)。

[1] D. H. Rouvray, Chem. Britain, 10, 11(1974).

· 中国化学会通讯 ·

第 24 届国际化学奥林匹克竞赛在美国举行

中国化学奥林匹克队,由领队程铁明教授、副领队 段连运副教授,观察员范景玮高级工程师及四名中学 生,应第 24 届国际化学奥林匹克组委会邀请,于 1992 年 7 月 11 日至 22 日赴美国匹兹堡、华盛顿参加了以 "地球化学日"为题的化学竞赛活动。

在美国国家科学院的礼堂里,举行了隆重地颁奖仪式,我国选手接受了三名诺贝尔奖金获得者授予的奖牌和奖状。中国队员郑页(北京四中)获金牌第一名,沈珺(女,上海华东师大二附中)获金牌第六名,汤志浩(安徽合肥六中)获金牌第九名,林熹晨(山东省实验中学)获银牌第一名。当他们走上领奖合时,全场掌声雷动。在欢快的乐曲声中全体起立,各国师生向中国队员致以热烈地祝贺。本届竞赛由33个国家的131名中学生选手参加。中国选手在激烈地角逐中,沉着冷

静、发挥自如,顺利地完成了5个小时的化学实验考试和4.5小时的化学理论答卷,以团体总分第一、金牌得数第一(金牌总数设16枚)的优秀成绩,为祖国赢得了荣誉。台湾队首次参加化学竞赛,喜获一金、一银、一铜的好成绩·海峡两岸师生相互勉励,共同祝贺。

自 1987 年以来我国共组织参赛六届,派出参赛队员 24 名、共获奖牌 24 枚,其中金牌 16 枚、银牌 5 枚和铜牌 3 枚。国际化学奥林匹克竞赛活动,为早期发现和培养人才、为智力超群、热爱化学的中学生提供了一个展示才华的舞台。我国各省、市近年来已从几千人发展到几十万人参加到化学竞赛活动中来,同时对促进化学教育方法、教育内容、教育结构和教育体制的改革起到了促进作用。

(范景玮)



论文写作,论文降重, 论文格式排版,论文发表, 专业硕博团队,十年论文服务经验



SCI期刊发表,论文润色, 英文翻译,提供全流程发表支持 全程美籍资深编辑顾问贴心服务

免费论文查重: http://free.paperyy.com

3亿免费文献下载: http://www.ixueshu.com

超值论文自动降重: http://www.paperyy.com/reduce_repetition

PPT免费模版下载: http://ppt.ixueshu.com

阅读此文的还阅读了:

- 1. 用ARMA模型预测深沪股市
- 2. 浅谈增塑剂的应用拓展及发展趋势
- 3. 后过渡金属催化剂催化乙烯制备超支化聚乙烯新进展
- 4. 例谈同位素标记法在生物学中的应用
- 5. 2014年中考化学大预测 (1)
- 6. 世界制药生产中应对药用辅料变异性挑战的方法开发应用新进展
- 7. RNA简约提取和无标记转化技术创新及应用
- 8. 用ARMA模型预测深沪股市
- 9. 乙二醇单甲醚的应用
- 10. 地层无机结垢预测技术研究与应用
- 11. 不动点定理及其应用
- 12. 指数递减法在凝析油气藏中的应用
- 13. EHF疫区带病毒鼠指数的时间分布理论拟合与趋势预测
- 14. R.Kooistra不等式的新推广
- 15. 高吸水树脂的新用途:灭火与防火
- 16. 拓扑递减方法预测周期注水开发动态

- 17. Characteristics of Hercynite and Its Application in Refractories
- 18. 运用时间序列对上证综合指数进行预测分析
- 19. 基于SCADA/EMS的系统拓扑"五防"技术应用
- 20. 成长性不如主板创业板风险显现
- 21. 拓扑指数与对饱和脂肪酯物理化学性质的预测
- 22. 甲壳质化学及应用
- 23. 不动点定理及其应用
- 24. 异麦芽酮糖醇的性质和用途
- 25. 叔碳酸在农药行业的市场发展现状及应用前景
- 26. 加入WTO后四川价格指数的预测
- 27. 新型表面活性剂双季铵盐的特性及应用
- 28. 拓扑指数与对饱和脂肪酯物理化学性质的预测
- 29. 应用拓扑指数预测化学性质
- 30. 运用SAS软件系统对上证综合指数的预测分析
- 31. 纳米技术在木制品工业中的新应用
- 32. BL67在催化剂厂PCS7系统中的应用
- 33. 利用水驱特征曲线确定活塞式驱程度指数的方法
- 34. 硅藻土在涂料油漆等行业中的应用
- 35. 一个不等式的指数推广及应用
- 36. 金属活动性顺序的应用
- 37. 叔碳酸在涂料市场应用前景看好
- 38. 拓扑和变分方法及其在非线性边值问题的应用
- 39. 平均值不等式的推广及应用
- 40. 六氟化硫断路器的应用及六氟化硫环保问题
- 41. 稀土元素的化学性质与稀土应用
- 42. 注水井效益动态评价图
- 43. 指数平滑法及其应用
- 44. 稀土在花木生产上的应用
- 45. 金属活动性顺序在中考中的应用
- 46. 例谈建构思维与拓扑思维在数学教学中的应用
- 47. 新型缓蚀杀菌剂双季铵盐在油田中的应用
- 48. 应用拓扑指数预测LAS的EACN值
- 49. 纳米材料的物理、化学性质及其应用
- 50. 甘油硬脂酸脂的化学性质和应用