

接続行列を用いたRandicの指数の再定式化

大 前 貴 之

Reexamination of the Connectivity Index by Using the Incidence Matrix

Takayuki OHMAE

Summary

In order to show the utility of the incidence matrix in the chemical graph theory, we have studied the connectivity index which was defined by Randic. A simple formula of the connectivity index has been found by using the incidence matrix. This formula is suitable to computer calculation.

1. 緒 言

分子の幾何学的性質を抽出した分子グラフ (G) を用いて化合物の化学的性質を研究する際、しばしばGの隣接行列 (the adjacency matrix, A(G)) が用いられる¹⁾。しかしながら、グラフ理論において良く知られているように、Gは各頂点と各辺の接続関係を記述した接続行列 (the incidence matrix, C(G)) を用いてもA(G)を用いた場合と同様に一意的に記述することができる²⁾。式(1), (2)に、Fig.1に図示したイソペンタンのGのA(G)とC(G)を示した。

$$A(G) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad (1)$$

$$C(G) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (2)$$

ここで、1, 2, 3などの数字はGの頂点を表

し、1, 2, 3などの数字はGの辺を示す。

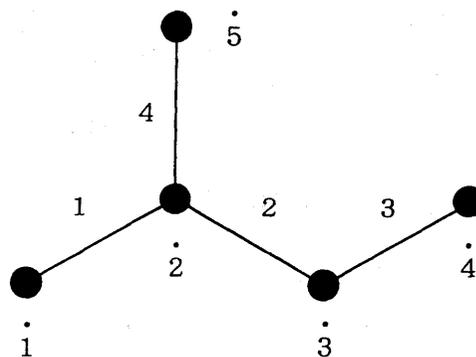


Fig.1 The molecular graph of isopentane.

1, 2, 3: Bonds; 1, 2, 3: Atoms.

Fig.1に示したようなGを用いて分子の幾何学的構造と物性との間の関係を議論するために、種々の指数が今までに提案されてきた^{3), 4)}。中でも分子の枝分れ構造、すなわちGにおける辺の相対的位置関係を重視した指数は種々の定量的構造活性相関の検討に利用され、その有用性が広く知られている⁴⁾。たとえば、Randicによって提案された指数 $\chi(G)$ は、各頂点に接続する辺の数、すなわち各頂点iの次数 d_i を用いて

$$\chi(G) \equiv \sum_{i_1, i_2, \dots, i_n} [d_{i_1} d_{i_2} \dots d_{i_n}]^{\frac{1}{2}} \quad (3)$$

で定義され^{4), 5)}、定量的構造活性相関における有

用性が詳しく検討されてきた^{4), 6)}。ただし、式(3)で p_n は長さ n の経路についてだけ和を実行することを意味する。このような辺の相対的位置関係を記述するような指数を求めるには、従来から広く利用されてきた $A(G)$ を用いるよりも辺と頂点の接続関係をあらわに表現した $C(G)$ を用いるほうが有利であり、計算機上で実際に指数を求める際にも単純な手続きによって計算が可能になることが期待できる。

そこで本報では、 $C(G)$ を用いてRandicの指標の中でも特に取り上げられることの多い0次、1次、2次の指標を記述することを試みる。

2. 接続行列を用いたRandicの指数の再検討

式(3)から明らかのように、Randicの指数は各次数に対応した長さの経路に属す頂点の次数によって定義される。したがって、 $C(G)$ を用いて任意の長さの経路上にある頂点の次数を求めることができれば、Randicの指数がただちに得られることになる。以下では、0次、1次、2次のそれぞれの場合に分けて、 $C(G)$ を用いたRandicの指数の表現法を検討する。

2.1 0次のRandic指数

式(3)から0次のRandicの指数 ${}^0\chi(G)$ は

$${}^0\chi(G) \equiv \sum_{i=1}^{p_0} \frac{1}{\sqrt{d_i}} \quad (4)$$

で定義される。ここで、 $C(G)$ に $C^T(G)$ を右側から乗じて得られる行列の対角項が各頂点の次数を与えるという良く知られた性質²⁾、すなわち

$$d_i = [C(G)C^T(G)]_{ii} \quad (5)$$

を思い出せば、式(4)で定義された ${}^0\chi(G)$ は式(5)の関係をjを用いて、ただちに

$${}^0\chi(G) \equiv \sum_{i=1} \frac{1}{\sqrt{[C(G)C^T(G)]_{ii}}} \quad (6)$$

のように $C(G)$ を用いて表現することができる。

例として、Fig. 1に示したイソペンタンを考え

ると、式(2)から

$$C(G)C^T(G) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \\ = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \quad (7)$$

が得られ、式(6)、(7)からただちにイソペンタンの ${}^0\chi(G)$ の値が

$${}^0\chi(G) = \frac{1}{\sqrt{1}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{1}} + \frac{1}{\sqrt{1}} \\ \equiv 4.2845 \quad (8)$$

のように、 $C(G)$ に単純な操作を加えることで正しく再現され

る⁴⁾。ここで、分子グラフの各頂点の次数を対角項の要素とし他の要素は0であるような、いわゆる次数行列 (the degree matrix, $D(G)$) を用いると、式(7)が

$$C(G)C^T(G) = A(G) + D(G) \quad (9)$$

と書くことができ、この関係がイソペンタンだけでなく一般的に成立する関係であることに注意しておこう²⁾。

${}^0\chi(G)$ を求める場合に限れば式(6)のような $C(G)$ による計算法を用いなくても、 $A(G)$ の各行または各列の和が頂点 i の次数 d_i を与えるという性質を用いて、簡単に計算することができる。しかしながら、 ${}^1\chi(G)$ や ${}^2\chi(G)$ を求める場合には、長さ1、2のあらゆる経路に属す頂点の次数の積を求めることが必要になり、このような場合には以下に示すように、 $D(G)$ の性質を利用することによって単純で系統的な計算法を考案することが可能になる。

2.2 1次のRandicの指数

式(3)から1次のRandicの指数 ${}^1\chi(G)$ は

$${}^1\chi(G) \equiv \sum_{(i,j)}^{p_1} \frac{1}{\sqrt{d_i d_j}} \quad (10)$$

で定義される。ただし、式(10)で p_1 はグラフの長さ1のあらゆる経路、すなわちグラフのすべての辺の両端の頂点 (i,j) についてだけ和を実行することを意味する。ここで、式(9)の関係を用いて $C(G)$ から簡単に求められる $D(G)$ を $C(G)$ の左側から掛けて得られる行列の列が各辺の両端になる頂点でだけ値を持ち、かつその値が頂点の次数に等しいという性質を持つこと、すなわち

$$\begin{aligned} \{D(G)C(G)\}_{ik} &= d_{ii}c_{ik} \\ &= d_i c_{ik} \end{aligned} \quad (11)$$

の関係に注意すると、 ${}^1\chi(G)$ が

$${}^1\chi(G) = \sum_j \frac{1}{\sqrt{\prod_i \{D(G)C(G)\}_{ij}}} \quad (12)$$

のように $C(G)$ を用いて表現できる。ただし、式(11)で i と k はそれぞれ頂点と辺を示すことと、式(12)の求積記号に*印をつけたのは0に関する積の規則を以下のように変更して実行することを示していることに注意しておこう。

$$0 \times 0 = 0 \quad (13a)$$

$$0 \times a = a \times 0 = a \quad (13b)$$

言うまでもないことではあるが、式(12)を計算機上で計算する際には式(13a)、(13b)のような0に関する積の規則を変更する必要はなく、行列要素が0のときには積を実行しないようにプログラムすればよい。

再びFig.1のイソペンタンで具体例を示すと、式(1)、(2)、(9)から

$$D(G) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (14)$$

が得られ、さらに式(2)、(14)から

$$D(G)C(G) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 3 & 0 & 3 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (15)$$

が求められるので、最終的に式(12)、(15)からイソペンタンの ${}^1\chi(G)$ の値が

$$\begin{aligned} {}^1\chi(G) &= \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} \\ &\cong 2.2701 \end{aligned} \quad (16)$$

のように正しく再現される⁴⁾。ここで示した方法に従えば、式(10)の定義を用いて計算する場合と異なり、どの頂点がどの辺に属しているかというような条件を考慮することなく $C(G)$ の単純な操作から ${}^1\chi(G)$ の値を求めることができることに注意しておこう。

2.3 2次のRandicの指数

式(3)からわかるように、2次のRandic指数 ${}^2\chi(G)$ は互いに隣接する2本の辺に属す3個の頂点の次数を用いて

$${}^2\chi(G) \equiv \sum_{(i,j,k)}^{p_2} \frac{1}{\sqrt{d_i d_j d_k}} \quad (17)$$

で定義される。ただし式(17)で i, j, k はグラフの頂点を表し、 p_2 は長さ2のすべての経路について和を実行することを意味する。ここで上述の和の条件をグラフ論的な見地から見ると、 ${}^2\chi(G)$ を計算するには3個の頂点を切り取る独立カットセット行列 $B(G)$ 求め

$${}^2\chi(G) = \sum_j \frac{1}{\sqrt{\prod_i \{D(G)B(G)\}_{ij}}} \quad (18)$$

とすればよいことがわかる。ただしここで、式(12)で ${}^1\chi(G)$ を定義する際に用いた $C(G)$ がグラフ論的には互いに隣接する頂点を切り取る独立カットセット行列であることと、単位行列 $(E(G))$ が1個の頂点を切り取る独立カットセット行列でもあることから ${}^0\chi(G)$ を定義する式(6)が

$${}^0\chi(G) = \sum_j \frac{1}{\sqrt{\prod_i \{D(G)E(G)\}_{ij}}} \quad (19)$$

と書き換えられることに注意した。式(19)は後に一般化を考える際には有用であるが、計算機上で計算を実行することを念頭に置く場合には余分な条件がない式(6)の定義のほうが使いやすいと考えられる。

話題を G の隣接する2本の辺に属す3個の頂点を切り取る $B(G)$ に戻すと、次に示す手順で $B(G)$ が $C(G)$ から容易に得られる。ただし以下では $B(G)$ を求める手順を具体的に示すために、Fig. 1のイソペンタンの場合を例に用いることにする。

まず $C(G)$ の各列どうしの積を列とする $C^*(G)$ を式(13a), (13b)の規則に従って作る。すなわち

$$C^*(G) = \begin{matrix} & \begin{matrix} 1 \times 2 & 1 \times 3 & 1 \times 4 & 2 \times 3 & 2 \times 4 & 3 \times 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \end{matrix} \quad (20)$$

次に $C^*(G)$ の列の要素の和がmod 2で0になる列、すなわち頂点を共有しない辺が存在する不連続な経路に対応する列を $C^*(G)$ から消去すれば、これが求める $B(G)$ になる。すなわち

$$B^*(G) = \begin{matrix} & \begin{matrix} 1 \times 2 & 1 \times 4 & 2 \times 3 & 2 \times 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \end{matrix} \quad (21)$$

ただし、 $C^*(G)$ を計算機上で求める際には積の規則を変更するよりも、 $C(G)$ の各列の要素の和を

要素とする行列をあらかじめ作ってからその行列の有限の要素を1に置き換えるという上述の手順と同等な別法で $C^*(G)$ を求めることもできる。こうして得られた $B(G)$ を $D(G)$ の右側から乗じて得られた行列の各列から、隣接する2本の辺に属す3個の頂点の次数が求められる。すなわち

$$D(G)B(G) = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 1 \times 2 & 1 \times 4 & 2 \times 3 & 2 \times 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 3 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 4 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 5 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \end{matrix} \quad (22)$$

ここで、式(20), (21), (22)の列を示す 1×2 のような記号の意味を模式的にFig. 2に示した。したがって、式(22)を式(18)に代入すれば目的の2次のRandicの指数が

$${}^2\chi(G) = \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}} \cong 1.8021 \quad (23)$$

のように求められる。

本論では、 $C(G)$ が $A(G)$ と同様に有用なグラフの行列表現であることを教育的に示すためにRandicの指数の中でも実際によく用いられる3個の指数だけを順次取り扱ったが、式(12), (18), (19)における $C(G)$, $B(G)$, $E(G)$ の役割が「長さ n の経路についてだけ和を取る組み合わせを抽出する」というRandic指数の定義と同等であることを考えれば、 n 次のRandic指数が一般に

$${}^n\chi(G) \equiv \sum_j \frac{1}{\sqrt{\prod_i \{D(G)K_n(G)\}_{ij}}} \quad (24)$$

で与えられることは自明であろう。ただし、式(24)で $K_n(G)$ は $n+1$ 個の頂点を切り取る独立カットセット行列を示す。

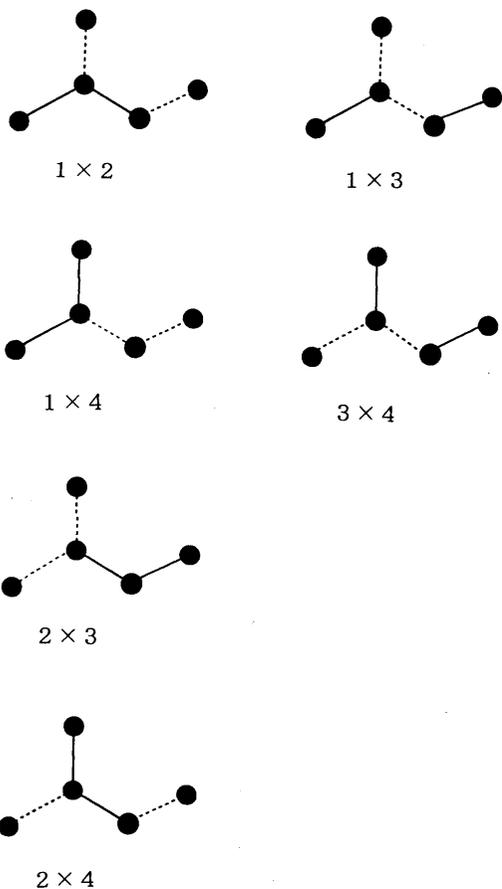


Fig.2 Schematic diagrams which indicate $n \times m$ of Eqs.(20),(21) and (22).

3. 結 論

A(G)と同様にGを記述するC(G)の有用性を示すために、Randicの指数のうち0次、1次、2次の指数をC(G)から求める手順を示した。C(G)を用いることによって、計算機上での計算に適した単純な規則に従ってRandicの指数が得られることが明らかになった。

文 献

- 1) N.Trinajstić, "CHEMICAL GRAPH THEORY", 2nd rev. ed., CRC Press, Inc., Boca Raton(1992), pp.41-52.
- 2) D.H.Rouvray, "CHEMICAL GRAPH THEORY", D.Bonchev, D.H.Rouvray ed., Gordon and Breach Science Publishers, New York(1990), pp.78-83.
- 3) H.Hosoya, *Bull. Chem. Soc. Jpn.*, 44, 2332(1971).

- 4) N.Trinajstić, "CHEMICAL GRAPH THEORY", 2nd rev. ed., CRC Press, Inc., Boca Raton(1992), pp.225-273. とその参考文献
- 5) M.Randic, *J. Am. Chem. Soc.*, 97, 6609(1975).
- 6) D.E.Needham, I-C.Wei, P.G.Seybold, *J. Am. Chem. Soc.*, 110,4186(1988).