DIRECT ESTIMATION OF THE HOMO-LUMO SEPARATION IN THE HÜCKEL MODEL OF ALTERNANT SYSTEMS

MAKOTO ISIHARA

This paper presents a simple formula for estimating the difference between the highest occupied and the lowest unoccupied molecular orbital energies in the Hückel model of alternant systems. The formula uses the traces of the $2k$-th and the $4k$-th powers of the inverse of the adjacency matrix without solving the eigenvalue problem. Numerical examples for typical alternant systems show that the formula works quite satisfactorily.

1. Introduction

The energy difference between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) is one of the most interesting quantities in molecular orbital theory. This quantity, conventionally called the HOMO-LUMO separation, is usually evaluated with solving the eigenvalue problem. In particular, for alternant systems having no nonbonding orbital, the quantity becomes just twice the absolute value of the energy of HOMO or LUMO. Let us confine ourselves to the Hückel model of such alternant systems in the following, where the graph-theoretical terminology is used throughout.

Suppose a bipartite graph having n vertices and assume the graph to have no zero eigenvalue. Since n is necessarily even, let us denote $n/2$ by μ and number the n eigenvalues $\{\lambda_j\}$ of the graph in nonincreasing order as

$$\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_\mu > 0 > \lambda_{\mu+1} \geq \cdots \geq \lambda_{n-1} > \lambda_n, \tag{1}$$

where we have

$$\lambda_j = \lambda_{n-j+1} \tag{2}$$

by virtue of the bipartite symmetry. The HOMO-LUMO separation δ is given as twice the smallest positive eigenvalue of the graph;

$$\delta = 2\lambda_\mu. \tag{3}$$

There exist already a few attempts to estimate the HOMO-LUMO separation directly without solving the eigenvalue problem (Gutman and Rouvray, 1979; Graovac and Gutman, 1980; Trinajstić, Mihalić, and Graovac, 1994). It is remarkable

above all that the Graovac-Gutman formula works very well considering its simple form

$$\delta \cong \frac{3n-2}{n}\sqrt{\frac{|a_n|}{|a_{n-2}|}}, \tag{4}$$

where a_n and a_{n-2} stand, respectively, for the constant term and the coefficient of the second power term in the characteristic polynomial (Graovac and Gutman, 1980). Note that the knowledge of a_n and a_{n-2} are almost of the same level as that of the inverse of the adjacency matrix; $|a_{n-2}/a_n|$ is really nothing other than the trace of the squared inverse of the adjacency matrix.

This paper presents a simple formula for estimating the HOMO-LUMO separation, that is, twice the smallest positive eigenvalue of a bipartite graph having no zero eigenvalue. The formula uses the traces of the $2k$-th and the $4k$-th powers of the inverse of the adjacency matrix. The idea essentially stems from the previous treatment for estimating the largest eigenvalue of a simple connected graph (Isihara, 2003). Numerical examples for typical bipartite Hückel graphs are given to compare the present formula with the Graovac-Gutman formula.

2. Formula

The adjacency matrix $\boldsymbol{A}$ of a bipartite graph consisting of μ black and μ white vertices can be written in the form

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{B} \\ {}^{t}\boldsymbol{B} & \boldsymbol{0} \end{pmatrix} \tag{5}$$

with a $\mu \times \mu$ submatrix $\boldsymbol{B}$ and its transpose ${}^{t}\boldsymbol{B}$, so that the assumption of no zero eigenvalue leads us to

$$\boldsymbol{A}^{-2k} = \begin{pmatrix} (\boldsymbol{B}\,{}^{t}\boldsymbol{B})^{-k} & \boldsymbol{0} \\ \boldsymbol{0} & ({}^{t}\boldsymbol{B}\boldsymbol{B})^{-k} \end{pmatrix} \tag{6}$$

with k an arbitrary positive integer. The quantity

$$F_{-2k} = \mathrm{Tr}((\boldsymbol{B}\,{}^{t}\boldsymbol{B})^{-k}) = \mathrm{Tr}(({}^{t}\boldsymbol{B}\boldsymbol{B})^{-k}) \tag{7}$$

is naturally equal to a half of the trace of $\boldsymbol{A}^{-2k}$ and, therefore, we have

$$F_{-2k} = \sum_{j=1}^{\mu} \lambda^{-2k}. \tag{8}$$

Since it is possible for the smallest positive eigenvalue to be degenerate, let us denote its multiplicity by ζ.

Applying the Cauchy inequality to the set of the positive numbers λ_j^{-2k}'s for $1 \leq j \leq \mu - \zeta$, we have

$$\left(\sum_{j=1}^{\mu-\zeta} \lambda_j^{-2k}\right)^2 \leq (\mu-\zeta) \sum_{j=1}^{\mu-\zeta} \lambda_j^{-4k} \tag{9}$$

or, in terms of F_{-2k} and F_{-4k},

$$(F_{-2k} - \zeta\lambda_\mu^{-2k})^2 \leq (\mu-\zeta)(F_{-4k} - \zeta\lambda_\mu^{-4k}), \tag{10}$$

which is converted by elementary calculation to

$$(\lambda_\mu^{-2k} - f_{-2k})^2 \leq (\mu_0 - 1)(f_{-4k} - f_{-2k}^2) \tag{11}$$

with the mean quantity

$$f_{-2k} = F_{-2k}/\mu \tag{12}$$

and

$$\mu_0 = \mu/\zeta. \tag{13}$$

Noting that f_{-2k} does not exceed λ_μ^{-2k}, we can solve the inequality for λ_μ to obtain

$$\delta \geq \frac{2}{\sqrt[2k]{f_{-2k} + \sqrt{(\mu_0 - 1)(f_{-4k} - f_{-2k}^2)}}}. \tag{14}$$

It is readily found that the right-hand side, denoted as $\delta(k)$ from now on, of this inequality tends monotonously to δ as k increases;

$$\delta(k) < \delta(k') \to \delta \qquad (k < k' \to \infty). \tag{15}$$

However, in actual computation, the calculation of higher-order powers of the inverse of the adjacency matrix suffers from serious difficulty in retaining accuracy. Here, let us focus on approximate estimation with small k's.

The multiplicity ζ remains unity except for highly symmetric graphs and, therefore, we can use μ as μ_0 for most Hückel graphs familiar in hydrocarbon chemistry. The inequality above itself is not violated by using μ in place of μ_0 even in the case of ζ off unity.

It should be noted that we need not calculate $(\boldsymbol{B}\,{}^t\boldsymbol{B})^{-2k}$ to evaluate F_{-4k}. Using the (i,j)-component w_{ij} of $(\boldsymbol{B}\,{}^t\boldsymbol{B})^{-k}$, we have

$$F_{-4k} = \sum_i w_{ii}^2 + 2\sum_{i<j} w_{ij}^2. \tag{16}$$

Anyway, the matrix multiplication and the root extraction are plain procedures. The essential task in the calculation process is to invert $\boldsymbol{B}\,{}^t\boldsymbol{B}$ or $\boldsymbol{B}$. Fortunately, this is not very difficult, if the matrix is sparse or not gigantic.

When the three coefficients a_{n-4}, a_{n-2}, and a_n of the characteristic polynomial are known ahead, they suffice to use the formula above with $k=1$ by the relationships

$$f_{-2} = -\frac{a_{n-2}}{\mu a_n} \tag{17}$$

and

$$f_{-4}-f_{-2}^2=(\mu-1)f_{-2}^2-\frac{2a_{n-4}}{\mu a_n}. \tag{18}$$

3. Numerical Examples

Table 1 shows the estimates $\delta(k)$'s, δ_G, and δ_H of the HOMO-LUMO separation, respectively, by the present formula with small k's, by the Graovac-Gutman formula, and by solving the Hückel eigenvalue problem for many typical bipartite Hückel graphs. The table includes the difference of $\delta(k)$'s and δ_G with δ_H as well for convenience.

It is obvious that the present formula works satisfactorily enough considering its simplicity. Surely, it is a little bulkier and uses a little more data for estimation than the Graovac-Gutman formula, but the essential task involved is to invert the adjacency matrix in each of the two; the present formula requires no serious additional task, compared with the Graovac-Gutman one. It may be regarded as an advantage of the present formula that the estimate is guaranteed to keep the inequality $\delta \geq \delta(k)$.

As far as the examples here are concerned, even the estimate $\delta(1)$ is almost always better than δ_G and the estimate $\delta(4)$ agrees with δ_H up to the third decimal place without a tiny exception. Unfortunately, the estimate $\delta(1)$ partially fails in the descrimination among the pentahex systems, though some improvement can be found compared with the Graovac-Gutman one. The estimate $\delta(3)$ can immediately overcome this situation. The use of the present formula with small k's requires no particular computational devising for retaining accuracy and can be a useful desk tool for estimating the HOMO-LUMO separation for alternant systems.

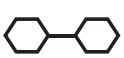
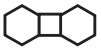
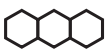
As already mentioned, we have much to be done to use the present formula for exact estimation. A crucial point is how to retain accuracy in computing higher-order powers of the inverse of the adjacency matrix.

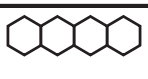
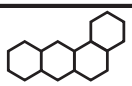
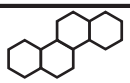
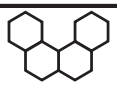
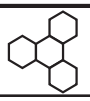
Accepted June, 1, 2003.

Table 1. Estimates of the HOMO-LUMO Separation in Bipartite Hückel Graphs

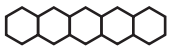
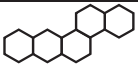
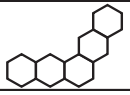
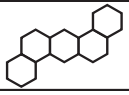
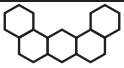
						
δ_G[a)]	1.443	1.089	1.778	1.008	0.870	1.375
$\delta(1)$	1.236	0.890	2.000[c)]	0.826	0.747	1.299
$\delta(2)$				0.828	0.763	1.321
$\delta(3)$					0.764	1.324
δ_H[b)]	1.236	0.890	2.000	0.828	0.764	1.324
$\delta_G-\delta_H$	0.207	0.199	– 0.222	0.179	0.106	0.051
$\delta(1)-\delta_H$	0.000	0.000	0.000	– 0.002	– 0.017	– 0.025
$\delta(2)-\delta_H$				0.000	– 0.001	– 0.003
$\delta(3)-\delta_H$					0.000	0.000

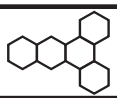
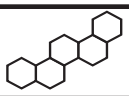
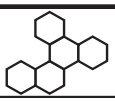
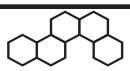
a) by the Graovac-Gutman formula. b) by solving the Hückel eigenvalue problem. c) with ζ=2. If ζ=1 is used, we have 1.789 (– 0.211) as $\delta(1)$ and 1.964(– 0.036) as $\delta(4)$.

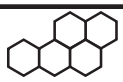
						
δ_G	0.688	1.281	1.259	1.002	0.939	1.109
$\delta(1)$	0.525	1.216	1.346	0.883	0.824	1.142
$\delta(2)$		1.234	1.398	0.890	0.828	1.194
$\delta(3)$		1.236	1.409			1.209
$\delta(4)$						1.210
δ_H	0.525	1.236	1.409	0.890	0.828	1.210
$\delta_G-\delta_H$	0.163	0.045	– 0.150	0.112	0.111	– 0.101
$\delta(1)-\delta_H$	0.000	– 0.020	– 0.063	– 0.007	– 0.004	– 0.068
$\delta(2)-\delta_H$		– 0.002	– 0.011	0.000	0.000	– 0.016
$\delta(3)-\delta_H$		0.000	0.000			– 0.001
$\delta(4)-\delta_H$						0.000

						
δ_G	0.945	0.709	0.900	0.962	0.970	1.021
$\delta(1)$	0.880	0.588	0.880	1.001	1.037	1.328[d)]
$\delta(2)$	0.890	0.590	0.902	1.036	1.104	1.359
$\delta(3)$			0.905	1.040	1.130	1.367
$\delta(4)$					1.135	1.368
δ_H	0.890	0.590	0.905	1.040	1.135	1.368
$\delta_G-\delta_H$	0.055	0.119	– 0.005	– 0.078	– 0.165	– 0.347
$\delta(1)-\delta_H$	– 0.010	– 0.002	– 0.025	– 0.039	– 0.098	– 0.040
$\delta(2)-\delta_H$	0.000	0.000	– 0.003	– 0.004	– 0.031	– 0.009
$\delta(3)-\delta_H$			0.000	0.000	– 0.005	– 0.001
$\delta(4)-\delta_H$					0.000	0.000

d) with ζ=2. If ζ=1 is used, we have 1.160 (– 0.208) as $\delta(1)$ and 1.340 (– 0.028) as $\delta(4)$.

					
δ_G	0.540	0.802	0.789	0.848	0.850
$\delta(1)$	0.438	0.791	0.801	0.900	0.912
$\delta(2)$	0.439	0.808	0.853	0.941	0.967
$\delta(3)$		0.810	0.871	0.947	0.982
$\delta(4)$			0.874		0.983
δ_H	0.439	0.810	0.874	0.947	0.983
$\delta_G-\delta_H$	0.110	– 0.008	– 0.085	– 0.099	– 0.133
$\delta(1)-\delta_H$	– 0.001	– 0.019	– 0.073	– 0.047	– 0.071
$\delta(2)-\delta_H$	0.000	– 0.002	– 0.021	– 0.006	–0.016
$\delta(3)-\delta_H$		0.000	– 0.003	0.000	– 0.001
$\delta(4)-\delta_H$			0.000		0.000

					
δ_G	0.871	0.871	0.894	0.878	0.876
$\delta(1)$	0.945	0.945	0.996	0.977	0.974
$\delta(2)$	0.991	0.994	1.052	1.047	1.050
$\delta(3)$	0.998	1.003	1.063	1.069	1.086
$\delta(4)$		1.004	1.064	1.071	1.098
δ_H	0.998	1.004	1.064	1.071	1.100
$\delta_G-\delta_H$	– 0.127	– 0.133	– 0.170	– 0.193	– 0.224
$\delta(1)-\delta_H$	– 0.054	– 0.059	– 0.068	– 0.094	– 0.126
$\delta(2)-\delta_H$	– 0.007	– 0.010	– 0.012	– 0.024	– 0.050
$\delta(3)-\delta_H$	0.000	– 0.001	– 0.001	– 0.002	– 0.014
$\delta(4)-\delta_H$		0.000	0.000	0.000	– 0.002

					
δ_G	0.804	0.911	0.782	0.834	0.655
$\delta(1)$	0.736	0.953	0.691	1.064[e)]	0.656
$\delta(2)$	0.742	0.989	0.695	1.078	0.671
$\delta(3)$		0.994			0.672
δ_H	0.742	0.994	0.695	1.078	0.672
$\delta_G-\delta_H$	0.062	– 0.083	0.087	– 0.244	– 0.017
$\delta(1)-\delta_H$	– 0.006	– 0.041	– 0.004	– 0.014	– 0.016
$\delta(2)-\delta_H$	0.000	– 0.005	0.000	0.000	– 0.001
$\delta(3)-\delta_H$		0.000			0.000

e) with ζ=2. If ζ=1 is used, we have 0.918 (– 0.160) as $\delta(1)$ and 1.056 (– 0.022) as $\delta(4)$.

References

Gutman, I. and D. H. Rouvray (1979) "An Approximate Topological Formula for the HOMO-LUMO Separation in Alternant Hydrocarbons", *Chemical Physics Letters*, Vol. 64. No. 2. pp.384-388.

Graovac, A. and I. Gutman (1980) "Estimation of the HOMO-LUMO Separation", *Croatica Chemica Acta*, Vol. 53, No. 1, pp.45-50.

Isihara, M. (2003) "Estimation of Bounds of Orbital Energies in the Hückel Model", *Review of Economics and Information Studies*, Vol. 3. No. 1-4, pp.137-145.

Trinajstić, N. and Z. Mihalić, and A. Graovac (1994) "The Interplay between Graph Theory and Molecular Orbital Theory" in "Graph Theoretical Approaches to Chemical Reactivity" ed. D. Bonchev and O. Mekenyan (Kluwer Academic Publishers, Dordrecht) pp.37-72.