J Math Chem (2011) 49:6–11
DOI 10.1007/s10910-010-9737-8

BRIEF COMMUNICATION

The relationship between the eigenvalues and eigenvectors of a similarity matrix and its associated Carbó index matrix

Paul W. Ayers · Ramon Carbó-Dorca

Received: 12 July 2010 / Accepted: 21 September 2010 / Published online: 10 October 2010

Abstract The eigenvalues and eigenvectors of a quantum similarity matrix are also generalized eigenvalues and eigenvectors of the associated matrix of Carbó indices. This establishes bounds on the spectrum of the Carbó index matrix; for example, a quantum similarity matrix is positive semidefinite if and only if the associated Carbó index matrix is also positive semidefinite. The generalized eigenvalue problem for the Carbó index matrix has a diagonal metric matrix on the right-hand-side. Every generalized eigenvalue problem can be written in this diagonal form (i.e., this form is not *special* to this application). This diagonally structure generalized eigenvalue problem is especially convenient because it can be converted to a conventional eigenvalue problem by a particularly simple partial Löwdin transformation.

Keywords Quantum similarity · Carbó index · Generalized eigenvalue problem

1 Introduction

The most prevalent mathematical measures of quantum similarity are the quantum similarity matrices and the matrix containing their related Carbó similarity indices. Both quantities appear already in the first paper that uses quantum similarity to make a molecular ordering [1] and have been used, together or separately, many times since then. (See refs. [2–6] for recent reviews.) Although the eigenvalues and eigenvectors

P. W. Ayers (✉)
Department of Chemistry and Chemical Biology, McMaster University, Hamilton, ON L8S 4M1, Canada
e-mail: ayers@mcmaster.ca

R. Carbó-Dorca
Institut de Química Computacional, Universitat de Girona, Girona, Catalonia 17071, Spain

of quantum similarity matrices have been analyzed through a connection to semi-stochastic matrices, [7] no analysis of the relationship between the eigenvalues and eigenvectors of a quantum similarity matrix and its associated matrix of Carbó similarity indices has been presented.

We denote the quantum similarity matrix as $\mathbf{Z}$ and denote its elements as $\{z_{ij}\}$. The elements of a quantum similarity matrix are usually chosen to have the form

$$z_{ij} = \iint p_i(\mathbf{r})\,\Omega\left(\mathbf{r},\mathbf{r}'\right)\,p_j\left(\mathbf{r}'\right)\,d\mathbf{r}d\mathbf{r}' \tag{1}$$

where $\Omega\left(\mathbf{r},\mathbf{r}'\right)$ is a positive-definite integral kernel and $p_i(\mathbf{r})$ is the property density for some quantum mechanical observable associated with molecule i. Usually the property of interest is the electron density or the valence electron density. In the simplest case, $\Omega\left(\mathbf{r},\mathbf{r}'\right) = \delta\left(\mathbf{r}-\mathbf{r}'\right)$. See, for example, refs. [8,9]. The quantum similarity matrix, $\mathbf{Z}$, is usually positive semidefinite, but it can have negative eigenvalues if, as is common, the molecules i and j are aligned to maximize each individual matrix element z_{ij} [10].

The matrix of Carbó indices, $\mathbf{R}$, is to the quantum similarity matrix what the correlation matrix is to the covariance matrix. I.e., the elements of the Carbó index matrix are

$$r_{ij} = \frac{z_{ij}}{\sqrt{z_{ii}z_{jj}}}. \tag{2}$$

For the purposes of this paper, it is useful to define the diagonal matrix,

$$\mathbf{D} = \mathrm{diag}\left(\mathbf{Z}\right)^{-\frac{1}{2}} \tag{3}$$

with elements

$$d_{ij} = \frac{\delta_{ij}}{\sqrt{z_{ii}}}. \tag{4}$$

Then

$$\mathbf{R} = \mathbf{DZD}. \tag{5}$$

2 Generalized eigensystem relationships between Z and R

Is there any relationship between the eigenvectors and eigenvalues of $\mathbf{Z}$ and $\mathbf{R}$?

Denote the eigenvalues and eigenvectors of $\mathbf{Z}$ as ζ_k and $\mathbf{y}_k$, respectively. Start with the defining relation of the eigenvalues and the eigenvectors,

$$\mathbf{Z}\mathbf{y}_k = \zeta_k\mathbf{y}_k \tag{6}$$

Insert the identity matrix, $\mathbf{I} = \mathbf{D}\mathbf{D}^{-1}$ before the eigenvector on both sides of this equation and left-multiply the equation by $\mathbf{D}$. This gives

$$\begin{aligned} \mathbf{D}\mathbf{Z}\left(\mathbf{D}\mathbf{D}^{-1}\right)\mathbf{y}_k &= \zeta_k \mathbf{D}\left(\mathbf{D}\mathbf{D}^{-1}\right)\mathbf{y}_k \\ \mathbf{R}\left(\mathbf{D}^{-1}\mathbf{y}_k\right) &= \zeta_k \mathbf{D}^2\left(\mathbf{D}^{-1}\mathbf{y}_k\right) \end{aligned} \tag{7}$$

This equation has the structure of a generalized eigenvalue problem. (See, for example, references [11–13].) Specifically,

$$\mathbf{R}\mathbf{x}_k = \zeta_k \mathbf{S}\mathbf{x}_k \tag{8}$$

with

$$\begin{aligned} \mathbf{x}_k &= \mathbf{D}^{-1}\mathbf{y}_k \\ \mathbf{S} &= \mathbf{D}^2 \end{aligned} \tag{9}$$

Each rescaled eigenvector of $\mathbf{Z}$, $\mathbf{D}^{-1}\mathbf{y}_k$, is a solution to the generalized eigenvalue problem for $\mathbf{R}$, with the same eigenvalue, ζ_k.

Similarly, starting from the conventional (*not generalized*) eigensystem of $\mathbf{R}$, denoting the eigenvalues and vectors ρ_k and $\mathbf{q}_k$, respectively, we can derive a generalized eigenvalue problem for $\mathbf{Z}$,

$$\begin{aligned} \mathbf{R}\mathbf{q}_k &= \rho_k \mathbf{q}_k \\ \mathbf{D}^{-1}\mathbf{R}\mathbf{D}^{-1}\mathbf{D}\mathbf{q}_k &= \rho_k \mathbf{D}^{-1}\mathbf{D}^{-1}\mathbf{D}\mathbf{q}_k \\ \mathbf{Z}\left(\mathbf{D}\mathbf{q}_k\right) &= \rho_k \mathbf{D}^{-2}\left(\mathbf{D}\mathbf{q}_k\right) \end{aligned} \tag{10}$$

Therefore,

$$\mathbf{Z}\mathbf{p}_k = \rho_k \mathbf{T}\mathbf{p}_k \tag{11}$$

with

$$\begin{aligned} \mathbf{p}_k &= \mathbf{D}\mathbf{q}_k \\ \mathbf{T} &= \mathbf{D}^{-2} \end{aligned} \tag{12}$$

Because $\mathbf{D}^2$ is a diagonal matrix, it is trivial to invert it and transform Eq. 7 into a conventional (but nonsymmetric) eigenvalue problem,

$$\mathbf{D}^{-2}\mathbf{R}\mathbf{x}_k = \zeta_k \mathbf{x}_k \tag{13}$$

This result emphasizes there is a simple relationship between the eigenvalues and eigenvectors of the quantum similarity matrix and the Carbó index matrix.

Using the variational principle for the extreme generalized eigenvalues, we can derive an upper bound on the spectrum of the Carbó index matrix,

$$\max(\rho_k) = \underbrace{\max}_{\mathbf{p}} \frac{\mathbf{p}^\dagger \mathbf{Z}\mathbf{p}}{\mathbf{p}^\dagger \mathbf{D}^{-2}\mathbf{p}} \leq \frac{\max(\zeta_k)}{\min(z_{kk})}. \tag{14}$$

Similarly, lower bounds can be derived. If the quantum similarity matrix is positive semidefinite, we have,

$$\min(\rho_k) = \underbrace{\min}_{\mathbf{p}} \frac{\mathbf{p}^\dagger \mathbf{Z}\mathbf{p}}{\mathbf{p}^\dagger \mathbf{D}^{-2}\mathbf{p}} \geq \frac{\min(\zeta_k)}{\max\left(z_{kk}^{-1}\right)} = \min(\zeta_k)\min(z_{kk}). \tag{15}$$

If $\mathbf{Z}$ has one or more negative eigenvalues, we have instead

$$\min(\rho_k) = \underbrace{\min}_{\mathbf{p}} \frac{\mathbf{p}^\dagger \mathbf{Z}\mathbf{p}}{\mathbf{p}^\dagger \mathbf{D}^{-2}\mathbf{p}} \geq \frac{\min(\zeta_k)}{\min\left(z_{kk}^{-1}\right)} = \min(\zeta_k)\max(z_{kk}). \tag{16}$$

Equation (15) establishes that if the quantum similarity matrix is positive semidefinite, so is the Carbó index matrix. The converse is also true: $\mathbf{R} \succeq 0$ implies $\mathbf{Z} \succeq 0$. Similar bounds can be derived for the spectrum of $\mathbf{Z}$ using the extreme eigenvalues of $\mathbf{R}$. Note that the bounds in Eqs. (14)–(16) are not expected to be especially tight.

3 Partial Löwdin transformation and the generalized eigenvalue problem for the Carbó index matrix

One might suspect that because the metric matrix in the generalized eigenvalue problem in Eq. 8, is diagonal, there might be especially appealing analytic relationships between $\mathbf{R}$ and $\mathbf{Z}$ matrices [14]. We have mined the literature and this does not seem to be the case. To understand why, it is useful to recall that one of us has previously studied generalized eigenvalue problems with this special form, and shown that this form is equivalent to the more general generalized eigenvalue problem up to a unitary transformation (amounting to the choice of basis set for the space) [15,16].

Consider Eq. 8 in generalized form,

$$\mathbf{R}\mathbf{x}_k = \zeta_k \mathbf{S}\mathbf{x}_k \tag{17}$$

but now allow $\mathbf{S}$ to be an arbitrary Hermitian matrix. $\mathbf{S}$ can be diagonalized by the unitary matrix $\mathbf{U}$. I.e.,

$$\begin{aligned} \mathbf{U}^\dagger \mathbf{S}\mathbf{U} &= \Sigma \\ \mathbf{U}\Sigma\mathbf{U}^\dagger &= \mathbf{S} \end{aligned} \tag{18}$$

This allows us to rewrite Eq. 17 in a form where the right-hand-side matrix is diagonal,

$$\begin{aligned} \mathbf{R}\mathbf{x}_k &= \zeta_k \mathbf{U}\Sigma\mathbf{U}^\dagger \mathbf{x}_k \\ \mathbf{U}^\dagger\mathbf{R}\mathbf{U}\left(\mathbf{U}^\dagger\mathbf{x}_k\right) &= \zeta_k \Sigma\left(\mathbf{U}^\dagger\mathbf{x}_k\right) \end{aligned} \tag{19}$$

Defining

$$\begin{aligned} \mathbf{w}_k &= \mathbf{U}^\dagger\mathbf{x}_k \\ \mathbf{R}_U &= \mathbf{U}^\dagger\mathbf{R}\mathbf{U} \end{aligned} \tag{20}$$

gives a generalized eigenvalue problem in the new basis defined by the unitary transformation,

$$\mathbf{R}_U\mathbf{w}_k = \zeta_k \Sigma\mathbf{w}_k \tag{21}$$

This equation has a diagonal matrix on the right hand side, just like the generalized eigensystems of interest here. If the initial **S** matrix is positive semidefinite, then we can write, $\Sigma = \mathbf{D}^2$, where $\mathbf{D} = \Sigma^{\frac{1}{2}}$ is a diagonal matrix. Otherwise, we can still use this form, but **D** will have some pure-imaginary components. One implication of this result is that generalized eigenvalue problems with diagonal right-hand-side matrix (like Eqs. 8 and 11) are not really mathematically distinct from the most general form of the generalized eigensystem, but a canonical form of it.

The specific form of the generalized eigenvalaue problem in Eq. 21 can be solved using a partial Löwdin transformation to obtain the simplified secular equation,

$$\Sigma^{-\frac{1}{2}}\mathbf{R}_U\Sigma^{-\frac{1}{2}}\Sigma^{\frac{1}{2}}\mathbf{w}_k = \zeta_k\Sigma^{\frac{1}{2}}\mathbf{w}_k \tag{22}$$

Defining

$$\begin{aligned} \mathbf{v}_k &= \Sigma^{\frac{1}{2}}\mathbf{w}_k = \Sigma^{\frac{1}{2}}\mathbf{U}^\dagger\mathbf{x}_k \\ \mathbf{R}_L &= \Sigma^{-\frac{1}{2}}\mathbf{R}_U\Sigma^{-\frac{1}{2}} = \Sigma^{-\frac{1}{2}}\mathbf{U}^\dagger\mathbf{R}\mathbf{U}\Sigma^{-\frac{1}{2}} \end{aligned} \tag{23}$$

allows us to express this as a Hermitian eigenvalue problem,

$$\mathbf{R}_L\mathbf{v}_k = \zeta_k\mathbf{v}_k. \tag{24}$$

The eigenvalues of this eigensystem are the same as that of the underlying generalized eigenvalue problem; only the eigenvectors have been altered. In the specific case of interest here, Eq. 24 is equivalent to the original problem because **U** equals the identity matrix, $\Sigma^{-\frac{1}{2}} = \mathbf{D}^{-1}$, and

$$\mathbf{R}_L = \mathbf{D}^{-1}\cdot\mathbf{I}\cdot\mathbf{R}\cdot\mathbf{I}\cdot\mathbf{D}^{-1} = \mathbf{D}^{-1}\mathbf{R}\mathbf{D}^{-1} = \mathbf{Z}. \tag{25}$$

(See Eq. 5).

4 Discussion

In addition to the Carbó index, several other similarity indices can be found in the literature. Some of the most prevalent indices employed in the quantum similarity field are: (a) the Hodgkin-Richards index [17] (which simply replaces the geometric mean in the denominator of Eq. 2 with the arithmetic mean), (b) the Tanimoto index [18] (which is based on comparisons between bit strings), and Petke's index [19] (which uses a maximal value distance). It seems difficult, and perhaps impossible, to find a simple algebraic relationship between the spectrum of these indices and the underlying quantum similar matrices. This is in contrast to the Carbó index matrix, which has a generalized eigensystem identical to the conventional eigensystem of its associated quantum similarity matrix. The appealing algebraic structure of the Carbó index matrix is not unexpected, because each Carbó index element has a sound geometrical interpretation, even in generalized Hilbert and Riemann spaces, as the cosine of the angle subtended by the two functions involved. Some work on the connection between various quantum molecular similarity matrices and the quantum molecular similarity indices, along with the relationships between different indices, has been done before, but that work did not address linear algebraic issues considered in this paper [20].

Acknowledgments RC-D acknowledges funding from the Spanish Ministerio de Educación y Ciencia, under the fellowship: CTQ2006-04410/BQU; PWA acknowledges funding and other support from NSERC, the Canada Research Chairs, the Sloan Foundation, and Sharcnet.

References

1. R. Carbó, L. Leyda, M. Arnau, Int. J. Quantum Chem. **17**, 1185 (1980)
2. E. Besalú, X. Girones, L. Amat, R. Carbó-Dorca, Acc. Chem. Res. **35**, 289 (2002)
3. R. Carbó-Dorca, L. Amat, E. Besalú, X. Girones, D. Robert, J. Mol Struct-Theochem **504**, 181 (2000)
4. P. Bultinck, X. Girones, R. Carbó-Dorca, Molecular quantum similarity: theory and applications. (2005)
5. R. Carbó-Dorca, L. Amat, E. Besalú, M. Lobato, in *Advances in Molecular Similarity*, Vol. 2, ed. by R. Carbó-Dorca, P.G. Mezey (1998), p. 1
6. R. Carbó-Dorca, E. Besalú, J. Mol. Struct.-Theochem **451**, 11 (1998)
7. R. Carbó-Dorca, J. Math. Chem. **27**, 357 (2000)
8. R. Carbó, B. Calabuig, Int. J. Quantum Chem. **42**, 1681 (1992)
9. R. Carbó, B. Calabuig, Int. J. Quantum Chem. **42**, 1695 (1992)
10. R. Carbó-Dorca, E. Besalu, L. D. Mercado, J. Comp. Chem. submitted (2010)
11. G.H. Golub, C.F. Van Loan, *Matrix computations*, 3rd edn. (Johns Hopkins UP, Baltimore, MD, U.S.A., 1996)
12. *Handbook for Automatic Computation, vol. 2, Linear Algebra*, ed. by J.H. Wilkinson, C. Reinsh (Springer, New York, 1971)
13. J.H. Wilkinson, *The algebraic eigenvalue problem* (Clarendon, Oxford, 1965)
14. It is, of course, computationally very convenient that the metric matrix is diagonal this makes evaluating its inverse computationally stable and efficient
15. R. Carbó, Anales de Física **85**, 67 (1971)
16. R. Carbó-Dorca, J. Mol. Struct.-Theochem **943**, 32 (2010)
17. E.E. Hodgkin, W.G. Richards, Int. J. Quantum Chem. **S14**, 105 (1987)
18. J.T. Tou, R.C. González, *Pattern recognition principles* (Addison, Reading, Mass, U.S.A., 1974)
19. J.D. Petke, J. Comput. Chem. **14**, 928 (1993)
20. R. Carbó, E. Besalú, L. Amat, X. Fradera, J. Math. Chem. **19**, 47 (1996)