Article

On the Raleigh-Ritz Variational Method. Non-orthogonal Basis Set

Francisco Fernández *

Posted Date: 15 May 2024

doi: 10.20944/preprints202405.1006.v1

Keywords: Rayleigh-RItz; variational method; matrix diagonalization; Hermitian matrix; eigenvalues; convergence

 doi:10.20944/preprints202405.1006.v1

Article

On the Raleigh-Ritz Variational Method, Non-Orthogonal Basis Set

Francisco M. Fernández

INIFTA, DQT, Sucursal 4, C. C. 16, 1900 La Plata, Argentina; fernande@quimica.unlp.edu.ar

Abstract: We overview the main equations of the Rayleigh-Ritz variational method and discuss their connection with the problem of simultaneous diagonalization of two Hermitian matrices.

Keywords: rayleigh-ritz; variational method; matrix diagonalization; hermitian matrix; eigenvalues; convergence

---

1. Introduction

The Rayleigh-Ritz variational method (RR) is one of the approximate methods most commonly used in the study of the electronic structure of atoms and molecules [1,2]. One of its main advantages is that it provides increasingly accurate upper bounds to all the eigenvalues of the Hamiltonian operator of the system [3,4]. In this paper we provide a comprehensible overview of the approach and illustrate some of its relevant points by means of a simple problem.

2. The Rayleigh-Ritz Variational Method

The starting point of our analysis is a linearly independent set of vectors $\mathcal{V} = \{f_1, f_2, \ldots\}$. Clearly, the only solution to the vector equation

$$\sum_{i=1}^{N} a_i f_i = 0, \tag{1}$$

is $a_i = 0$ for all $i = 1, 2, \ldots, N$. If we apply the bras $\langle f_j|$, $j = 1, 2, \ldots, N$, to this equation from the left we obtain

$$\sum_{i=1}^{N} S_{ji} a_i = 0, \; j = 1, 2, \ldots, N, \tag{2}$$

where $S_{ij} = \langle f_i | f_j \rangle$. We have an homogeneous system of N linear equations with N unknowns a_i with the only solution $a_i = 0$. Consequently, $|\mathbf{S}| \neq 0$ where $\mathbf{S} = \left(S_{ij}\right)_{i,j=1}^{N}$ is an $N \times N$ Hermitian matrix and $|...|$ stands for determinant. Note that $S_{ij} = S_{ji}^{*}$ so that $\mathbf{S}^{\dagger} = \mathbf{S}$ where $\dagger$ stands for adjoint. The matrix $\mathbf{S}$ is commonly called overlap matrix[1].

Let $\mathbf{v}$ be an eigenvector of $\mathbf{S}$ with eigenvalue s, $\mathbf{S}\mathbf{v} = s\mathbf{v}$, then $\mathbf{v}^{\dagger}\mathbf{S}\mathbf{v} = s\mathbf{v}^{\dagger}\mathbf{v}$. If v_i, $i = 1, 2, \ldots, N$, are the elements of the $N \times 1$ column vector $\mathbf{v}$ then

$$\mathbf{v}^{\dagger}\mathbf{S}\mathbf{v} = \left\langle \sum_{i=1}^{N} v_i f_i \middle| \sum_{j=1}^{N} v_j f_j \right\rangle > 0, \tag{3}$$

and we conclude that $s > 0$. In other words, the overlap matrix $\mathbf{S}$ is positive definite.

We are interested in the eigenvalue equation

$$\begin{aligned} H\psi_n &= E_n \psi_n, \; n = 1, 2, \ldots, \\ E_1 &\leq E_2 \leq \ldots, \; \langle \psi_i | \psi_j \rangle = \delta_{ij}, \end{aligned} \tag{4}$$

for an Hermitian operator H. In order to solve it approximately we propose and ansatz of the form

$$\varphi = \sum_{j=1}^{N} c_j f_j, \tag{5}$$

 doi:10.20944/preprints202405.1006.v1

where $\mathcal{V} = \{f_1, f_2, \ldots\}$ is not only assumed to be linearly independent but also complete.

The RR variational method consists of minimizing the integral

$$W = \frac{\langle \varphi | H | \varphi \rangle}{\langle \varphi | \varphi \rangle}, \tag{6}$$

with respect to the expansion coefficients c_j

$$\frac{\partial W}{\partial c_j} = 0, \; j = 1, 2, \ldots, N\,. \tag{7}$$

This equation leads to the so-called secular equation[1,2]

$$\sum_{j=1}^{N} (H_{ij} - W S_{ij}) c_j = 0, \; i = 1, 2, \ldots, N, \tag{8}$$

where, $H_{ij} = \langle f_i | H | f_j \rangle$. There are nontrivial solutions c_j, $j = 1, 2, \ldots, N$, provided that the secular determinant vanishes

$$|\mathbf{H} - W\mathbf{S}| = 0, \tag{9}$$

where $\mathbf{H} = (H_{ij})_{i,j=1}^{N}$ is an $N \times N$ Hermitian matrix.

For each of the roots of the secular determinant (9), $W_1 \leq W_2 \leq \ldots \leq W_N$, we derive an approximate solution; for example, when $W = W_k$ we have

$$\varphi_k = \sum_{j=1}^{N} c_{jk} f_j, \tag{10}$$

and the secular equation (8) can be rewritten

$$\sum_{j=1}^{N} H_{ij} c_{jk} = \sum_{j=1}^{N} W_k S_{ij} c_{jk} = \sum_{j=1}^{N} \sum_{m=1}^{N} S_{ij} W_m \delta_{mk} c_{jm}. \tag{11}$$

If we define the $N \times N$ matrices $\mathbf{W} = (W_i \delta_{ij})_{i,j=1}^{N}$ and $\mathbf{C} = (c_{ij})_{i,j=1}^{N}$ then this equation can be rewritten in matrix form as

$$\mathbf{HC} = \mathbf{SCW}, \tag{12}$$

which is equivalent to

$$\mathbf{C}^{-1}\mathbf{S}^{-1}\mathbf{HC} = \mathbf{W}, \tag{13}$$

and the procedure reduces to the diagonalization of the matrix $\mathbf{S}^{-1}\mathbf{H}$ by means of the invertible matrix $\mathbf{C}$. Note that $\mathbf{S}^{-1}$ exists because $\mathbf{S}$ is positive definite as argued above.

In order to determine the coefficients c_{jk} completely, we require that $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$ that leads to

$$\langle \varphi_i | \varphi_j \rangle = \sum_{k=1}^{N} \sum_{m=1}^{N} c_{ki}^{*} c_{mj} \langle f_k | f_m \rangle = \delta_{ij}, \tag{14}$$

that in matrix form reads

$$\mathbf{C}^{\dagger}\mathbf{SC} = \mathbf{I}, \tag{15}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix. It follows from equations (15) and (12) that

$$\mathbf{C}^{\dagger}\mathbf{HC} = \mathbf{W}. \tag{16}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 May 2024 doi:10.20944/preprints202405.1006.v1

It is clear that there exists an invertible matrix ($\mathbf{C}$) that transforms two Hermitian matrices ($\mathbf{H}$ and $\mathbf{S}$), one of them positive definite ($\mathbf{S}$), into diagonal form. This procedure is well known in the mathematical literature[5]. However, it is most important to note that equations (15) and (16) are not what we commonly know as matrix diagonalization. In fact, the eigenvalues of $\mathbf{S}$ are not unity and the eigenvalues of $\mathbf{H}$ are not the RR eigenvalues W_i. We will illustrate this point in section 3 by means of a simple example. It is also worth noting that that we cannot obtain $\mathbf{C}$ neither from (15) or (16). One obtains the matrix $\mathbf{C}$ in the process of diagonalizing $\mathbf{S}^{-1}\mathbf{H}$ as in equation (13) and the remaining undefined matrix elements c_{ij} from equation (15).

Since $\mathbf{S}$ is positive definite, we can define $\mathbf{S}^{1/2}$. The matrix $\mathbf{U} = \mathbf{S}^{1/2}\mathbf{C}$ is unitary as shown by

$$\mathbf{U}^{\dagger}\mathbf{U} = \mathbf{C}^{\dagger}\mathbf{S}^{1/2}\mathbf{S}^{1/2}\mathbf{C} = \mathbf{I}. \tag{17}$$

On substituting $\mathbf{C} = \mathbf{S}^{-1/2}\mathbf{U}$ into equation (16) we obtain

$$\mathbf{U}^{\dagger}\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}\mathbf{U} = \mathbf{W}. \tag{18}$$

This equation is just the standard diagonalization of the Hermitian matrix $\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}$.

If the basis set is orthonormal, $\langle f_i | f_j \rangle = \delta_{ij}$, then $\mathbf{S} = \mathbf{I}$, $\mathbf{C}^{\dagger} = \mathbf{C}^{-1}$ and the secular equation (13) becomes

$$\mathbf{C}^{\dagger}\mathbf{H}\mathbf{C} = \mathbf{W}. \tag{19}$$

In this particular case, the eigenvalues of the matrix $\mathbf{H}$ are the RR eigenvalues W_i. Note that equations (16) and (19) look identical but were derived under different assumptions (they agree only when $\mathbf{S} = \mathbf{I}$).

3. Simple Example

As a simple example we consider the dimensionless eigenvalue equation

$$H\psi = E\psi,\ H = -\frac{1}{2}\frac{d^2}{dx^2} + \lambda x,\ \psi(0) = \psi(1) = 0. \tag{20}$$

In order to illustrate the RR variational method with a non-orthogonal basis set we choose $f_i(x) = x^i(1-x)$, $i = 1, 2, \ldots$, that satisfy the boundary conditions at $x = 0$ and $x = 1$.

A straightforward calculation shows that

$$S_{ij} = \frac{2}{(i+j+1)(i+j+2)(i+j+3)}, \tag{21}$$

and

$$H_{ij} = \frac{ij}{(i+j)(i+j+1)(i+j-1)} + \frac{2\lambda}{(i+j+2)(i+j+3)(i+j+4)}. \tag{22}$$

Tables 1 and 2 show the RR eigenvalues W_i, $i = 1, 2, 3, 4$, for $\lambda = 0$ and $\lambda = 1$, respectively. We appreciate that the approximate eigenvalues converge from above as expected[3,4].

In what follows, we illustrate some of the general results of Section 2 for the simplest case $N = 2$ when $\lambda = 0$. The matrices are

$$\mathbf{S} = \frac{1}{60}\begin{pmatrix} 2 & 1 \\ 1 & \frac{4}{7} \end{pmatrix},\ \mathbf{H} = \frac{1}{12}\begin{pmatrix} 2 & 1 \\ 1 & \frac{4}{5} \end{pmatrix}, \tag{23}$$

and we obtain

$$\mathbf{C}^{-1}\mathbf{S}^{-1}\mathbf{H}\mathbf{C} = \mathbf{W} = \begin{pmatrix} 5 & 0 \\ 0 & 21 \end{pmatrix},\ \mathbf{C} = \sqrt{30}\begin{pmatrix} 1 & \sqrt{7} \\ 0 & -2\sqrt{7} \end{pmatrix}. \tag{24}$$

One can easily verify that these matrices already satisfy equations (15) and (16). On the other hand, the symmetric matrices **S** and **H** can be diagonalized in the usual way by orthogonal matrices that we call $\mathbf{U}_S$ and $\mathbf{U}_H$, respectively.

$$
\begin{aligned}
\mathbf{U}_S^\dagger \mathbf{S} \mathbf{U}_S &= \frac{1}{420}\begin{pmatrix} 9-\sqrt{74} & 0 \\ 0 & 9+\sqrt{74} \end{pmatrix}, \\
\mathbf{U}_S &= \begin{pmatrix} \sqrt{\frac{1}{2}-\frac{5\sqrt{174}}{148}} & \sqrt{\frac{1}{2}+\frac{5\sqrt{174}}{148}} \\ -\sqrt{\frac{1}{2}+\frac{5\sqrt{174}}{148}} & \sqrt{\frac{1}{2}-\frac{5\sqrt{174}}{148}} \end{pmatrix}, \\
\mathbf{U}_H^\dagger \mathbf{H} \mathbf{U}_H &= \frac{1}{60}\begin{pmatrix} 7-\sqrt{34} & 0 \\ 0 & 70+\sqrt{34} \end{pmatrix}, \\
\mathbf{U}_H &= \begin{pmatrix} \sqrt{\frac{1}{2}-\frac{3\sqrt{34}}{68}} & \sqrt{\frac{1}{2}+\frac{3\sqrt{34}}{68}} \\ -\sqrt{\frac{1}{2}+\frac{3\sqrt{34}}{68}} & \sqrt{\frac{1}{2}-\frac{3\sqrt{34}}{68}} \end{pmatrix}
\end{aligned} \tag{25}
$$

We clearly see that the eigenvalues of **S** are not unity and those of **H** are not the RR eigenvalues W_i as argued in Section 2.

Using equation (25) one can easily obtain

$$
\mathbf{S}^{1/2} = \begin{pmatrix} \sqrt{\frac{233}{8880}+\frac{7\sqrt{7}}{8880}} & \sqrt{\frac{21}{2960}-\frac{7\sqrt{7}}{8880}} \\ \sqrt{\frac{21}{2960}-\frac{7\sqrt{7}}{8880}} & \sqrt{\frac{151}{62160}+\frac{7\sqrt{7}}{8880}} \end{pmatrix}. \tag{26}
$$

Table 1. Convergence of the Rayleigh-Ritz variational method with a non-orthogonal basis set for $\lambda = 0$

N	E_1	E_2	E_3	E_4
4	4.934874810	19.75077640	51.06512518	100.2492235
5	4.934802217	19.75077640	44.58681182	100.2492235
6	4.934802217	19.73923669	44.58681182	79.99595777
7	4.934802200	19.73923669	44.41473408	79.99595777
8	4.934802200	19.73920882	44.41473408	78.97848206
9	4.934802200	19.73920882	44.41322468	78.97848206
10	4.934802200	19.73920880	44.41322468	78.95700917
11	4.934802200	19.73920880	44.41321981	78.95700917
12	4.934802200	19.73920880	44.41321981	78.95683586
13	4.934802200	19.73920880	44.41321980	78.95683586
14	4.934802200	19.73920880	44.41321980	78.95683521
15	4.934802200	19.73920880	44.41321980	78.95683521
16	4.934802200	19.73920880	44.41321980	78.95683520
17	4.934802200	19.73920880	44.41321980	78.95683520
18	4.934802200	19.73920880	44.41321980	78.95683520
19	4.934802200	19.73920880	44.41321980	78.95683520
20	4.934802200	19.73920880	44.41321980	78.95683520

 doi:10.20944/preprints202405.1006.v1

Table 2. Convergence of the Rayleigh-Ritz variational method with a non-orthogonal basis set for $\lambda = 1$

N	E_1	E_2	E_3	E_4
4	5.432678349	20.25175971	51.56499993	100.7505620
5	5.432608286	20.25141191	45.08766430	100.7488422
6	5.432607868	20.23989706	45.08714181	80.49674963
7	5.432607855	20.23989074	44.91514957	80.49606992
8	5.432607855	20.23986309	44.91512224	79.47878520
9	5.432607855	20.23986306	44.91361487	79.47871372
10	5.432607855	20.23986304	44.91361453	79.45724985
11	5.432607855	20.23986304	44.91360967	79.45724783
12	5.432607855	20.23986304	44.91360967	79.45707467
13	5.432607855	20.23986304	44.91360966	79.45707465
14	5.432607855	20.23986304	44.91360966	79.45707400
15	5.432607855	20.23986304	44.91360966	79.45707400
16	5.432607855	20.23986304	44.91360966	79.45707400
17	5.432607855	20.23986304	44.91360966	79.45707400
18	5.432607855	20.23986304	44.91360966	79.45707400
19	5.432607855	20.23986304	44.91360966	79.45707400
20	5.432607855	20.23986304	44.91360966	79.45707400

4. Conclusions

We have shown that the main equations of the Rayleigh-Ritz variational method [1,2] lead to the mathematical problem of diagonalization of two Hermitian matrices[5]. Although equations (15) and (16) are discussed in some textbooks on quantum chemistry, the latter does not appear to be correctly interpreted[1].

References

1. F. L. Pilar, Elementary Quantum Chemistry, McGraw-Hill, New York, (1968).
2. A. Szabo and N. S. Ostlund, Modern Quantum Chemistry, Dover Publications, Inc., Mineola, New York, (1996).
3. J. K. L. MacDonald, Successive approximations by the Rayleigh-Ritz variation method, Phys Rev. 43 (1933) 830-833.
4. F. M. Fernández, On the Rayleigh-Ritz variational method, 2022. arXiv:2206.05122 [quant-ph].
5. R. Benedetti and P. Cragnolini, On simultaneous diagonallzation of one Hermitlan and one symmetric form, Lin Algebra Appl. 57 (1984) 215-226.