APPENDIX K

Diagonalization of a Matrix

In quantum chemistry, we often encounter the following mathematical problem.

We have a Hermitian[1] matrix $\boldsymbol{A}$ of dimension n, and we are interested in all numbers λ (called "*eigenvalues*"[2]) and the corresponding column vectors ("*eigenvectors*" of dimension n) $\boldsymbol{L}$, that satisfy the following equation:

$$(\boldsymbol{A} - \lambda\, \mathbf{1})\boldsymbol{L} = \mathbf{0}, \tag{K.1}$$

where $\mathbf{1}$ is the unit matrix (of dimension n). There is n solutions to the last equation: n eigenvalues of λ and n eigenvectors $\boldsymbol{L}$. Some eigenvalues λ may be equal (degeneracy); i.e., two or more linearly independent eigenvectors $\boldsymbol{L}$ correspond to a single eigenvalue λ. From Eq. (K.1), it is shown that any vector $\boldsymbol{L}$ is determined only to the accuracy of a multiplicative factor.[3] This is why in the future, it will be justified to normalize them to unity.

In quantum chemistry, the eigenvalue problem is solved in two ways: one is easy for $n \leq 2$, but more and more difficult for larger n, while the second way (with computers) treats all cases uniformly.

- The first way sees the eigenvalue equation as a set of linear homogeneous equations for the unknown components of the vector $\boldsymbol{L}$. Then, the condition for the non-trivial solution[4] to exist is $\det(\boldsymbol{A} - \lambda\, \mathbf{1}) = 0$. This condition can be fulfilled only for some particular values of λ, which are to be found by expanding the determinant and solving the resulting nth degree polynomial equation for λ. Then, each solution λ is inserted into Eq. (K.1) and the components of the corresponding vector $\boldsymbol{L}$ are found by using any method applicable to linear equations. Thus, we end up with λ_k and $\boldsymbol{L}_k$ for $k = 1, 2, 3, \ldots n$.
- The second way is based on diagonalization of $\boldsymbol{A}$.

First, let us show that *the same* λ satisfy the eigenvalue Eq. (K.1), but with a much simpler matrix. To this end, let us multiply Eq. (K.1) by (at the moment) *arbitrary* non-singular[5] square

[1] In practice, matrix $\boldsymbol{A}$ is usually real, and therefore it satisfies $(\boldsymbol{A}^T)^* = \boldsymbol{A}^T = \boldsymbol{A}$; i.e., $\boldsymbol{A}$ is symmetric.

[2] They are real.

[3] In other words, a unnormalized wave function still satisfies the Schrödinger equation, or to any normal mode can be assigned an arbitrary amplitude.

[4] The trivial one is obviously $\boldsymbol{L} = \boldsymbol{0}$, which is inacceptable, since the wave function cannot vanish everywhere, or atoms have to vibrate, etc.

[5] That is, its inverse matrix exists.

Ideas of Quantum Chemistry, Second Edition. http://dx.doi.org/10.1016/B978-0-444-59436-5.00031-3

matrix[6] $\boldsymbol{B}$. We obtain the following chain of transformations: $\boldsymbol{B}^{-1}(\boldsymbol{A}-\lambda\,\mathbf{1})\,\boldsymbol{L} = \boldsymbol{B}^{-1}(\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}^{-1} - \lambda\,\mathbf{1})\,\boldsymbol{L} = (\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B} - \lambda\,\mathbf{1})\,\boldsymbol{B}^{-1}\boldsymbol{L} = (\tilde{\boldsymbol{A}} - \lambda\,\mathbf{1})\tilde{\boldsymbol{L}} = \mathbf{0}$, where[7] $\tilde{\boldsymbol{A}} = \boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}$, and $\tilde{\boldsymbol{L}} = \boldsymbol{B}^{-1}\boldsymbol{L}$. Thus, another matrix and other eigenvectors, but *the same* λ! Now, let us choose such a special $\boldsymbol{B}$ that the resulting equation is as simple as possible (i.e., with a diagonal $\tilde{\boldsymbol{A}}$). Then, we will know (just by looking at) what the λ values have to be in order to satisfy the equation $(\tilde{\boldsymbol{A}} - \lambda\,\mathbf{1})\tilde{\boldsymbol{L}} = \mathbf{0}$.

Indeed, if $\tilde{\boldsymbol{A}}$ were diagonal, then $\det(\tilde{\boldsymbol{A}} - \lambda\,\mathbf{1}) = \Pi_{k=1}^{n}(\tilde{A}_{kk} - \lambda) = 0$, which gives the solutions $\lambda_k = \tilde{A}_{kk}$. Then it is easy to find the corresponding vector $\tilde{\boldsymbol{L}}_k$. For example, $\tilde{\boldsymbol{L}}_1$ we find from equation $(\tilde{\boldsymbol{A}} - \lambda_1\,\mathbf{1})\tilde{\boldsymbol{L}}_1 = \mathbf{0}$ in the following way[8]:

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & \tilde{A}_{22}-\lambda_1 & 0 & \dots & 0 \\ 0 & 0 & \tilde{A}_{33}-\lambda_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \tilde{A}_{nn}-\lambda_1 \end{pmatrix} \begin{pmatrix} \tilde{L}_{1,1} \\ \tilde{L}_{1,2} \\ \tilde{L}_{1,3} \\ \vdots \\ \tilde{L}_{1,n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

which means, that in order to get $\mathbf{0}$ on the right side, one has to have an *arbitrary* $\tilde{L}_{1,1}$, while the other $\tilde{L}_{1,j} = 0$ for $j = 2, 3, \ldots n$.

In order to have the length of $\tilde{\boldsymbol{L}}_1$ equal to 1, it is sufficient to put $\tilde{L}_{1,1} = 1$. Similarly, we find easily that the vectors $\tilde{\boldsymbol{L}}_k$ corresponding to λ_k represent nothing else but the column vectors with all components equal to 0 except the component k, that equals to 1. We are interested in vectors $\boldsymbol{L}$, rather than $\tilde{\boldsymbol{L}}$. We get these vectors from $\boldsymbol{L} = \boldsymbol{B}\tilde{\boldsymbol{L}}$, and when taking into account the form of $\tilde{\boldsymbol{L}}$, this means that $\boldsymbol{L}_k$ is nothing else but the kth column of the matrix $\boldsymbol{B}$. Since $\boldsymbol{B}$ is known, because precisely this matrix led to the diagonalization, there is no problem with $\boldsymbol{L}$:

> The columns of $\boldsymbol{B}$ represent the eigenvectors $\boldsymbol{L}$ of the equation $(\boldsymbol{A} - \lambda\,\mathbf{1})\boldsymbol{L} = \mathbf{0}$.

Our task is over!

[6] This matrix is to be found.

[7] Then, such a *unitary matrix* $\boldsymbol{B}$ (i.e., satisfying $(\boldsymbol{B}^T)^* = \boldsymbol{B}^{-1}$) can be found, that $\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}$ is *real and diagonal*. When (as it is the case in most applications) we are dealing with real matrices, then instead of unitary and Hermitian matrices, we are dealing with orthogonal and symmetric matrices, respectively.

[8] The λ has been replaced by λ_1 because one is interested by getting $\tilde{\boldsymbol{L}}_1$.