Mathematics of Linear Algebra

28th January 2010

Elements of linear algebra play a dominant role in chemical applications. For the purposes of undergraduate physical chemistry courses, quantum mechanics and select areas of thermodynamics can be formulated in terms of the elements of linear algebra. Thus, we present here a brief review of concepts of linear algebra. we first begin with vectors in Cartesian space (which everyone can conceptualize easily), and then we will generalize to generic vector spaces in anticipation of the use of linear algebra in quantum mechanics (think about how we would consider a function in quantum mechanics in terms of vectors?).

1 Linear Algebra in Cartesian Space

1.1 Vectors and Operations

Vectors in Cartesian space are treated as:

$$\begin{aligned} \vec{a} = & \quad \vec{e}_1 a_1 + \vec{e}_2 a_2 + \vec{e}_3 a_3 \\ = & \quad \textstyle\sum_i \vec{e}_i a_i \end{aligned}$$

The vectors $\vec{e}_i$ are termed a *basis*, and they represent a complete set of elements that can be used to describe all vectors. We are normally used to seeing them as the coordinate -x, -y, -z axes, but they can be any general **mutually perpendicular** unit vectors. Then, any vector in Cartesian space (3-space) can be written as a *linear combination* of the general basis vectors, $\epsilon_i, i = 1, 2, 3$.

$$\begin{aligned} \vec{a} = & \quad \vec{\epsilon}_1 a_1^{'} + \vec{\epsilon}_2 a_2^{'} + \vec{\epsilon}_3 a_3^{'} \\ = & \quad \textstyle\sum_i \vec{\epsilon}_i a_i^{'} \end{aligned}$$

Note that the coefficients have special meaning, and this will be reserved for later discussion.

A vector is thus represented by the **three components with respect to a basis**. Using the 2 general basis presented above, we can write a vector $\vec{a}$ as a **column matrix** as:

$$\vec{a} = \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

$$\vec{a}^{'} = \mathbf{a}^{'} = \begin{pmatrix} a_1^{'} \\ a_2^{'} \\ a_3^{'} \end{pmatrix}$$

- The **scalar or dot product** of two vectors $\vec{a}$ and $\vec{b}$ is defined as

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 = \sum_i a_i b_i$$

- Note that

$$\vec{a} \cdot \vec{a} = a_1^2 + a_2^2 + a_3^2 = |\vec{a}|^2$$

Now let's use our general definition of vectors to define the scalar product.

$$\vec{a} \cdot \vec{b} = \sum_i \sum_j \vec{e}_i \cdot \vec{e}_j a_i b_j$$

Let's expand this out. What do we obtain?
In order for the full sum to equate to the operational definition of the dot product, we arrive at a condition we have to require of the basis vectors. This is namely:

$$\vec{e}_i \vec{e}_j = \delta_{ij} = \delta_{ji} = \begin{matrix} 1 & if\ i = j \\ 0 & otherwise \end{matrix}$$

This is one way to state that the **basis vectors are mutually perpendicular (orthogonal) and have unit length (normal)**. They are **orthonormal**.

What is the projection of a vector $\vec{a}$ along one of the basis vectors $\vec{e}_j$? The scalar product (dot product) is the operational equivalent of projecting a vector onto another vector (in this case projecting $\vec{a}$ onto $\vec{e}_j$) as:

$$\vec{e}_j \cdot \vec{a} = \sum_i \vec{e}_j a_i = \sum_i \delta_{ij} a_i = a_j$$

We have used the orthonormality condition in the last step.

Employing the concept of the scalar product as a projection of a vector onto a vector, we can write the general form of a vector as:

$$\begin{aligned} \vec{a} &= \textstyle\sum_i \vec{e}_i \vec{e}_i \cdot \vec{a} \\ &= 1 \cdot \vec{a} \end{aligned}$$

The notation,

$$1 \quad = \textstyle\sum_i \vec{e}_i \vec{e}_i$$

is the *unit dyadic*. A dyadic is an entity that when dotted into a vector, leads to another vector. Ordinarily, a dot product of two vectors leads to a scalar. This is the distinction, and an important one.

1.2 Matrices, Operators, Operations

In the last section (Section 1), we defined vectors and relations between them. Keep in mind that each vector is defined in terms of more fundamental units of the relevant space. In Cartesian space, these are the 3 orthonormal coordinate axes. Let's consider an abstraction called an **operator** which acts, or operates, on a vector, and results in another vector.

Let's define an operator $\hat{O}$ as an entity which when acting on a vector $\vec{a}$ converts it into a vector $\vec{b}$,

$$\hat{O}\vec{a} = \vec{b}$$

The operator is said to be *linear* if for any numbers x and y,

$$\hat{O}(x\vec{a} + y\vec{b}) = x\hat{O}\vec{a} + y\hat{O}\vec{b}$$

A linear operator is completely determined if its effect on every possible vector is known. Since any vector can be represented as a linear combination of basis vectors, the operator is nicely determined by its affect on the basis. Since the operator acting on a basis vector is another vector that can be represented as a linear combination of the basis vectors, we can write:

$$\hat{O}\vec{e}_i = \sum_{j=1}^{3} \vec{e}_j O_{ji} \tag{1}$$

$$i = 1, 2, 3 \tag{2}$$

The number O_{ji} is the component of the vector $\hat{O}\vec{e}_i$ along the basis vector $\vec{e}_j$. The operator can thus be represented as a *matrix*:

$$\hat{O} = \begin{pmatrix} O_{11} & O_{12} & O_{13} \\ O_{21} & O_{22} & O_{23} \\ O_{31} & O_{32} & O_{33} \end{pmatrix}$$

The nomenclature one needs to be aware of is that the *matrix representation* of the operator $\hat{O}$ **in the basis** $\vec{e}_i$ is given by the matrix just defined immediately above. This is an important connection. The matrix representation completely specifies how the operator acts on **any** arbitrary vector since this vector can be expressed as a linear combination of the basis vectors.
If we have the matrix represenions of two operators, we can determine the matrix representation of an operator that is the product of the two known operators as:

$$\begin{aligned} \hat{C}\vec{e}_j &= \textstyle\sum_i \vec{e}_i C_{ij} \\ &= \hat{A}\hat{B}\vec{e}_j \\ &= \hat{A} \textstyle\sum_k \vec{e}_k B_{kj} \\ &= \textstyle\sum_{ik} \vec{e}_i A_{ik} B_{kj} \end{aligned}$$

Thus,

$$C_{ij} = \textstyle\sum_k A_{ik} B_{kj}$$

This last expression is the definition of matrix multiplication, so we see that the matrix representation of $\hat{C}$ is the product of the matrix representation of the two operators $\hat{A}$ and $\hat{B}$.

The order in which two operators or 2 matrices are multiplied is crucial. In general, two operators or two matrices do not **commute**. To understand the concept of observable properties in the context of quantum mechanics, we introduce the idea of a **commutator** of two operators or matrices as:

$$\begin{aligned}
\left[\hat{A}, \hat{B}\right] &= \hat{A}\hat{B} - \hat{B}\hat{A} \\
\left[\hat{A}_{matrix}, \hat{B}_{matrix}\right] &= \hat{A}_{matrix}\hat{B}_{matrix} - \hat{B}_{matrix}\hat{A}_{matrix}
\end{aligned}$$

The **anticommutator** is analogously defined as:

$$\begin{aligned}
\left[\hat{A}, \hat{B}\right] &= \hat{A}\hat{B} + \hat{B}\hat{A} \\
\left[\hat{A}_{matrix}, \hat{B}_{matrix}\right] &= \hat{A}_{matrix}\hat{B}_{matrix} + \hat{B}_{matrix}\hat{A}_{matrix}
\end{aligned}$$

1.3 More on Matrices

Since we have introduced vectors, operators, and matrices in the context of 3D space, we begin to generalize these results. A general matrix with N rows and M columns (an NXM matrix) is represented as the familiar:

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1M} \\ A_{21} & A_{22} & \cdots & A_{2M} \\ \vdots & \vdots & & A_{1M} \\ A_{N1} & A_{N2} & \cdots & A_{NM} \end{pmatrix}$$

If the $N = M$, the matrix is square. As we have seen, matrix multiplication of an NXM and MXP matrix is

$$C_{ij} = \sum_{k=1}^{M} A_{ik}B_{kj} \qquad i = 1, \cdots, N \; ; j = 1, \cdots, P$$

The set of M numbers a_i (i=1,...,M) can be represented as a column matrix:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{pmatrix}$$

We note that for an NXM matrix, $\mathbf{A}$,

$$\mathbf{A}\mathbf{a} = \mathbf{b} \tag{3}$$

and $\mathbf{b}$ is a column matrix with N elements,

$$b_i = \sum_{j=1}^{M} A_{ij} a_j \qquad i = 1, 2, ...N$$

The **adjoint** of an NXM matrix $\mathbf{A}$ is an MXN matrix with elements defined as:

$$\left(\mathbf{A}^{\dagger}\right)_{ij} = A^{*}_{ji}$$

where the A^{*}_{ji} notation signifies that we are taking the **complex conjugate** of the matrix elements of $\mathbf{A}$ and *interchanging rows and columns.* This interchange is known as the **transpose** if the elements of the matrix are *real* (show yourselves this).
The adjoint of a column matrix is a *row matrix* with the *complex conjugates* of the elements of hate column matrix:

$$\mathbf{a} \quad = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{pmatrix}$$

$$\mathbf{a}^{\dagger} \quad = \begin{pmatrix} a_1^* & a_2^* & a_3^* & \cdots & a_M^* \end{pmatrix}$$

We can determine the product of the two column matrices just described as (note each has dimension of M):

$$\begin{aligned} \mathbf{a}^{\dagger}\mathbf{b} \quad &= \begin{pmatrix} a_1^* & a_2^* & a_3^* & \cdots & a_M^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_M \end{pmatrix} \\ &= \textstyle\sum_{i=1}^{M} a_i^* b_i \end{aligned}$$

NOTE: If $\mathbf{a}$ and $\mathbf{b}$ are **real** and $M = 3$, we obtain the 3D scalar product of two vectors.

Here we list some useful properties relevant to **square** matrices (NXN).

- A **diagonal** matrix has zero off-diagonal elements: $A_{ij} = A_{ii}\ \delta_{ij}$
- The **trace** of a matrix is the sum of its diagonal elements $tr\mathbf{A} = \sum_i\ A_{ii}$

- The unit matrix is $\mathbf{1A} = \mathbf{A1} = \mathbf{A}$. where $(1)_{ij} = \delta_{ij}$
- The inverse of a matrix is such that $\mathbf{A^{-1}A} = \mathbf{AA^{-1}} = \mathbf{1}$
- A **unitary matrix** is one whose inverse is its adjoint $\mathbf{A^{-1}} = \mathbf{A}^\dagger$
- A **Hermitian matrix** is self-adjoint $\mathbf{A}^\dagger = \mathbf{A}$
- A real Hermitian matrix is **symmetric**

2 Generalization to N-Dimensional Complex Vector Spaces

We have outlined some fundamental ideas of 3-dimensional vector spaces familiar to everyone. For laying some of the conceptual foundations of quantum mechanics, we now consider generalizing those ideas to spaces of N-dimensions. We will also introduce Dirac's convenient bra/ket notation for representing vectors.

By analogy to 3 basis vectors in 3-d space, there are N basis vectors in N-dimensional complex vector spaces. They are mutually orthogonal and normalized, hence orthonormal. We represent the basis e_i $(i = 1, \cdots, N)$ by the symbol $\mid i\rangle$. These are the **ket** vectors, or kets.

A general vector is now written:

$$\mid a\rangle = \sum_{i=1}^{N} \mid i\rangle a_i$$

This is a generalization of what we have seen for the simple $N = 3$ case above. The vector is completely described once we supply the N a_i that give the projection of the vector onto the N basis vectors of the space.

We can write the matrix form of the vector as:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}$$

The adjoint of the column matrix is a row matrix:

$$\mathbf{a}^\dagger = (a_1^* \; a_2^* \; \cdots a_N^*)$$

The abstract **bra** vector $\langle a \mid$ is that whose matrix representation is $\mathbf{a}^\dagger$. The scalar product between bra and ket is defined as:

$$\begin{aligned}\langle a \mid\mid b\rangle = \langle a \mid b\rangle = \mathbf{a}^{\dagger}\mathbf{b} \\ = (a_1^* \ a_2^* \ \cdots a_N^*)\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} \\ = \textstyle\sum_{i=1}^{N} a_i^* b_i\end{aligned}$$

Keep in mind, the square of the length of a N-dimension vector is:

$$\langle a \mid a\rangle = \sum_{i=1}^{N} a_i^* a_i = \sum_{i=1}^{N} |a_i|^2$$

We can put forth similar mathematical statements for bra vectors:

$$\langle a \mid = \sum_i a_i^* \langle i \mid$$

The scalar product is then:

$$\langle a \mid b\rangle = \sum_j \sum_i a_i^* \langle i \mid j\rangle b_j$$

For this to be equivalent to our notion of a scalar product, what result naturally arises from the last expression?

$$\langle i \mid j\rangle = \delta_{ij}$$

This last is a statement of orthonormality of the basis.

2.1 Projections, Completeness, Operators

The projection elements of a vector onto the basis vectors are determined as we have seen before:

$$\begin{aligned}\langle j \mid a\rangle = \sum_i \langle j \mid i\rangle a_i = \sum_i \delta_{ji} a_i = a_j \\ \langle a \mid j\rangle = \sum_i a_i^* \langle i \mid j\rangle = \sum_i a_i^* \delta_{ij} = a_j^*\end{aligned}$$

Now we can determine the completeness relations:

$$| a\rangle = \sum_i | i\rangle a_i = \sum_i | i\rangle\langle i | a\rangle$$
$$\langle a |= \sum_i a_i^* \langle i |= \sum_i \langle a | i\rangle\langle i |$$

Thus, we obtain the completeness relation for the basis:

$$1 = \sum_i | i\rangle\langle i |$$

This is a powerful way to derive many useful relations in quantum mechanics. Keep in mind that that the result of the summation is a **scalar**.

An operator can be defined as an entity that acts on a ket to convert it to another ket:

$$\hat{O} | a\rangle =| b\rangle$$

The operator is completely determined if we know what it does to the basis $| i\rangle$:

$$\hat{O} | i\rangle = \sum_j | j\rangle O_{ji} = \sum_j | j\rangle \left(\mathbf{O}\right)_{ji}$$

O is the matrix representation of the operator $\hat{O}$ in the basis $| i\rangle$. The matrix elements are determined as:

$$\langle k | \hat{O} | i\langle= \sum_j \langle k | j\rangle \left(\mathbf{O}\right)_{ji} = \sum_j \delta_{kj} \left(\mathbf{O}\right)_{ji} = \left(\mathbf{O}\right)_{ki}$$

What is an alternate way to obtain the elements of the matrix representation **O** of the operator $\hat{O}$

$$\hat{O} | i\rangle = 1\hat{O} | i\rangle = sum_j | j\rangle\langle j | \hat{O} | i\rangle = sum_j | j\rangle$$

Thus

$$\langle| \hat{O} | i\rangle = \left(\mathbf{O}\right)_{ji} = O_{ji}$$

Let's consider another example of the usefulness of Dirac notation as well as the completeness of a basis. Take the matrix representation of a product of operators:

$$
\begin{aligned}
\langle i \mid \hat{C} \mid j\rangle = (\mathbf{C})_{ij} = \langle i \mid \hat{A}\hat{B} \mid j\rangle = \langle i \mid \hat{A}1\hat{B} \mid j\rangle \\
= \sum_k \langle i \mid \hat{A} \mid k\rangle\langle k \mid \hat{B} \mid j\rangle \\
= \sum_k (\mathbf{A})_{ik} (\mathbf{B})_{kj}
\end{aligned}
$$

The action of an adjoint of an operator acting on bra/ket vectors is analogous to the action of the vector as follows:

$$
\begin{aligned}
\hat{O} \mid a\rangle = \mid b\rangle \\
\langle a \mid \hat{O}^\dagger = \langle b \mid
\end{aligned}
$$

Show that:

$$
\langle i \mid \hat{O}^\dagger \mid j\rangle = \left(\mathbf{O}^\dagger\right)_{ij} = \langle j \mid \hat{O} \mid i\rangle^* = (\mathbf{O}^*)_{ji}
$$

Hermitian operators are **self-adjoint**:

$$
\hat{O} = \hat{O}^\dagger
$$

3 Change of Basis

We have introduced the idea of a **basis**. Recall that the choice of basis is not unique. We now consider how to convert a representation of a vector in one basis to its representation in another basis. This is relevant because certain operators may not have *diagonal* matrix representations in one basis, but we can transform the matrix to another basis, one in which the representation is diagonal. This is quite useful. This requires finding the relationship between the two bases. We now turn to this.

Consider that we have two bases $\mid i\rangle$ and $\mid \alpha\rangle$. We know:

$$
\begin{aligned}
\langle i \mid j\rangle = \delta_{ij} \\
\sum_i \mid i\rangle\langle i \mid = 1
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
\langle \alpha \mid \beta\rangle = \delta_{\alpha\beta} \\
\sum_\alpha \mid \alpha\rangle\langle \alpha \mid = 1
\end{aligned}
\tag{5}
$$

Since the completeness relations hold, we can represent any ket vector in the basis $\mid \alpha\rangle$ as a linear combination of kets in the basis $\mid i\rangle$ and vice versa. Thus,

$$\mid \alpha\rangle = 1 \mid \alpha\rangle = \sum_i \mid i\rangle\langle i \mid \alpha\rangle = \sum_i \mid i\rangle U_{i\alpha} = \sum_i \mid i\rangle(\mathbf{U})_{i\alpha}$$

The elements of the transformation matrix, $\mathbf{U}$, are evidently:

$$\langle i \mid \alpha\rangle = U_{i\alpha} = (\mathbf{U})_{i\alpha}$$

The analogous relation for representing the $\mid i\rangle$ basis in the $\mid \alpha\rangle$ basis is:

$$\mid i\rangle = 1 \mid i\rangle = \sum_\alpha \mid \alpha\rangle\langle \alpha \mid i\rangle = \sum_\alpha \mid \alpha\rangle U^*_{i\alpha} = \sum_\alpha \mid \alpha\rangle(\mathbf{U}^\dagger)_{\alpha i}$$

where,

$$\langle \alpha \mid i\rangle = \langle i \mid \alpha\rangle^* = U^*_{i\alpha} = (\mathbf{U}^\dagger)_{\alpha i}$$

NOTE *Because of the way* $\mathbf{U}$ *is defined,* $\langle \alpha \mid i\rangle \neq U_{\alpha i}$ *!*

We now show that an important property of this transformation matrix $\mathbf{U}$ is that is is **unitary**, $\mathbf{U^{-1}} = \mathbf{U}^\dagger$.

$$\begin{aligned} \delta_{ij} &= \langle i \mid j\rangle \\ &= \sum_\alpha \langle i \mid \alpha\rangle\langle \alpha \mid j\rangle \\ &= \sum_\alpha (\mathbf{U})_{i\alpha}(\mathbf{U}^\dagger)_{\alpha j} \\ &= (\mathbf{U}\mathbf{U}^\dagger)_{ij} \end{aligned}$$

This last is just another way of saying:

$$\mathbf{1} = \mathbf{U}\mathbf{U}^\dagger$$

This is just the definition of a unitary matrix!. We can start with $\langle \alpha \mid \beta\rangle = \delta_{\alpha\beta}$ relation to arrive at

$$\mathbf{1} = \mathbf{U}^\dagger\mathbf{U}$$

Now let's go back and see how we can relate the matrix representation of 2 operators in the 2 different bases. This has relevance to the Eigenvalue problem of the next section, as well as to quantum mechanics as well.

$$\hat{O} \mid i\rangle = \sum_j \mid j\rangle\langle j \mid \hat{O} \mid i\rangle = \sum_j \mid j\rangle O_{ji}$$
$$\hat{O} \mid \alpha\rangle = \sum_\beta \mid \beta\rangle\langle \beta \mid \hat{O} \mid \alpha\rangle = \sum_\beta \mid \beta\rangle \Omega_{\beta\alpha}$$

This obviously suggests:

$$\langle j \mid \hat{O} \mid i\rangle = O_{ji}$$
$$\langle \beta \mid \hat{O} \mid \alpha\rangle = \Omega_{\beta\alpha}$$

To find the explicit relationship between $\mathbf{O}$ and Ω,

$$\begin{aligned}
\Omega_{\alpha\beta} &= \langle \alpha \mid \hat{O} \mid \beta = \langle \alpha \mid 1\hat{O}1 \mid \beta\rangle \\
&= \textstyle\sum_{ij} \langle \alpha \mid i\rangle\langle i \mid \hat{O} \mid j\rangle\langle j \mid \beta\rangle \\
&= \textstyle\sum_{ij} (\mathbf{U}^\dagger)_{\alpha i} (\mathbf{O})_{ij} (\mathbf{U})_{j\beta}
\end{aligned}$$

Thus,

$$\mathbf{\Omega} = \mathbf{U}^\dagger \mathbf{O} \mathbf{U}$$
$$\mathbf{O} = \mathbf{U} \mathbf{\Omega} \mathbf{U}^\dagger$$

This shows that the matrix representations of the operators in the two bases are related by unitary transformation. The importance of these transformations is that **any Hermitian operator whose matrix representation in one basis is not diagonal, it is always possible to find another basis in which the matrix representation is diagonal**

$$\Omega_{\alpha\beta} = \omega_\alpha \delta_{\alpha\beta}$$

Problem Show that the *trace* of a matrix is invariant under a unitary transformation, that is

$$\mathbf{\Omega} = \mathbf{U}^\dagger \mathbf{O} \mathbf{U}$$

then show that $tr\mathbf{\Omega} = tr\mathbf{O}$.

4 The Eigenvalue Problem I: Definition

The Schroedinger equation is an eigenvalue problem. Here we discuss what this is. When an operator acts on a ket vector $\mid \alpha\rangle$, the result is in general another vector that is distinct from the original. If the result is *simply* a constant times the original vector, we have:

$$\hat{O} \mid \alpha\rangle = \omega_\alpha \mid \alpha\rangle$$

and we state that $\mid \alpha\rangle$ is an *eigenvector* of the operator $\hat{O}$ with an eigenvalue ω_α. We can choose the eigenvectors to be normalized ($\langle\alpha \mid \alpha\rangle = 1$).

The eigenvalues and eigenvectors of Hermitian operators ($\hat{O} = \hat{O}^\dagger$) have useful properties that are exploited in quantum mechanics (or rather fit nicely for the purposes of quantum treatment of matter):

- **The eigenvalues of a Hermitian operator are real**.

  $$\omega_\alpha = \langle\alpha \mid \hat{O} \mid \alpha\rangle = \langle\alpha \mid \hat{O}^\dagger \mid \alpha\rangle = \langle\alpha \mid \hat{O} \mid \alpha\rangle^* = \omega_\alpha^*$$

  For this to hold, the eigenvalues have to be real.

- **The eigenvectors of a Hermitian operator are orthogonal**

  $$\hat{O} \mid \beta\rangle = \omega_\beta \mid \beta\rangle$$

  We also have the adjoint relation:

  $$\begin{aligned} \langle\beta \mid \hat{O}^\dagger &= \langle\beta \mid \omega_\beta^* \\ \langle\beta \mid \hat{O} &= \langle\beta \mid \omega_\beta \end{aligned} \tag{6}$$

  Because of the Hermitian property of $\mathbf{O}$:

  $$\langle\beta \mid \hat{O} \mid \alpha\rangle - \langle\beta \mid \hat{O} \mid \alpha\rangle = (\omega_\beta - \omega_\alpha)\langle\beta \mid \alpha\rangle = 0$$

  Since we take the eigenvalues to be non-equivalent for the non-degenerate case, we see that the non-degenerate eigenvectors are orthogonal.

5 The Eigenvalue Problem II: Diagonalization

The matrix representation of a Hermitian operator in an **arbitrary** basis $\mid i\rangle$ is generally **not** diagonal. However, its matrix representation *in the basis formed by its eigenvectors* is *diagonal*. To show this:

$$\hat{O} \ \mid \alpha\rangle \ = \ \omega_\alpha \ \mid \alpha\rangle \qquad \textit{eigenvalue equation}$$

$$\langle\beta \mid \hat{O} \mid \alpha\rangle = \omega_\alpha \ \langle\beta \mid \alpha\rangle$$

$$\langle\beta \mid \hat{O} \mid \alpha\rangle = \omega_\alpha \ \delta_{\alpha\ \beta}$$

Thus, we pose the following problem as an *eigenvalue problem*:

Given the matrix representation of a Hermitian operator $\hat{O}$ in the orthonormal basis $\mid i\rangle, i = 1, 2, \cdots, N$ we wish to find the orthonormal basis $\mid \alpha\rangle, \alpha = 1, 2, \cdots, N$ in which the matrix representation of $\hat{\Omega}$ is *diagonal*, ($\Omega_{\alpha\beta} \ = \ \omega_{\alpha\beta} \ \delta_{\alpha\beta}$). That is, we would like to **diagonalize** the matrix representation of $\hat{O}$. Earlier we saw the relation between two representations of the operator $\hat{O}$ in two different bases through a unitary transformation:

$$\Omega = U^\dagger \ O \ U$$

The problem of diagonalizing the Hermitian matrix $\mathbf{O}$ is equivalent to the problem of *finding* the unitary matrix $\mathbf{U}$ that converts the matrix $\mathbf{O}$ into a diagonal matrix:

$$U^\dagger \ O \ U \ = \begin{pmatrix} \omega_1 & 0 & 0 & \cdots & 0 \\ 0 & \omega_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \cdots & 0 \\ 0 & \vdots & \vdots & \omega_{N-1} & \vdots \\ 0 & \cdots & 0 & 0 & \omega_N \end{pmatrix}$$

There exit numerous efficient methods to do this diagonalization. Here we dicusss the mechanics of the process using a more brute force approach, strictly for pedagogical purposes in the context of understanding diagonalization, not for optimizing computational algorithms.

Let's restate the eigenvalue problem as follows: *Given an N x N Hermitian matrix* $\mathbf{O}$*, we wish to find all distinct column vectors* $\mathbf{c}$ *(the eigenvectors of* $\mathbf{O}$*) and the corresponding numbers* ω *(the eigenvalues of* $\mathbf{O}$*) such that:*

$$\mathbf{O} \ \mathbf{c} \ = \ \omega \ \mathbf{c}(\mathbf{O} \ - \ \omega \ \mathbf{1}) \ \mathbf{c} \ = \ 0$$

The last equation has a nontrivial solution ($\mathbf{c} \neq \mathbf{0}$) only when the following holds:

$$|\mathbf{O} \; - \; \omega \; \mathbf{1}| \; = \; 0$$

where the above expression signifies that the determinant of the expression is equal to zero. This is called the **secular** determinant, and for the resulting polynomial of degree N (since we are dealing with an $N \; x \; N$ Hermitian operator and matrix), there will be N roots, $\omega_\alpha, \alpha = 1, 2, \cdots, N$, which are the eigenvalues of the matrix $\mathbf{O}$. The corresponding eigenvectors result from subustituting individually the eigenvalues into the defining equations of the eigenvalue problem. This leads to the eigenvectors, $\mathbf{c}^\alpha$ determined to within a multiplicative constant. The unique vectors are determined by normalization

$$\sum_i \; = \; (c_i^\alpha)^* \; c_i^\alpha \; = \; 1$$

Thus, we have our solutions for the eigenvectors and eigenvalues of the original matrix $\mathbf{O}$ representing the operator $\hat{O}$ in the original basis $\mid i\rangle, i = 1, 2, 3, \cdots, N$:

$$\mathbf{O} \; \mathbf{c}^\alpha \; = \; \omega_\alpha \; \mathbf{c}^\alpha$$

Since $\mathbf{O}$ is Hermitian, the eigenvalues are real and the eigenvectors are orthogonal

$$\sum_i \; = \; (c_i^\alpha)^* \; c_i^\beta \; = \; \delta_{\alpha\beta}$$

Now, let's finish the connection to the original discussion about the relation of the eigenvalues and eigenvectors of the arbitrary Hermitian matrix, $\mathbf{O}$ to unitary transformations.

Let's construct a matrix whose columns are the eigenvectors we have just determined for the matrix $\mathbf{O}$. We will call this matrix $\mathbf{U}$:

$$\mathbf{U} \; = \begin{pmatrix} c_1^1 & c_1^2 & \cdots & c_1^N \\ c_2^1 & c_2^2 & \cdots & c_2^N \\ \vdots & \vdots & \ddots & \vdots \\ c_N^1 & c_N^2 & \cdots & c_N^N \end{pmatrix} = \begin{pmatrix} \mathbf{c}^1 \; \mathbf{c}^2 \; \cdots \; \mathbf{c}^N \end{pmatrix}$$

It is clear that the α^{th} column of $\mathbf{U}$ is the eigenvector (as a column matrix), $\mathbf{c}^\alpha$, and so we can write:

$$(\mathbf{U})_{i\ \alpha} \ = \ U_{i\ \alpha} \ = \ c_i^{\alpha}$$

This allows us to write straightforwardly:

$$\mathbf{O}\ \mathbf{U} \ = \ \mathbf{U} \begin{pmatrix} \omega_1 & 0 & 0 & \cdots & 0 \\ 0 & \omega_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \cdots & 0 \\ 0 & \vdots & \vdots & \omega_{N-1} & \vdots \\ 0 & \cdots & 0 & 0 & \omega_N \end{pmatrix} \ = \ \mathbf{U}\ \omega$$

We also note that the orthonormality condition becomes:

$$\sum_i \ = \ (c_i^{\alpha})^* \ c_i^{\beta} \ = \ \sum_i U_{i\ \alpha}^* \, U_{i\ \beta} \ = \ \sum_i \left(\mathbf{U}^{\dagger}\right)_{\alpha\ i} (\mathbf{U})_{i\ \beta} \ = \ \delta_{\alpha\beta}$$

But the last relation is simply:

$$\mathbf{U}^{\dagger}\ \mathbf{U} \ = \ \mathbf{1}$$

The last relation allows us to write:

$$\mathbf{U}^{\dagger}\ \mathbf{O}\ \mathbf{U} \ = \ \omega$$

This gives the relation between the unitary transformation ($\mathbf{U}$) which diagonalizes the matrix $\mathbf{O}$ and the eigenvectors ($\mathbf{c}^{\alpha}$) of $\mathbf{O}$.

6 Exercise

Let's consider the following matrix and find the eigenvalues and eigenvectors for this arbitrary matrix representation of some operator.

$$\mathbf{O} \ = \ \begin{pmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{pmatrix}$$

6.1 Solution

We are looking for the eigenvalues and eigenvectors of the matrix. The restatement of the problem is:

$$\begin{pmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \omega \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

The secular determinant and associated second-order polynomial equation is:

$$\begin{vmatrix} O_{11} - \omega & O_{12} \\ O_{21} & O_{22} - \omega \end{vmatrix} = \omega^2 - \omega\,(O_{22} + O_{11}) + O_{11}O_{22} - O_{12}O_{21} = 0$$

The quadratic equation has two solutions; these are the **eigenvalues** of the matrix bf O:

$$\omega_1 = \frac{1}{2}\left[O_{11} + O_{22} - ((O_{22} - O_{11})^2 + 4O_{12}O_{21})^{1/2}\right]$$
$$\omega_2 = \frac{1}{2}\left[O_{11} + O_{22} + ((O_{22} - O_{11})^2 + 4O_{12}O_{21})^{1/2}\right]$$

Since we have two eigenvalues, there will be two eigenvectors (assuming non-degenerate eigenvalues). This requires us to determine the values of c_1 and c_2 for the cases associated with the two eigenvalues. For the case of ω_2, consider:

$$O_{11}\ c_1^2 + O_{12}\ c_2^2 = \omega_2\ c_1^2$$
$$O_{21}\ c_1^2 + O_{22}\ c_2^2 = \omega_2\ c_2^2$$

We can write analogous expressions for ω_1 (not shown here). One of the above equations coupled with the normalization condition:

$$(c_1^2)(c_1^2) + (c_2^2)(c_2^2) = 1$$

allows us to solve for the coefficients for the eigenvector associated with ω_2. For the particular case where $O_{11} = O_{22} = a$ and $O_{12} = O_{21} = b$, the eigenvalues are :

$$\omega_1 = a - b$$
$$\omega_2 = a + b$$

To find the eigenvector, we use :

$$O_{11}\ c_1^2 + O_{12}\ c_2^2 = \omega_2\ c_1^2$$
$$a\ c_1^2 + b\ c_2^2 = (a + b)\ c_1^2$$

which gives $c_1^2 = c_2^2$. The normalization condition gives $c_1^2 = c_2^2 = \frac{1}{\sqrt{2}}$. One can also solve this eigenvalue equation by directly finding the unitary transformation matrix, $\mathbf{U}$. Try this approach to see what you can come up with.

7 Orthogonal Functions, Eigenfunctions, and Operators

We have so far encountered the idea that vectors in N-dimensional spaces can be represented as linear combinations of basis vectors. By analogy to Fourier series, we now see how we can set up the formalism to treat *functions* as bases, and describe arbitrary functions as linear combinations of these bases functions. You have already encountered this in the idea of Fourier sine and cosine series representations of functions (recall that one can represent a well-behaved function by infinite sums of sine and cosine functions, or truncated sums to sufficient accuracy as dictated by the application or individual preference).

Consider an *infinite* set of functions $(\psi_i(x), = 1, 2, ...)$ that satisfy the orthonormality conditions on some interval $[x_1, x_2]$. (In quantum mechanics, you will encounter several examples of such mathematical functions):

$$\int_{x_1}^{x_2} dx\ \psi_i^*(x)\psi_j(x) = \delta_{ij}$$

In the following, we drop the integration limits. As in the case of a Fourier expansion, we now suppose that any function $a(x)$ can be expressed as a linear combination of the set of functions (ψ_i):

$$a(x) = \sum_i \psi_i(x) a_i$$

The basis $(\psi_i(x), = 1, 2, ...)$ is thus complete. The components of the basis functions, a_j are determined from:

$$\int dx \psi_j^*(x)\ a(x) = \sum_i \int dx\ \psi_j^*(x)\ \psi_i(x)\ a_i\ =\ \sum_i \delta_{ij}\ a_i\ =\ a_j$$

Substituting our expressions for the coefficients a_j in the expansion, we obtain:

$$a(x) = \int dx' \left[\sum_i \psi_i(x)\ \psi_i^*(x') \right] a(x')$$

The quantity in brackets is a functions of x and $x^{'}$. It serves to pick out $a(x)$ if it is multiplied by $a(x^{'})$ and integrated over all values of $x^{'}$. This is commonly referred to as the *Dirac delta function* $\delta(x - x^{'})$.

$$\sum_i \psi_i(x)\ \psi_i^*(x^{'}) = \delta(x - x^{'})$$

The Dirac delta **function** is a generalization of the discrete Kronecker delta as:

$$a_i = \sum_j \delta_{ij} a_j \leftrightarrow \ a(x) = \int\ dx^{'}\ \delta(x - x^{'}) a(x^{'})$$

also, as one might expect,

$$\delta(x^{'} - x) = \delta(x - x^{'})$$

We can get a little further insight into a useful property of the Dirac delta function via the following. If we let $x = 0$ in our definition of the Dirac delta function,

$$a(0) = \int\ dx^{'}\ \delta(x^{'}) a(x^{'})$$

Finally, taking $a(x^{'})$ to be 1, one obtains:

$$1 = \int\ dx^{'}\ \delta(x^{'})$$

This shows the familiar property of the Dirac delta function of having identically **unit area**. Furthermore, by multiplying a function $a(x)$ and integrating over any interval *containing* $x = 0$, it picks out the value of the function at $x = 0$.

7.1 Functions in the Language of Linear Algebra

To make the connection between the linear algebra of vectors and complete orthonormal functions, we first introduce the *shorthand notation*

$$\psi_i(x)\ = |\ i\rangle \quad \psi_i^*(x)\ =\ \langle i\ |$$

$$a(x)\ = |\ a\rangle \quad a^*(x)\ =\ \langle a\ |$$

The scalar product of 2 functions is :

$$\int dx\ a^*(x)\ b(x)\ =\ \langle a \mid b \rangle$$

We also have the following relations which should be fairly evident at this point:

$$\langle i \mid j \rangle = \delta_{ij}$$

$$\langle j \mid a \rangle = a_j$$

$$\mid a \rangle = \sum_i \ \mid i \rangle \langle i \mid a \rangle$$

Operators serve to transform functions into other functions:

$$\hat{O}\ a(x)\ =\ b(x)$$

$$\hat{O}\ \ \mid a \rangle\ = \mid b \rangle$$

The idea of eigenfunctions and eigenvalues is also evident for functions.

$$\hat{O}\ \phi_\alpha(x) = \omega_\alpha\ \phi_\alpha(x)$$

$$\hat{O}\ \ \mid \alpha \rangle\ =\ \omega_\alpha \mid \alpha \rangle$$

The ϕ_α are eigenfunctions of the operator, and the ω_α are eigenvalues. The eigenfunctions are normalized:

$$\int dx\ \phi^*_\alpha(x)\ \phi_\alpha(x)\ =\ \langle \alpha \mid \alpha \rangle = 1$$

and the eigenvalues can be represented as:

$$\omega_\alpha\ =\ \int\ dx\ \phi^*_\alpha(x)\ \hat{O}\ \phi_\alpha(x)\ =\ \langle \alpha \mid \hat{O} \mid \alpha \rangle$$

In general,

$$\omega_\alpha\ =\ \int\ dx\ \phi^*_\alpha(x)\ \hat{O}\ \phi_\beta(x)\ =\ \langle \alpha \mid \hat{O} \mid \beta \rangle$$

We finally comment on the properties of Hermitian operators acting on functions:

$$\int dx\ a^*(x)\ \hat{O}\ b(x) = \int dx\ b(x)(\hat{O}\ a(x))^* \ = \ \left(\int dx\ b^*(x)\ \hat{O}\ a(x)\right)^*$$
$$\langle a \mid \hat{O} \mid b\rangle \ = \ \langle b \mid \hat{O} \mid a\rangle^*$$