7

Eigenvectors and Hermitian Operators

7.1 Eigenvalues and Eigenvectors

Basic Definitions

Let $\mathcal{L}$ be a linear operator on some given vector space $\mathcal{V}$. A scalar λ and a nonzero vector $\mathbf{v}$ are referred to, respectively, as an *eigenvalue* and corresponding *eigenvector* for $\mathcal{L}$ if and only if

$$\mathcal{L}(\mathbf{v}) = \lambda\mathbf{v} \quad .$$

Equivalently, we can refer to an eigenvector $\mathbf{v}$ and its corresponding eigenvalue λ, or to the *eigen-pair* $(\lambda, \mathbf{v})$.

Do note that an eigenvector is required to be nonzero.[1] An eigenvalue, however, can be zero.

!►Example 7.1: *If $\mathcal{L}$ is the projection onto the vector $\mathbf{i}$, $\mathcal{L}(\mathbf{v}) = \overrightarrow{\text{pr}}_{\mathbf{i}}(\mathbf{v})$, then*

$$\mathcal{L}(\mathbf{i}) = \overrightarrow{\text{pr}}_{\mathbf{i}}(\mathbf{i}) = \mathbf{i} = 1 \cdot \mathbf{i} \qquad \text{and} \qquad \mathcal{L}(\mathbf{j}) = \overrightarrow{\text{pr}}_{\mathbf{i}}(\mathbf{j}) = \mathbf{0} = 0 \cdot \mathbf{j} \quad .$$

So, for this operator, $\mathbf{i}$ is an eigenvector with corresponding eigenvalue 1, and $\mathbf{j}$ is an eigenvector with corresponding eigenvalue 0.

!►Example 7.2: *Suppose $\mathcal{V}$ is a two-dimensional vector space with basis $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2\}$, and let $\mathcal{L}$ be the linear operator whose matrix with respect to $\mathcal{A}$ is*

$$\mathbf{L} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \quad .$$

Letting $\mathbf{v} = 2\mathbf{a}_1 + 3\mathbf{a}_2$, we see that

$$|\mathcal{L}(\mathbf{v})\rangle = \mathbf{L}\,|\mathbf{v}\rangle = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2+6 \\ 6+6 \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 3 \end{bmatrix} = 4\,|\mathbf{v}\rangle = |4\mathbf{v}\rangle \quad .$$

This shows that

$$\mathcal{L}(\mathbf{v}) = 4\mathbf{v} \quad ,$$

so $\mathbf{v}$ is an eigenvector for $\mathcal{L}$ with corresponding eigenvalue 4.

[1] This simply is to avoid silliness. After all, $\mathcal{L}(\mathbf{0}) = \mathbf{0} = \lambda\mathbf{0}$ for *every* scalar λ.

!►Example 7.3: *Let $\mathcal{V}$ be the vector space of all infinitely-differentiable functions, and let $\triangle$ be the differential operator $\triangle(f) = f''$. Observe that*

$$\triangle(\sin(2\pi x)) \;=\; \frac{d^2}{dx^2}\sin(2\pi x) \;=\; -4\pi^2 \sin(2\pi x) \quad .$$

Thus, for this operator, $-4\pi^2$ is an eigenvalue with corresponding eigenvector $\sin(2\pi x)$.[2]

?►Exercise 7.1: *Find other eigenpairs for $\triangle$.*

In practice, eigenvalues and eigenvectors are often associated with square matrices, with a scalar λ and a column matrix $\mathbf{V}$ being called an eigenvalue and corresponding eigenvector for a square matrix $\mathbf{L}$ if and only if

$$\mathbf{LV} \;=\; \lambda\mathbf{V} \quad .$$

For example, in example 7.2 we saw that

$$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} \;=\; 4\begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad .$$

Thus, we would say that matrix

$$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$$

has

$$\text{eigenvalue } \lambda = 4 \qquad \text{and} \qquad \text{corresponding eigenvector } \mathbf{V} \;=\; \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad .$$

This approach to eigenvalues and eigenvectors is favored by instructors and texts whose main interest is in getting their students to compute a bunch of eigenvalues and eigenvectors. Unfortunately, it also obscures the basis-independent nature of the theory, as well as the basic reasons we may be interested in eigen-thingies. It also pretty well limits us to the cases where the operators are only defined on finite-dimensional vector spaces. Admittedly, every linear operator on a finite-dimensional space can be described in terms of a square matrix with respect to some basis, and every square matrix can be viewed as a matrix with respect to some basis for some linear operator on some finite-dimensional vector space. So associating eigenpairs with matrices could be considered "the same" as associating eigenpairs with operators — provided we only consider finite-dimensional spaces. Ultimately, though, we want to consider infinite-dimensional spaces (when solving partial differential equations and generating generalized Fourier series). Also, it really is a good idea to associate eigenpairs with the single operator generating the matrices, rather than individual matrices, if only so that we can avoid having to prove such facts as

> *Assume $\mathbf{L}$ and $\mathbf{L}'$ are two square matrices related by a similarity transform. Then $\mathbf{L}$ and $\mathbf{L}'$ have the same eigenvalues, and their eigenvectors are related by the matrices used in the similarity transform.*

If we associate eigenpairs with operators, then the above will not need be proven. It will follow automatically from the change of basis formulas for the matrices of linear operators.

---

[2] When the eigenvector is actually a function, we often use the term *eigenfunction* instead of eigenvector.

All that said, you will often (even in this class) be asked to "compute the eigenvalues and corresponding eigenvectors for some $N \times N$ matrix." Just remember that we are really looking for the eigenpairs for the operator described by that matrix using some basis for some N-dimensional vector space, with the column matrix being called an eigenvector really being the components of a eigenvector with respect to that basis. And unless otherwise specified, assume the scalars for this unspecified vector space are complex.

Eigenspaces and Multiple Eigenvalues

Now suppose we have two eigenvectors $\mathbf{v}$ and $\mathbf{w}$ for some linear operator $\mathcal{L}$, with both corresponding to the same eigenvalue λ. If α and β are any two scalars, then

$$\mathcal{L}(\alpha\mathbf{v} + \beta\mathbf{w}) \;=\; \alpha\mathcal{L}(\mathbf{v}) \;+\; \beta\mathcal{L}(\mathbf{w}) \;=\; \alpha\lambda\mathbf{v} \;+\; \beta\lambda\mathbf{w} \;=\; \lambda\,[\alpha\mathbf{v} \;+\; \beta\mathbf{w}] \quad .$$

This shows that any linear combination of eigenvectors corresponding to a single eigenvalue is also an eigenvector corresponding to that eigenvalue (provided the linear combination doesn't happen to yield the zero vector). Consequently, the set of all eigenvectors corresponding to a single eigenvalue is a vector space (after tossing in the zero vector). We call this space the *eigenspace* corresponding to the given eigenvalue.

?►*Exercise 7.2:* *Consider the projection operator in example 7.1. What is the eigenspace corresponding to eigenvalue* 1 *? What is the eigenspace corresponding to eigenvalue* 0 *?*

In practice, this means that we can describe all eigenvectors for a given eigenvalue λ by simply giving a basis for its corresponding eigenspace. This also means that we can associate with each eigenvalue the dimension of its eigenspace. This leads to more terminology, namely:

1. The *geometric multiplicity* of an eigenvalue λ is the dimension of the corresponding eigenspace.[3]

2. An eigenvalue λ is *(geometrically) simple* if its corresponding eigenspace is one dimensional. Otherwise, λ is called a *multiple* or *repeated* eigenvalue.

If λ is a simple eigenvalue, then a basis for its eigenspace consists of a single vector $\mathbf{v}$, and its eigenspace is the set of all possible scalar multiples $\alpha\mathbf{v}$ of this vector. If λ is a repeated eigenvalue with multiplicity N, then any basis for its eigenspace will have N vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$. Naturally, if we are actually going to do something with these basis vectors, we will want the basis for each eigenspace to be orthonormal.

Finding Eigenpairs (Finite-Dimensional Case)

The goal is to find every scalar λ and every corresponding nonzero vector $\mathbf{v}$ satisfying

$$\mathcal{L}(\mathbf{v}) \;=\; \lambda\mathbf{v} \tag{7.1}$$

where $\mathcal{L}$ is some linear transformation. Note that this equation is completely equivalent to the equation

$$\mathcal{L}(\mathbf{v}) \;-\; \lambda\mathbf{v} \;=\; \mathbf{0} \quad . \tag{7.1$'$}$$

[3] We'll discuss a slightly different notion of "multiplicity" for eigenvalues ("algebraic multiplicity") in a few pages.

We will assume the underlying vector space is N-dimensional, and that we've already picked out some basis for this space. Letting $\mathbf{L}$ be the matrix for $\mathcal{L}$ with respect to this basis, we have that

$$|\mathcal{L}(\mathbf{v})\rangle = \mathbf{L}\,|\mathbf{v}\rangle \qquad \text{and} \qquad |\lambda\mathbf{v}\rangle = \lambda\,|\mathbf{v}\rangle = \lambda\mathbf{I}\,|\mathbf{v}\rangle$$

where, as you should also recall, $\mathbf{I}$ is the $N \times N$ identity matrix. So, in terms of components with respect to our basis, equations (7.1) and (7.1′) become

$$\mathbf{L}\,|\mathbf{v}\rangle = \lambda\,|\mathbf{v}\rangle \tag{7.2}$$

and

$$\mathbf{L}\,|\mathbf{v}\rangle - \lambda\mathbf{I}\,|\mathbf{v}\rangle = |\mathbf{0}\rangle \quad . \tag{7.2′}$$

If we have the value of λ, the first matrix equation, $\mathbf{L}\,|\mathbf{v}\rangle = \lambda\,|\mathbf{v}\rangle$, turns out to be nothing more that a system of N linear equations (the rows of the matrix equation) with N unknowns (the components of $\mathbf{v}$). That is easily solved by various methods you should already know.

To see how to find the possible values of λ, look at the second matrix equation, equation (7.2′). Factoring out the $|\mathbf{v}\rangle$ on the right, this becomes

$$[\mathbf{L} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle = |\mathbf{0}\rangle \quad .$$

If the matrix $\mathbf{L} - \lambda\mathbf{I}$ were invertible, then

$$|\mathbf{v}\rangle = [\mathbf{L} - \lambda\mathbf{I}]^{-1}\,|\mathbf{0}\rangle = |\mathbf{0}\rangle \quad ,$$

implying that $\mathbf{v}$ is the zero vector, contrary to the initial requirement that $\mathbf{v}$ be a nonzero vector. So, recalling an old test for the invertibility of a matrix (see page 4–11), we see that, for a scalar λ and nonzero vector $\mathbf{v}$,

$$\begin{aligned} & \mathcal{L}(\mathbf{v}) = \lambda\mathbf{v} \\ \iff \quad & [\mathbf{L} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle = |\mathbf{0}\rangle \\ \iff \quad & \mathbf{L} - \lambda\mathbf{I} \text{ is not invertible} \\ \iff \quad & \det(\mathbf{L} - \lambda\mathbf{I}) = 0 \quad . \end{aligned}$$

Computing out the last equation (involving the determinant) yields an N^{th} degree polynomial equation with λ as the unknown. This equation is called the *characteristic equation* for $\mathbf{L}$, and the polynomial obtained by computing out $\det(\mathbf{L} - \lambda\mathbf{I})$ is called the *characteristic polynomial* for $\mathbf{L}$. The possible values of λ, then, are the solutions to this equation.

All this leads to the following general procedure for finding all eigenpairs $(\lambda, \mathbf{v})$ for $\mathcal{L}$:

1. Pick a convenient basis for the space and find the corresponding matrix $\mathbf{L}$ for $\mathcal{L}$.

2. Solve the characteristic equation

$$\det(\mathbf{L} - \lambda\mathbf{I}) = 0$$

for all possible values of λ. Remember, unless otherwise stated, we are assuming scalars can be complex; so don't forget to also find the possible complex values for λ.

3. For each λ found, set

$$|\mathbf{v}\rangle = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} ,$$

where the v_k's are scalars "to be determined", and then solve the " $N \times N$ " linear system

$$\mathbf{L}\,|\mathbf{v}\rangle = \lambda\,|\mathbf{v}\rangle$$

for the possible v_k's. Equivalently, you could also solve

$$[\mathbf{L} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle = |\mathbf{0}\rangle \quad .$$

Once you have found the possible v_k's, write out the corresponding column matrices for $|\mathbf{v}\rangle$.

(Note 1: Because the matrix $\mathbf{L}-\lambda\mathbf{I}$ is not invertible, the above systems are degenerate, and you will get arbitrary constants in your solution. The total number of arbitrary constants will end up being the dimension of the corresponding eigenspace.)

(Note 2: In doing your computations, you may want to use symbols like x, y and z instead of v_1, v_2 and v_3.)

4. If desired (or demanded), find a basis — or even an orthonormal basis — for the eigenspace corresponding to each eigenvalue.

In practice (in assigned problems, at least), the first step (finding the matrix for the operator with respect to some basis) has usually already been done.

!►*Example 7.4:* *Assume we have a three-dimensional vector space, a basis* $\mathcal{A} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *for that space, and a linear transformation on that space whose matrix is*

$$\mathbf{L}_{\mathcal{A}} = \mathbf{L} = \begin{bmatrix} 1 & 0 & -5 \\ 0 & 6 & 0 \\ 1 & 0 & 7 \end{bmatrix} \quad .$$

Suppose, naturally, that we want to find all eigenvalues and corresponding eigenvectors for this operator.

First, we find the characteristic equation:

$$\det[\mathbf{L} - \lambda\mathbf{I}] = 0$$

$$\Longrightarrow \qquad \det\left(\begin{bmatrix} 1 & 0 & -5 \\ 0 & 6 & 0 \\ 1 & 0 & 7 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = 0$$

$$\Longrightarrow \qquad \det\left(\begin{bmatrix} 1 & 0 & -5 \\ 0 & 6 & 0 \\ 1 & 0 & 7 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}\right) = 0$$

$$\Longrightarrow \qquad \det\begin{bmatrix} 1-\lambda & 0 & -5 \\ 0 & 6-\lambda & 0 \\ 1 & 0 & 7-\lambda \end{bmatrix} = 0$$

$$\Longrightarrow \qquad (1-\lambda)(6-\lambda)(7-\lambda) + 5(6-\lambda) = 0$$

$$\Longrightarrow \qquad (6-\lambda)\left[(1-\lambda)(7-\lambda) + 5\right] = 0$$

$$\Longrightarrow \qquad (6-\lambda)\left[\lambda^2 - 8\lambda + 12\right] = 0 \quad .$$

Note that we did not completely multiply out the terms. Instead, we factored out the common factor $6-\lambda$ *to simplify the next step, which is to find all solutions to the characteristic equation. Fortunately for us, we can easily factor the rest of the polynomial, obtaining*

$$(6-\lambda)\left[\lambda^2 - 8\lambda + 12\right] = (6-\lambda)\left[(\lambda-2)(\lambda-6)\right] = -(\lambda-2)(\lambda-6)^2$$

as the completely factored form for our characteristic polynomial. Thus our characteristic equation $\det[\mathbf{L} - \lambda\mathbf{I}]$ *reduces to*

$$-(\lambda-2)(\lambda-6)^2 = 0 \quad ,$$

which we can solve by inspection. We have two eigenvalues,

$$\lambda = 2 \qquad \text{and} \qquad \lambda = 6$$

(with $\lambda = 6$ *being a "repeated" root of the characteristic polynomial).*

To find the eigenvectors corresponding to eigenvalue $\lambda = 2$ *, we first write out*

$$[\mathbf{L} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle = \mathbf{0}$$

with $\lambda = 2$ *and*

$$|\mathbf{v}\rangle = |\mathbf{v}\rangle_{\mathcal{A}} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad .$$

Then

$$[\mathbf{L} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle = \mathbf{0}$$

$$\Longrightarrow \qquad \left(\begin{bmatrix} 1 & 0 & -5 \\ 0 & 6 & 0 \\ 1 & 0 & 7 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \qquad \begin{bmatrix} 1-2 & 0 & -5 \\ 0 & 6-2 & 0 \\ 1 & 0 & 7-2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \qquad \begin{bmatrix} -1 & 0 & -5 \\ 0 & 4 & 0 \\ 1 & 0 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad .$$

So the unknowns x , y *and* z *must satisfy*

$$\begin{array}{rcrcrcl} -x & + & 0y & - & 5z & = & 0 \\ 0x & + & 4y & + & 0z & = & 0 \\ 1x & + & 0y & + & 5z & = & 0 \end{array} \quad .$$

Using whichever method you like, this is easily reduced to

$$y = 0 \qquad \text{and} \qquad x + 5z = 0 \quad .$$

We can choose either x *or* z *to be an arbitrary constant. Choosing* z *, we then must have* $x = -5z$ *to satisfy the last equation above. Thus, the eigenvectors corresponding to eigenvalue* $r = 2$ *are given by*

$$|\mathbf{v}\rangle = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -5z \\ 0 \\ z \end{bmatrix} = z \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}$$

where z *can be any nonzero value. That is, in terms of our basis vectors, the one-dimensional eigenspace corresponding to eigenvalue* $\lambda = 2$ *is the set of all constant multiples of*

$$\mathbf{v}_1 = -5\mathbf{e}_1 + \mathbf{e}_3 \quad .$$

To find the eigenvectors corresponding to eigenvalue $\lambda = 6$ *, we write out*

$$[\mathbf{L} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle = \mathbf{0}$$

with $\lambda = 6$ *and*

$$|\mathbf{v}\rangle = |\mathbf{v}\rangle_{\mathcal{A}} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad .$$

Then

$$[\mathbf{L} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle = \mathbf{0}$$

$$\Longrightarrow \qquad \left(\begin{bmatrix} 1 & 0 & -5 \\ 0 & 6 & 0 \\ 1 & 0 & 7 \end{bmatrix} - 6 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \qquad \begin{bmatrix} -5 & 0 & -5 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad ,$$

which reduces to

$$x + z = 0 \qquad \text{and} \qquad 0y = 0 \quad .$$

In this case, let us take x *to be an arbitrary constant with* $z = -x$ *so that the first equation is satisfied. The second equation is satisfied no matter what* y *is, so* y *is another arbitrary constant. Thus, the eigenvectors corresponding to eigenvalue* $\lambda = 2$ *are given by*

$$|\mathbf{v}\rangle = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ -x \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ -x \end{bmatrix} + \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

where x and y are arbitrary constants. And thus, also, we see that the eigenspace corresponding to eigenvalue $\lambda = 6$ is two-dimensional, and has basis $\{\mathbf{v}_2, \mathbf{v}_2\}$ with

$$\mathbf{v}_2 = \mathbf{e}_1 - \mathbf{e}_3 \qquad \textit{and} \qquad \mathbf{v}_3 = \mathbf{e}_2 \quad .$$

7.2 More Basic "Eigen-Theory"

Basis Independence of the Characteristic Polynomial

Before looking at the special cases where the operators are self adjoint, it is worthwhile to note that the characteristic polynomial and the corresponding characteristic equation described above are actually "basis independent". To see this, let $\mathcal{A}$ and $\mathcal{B}$ be any two bases for our (finite-dimensional) vector space $\mathcal{V}$, and let, as usual, $\mathbf{L}_{\mathcal{A}}$ and $\mathbf{L}_{\mathcal{B}}$ be the matrices of a linear operator $\mathcal{L}$ with respect to these two respective bases. From our discussion on "change of basis", we know these two matrices are related by

$$\mathbf{L}_{\mathcal{B}} = \mathbf{M}^{-1}\mathbf{L}_{\mathcal{A}}\mathbf{M}$$

where $\mathbf{M}$ is the matrix for the change of basis formulas ($\mathbf{M} = \mathbf{M}_{\mathcal{A}\mathcal{B}}$ and $\mathbf{M}^{-1} = \mathbf{M}_{\mathcal{B}\mathcal{A}} = \mathbf{M}_{\mathcal{A}\mathcal{B}}{}^{\dagger}$ if $\mathcal{A}$ and $\mathcal{B}$ are both orthonormal). Observe that

$$\begin{aligned}\mathbf{M}^{-1}\left[\mathbf{L}_{\mathcal{A}} - \lambda\mathbf{I}\right]\mathbf{M} &= \mathbf{M}^{-1}\left[\mathbf{L}_{\mathcal{A}}\mathbf{M} - \lambda\mathbf{M}\right] \\ &= \mathbf{M}^{-1}\mathbf{L}_{\mathcal{A}}\mathbf{M} - \lambda\mathbf{M}^{-1}\mathbf{M} = \mathbf{L}_{\mathcal{B}} - \lambda\mathbf{I} \quad .\end{aligned}$$

Using this, along with properties of determinants and the fact that the value of a determinant is just a number, we get

$$\begin{aligned}\det(\mathbf{L}_{\mathcal{B}} - \lambda\mathbf{I}) &= \det\!\left(\mathbf{M}^{-1}\left[\mathbf{L}_{\mathcal{A}} - \lambda\mathbf{I}\right]\mathbf{M}\right) \\ &= \det\!\left(\mathbf{M}^{-1}\right)\det(\mathbf{L}_{\mathcal{A}} - \lambda\mathbf{I})\det(\mathbf{M}) \\ &= \det(\mathbf{M})^{-1}\det(\mathbf{L}_{\mathcal{A}} - \lambda\mathbf{I})\det(\mathbf{M}) \\ &= \det(\mathbf{M})^{-1}\det(\mathbf{M})\det(\mathbf{L}_{\mathcal{A}} - \lambda\mathbf{I}) = \det(\mathbf{L}_{\mathcal{A}} - \lambda\mathbf{I}) \quad .\end{aligned}$$

Cutting out the distracting middle leaves us with

$$\det(\mathbf{L}_{\mathcal{A}} - \lambda\mathbf{I}) = \det(\mathbf{L}_{\mathcal{B}} - \lambda\mathbf{I}) \quad .$$

The determinants on the left and right of the last equation are the characteristic polynomials for $\mathbf{L}_{\mathcal{A}}$ and $\mathbf{L}_{\mathcal{B}}$, respectively, and the "$=$" between them tells us that these two polynomials are the same.

What this means is that we have the following theorem:

Theorem 7.1
Each linear operator $\mathcal{L}$ on an N-dimensional vector space $\mathcal{V}$ has a single associated N^{th} degree polynomial $P_{\mathcal{L}}$, and this polynomial is given by

$$P_{\mathcal{L}}(\lambda) = \det(\mathbf{L} - \lambda\mathbf{I})$$

where $\mathbf{L}$ *is the matrix for* $\mathcal{L}$ *with respect to any desired basis for* $\mathcal{V}$. *Moreover,* λ *is an eigenvalue* λ *for* $\mathcal{L}$ *if and only if*

$$P_{\mathcal{L}}(\lambda) = 0 \quad .$$

The polynomial $P_{\mathcal{L}}(\lambda)$ and the equation $P_{\mathcal{L}}(\lambda) = 0$ are, of course, referred to as the *characteristic polynomial* and *characteristic equation* for the operator $\mathcal{L}$. Changing the basis, and hence the matrix $\mathbf{L}$, does not change the polynomial $P_{\mathcal{L}}$.[4]

You should be aware that any N^{th} degree polynomial

$$P(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_N x^N$$

can be rewritten in "completely factored form"

$$P(x) = a_N (x - r_1)^{m_1} (x - r_2)^{m_2} \cdots (x - r_K)^{m_K} \quad .$$

The r_1, r_2, …, and r_K are all the different roots of P (i.e., solutions to $P(r) = 0$). Some, or all, of these roots may be complex numbers. Each m_j is a positive integer, called the (algebriac) multiplicity of r_j. Since each of the above expressions is supposed to represent an N^{th} degree polyomial, the number of roots, K, must be no larger than N with

$$m_1 + m_2 + \cdots + m_K = N \quad .$$

On occassion, it turns out to be more convenient to not "gather like terms" and to write the polynomial as

$$P(x) = a_N (x - x_1)(x - x_2)(x - x_3) \cdots (x - x_N)$$

where m_1 of the x_j's are r_1, m_2 of the x_j's are r_2, and so on.

Keeping in mind that $\mathbf{L}$ is an $N \times N$ matrix and $\mathbf{I}$ is the $N \times N$ identity matrix, it is easy to verify that the "completely factored form" for our characteristic polynomial

$$P_{\mathcal{L}}(\lambda) = \det(\mathbf{L} - \lambda\mathbf{I})$$

is

$$P_{\mathcal{L}}(\lambda) = \pm(\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_K)^{m_K}$$

where the λ_j's are all the different eigenvalues for $\mathcal{L}$, and the "$\pm$" will be "$+$" if the dimension N of the space is even, and "$-$" otherwise (in practice, though, the $\pm$ is irrelevant and usually dropped). For example, the completely factored form of the characteristic polynomial from our example is

$$-(\lambda - 2)(\lambda - 6)^2 \quad .$$

Not gathering like terms together, this is

$$-(\lambda - 2)(\lambda - 6)(\lambda - 6) \quad .$$

[4] Thus, the polynomial, itself, is a basis-free formula.

Eigenvectors as Basis Vectors

Let us now suppose we were clever (or lucky) enough so that we have a basis

$$\mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N \}$$

in which as least some of the $\mathbf{b}_k$'s are eigenvectors for our linear operator $\mathcal{L}$, and let us see what we can say about

$$\mathbf{L}_{\mathcal{B}} = \begin{bmatrix} L_{11} & L_{12} & L_{13} & \cdots & L_{1N} \\ L_{21} & L_{22} & L_{23} & \cdots & L_{2N} \\ L_{31} & L_{32} & L_{33} & \cdots & L_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_{N1} & L_{N2} & L_{N3} & \cdots & L_{NN} \end{bmatrix} ,$$

the matrix for $\mathcal{L}$ with respect to basis $\mathcal{B}$.

Remember, the components of the $\mathbf{b}_k$'s with respect to $\mathcal{B}$ end up simply being

$$|\mathbf{b}_1\rangle_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} , \quad |\mathbf{b}_2\rangle_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} , \quad |\mathbf{b}_3\rangle_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} , \quad \ldots .$$

Now observe:

$$\mathbf{b}_1 \text{ is an eigenvector for } \mathcal{L} \text{ with corresponding eigenvalue } \lambda_1$$

$$\iff \quad \mathcal{L}(\mathbf{b}_1) = \lambda_1 \mathbf{b}_1$$

$$\iff \quad \mathbf{L}_{\mathcal{B}} |\mathbf{b}_1\rangle_{\mathcal{B}} = \lambda_1 |\mathbf{b}_1\rangle_{\mathcal{B}}$$

$$\iff \quad \begin{bmatrix} L_{11} & L_{12} & L_{13} & \cdots & L_{1N} \\ L_{21} & L_{22} & L_{23} & \cdots & L_{2N} \\ L_{31} & L_{32} & L_{33} & \cdots & L_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L_{N1} & L_{N2} & L_{N3} & \cdots & L_{NN} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\iff \quad \begin{bmatrix} L_{11} \\ L_{21} \\ L_{31} \\ \vdots \\ L_{N1} \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} .$$

So,

$$\text{Basis vector } \mathbf{b}_1 \text{ is an eigenvector for } \mathcal{L} \text{ with corresponding eigenvalue } \lambda_1 \iff \begin{bmatrix} L_{11} \\ L_{21} \\ L_{31} \\ \vdots \\ L_{N1} \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} .$$

By similar computations, you can easily derive

$$\text{Basis vector } \mathbf{b}_2 \text{ is an eigenvector for } \mathcal{L} \text{ with corresponding eigenvalue } \lambda_2 \iff \begin{bmatrix} L_{12} \\ L_{22} \\ L_{32} \\ \vdots \\ L_{N2} \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} ,$$

$$\text{Basis vector } \mathbf{b}_3 \text{ is an eigenvector for } \mathcal{L} \text{ with corresponding eigenvalue } \lambda_3 \iff \begin{bmatrix} L_{13} \\ L_{23} \\ L_{33} \\ \vdots \\ L_{N3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \lambda_3 \\ \vdots \\ 0 \end{bmatrix} ,$$

and so on.

In summary,

Lemma 7.2

Let $\mathcal{L}$ be a linear operator on a finite dimensional vector space, and let $\mathbf{L}_{\mathcal{B}}$ be its matrix with respect to some basis $\mathcal{B}$. If the k^{th} vector in basis $\mathcal{B}$ is an eigenvector for $\mathcal{L}$ with corresponding eigenvalue λ_k, then the k^{th} column of $\mathbf{L}_{\mathcal{B}}$ is given by

$$[\mathbf{L}_{\mathcal{B}}]_{jk} = \lambda_k \delta_{jk} \qquad \text{for} \quad j = 1, 2, \ldots, N \quad .$$

All this shows that the matrix for a linear operator $\mathcal{L}$ is particularly simple when as many basis vectors as possible are also eigenvectors for the operator. Such a basis would be considered a "natural basis" for the operator. If we are very lucky, there will be N independent eigenvectors (with N being the dimension of the vector space). Using those eigenvectors as the basis $\mathcal{B}$, the matrix for $\mathcal{L}$ simplifies to the diagonal matrix

$$\mathbf{L}_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_N \end{bmatrix} .$$

When this happens, we sometimes say that we have a *complete set of eigenvectors*. This will certainly be the case if we have N *different* eigenvalues. Then each eigenvalue λ_k will have a corresponding eigenvector $\mathbf{b}_k$, and it is easy to verify that $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ will be a linearly independent set — and hence a basis for our N-dimensional vector space. On the other hand, if an eigenvalue λ_k is a repeated root of the characteristic equation, then the most one can generally prove is that

$$\text{dimension of the eigenspace corresponding to } \lambda_j \leq \text{ algebraic multiplicity of } \lambda_j \text{ as a root of the characteristic polynomial.}$$

As you should have seen in the homework, it is quite possible for

$$\text{dimension of the eigenspace corresponding to } \lambda_j \lneq \text{ algebraic multiplicity of } \lambda_j \text{ as a root of the characteristic polynomial.}$$

So there are linear operators and square matrices that do not have complete sets of eigenvectors (such operators or matrices are sometimes cruelly said to be *defective*).

At this point, it is worth noting that things simplify even more if our operator/matrix is self adjoint (i.e., Hermitian). Using the above lemma and the basic idea of "self adjointness", you can easily verify the following corollary to the above lemma:

Corollary 7.3
Let $\mathcal{H}$ be a self-adjoint linear operator on a finite dimensional vector space, and let $\mathbf{H}_{\mathcal{B}}$ be its matrix with respect to any orthonormal basis $\mathcal{B}$. If the k^{th} vector in basis $\mathcal{B}$ is an eigenvector for $\mathcal{H}$ with corresponding eigenvalue λ_k, then the k^{th} column of $\mathbf{H}_{\mathcal{B}}$ is given by

$$[\mathbf{H}_{\mathcal{B}}]_{jk} = \lambda_k \delta_{jk} \qquad \text{for} \quad j = 1, 2, \ldots, N \quad ,$$

and the k^{th} row of $\mathbf{H}_{\mathcal{B}}$ is given by

$$[\mathbf{H}_{\mathcal{B}}]_{kj} = \lambda_k \delta_{kj} \qquad \text{for} \quad j = 1, 2, \ldots, N \quad .$$

Since we will be using this corollary in a few pages, you really should verify it.

?►*Exercise 7.3:* *Verify the above corollary.*

"Diagonalizing" a Matrix

It is a fairly common exercise to "diagonalize" a given $N \times N$ matrix $\mathbf{L}$. Here are the basic ideas: You view the given matrix as the matrix of some linear operator $\mathcal{L}$ with respect to whatever orthonormal basis $\mathcal{A}$ we are using as a default for our N-dimensional vector space $\mathcal{V}$. So, in our more explicit notation

$$\mathbf{L} = \mathbf{L}_{\mathcal{A}} \quad .$$

We then find a linearly independent set of N eigenvectors $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ along with their corresponding eigenvalues λ_1, λ_2, …, and λ_N. Then, by the above, we know $\mathbf{L}_{\mathcal{B}}$ is a diagonal matrix,

$$\mathbf{L}_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_N \end{bmatrix} \quad ,$$

which, for possibly obvious reasons, we may wish to denote by $\mathbf{D}$. Part of the exercise is to also find the "change of basis" matrix $\mathbf{M}$ for computing $\mathbf{D}$ from $\mathbf{L}$ (i.e., computing $\mathbf{L}_{\mathcal{B}}$ from $\mathbf{L}_{\mathcal{A}}$) via

$$\mathbf{D} = \mathbf{M}^{-1}\mathbf{L}\mathbf{M} \quad .$$

This gives us

1. a natural basis for the operator,

2. the simplest possible matrix for this operator,

3. and the matrices for converting from the original basis to this natural basis and back again.

Of course, getting that diagonal matrix requires that the original matrix/operator has a complete set of eigenvectors to serve as that basis of eigenvectors for the operator. If we are lucky, the matrix/operator will have a complete set. If we are very lucky, we can even find an orthonormal set. Then the change of basis matrix $\mathbf{M}$ will simply be the unitary matrix $\mathbf{M}_{\mathcal{AB}}$ (and $\mathbf{M}^{-1} = \mathbf{M}_{\mathcal{BA}} = \mathbf{M}_{\mathcal{AB}}{}^{\dagger}$).

Just why we would want all this depends on the application. If, for example, this operator is the "moment of inertia tensor" for some physical object, then the natural basis gives you the principle axes ("axes of symmetry") through the center of mass of the object.

In any case, of course, it's just plain easier to do matrix computations with diagonal matrices.

7.3 Self-Adjoint/Hermitian Operators and Matrices
Some General Background Stuff

Let us quickly recall a few definitions and facts:

1. If $\mathbf{A}$ is a matrix, then its adjoint $\mathbf{A}^{\dagger}$ is the transpose of its conjugate. So

$$\left[\mathbf{A}^{\dagger}\right]_{jk} = \left([\mathbf{A}]_{kj}\right)^{*} \quad .$$

2. Any matrix $\mathbf{A}$ that equals its own adjoint (i.e., $\mathbf{A} = \mathbf{A}^{\dagger}$) is said to be self adjoint or, equivalently, Hermitian.

3. More generally, if $\mathcal{L}$ is a linear operator on some vector space $\mathcal{V}$, then its adjoint $\mathcal{L}^{\dagger}$ is defined to be the linear operator satisfying

$$\left\langle \mathcal{L}^{\dagger}(\mathbf{v}) \,\middle|\, \mathbf{a} \right\rangle = \langle \mathbf{v} \mid \mathcal{L}(\mathbf{a}) \rangle \qquad \text{for every} \quad \mathbf{a}, \mathbf{v} \in \mathcal{V} \quad .$$

It should be noted that, since $\langle \mathbf{b} \,|\, \mathbf{a} \rangle = \langle \mathbf{a} \,|\, \mathbf{b} \rangle^{*}$, the above equation is completely equivalent to

$$\left\langle \mathbf{a} \,\middle|\, \mathcal{L}^{\dagger}(\mathbf{v}) \right\rangle = \langle \mathcal{L}(\mathbf{a}) \mid \mathbf{v} \rangle \qquad \text{for every} \quad \mathbf{a}, \mathbf{v} \in \mathcal{V} \quad .$$

If $\mathcal{V}$ is finite dimensional and $\mathcal{B}$ is an orthonormal basis, then the matrix for $\mathcal{L}^{\dagger}$ with respect to $\mathcal{B}$ is simply $\mathbf{L}^{\dagger}$ where $\mathbf{L}$ is the matrix for $\mathcal{L}$ with respect to $\mathcal{B}$.

4. Any linear operator $\mathscr{L}$ that equals its own adjoint (i.e., $\mathscr{L} = \mathscr{L}^\dagger$) is said to be self adjoint or, equivalently, Hermitian.

5. If $\mathscr{L}$ is a linear operator on some finite dimensional vector space, and $\mathbf{L}$ is its matrix with respect to some orthonormal basis, then

$$\mathscr{L} \text{ is a Hermitian operator} \iff \mathbf{L} \text{ is a Hermitian matrix} \quad .$$

6. If the scalars involved are simply the real numbers, then "adjoints" reduce to "transposes", and Hermitian/self-adjoint matrices and operators can be referred to as being "symmetric".

There are several reasons to be interested in Hermitian operators and matrices. Here are a few:

1. There are some nice (and useful) "mathematics" we can derive regarding the eigenvalues and eigenvectors of Hermitian operators and matrices.

2. Matrices and transformations arising in many applications are Hermitian (or even symmetric) because of natural symmetries. As a pseudo-example, consider a matrix $\mathbf{A}$ defined by

$$[\mathbf{A}]_{jk} = \frac{\partial^2 \phi}{\partial x_j \partial x_k}$$

where ϕ is some (sufficiently differentiable) real-valued function of several variables $(x_1, x_2, \ldots, x_N)$. Because the order in which partial derivatives are computed does not affect the final result, we have that

$$[\mathbf{A}]_{jk} = \frac{\partial^2 \phi}{\partial x_j \partial x_k} = \frac{\partial^2 \phi}{\partial x_k \partial x_j} = [\mathbf{A}]_{kj} \quad .$$

A better example is given by the moment of inertia tensor/matrix for any given object — see example 2 on page 299 of Arfken, Weber and Harris. Unfortunately, they treat the moment of inertia as a matrix (and use unfortunate notation). While reading this example, try to mentally replace their moment of inertia matrix $\mathbf{I}$ with an operator $\mathscr{I}$ such that, if the given body is rotating about its center of mass with angular velocity $\boldsymbol{\omega}$, then its angular momentum and the energy in the rotation are given by

$$\mathscr{I}(\boldsymbol{\omega}) \qquad \text{and} \qquad \langle\, \boldsymbol{\omega} \mid \mathscr{I}(\boldsymbol{\omega})\, \rangle \quad ,$$

respectively. The matrix $\mathbf{I}$ they have is the matrix of $\mathscr{I}$ with respect to some convenient orthonormal basis. Change the basis (say, to one based on the major axes of rotation of the object) and you have to change the matrix, but not the operator.

3. The Sturm-Liouville theory we will develop for solving partial differential equations is the infinite dimensional version of what we will develop here. In fact, we will use some of what we develop here.

Eigenpairs for Hermitian Operators and Matrices

Let $\mathscr{H}$ be an Hermitian operator on some vector space $\mathcal{V}$. Remember, this means $\mathscr{H}^\dagger = \mathscr{H}$. So, for every pair of vectors $\mathbf{v}$ and $\mathbf{w}$,

$$\langle\, \mathbf{v} \mid \mathscr{H}(\mathbf{w})\, \rangle = \left\langle\, \mathscr{H}^\dagger(\mathbf{v}) \,\middle|\, \mathbf{w}\, \right\rangle = \langle\, \mathscr{H}(\mathbf{v}) \mid \mathbf{w}\, \rangle \quad .$$

Cutting out the middle, gives

$$\langle \mathbf{v} \mid \mathcal{H}(\mathbf{w}) \rangle = \langle \mathcal{H}(\mathbf{v}) \mid \mathbf{w} \rangle \quad , \tag{7.3}$$

an equation important enough to be given a number.

In particular, now, let's assume $\mathbf{v}$ is an eigenvector with corresponding eigenvalue λ. Replacing $\mathbf{w}$ in the last equation with $\mathbf{v}$, using the fact that $\mathcal{H}(\mathbf{v}) = \lambda\mathbf{v}$, and recalling the properties of inner products, we have the following:

$$\begin{aligned} & \langle \mathbf{v} \mid \mathcal{H}(\mathbf{v}) \rangle = \langle \mathcal{H}(\mathbf{v}) \mid \mathbf{v} \rangle \\ \Longrightarrow \quad & \langle \mathbf{v} \mid \lambda\mathbf{v} \rangle = \langle \lambda\mathbf{v} \mid \mathbf{v} \rangle \\ \Longrightarrow \quad & \lambda \langle \mathbf{v} \mid \mathbf{v} \rangle = \lambda^* \langle \mathbf{v} \mid \mathbf{v} \rangle \\ \Longrightarrow \quad & \lambda = \lambda^* \quad . \end{aligned}$$

But the only way a scalar can be equal to its complex conjugate is for that scalar to be just a real number. So we have just (rigorously) derived

Lemma 7.4 ('realness' of Hermitian eigenvalues)
All the eigenvalues for a Hermitian operator must be real values.

Now let us suppose $\mathbf{v}$ and $\mathbf{w}$ are two eigenvectors corresponding to two different eigenvalues λ and μ with $\mathcal{H}(\mathbf{v}) = \lambda\mathbf{v}$ and $\mathcal{H}(\mathbf{w}) = \mu\mathbf{w}$. This, combined with equation (7.3) gives the following:

$$\begin{aligned} & \langle \mathbf{v} \mid \mathcal{H}(\mathbf{w}) \rangle = \langle \mathcal{H}(\mathbf{v}) \mid \mathbf{w} \rangle \\ \Longrightarrow \quad & \langle \mathbf{v} \mid \mu\mathbf{w} \rangle = \langle \lambda\mathbf{v} \mid \mathbf{w} \rangle \\ \Longrightarrow \quad & \mu \langle \mathbf{v} \mid \mathbf{w} \rangle = \lambda^* \langle \mathbf{v} \mid \mathbf{w} \rangle \\ \Longrightarrow \quad & (\mu - \lambda^*) \langle \mathbf{v} \mid \mathbf{w} \rangle = 0 \quad . \end{aligned}$$

Since we just saw that the eigenvalues of $\mathcal{H}$ are real, $\lambda^* = \lambda$ and the last equation becomes

$$(\mu - \lambda) \langle \mathbf{v} \mid \mathbf{w} \rangle = 0 \quad ,$$

which means that either

$$\mu - \lambda = 0 \qquad \text{or} \qquad \langle \mathbf{v} \mid \mathbf{w} \rangle = 0 \quad .$$

The first must be ruled out because we are assuming μ and λ are different values. This leaves us with

$$\langle \mathbf{v} \mid \mathbf{w} \rangle = 0$$

verifying

Lemma 7.5 (orthogonality of eigenvectors for Hermitian operators)
Any pair of eigenectors corresponding to different eigenvalues for a Hermitian operator must be orthogonal.

For future applications, note that we did not assume the vector space is finite dimensional in the above.

The Completeness of the Set of Eigenvectors

Now let consider a Hermitian operator $\mathcal{H}$ on a vector space $\mathcal{V}$ of finite dimension N.

If $\mathcal{H}$ has N different eigenvalues, then it has N different corresponding eigenvectors

$$\{ \mathbf{v}_1 , \mathbf{v}_2 , \ldots , \mathbf{v}_N \}$$

As we just saw (lemma 7.5), this will be an orthogonal set of nonzero vectors. Hence, this set will be a basis for our N-dimensional vector space. Dividing each $\mathbf{v}_j$ by its length then yields an orthonormal basis,

$$\{ \mathbf{u}_1 , \mathbf{u}_2 , \ldots , \mathbf{u}_N \} \qquad \text{where} \quad \mathbf{u}_j = \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|} \quad .$$

But what if one or more of the eigenvalues are repeated?

Two-Dimensional Case

Assume $\mathcal{H}$ is an Hermitian operator on a two-dimensional vector space $\mathcal{V}$. The characteristic polynomial for $\mathcal{H}$ can be written

$$P_{\mathcal{H}}(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)$$

where λ_1 and λ_2 are the eigenvalues of $\mathcal{H}$. Let $\mathbf{v}_1$ be an eigenvector corresponding to λ_1. Since $\mathcal{V}$ is two dimensional, we can find another vector $\mathbf{v}_2$ perpendicular to $\mathbf{v}_1$, and

$$\mathcal{B} = \{ \mathbf{u}_1 , \mathbf{u}_2 \} \qquad \text{where} \quad \mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \quad \text{and} \quad \mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}$$

will be an orthonormal basis for our two-dimensional space, with $(\lambda_1, \mathbf{u}_1)$ being an eigenpair for $\mathcal{H}$.

Now let $\mathbf{H}_{\mathcal{B}}$ be the matrix for $\mathcal{H}$ with respect to this basis. Since $(\lambda_1, \mathbf{u}_1)$ is an eigenpair for $\mathcal{H}$ we know (corollary 7.3 on page 7–12) that

$$\mathbf{H}_{\mathcal{B}} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & H_{22} \end{bmatrix} \quad .$$

which means that

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = P_{\mathcal{H}}(\lambda) = \det \begin{bmatrix} \lambda_1 - \lambda & 0 \\ 0 & H_{22} - \lambda \end{bmatrix} = (\lambda - \lambda_1)(\lambda - H_{22}) \quad .$$

This tells us that $H_{22} = \lambda_2$. So, in fact,

$$\mathbf{H}_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad ,$$

which, in turn, means that

$$|\mathcal{H}(\mathbf{u}_2)\rangle_{\mathcal{B}} = \mathbf{H}_{\mathcal{B}} \, |\mathbf{u}_2\rangle_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \lambda_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |\lambda_2 \mathbf{u}_2\rangle_{\mathcal{B}} \quad ,$$

telling us that $\mathbf{u}_2$ is an eigenvector corresponding to λ_2.

Thus,

If $\mathcal{H}$ is a Hermitian operator on a two-dimensional vector space $\mathcal{V}$, then

$$P_{\mathcal{H}}(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)$$

where λ_1 and λ_2 are the two eigenvalues for $\mathcal{H}$ (not necessarily different), and $\mathcal{V}$ has an orthonormal basis

$$\{ \mathbf{u}_1, \mathbf{u}_2 \}$$

with each $\mathbf{u}_k$ being an eigenvector for $\mathcal{H}$ corresponding to eignevalue λ_k.

Now let us verify that a similar statement can be made when $\mathcal{V}$ is three-dimensional. Many of the details will be left to you.

Three-Dimensional Case

Assume $\mathcal{H}$ is an Hermitian operator on a three-dimensional vector space $\mathcal{V}$. We know the characteristic polynomial for $\mathcal{H}$ can be written

$$P_{\mathcal{H}}(\lambda) = -(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$$

where λ_1, λ_2 and λ_3 are the (not necessarily different) eigenvalues of $\mathcal{H}$. Let $\mathbf{v}_3$ be an eigenvector corresponding to λ_3. By a straightforward modification of the Gram-Schmidt procedure, we can find an orthonormal basis for $\mathcal{V}$

$$\mathcal{B} = \{ \mathbf{w}_1, \mathbf{w}_2, \mathbf{u}_3 \}$$

with

$$\mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} \quad .$$

So $(\lambda_3, \mathbf{u}_3)$ will be an eigen-pair for $\mathcal{H}$.

Now let $\mathbf{H}_{\mathcal{B}}$ be the matrix for $\mathcal{H}$ with respect to this basis.

?►*Exercise 7.4:* *Show that the matrix for $\mathcal{H}$ with respect to $\mathcal{B}$ is*

$$\mathbf{H}_{\mathcal{B}} = \begin{bmatrix} H_{11} & H_{12} & 0 \\ H_{21} & H_{22} & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad ,$$

with

$$\mathbf{H}_0 = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$$

being a 2×2 Hermitian matrix.

Now let $\mathcal{V}_0$ be the two-dimensional subspace of $\mathcal{V}$ with basis

$$\mathcal{B}_0 = \{ \mathbf{w}_1, \mathbf{w}_2 \} \quad ,$$

and let $\mathcal{H}_0$ be the linear operator on $\mathcal{V}_0$ whose matrix with respect to $\mathcal{B}_0$ is $\mathbf{H}_0$. Since $\mathbf{H}_0$ is Hermitian, so is $\mathcal{H}_0$.

?►Exercise 7.5: *Starting with the fact that*

$$P_{\mathcal{H}}(\lambda) \;=\; \det(\mathbf{H}_{\mathcal{B}} - \lambda\mathbf{I}) \quad ,$$

show that the characteristic polynomial for $\mathcal{H}_0$ *is*

$$P_{\mathcal{H}_0}(\lambda) \;=\; (\lambda - \lambda_1)(\lambda - \lambda_2) \quad .$$

?►Exercise 7.6: *Let* $\mathbf{a}$ *be any vector in the subspace* $\mathcal{V}_0$ *. Show that:*

a: *The set* $\{\mathbf{a}, \mathbf{u}_3\}$ *is orthogonal.*

b: $\mathcal{H}(\mathbf{a}) = \mathcal{H}_0(\mathbf{a})$ *. (Hint: Use the matrices* $\mathcal{H}_{\mathcal{B}}$ *and* $\mathcal{H}_0$ *along with the components of* $\mathbf{a}$ *with respect to basis* $\mathcal{B}$ *.)*

c: *If* $\mathbf{a}$ *is an eigenvector for* $\mathcal{H}_0$ *with eigenvalue* λ *, then* $\mathbf{a}$ *is also an eigenvector for* $\mathcal{H}$ *with eigenvalue* λ *.*

Now, by our discussion of the "two-dimensional case", above, we know there is an orthonormal basis for $\mathcal{V}_0$

$$\{\, \mathbf{u}_1 ,\ \mathbf{u}_2 \,\}$$

such that $(\lambda_1, \mathbf{u}_1)$ and $(\lambda_2, \mathbf{u}_2)$ are eigen-pairs for $\mathcal{H}_0$. As you just verified in the last exercise, $(\lambda_1, \mathbf{u}_1)$ and $(\lambda_2, \mathbf{u}_2)$ are then eigen-pairs for $\mathcal{H}$ as well. Moreover, from the first part of the last exercise, it also follows that

$$\{\, \mathbf{u}_1 ,\ \mathbf{u}_2 ,\ \mathbf{u}_3 \,\} \quad .$$

is an orthonormal basis for $\mathcal{V}$. And since we started with $(\lambda_3, \mathbf{u}_3)$ being an eigen-pair for $\mathcal{H}$, it should be clear that we have the following:

> *If* $\mathcal{H}$ *is a Hermitian operator on a three-dimensional vector space* $\mathcal{V}$ *, then*
>
> $$P_{\mathcal{H}}(\lambda) \;=\; -(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$$
>
> *where* λ_1 *,* λ_2 *and* λ_3 *are the three eigenvalues for* $\mathcal{H}$ *(not necessarily different), and* $\mathcal{V}$ *has an orthonormal basis*
>
> $$\{\, \mathbf{u}_1 ,\ \mathbf{u}_2 ,\ \mathbf{u}_3 \,\}$$
>
> *with each* $\mathbf{u}_k$ *being an eigenvector for* $\mathcal{H}$ *corresponding to eignevalue* λ_k *.*

The General Case

Continuing the above yields:

Lemma 7.6
If $\mathcal{H}$ *is a Hermitian operator on an* N*-dimensional vector space* $\mathcal{V}$ *, then*

$$P_{\mathcal{H}}(\lambda) \;=\; \pm(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)\cdots(\lambda - \lambda_N)$$

where λ_1 *,* λ_2 *,* $\ldots$ *and* λ_N *are the N eigenvalues for* $\mathcal{H}$ *(not necessarily different), and* $\mathcal{V}$ *has an orthonormal basis*

$$\{\, \mathbf{u}_1 ,\ \mathbf{u}_2 ,\ \ldots ,\ \mathbf{u}_N \,\}$$

with each $\mathbf{u}_k$ being an eigenvector for $\mathcal{H}$ corresponding to eigenvalue λ_k .

It is worth noting that if, say, m of the above λ_j's are the same value, then the lemma assures us that there are m corresponding $\mathbf{u}_j$'s in a basis for $\mathcal{V}$. Hence:

Corollary 7.7
If $\mathcal{H}$ is a Hermitian operator on an N-dimensional vector space $\mathcal{V}$ with characteristic polynomial

$$P_{\mathcal{H}}(\lambda) = \pm(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}(\lambda - \lambda_3)^{m_3} \cdots (\lambda - \lambda_K)^{m_K}$$

where λ_1 , λ_2 , ... and λ_K are all the different eigenvalues $\mathcal{H}$, then the dimension of the eigenspace for each λ_j is the algebraic multiplicity, m_j , of that eigenvalue.

In other words, if λ is an eigenvalue for a Hermitian operator, then

The dimension of the eigenspace corresponding to λ
$=$ the algebraic multiplicity of λ as a root of the characteristic polynomial.

Compare this to the corresponding statement made a few pages ago (page 7–12) for the case where λ is an eigenvalue for an arbitrary linear operator on some vector space:

The dimension of the eigenspace corresponding to λ_j
$\leq$ the algebraic multiplicity of λ_j as a root of the characteristic polynomial.

Summary of the Big Results

We have too many lemmas regarding Hermitian operators. Let's summarize them in one theorem so that, in the future, we can forget the lemmas and just refer to the one theorem.

Theorem 7.8 (Big Theorem on Hermitian Operators)
Let $\mathcal{H}$ be a Hermitian (i.e., self-adjoint) operator on a vector space $\mathcal{V}$. Then:

1. *All eigenvalues of $\mathcal{H}$ are real.*
2. *Any pair of eigenvectors corresponding to different eigenvalues is orthogonal.*

Moreover, if $\mathcal{V}$ is finite dimensional, then

1. *If λ is an eigenvalue of algebraic multiplicity m in the characteristic polynomial, then we can find an orthonormal set of exactly m eigenvectors whose linear combinations generate all other eigenvectors corresponding to λ .*
2. *$\mathcal{V}$ has an orthonormal basis consisting of eigenvectors for $\mathcal{H}$.*

7.4 Diagonalizing Hermitian Matrices and Operators
General Ideas and Procedures

The above big theorem on Hermitian operators, along with what we know know about "change of basis" (especially theorem 6.1 on page 6–8) and the matrix of an operator with respect to a basis of eigenvectors (lemma 7.2 on page 7–11 or corollary 7.3 on page 7–12), gives us the following result on diagonalizing Hermitian operators/matrices.

Corollary 7.9
Let $\mathcal{H}$ be a Hermitian operator on a vector space $\mathcal{V}$ of finite dimension N, and let $\mathbf{H}_{\mathcal{A}}$ be the matrix for $\mathcal{H}$ with respect to some orthonormal basis $\mathcal{A}$. Let

$$\mathcal{U} = \{ \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N \}$$

be an orthonormal basis for $\mathcal{V}$ consisting of eigenvectors for H. Then the matrix for $\mathcal{H}$ with respect to this basis of eigenvectors is given by

$$\mathbf{H}_{\mathcal{U}} = \mathbf{M}_{\mathcal{U}\mathcal{A}}\mathbf{H}_{\mathcal{A}}\mathbf{M}_{\mathcal{A}\mathcal{U}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

where each λ_j is the eigenvalue corresponding to eigenvector $\mathbf{u}_j$, and $\mathbf{M}_{\mathcal{A}\mathcal{U}}$ and $\mathbf{M}_{\mathcal{U}\mathcal{A}}$ are the unitary change of bases matrices described in theorem 5.3 on page 5–9.

In light of this corollary, "diagonalizing a Hermitian operator $\mathcal{H}$" simply means the finding of an orthonormal basis of eigenvectors $\mathcal{U}$ so that the matrix of this operator with respect to this basis is the diagonal matrix $\mathbf{H}_{\mathcal{U}}$ given in the corollary.

If on the other hand, we are asked to "diagonalize a Hermitian matrix $\mathbf{H}$", then we are to find a unitary matrix $\mathbf{U}$ and a diagonal matrix $\mathbf{D}$ so that

$$\mathbf{D} = \mathbf{U}^{\dagger}\mathbf{H}\mathbf{U} \quad .$$

The above corollary tells us that, treating $\mathbf{H}$ as $\mathbf{H}_{\mathcal{A}}$, the matrix for a corresponding operator with respect to the "original basis" , we can use

$$\mathbf{D} = \mathbf{H}_{\mathcal{U}} \qquad \text{and} \qquad \mathbf{U} = \mathbf{M}_{\mathcal{A}\mathcal{U}} \quad .$$

Let us now consider the mechanics of "diagonalizing a Hermitian matrix/operator". Assume we have either a Hermitian operator $\mathcal{H}$ on an N-dimensional vector space $\mathcal{V}$, or an $N \times N$ Hermitian matrix $\mathbf{H}$.

If it is an operator $\mathcal{H}$ that we have, then the first thing is to find its matrix $\mathbf{H} = \mathbf{H}_{\mathcal{A}}$ with respect some convenient orthonormal basis

$$\mathcal{A} = \{ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N \} \quad .$$

If, instead, we are starting with a matrix $\mathbf{H}$, then the first thing is to view it as $\mathbf{H} = \mathbf{H}_{\mathcal{A}}$, the matrix of a Hermitian operator $\mathcal{H}$ with respect to some convenient orthonormal basis

$$\mathcal{A} = \{ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N \} \quad .$$

In practice, of course, the vectors in our convenient orthonormal basis $\mathcal{A}$ will usually be $\mathbf{a}_1 = \mathbf{i}$, $\mathbf{a}_2 = \mathbf{j}$, ...

For example, let's suppose we've been asked to diagonalize the Hermitian matrix

$$\mathbf{H} = \begin{bmatrix} 8 & -2 & -2 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix} \quad .$$

Here, we have a 3×3 *Hermitian matrix* $\mathbf{H}$ *which we will view as the matrix* $\mathbf{H}_{\mathcal{A}}$ *of a Hermitian operator* $\mathcal{H}$ *with respect to*

$$\mathcal{A} = \{ \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \} = \{ \mathbf{i}, \mathbf{j}, \mathbf{k} \} \quad .$$

The first step after getting the matrix $\mathbf{H}$ is to find all of its eigenvalues (and their multiplicities) by solving

$$\det(\mathbf{H} - \lambda \mathbf{I}) = 0 \quad .$$

(If $N \geq 3$, this can be a challenge.)

For our example,

$$\det(\mathbf{H} - \lambda \mathbf{I}) = 0$$

$$\Longrightarrow \quad \det\left(\begin{bmatrix} 8 & -2 & -2 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) = 0$$

$$\Longrightarrow \quad \det \begin{bmatrix} 8-\lambda & -2 & -2 \\ -2 & 5-\lambda & -4 \\ -2 & -4 & 5-\lambda \end{bmatrix} = 0$$

$$\Longrightarrow \quad \cdots = 0$$

$$\Longrightarrow \quad -\left[\lambda^3 - 18\lambda^2 + 81\lambda \right] = 0$$

$$\Longrightarrow \quad -\lambda(\lambda - 9)^2 = 0$$

$$\Longrightarrow \quad -(\lambda - 0)(\lambda - 9)^2 = 0 \quad .$$

So,

$$\lambda = 0 \text{ is a simple eigenvalue for } \mathbf{H} \quad ,$$

and

$$\lambda = 9 \text{ is an eigenvalue of multiplicity } 2 \text{ for } \mathbf{H} \quad .$$

For each simple eigenvalue λ (i.e., each eigenvalue with multiplicity 1), we merely need to solve either

$$\mathbf{H}\,|\mathbf{v}\rangle_{\mathcal{A}} = \lambda\,|\mathbf{v}\rangle_{\mathcal{A}} \qquad \text{or} \qquad [\mathbf{H} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle_{\mathcal{A}} = |\mathbf{0}\rangle_{\mathcal{A}}$$

(our choice) for the components with respect to $\mathcal{A}$ of any eigenvector $\mathbf{v}$ corresponding to that eigenvalue λ. Since this eigenvalue has multiplicity 1 , its eigenspace will be one-dimensional, which means that you should get an expression for $|\mathbf{v}\rangle_{\mathcal{A}}$ involving exactly one arbitrary constant. Set that constant equal to any convenient (nonzero) value to get a eigenvector $\mathbf{v}$, and then normalize $\mathbf{v}$ (i.e., divide by its length) to get a unit eigenvector $\mathbf{u}$ corresponding to eigenvalue λ.

For our example, the only simple eigenvalue is $\lambda = 0$. Let v_1, v_2 and v_3 be the components with respect to basis $\mathcal{A}$ of any corresponding eigenvector $\mathbf{v}$. Then

$$\mathbf{H}\,|\mathbf{v}\rangle = \lambda\,|\mathbf{v}\rangle \quad \Longrightarrow \quad \begin{bmatrix} 8 & -2 & -2 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0 \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad .$$

Multiplying through then yields the system

$$\begin{aligned} 8v_1 - 2v_2 - 2v_3 &= 0 \\ -2v_1 + 5v_2 - 4v_3 &= 0 \\ -2v_1 - 4v_2 + 5v_3 &= 0 \quad , \end{aligned}$$

which, after appropriately multiplying the equations by appropriate scalars and adding/subtracting various equations, reduces to the system

$$\begin{aligned} 2v_1 + 0v_2 - 1v_3 &= 0 \\ 1v_2 - 1v_3 &= 0 \\ 0 &= 0 \quad . \end{aligned}$$

As noted above, there will be exactly one arbitrary constant since $\lambda = 0$ is a simple eigenvalue. We can use v_1 or v_2 or v_3 as that constant in this case. Choosing v_1 will be convenient. Then, by the last set of equations, we have that

$$v_3 = 2v_1 \qquad \text{and} \qquad v_2 = v_3 = 2v_1 \quad .$$

So

$$|\mathbf{v}\rangle = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ 2v_1 \\ 2v_1 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \quad .$$

In particular, (with $v_1 = 1$) we have that

$$\mathbf{v} = 1\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$$

is an eigenvector corresponding to $\lambda = 0$. Normalizing this eigenvector then gives our first unit eigenvector,

$$\mathbf{u}_1 = \frac{1}{3}\big[1\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}\big] \quad .$$

Remember, it corresponds to eigenvalue $\lambda_1 = 0$. (It's a good idea to start indexing the unit eigenvectors and corresponding eigenvalues.)

For each eigenvalue λ with multiplicity $m > 1$, we still need to solve either

$$\mathbf{H}\,|\mathbf{v}\rangle_{\mathcal{A}} \;=\; \lambda\,|\mathbf{v}\rangle_{\mathcal{A}} \qquad \text{or} \qquad [\mathbf{H} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle_{\mathcal{A}} \;=\; |\mathbf{0}\rangle_{\mathcal{A}}$$

(our choice) for the components with respect to $\mathcal{A}$ of any eigenvector $\mathbf{v}$ corresponding to that eigenvalue λ.[5] Since this eigenvalue has multiplicity m, its eigenspace will be m-dimensional. Thus, any basis for this eigenspace will contain m vectors. This means that you should get an expression for $|\mathbf{v}\rangle_{\mathcal{A}}$ involving exactly m arbitrary constants — call them α_1, α_2, ..., and α_m, for now. Now choose m sets of values for these constants to generate a linearly independent set of m eigenvectors

$$\{\, \mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_m \,\} \quad .$$

For example, $\mathbf{v}_1$ could be the result of choosing

$$\alpha_k \;=\; \begin{cases} 1 & \text{if} \quad k = 1 \\ 0 & \text{otherwise} \end{cases} \quad ,$$

$\mathbf{v}_2$ could be the result of choosing

$$\alpha_k \;=\; \begin{cases} 1 & \text{if} \quad k = 2 \\ 0 & \text{otherwise} \end{cases} \quad ,$$

and so on. With great luck (or skill on your part) the resulting set of m eigenvectors will be orthonormal. If not, construct an orthonormal set from the $\mathbf{v}_j$'s using the Gram-Schmidt procedure (or any other legitimate trick).

For our example, $\lambda = 9$ is an eigenvalue of multiplicity 2. Let v_1, v_2 and v_3 be the components with respect to basis $\mathcal{A}$ of any eigenvector $\mathbf{v}$ corresponding to this eigenvalue. Then

$$[\mathbf{H} - \lambda\mathbf{I}]\,|\mathbf{v}\rangle_{\mathcal{A}} \;=\; |\mathbf{0}\rangle_{\mathcal{A}}$$

$$\Longrightarrow \quad \left(\begin{bmatrix} 8 & -2 & -2 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix} - 9 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \;=\; \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Longrightarrow \quad \begin{bmatrix} -1 & -2 & -2 \\ -2 & -4 & -4 \\ -2 & -4 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \;=\; \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad .$$

Multiplying through gives the system

$$\begin{aligned} -1v_1 \;-\; 2v_2 \;-\; 2v_3 \;&=\; 0 \\ -2v_1 \;-\; 4v_2 \;-\; 4v_3 \;&=\; 0 \\ -2v_1 \;-\; 4v_2 \;-\; 4v_3 \;&=\; 0 \quad . \end{aligned}$$

But each of these equations is just

$$v_1 \;+\; 2v_2 \;+\; 2v_3 \;=\; 0$$

[5] Unless you can use one or more of the 'cheap tricks' discussed later.

multiplied by either -1 *or* -2*. Since the eigenvalue here has multiplicity* 2*, we can choose any two of the components to be arbitrary constants. Choosing* v_2 *and* v_3 *to be our two arbitrary constants means that*

$$v_1 = -2v_2 - 2v_3 \quad ,$$

and that every eigenvector $\mathbf{v}$ *corresponding to* $\lambda = 9$ *can be written*

$$|\mathbf{v}\rangle_{\mathcal{A}} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -2v_2 - 2v_3 \\ v_2 \\ v_3 \end{bmatrix} = v_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + v_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \quad .$$

Setting $(v_2, v_3) = (1, 0)$ *and then* $(v_2, v_3) = (0, 1)$ *gives two eigenvectors* $\mathbf{v}_2$ *and* $\mathbf{v}_3$ *with*

$$|\mathbf{v}_2\rangle_{\mathcal{A}} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad |\mathbf{v}_3\rangle_{\mathcal{A}} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \quad .$$

That is,

$$\mathbf{v}_2 = -2\mathbf{i} + \mathbf{j} \qquad \text{and} \qquad \mathbf{v}_3 = -2\mathbf{i} + \mathbf{k} \quad .$$

But $\{\mathbf{v}_2, \mathbf{v}_3\}$ *is not orthonormal. To get a corresponding orthonormal pair* $\{\mathbf{u}_2, \mathbf{u}_3\}$ *of eigenvectors corresponding to* $\lambda = 9$*, we'll use the Gram-Schmidt procedure. Hence, we'll set*

$$\mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \cdots = \frac{1}{\sqrt{5}}\left[-2\mathbf{i} + \mathbf{j}\right]$$

and

$$\mathbf{u}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|}$$

where

$$\mathbf{w}_3 = \mathbf{v}_3 - \langle\, \mathbf{u}_2 \mid \mathbf{v}_3 \,\rangle\, \mathbf{u}_2 = \cdots = \frac{1}{5}\left[-2\mathbf{i} - 4\mathbf{j} + 5\mathbf{k}\right] \quad .$$

Thus,

$$\mathbf{u}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \cdots = \frac{1}{3\sqrt{5}}\left[-2\mathbf{i} - 4\mathbf{j} + 5\mathbf{k}\right] \quad .$$

Eventually, you will have used up all the eigenvalues, and found an orthonormal set of N eigenvectors

$$\mathcal{U} = \{\, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N \,\} \quad .$$

This is a basis for our vector space. Let

$$\{\, \lambda_1, \lambda_2, \ldots, \lambda_N \,\}$$

be the corresponding set of eigenvalues (with λ_k being the eigenvalue corresponding to eigenvector $\mathbf{u}_k$).

If we are "diagonalizing the Hermitian operator $\mathcal{H}$", then, as we already noted, $\mathcal{U}$ is the desired basis for our vector space, and

$$\mathbf{H}_{\mathcal{U}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

is our desired diagonal matrix. Writing these facts down (with the values found) completes the task.

If we are "diagonalizing the Hermitian matrix $\mathbf{H}$", then, as we already noted, we have a diagonal matrix $\mathbf{D}$ and a unitary matrix $\mathbf{U}$ such that

$$\mathbf{D} = \mathbf{U}^\dagger \mathbf{H} \mathbf{U}$$

by just setting $\mathbf{U} = \mathbf{M}_{\mathcal{A}\mathcal{U}}$ and

$$\mathbf{D} = \mathbf{H}_{\mathcal{U}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_N \end{bmatrix} \quad .$$

Computing $\mathbf{M}_{\mathcal{A}\mathcal{U}}$, and writing out the resulting $\mathbf{U}$ and $\mathbf{D}$ completes the task.

Don't forget the true reason for doing all this: $\mathbf{H}_{\mathcal{U}}$ is the matrix for the operator $\mathcal{H}$ with respect to the orthonormal basis $\mathcal{U}$. This basis is a "natural basis" for the operator, and $\mathbf{H}_{\mathcal{U}}$ is the simplest description for this operator.

For our example, we have obtained the orthonormal basis of eigenvectors

$$\mathcal{B} = \{ \mathbf{u}_1 , \mathbf{u}_2 , \mathbf{u}_3 \}$$

with corresponding eigenvalues

$$\lambda_1 = 0 \quad , \quad \lambda_2 = 9 \quad \textit{and} \quad \lambda_3 = 9 \quad .$$

We can compute $\mathbf{M}_{\mathcal{A}\mathcal{U}}$ *either using the basic definition,*

$$[\mathbf{M}_{\mathcal{A}\mathcal{U}}]_{jk} = \left\langle \mathbf{a}_j \,\middle|\, \mathbf{u}_k \right\rangle \quad ,$$

or, more easily, by observing that our formulas for the $\mathbf{u}_k$*'s can be written as*

$$[\,\mathbf{u}_1 , \mathbf{u}_2 , \mathbf{u}_3\,] = \underbrace{\left[\mathbf{i}, \mathbf{j}, \mathbf{k}\right]}_{[\,\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\,]} \underbrace{\begin{bmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{bmatrix}}_{\mathbf{M}_{\mathcal{A}\mathcal{U}}} \quad .$$

So $\mathbf{D} = \mathbf{U}^\dagger \mathbf{H} \mathbf{U}$ *with*

$$\mathbf{D} = \mathbf{H}_{\mathcal{U}} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

and

$$\mathbf{U} = \mathbf{M}_{\mathcal{A}\mathcal{U}} = \begin{bmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{bmatrix} \quad .$$

Cheap Tricks

Don't forget that the $\mathbf{u}_k$'s discussed above form an orthonormal basis. You can often use this fact to shortcut your computations by using one or more of the following 'cheap tricks':

1. If $\mathcal{V}$ can be viewed as a three-dimensional space of traditional vectors, and an orthonormal pair of eigenvectors $\{\mathbf{u}_1, \mathbf{u}_2\}$ has already been found, then the third vector in the orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ can be found via the cross product,
$$\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2 \quad .$$
2. If the $\mathbf{u}_k$'s have been found for all but one value of the eigenvalues, then the remaining eigenvectors can be chosen by just picking vectors orthogonal to those already found. By the theory we've developed, we know these will be eigenvectors corresponding to our last eigenvalue.

In our example above, we found that the simple eigenvalue $\lambda_1 = 0$ has corresponding unit eigenvector
$$\mathbf{u}_1 = \frac{1}{3}\left[1\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}\right] \quad .$$
For $\{\mathbf{u}_2, \mathbf{u}_3\}$, the orthonormal pair corresponding to the double eigenvalue $\lambda = 9$, we can first pick any unit vector orthogonal to $\mathbf{u}_1$, say,
$$\mathbf{u}_2 = \frac{1}{\sqrt{2}}\left[\mathbf{j} - \mathbf{k}\right]$$
(you should verify that this is a unit vector orthogonal to $\mathbf{u}_1$), and then set
$$\mathbf{u}_3 = \mathbf{u}_1 \times \mathbf{u}_2 = \cdots = \frac{1}{3\sqrt{2}}\left[-4\mathbf{i} + \mathbf{j} + \mathbf{k}\right] \quad .$$

7.5 Spectral Decomposition

Let $\mathcal{V}$ be an N-dimensional vector spacehaving an orthonormal basis
$$\mathcal{B} = \{\, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N \,\} \quad .$$
We now know that, given any Hermitian operator $\mathcal{H}$ on $\mathcal{V}$, there is an orthonormal basis of eigenvectors for $\mathcal{H}$
$$\mathcal{U} = \{\, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N \,\} \quad .$$
We also know that the matrix of $\mathcal{H}$ with respect to $\mathcal{U}$ is simply
$$\mathbf{H}_{\mathcal{U}} = \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

where

$$\{\, \lambda_1 ,\ \lambda_2 ,\ \ldots ,\ \lambda_N \,\}$$

is the set of eigenvalues corresponding to $\mathcal{U}$. We even know that this matrix is related to the matrix of $\mathcal{H}$ with respect to basis $\mathcal{B}$ via

$$\mathbf{D} \;=\; \mathbf{M}_{\mathcal{U}\mathcal{B}}\mathbf{H}_{\mathcal{B}}\mathbf{M}_{\mathcal{B}\mathcal{U}} \quad .$$

Equivalently,

$$\mathbf{H}_{\mathcal{B}} \;=\; \mathbf{M}_{\mathcal{B}\mathcal{U}}\mathbf{D}\mathbf{M}_{\mathcal{U}\mathcal{B}} \quad .$$

Now, notice that the entries in $\mathbf{D}$ are given by

$$[\mathbf{D}]_{mn} \;=\; \lambda_m \delta_{mn} \quad .$$

Using this, and various things from previous chapters,

$$\begin{aligned}
[\mathbf{H}_{\mathcal{B}}]_{jk} \;&=\; [\mathbf{M}_{\mathcal{B}\mathcal{U}}\mathbf{D}\mathbf{M}_{\mathcal{U}\mathcal{B}}]_{jk} \\
&=\; \sum_{m=1}^{N}\sum_{n=1}^{N} [\mathbf{M}_{\mathcal{B}\mathcal{U}}]_{jm}\, [\mathbf{D}]_{mn}\, [\mathbf{M}_{\mathcal{U}\mathcal{B}}]_{mk} \\
&=\; \sum_{m=1}^{N}\sum_{n=1}^{N} \left\langle\, \mathbf{e}_j \;\middle|\; \mathbf{u}_m \,\right\rangle \lambda_m \delta_{mn} \left\langle\, \mathbf{u}_n \;\middle|\; \mathbf{e}_k \,\right\rangle \\
&=\; \sum_{m=1}^{N} \left\langle\, \mathbf{e}_j \;\middle|\; \mathbf{u}_m \,\right\rangle \lambda_m \left\langle\, \mathbf{u}_m \;\middle|\; \mathbf{e}_k \,\right\rangle \quad .
\end{aligned}$$

Rewritten in terms of the bra and ket matrices of components, this is

$$\begin{aligned}
[\mathbf{H}_{\mathcal{B}}]_{jk} \;&=\; \sum_{m=1}^{N} \langle \mathbf{e}_j|_{\mathcal{B}}\, |\mathbf{u}_m\rangle_{\mathcal{B}}\, \lambda_m\, \langle \mathbf{u}_m|_{\mathcal{B}}\, |\mathbf{e}_k\rangle_{\mathcal{B}} \\
&=\; \langle \mathbf{e}_j|_{\mathcal{B}} \left(\sum_{m=1}^{N} |\mathbf{u}_m\rangle_{\mathcal{B}}\, \lambda_m\, \langle \mathbf{u}_m|_{\mathcal{B}} \right) |\mathbf{e}_k\rangle_{\mathcal{B}} \;=\; \left[\sum_{m=1}^{N} |\mathbf{u}_m\rangle_{\mathcal{B}}\, \lambda_m\, \langle \mathbf{u}_m|_{\mathcal{B}} \right]_{jk} \quad .
\end{aligned}$$

In other words,

$$\mathbf{H}_{\mathcal{B}} \;=\; \sum_{m=1}^{N} |\mathbf{u}_m\rangle_{\mathcal{B}}\, \lambda_m\, \langle \mathbf{u}_m|_{\mathcal{B}} \quad . \tag{7.4}$$

This equation is just another way of writing

$$\mathbf{H}_{\mathcal{B}} \;=\; \mathbf{M}_{\mathcal{B}\mathcal{U}}\mathbf{D}\mathbf{M}_{\mathcal{U}\mathcal{B}} \quad .$$

Do note that both this and (7.4) are formulas that are independent of the choice of original orthonormal basis $\mathcal{B}$.

By the way, the set of eigenvalues of an operator is said to be the *spectrum* of that operator, and formula (7.4) is called the *spectral decomposition* of Hermitian operator $\mathcal{H}$.