MATHCHAPTER H

Matrix Eigenvalue Problems

The Schrödinger equation

$$\hat{H}\psi = E\psi \tag{H.1}$$

is an eigenvalue problem; ψ is the eigenfunction and E is the corresponding eigenvalue. We've seen in MathChapter G that operators can be represented by matrices, and so the matrix equation

$$\mathsf{Ac} = \lambda\mathsf{c} \tag{H.2}$$

which is analogous to Equation H.1, is called a *matrix eigenvalue problem*, where c is an *eigenvector* of the matrix A and λ is the corresponding eigenvalue. Equations H.1 and H.2 suggest that there is a strong relationship between the Schrödinger equation and a matrix eigenvalue problem. We have seen this relationship in Chapter 8, but we didn't develop it there. In fact, quantum mechanics can be presented entirely in terms of matrices instead of differential equations as we have done in this book. It's traditional for quantum chemistry to be presented in terms of differential equations because chemistry students are presumably more comfortable or familiar with differential equations than with matrices; but, in fact, matrix algebra is much easier than differential equations, and most research in molecular quantum mechanics is couched in terms of matrices and matrix eigenvalue problems.

To see explicitly the relation between the Schrödinger equation and a matrix eigenvalue problem, we expand the (unknown) eigenfunction ψ in Equation H.1 in terms of some convenient set of (real and normalized) functions ϕ_i:

$$\psi = \sum_{i=1}^{N} c_i \phi_i \tag{H.3}$$

As N gets larger and larger, we expect Equation H.3 to become more and more exact if we choose the ϕ_i well. The unknown nature of ψ is now represented by the set of unknown coefficients $\{c_i\}$. We substitute Equation H.3 into Equation H.1, multiply by ϕ_j^*, and then integrate over all the coordinates to obtain the set of algebraic equations

$$\begin{aligned} H_{11}c_1 + H_{12}c_2 + \cdots + H_{1N}c_N &= E(c_1 + c_2S_{12} + \cdots + c_NS_{1N}) \\ H_{21}c_1 + H_{22}c_2 + \cdots + H_{2N}c_N &= E(c_1S_{21} + c_2 + \cdots + c_NS_{1N}) \\ \vdots \qquad \vdots \qquad\qquad \vdots \;\; &= \qquad\qquad \vdots \\ H_{N1}c_1 + H_{N2}c_2 + \cdots + H_{NN}c_N &= E(c_1S_{N1} + c_2S_{N2} + \cdots + c_N) \end{aligned} \tag{H.4}$$

where the

$$H_{ij} = \int d\tau\, \phi_i \hat{H} \phi_j \tag{H.5}$$

and the

$$S_{ij} = \int d\tau\, \phi_i \phi_j$$

are called *matrix elements*. We have used the fact that the ϕ_i are normalized ($S_{ii} = 1$) in writing Equation H.4. We can write Equation H.4 as a matrix eigenvalue problem

$$\mathsf{H}\mathsf{c} = E\mathsf{S}\mathsf{c} \tag{H.6}$$

Equation H.6 is equivalent to Equation 8.38. This type of equation appears often in quantum chemistry, and will appear repeatedly in later chapters. Equation H.6 becomes the same as Equation H.2 (with $\mathsf{A} = \mathsf{S}^{-1}\mathsf{H}$) if we multiply Equation H.6 from the left by S^{-1}. Thus, we see that the Schrödinger equation can be expressed as a matrix eigenvalue problem.

Let's look at Equation H.2 more closely. Equation H.2 represents the system of homogeneous linear equations

$$\begin{aligned} (a_{11} - \lambda)c_1 + a_{12}c_2 + \cdots + a_{1N}c_N &= 0 \\ a_{21}c_1 + (a_{22} - \lambda)c_2 + \cdots + a_{2N}c_N &= 0 \\ \vdots \qquad\qquad \vdots \qquad\qquad\quad \vdots \quad &\;\; \vdots \\ a_{N1}c_1 + a_{N2}c_2 + \cdots + (a_{NN} - \lambda)c_N &= 0 \end{aligned} \tag{H.7}$$

As we have seen a number of times before, the determinant of the c_j's must be equal to zero in order to have a nontrivial solution; in other words, a solution where not all the $c_j = 0$. Thus, we write

$$\det(\mathsf{A} - \lambda\mathsf{I}) = 0 \tag{H.8}$$

which leads to the secular equation, which is an Nth-degree polynomial equation in λ. The solution to this equation gives us N eigenvalues in Equation H.2. Associated with each eigenvalue is an eigenvector. We obtain each eigenvector by substituting one of the values of λ into Equation H.4 and then solving for the c_j's. We did this repeatedly in Chapter 8.

EXAMPLE H–1

Find the eigenvalues and eigenvectors of

$$\mathsf{A} = \begin{pmatrix} a & 1 \\ 1 & a \end{pmatrix}$$

where a is a constant.

SOLUTION: The determinant of $\mathsf{A} - \lambda \mathsf{I}$ is given by

$$\det(\mathsf{A} - \lambda \mathsf{I}) = \begin{vmatrix} a-\lambda & 1 \\ 1 & a-\lambda \end{vmatrix} = (a-\lambda)^2 - 1 = 0$$

and so the eigenvalues are given by the solution to $(a-\lambda)^2 - 1 = 0$, or $\lambda = a \pm 1$. The equations for the eigenvectors are (see Equation H.7)

$$(a-\lambda)c_1 + c_2 = 0$$
$$c_1 + (a-\lambda)c_2 = 0$$

If we substitute $\lambda = a + 1$ into these equations, we obtain

$$-c_1 + c_2 = 0$$
$$c_1 - c_2 = 0$$

or $c_1 = c_2$. Thus, the eigenvector is (c_1, c_1), where c_1 is an arbitrary constant. We can fix the value of c_1 by requiring that the eigenvector be normalized, in which case we have

$$\mathsf{c}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

The other normalized eigenvector is given by

$$\mathsf{c}_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

It's an easy exercise to verify that $\mathsf{A}\mathsf{c}_1 = \lambda_1 \mathsf{c}_1$ and that $\mathsf{A}\mathsf{c}_2 = \lambda_2 \mathsf{c}_2$.

In Example H–1, we solved a 2×2 eigenvalue problem. The algebra was simple because we had to solve only a quadratic equation to find the two eigenvalues. The algebra increases drastically as we go on to problems of dimension greater than two; even a

3×3 system leads to a cubic equation for λ, which is usually quite tedious to solve, and the tedium grows rapidly with the size of the matrix. There are a number of user-friendly mathematical computer programs available nowadays that can easily handle very large matrices. Three such programs are *MathCad, Maple,* and *Mathematica,* each of which can perform algebraic manipulations as well as do numerical calculations. At least one of these programs is available in most chemistry departments, and you should learn how to use one of these programs. Any of these programs, as well as others, can solve for all the eigenvalues and corresponding eigenvectors of a sizable matrix in seconds.

Note that the two eigenvectors in Example H–1 are orthonormal because

$$\mathsf{c}_1 \cdot \mathsf{c}_1 = \left(\frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \right) = 1$$

$$\mathsf{c}_2 \cdot \mathsf{c}_2 = \left[\frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} + \left(-\frac{1}{\sqrt{2}} \right) \left(-\frac{1}{\sqrt{2}} \right) \right] = 1$$

and

$$\mathsf{c}_1 \cdot \mathsf{c}_2 = \left[\frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \left(-\frac{1}{\sqrt{2}} \right) \right] = 0$$

This is generally true for eigenvectors of distinct eigenvalues of a symmetric matrix; in other words, one for which $\mathsf{A} = \mathsf{A}^{\mathsf{T}}$. This result is completely analogous to the fact that the nondegenerate eigenfunctions of Hermitian operators are orthonormal (see Section 4.6). Recall that a definition of a Hermitian operator $\hat{A}$ is

$$\int d\tau \ \psi_i^* \hat{A} \psi_j = \int d\tau \ (\hat{A}\psi_i)^* \psi_j \tag{H.9}$$

If we let $A_{ij} = \int d\tau \psi_i^* \hat{A} \psi_j$, then Equation H.9 says that

$$A_{ij} = A_{ji}^* \qquad \text{(Hermitian matrix)} \tag{H.10}$$

A symmetrical matrix would have $A_{ij} = A_{ji}$. Equation H.10 is the extension of the definition of a symmetric matrix to a complex space, where the elements of the matrices may be complex. If A satisfies Equation H.10, it is said to be a *Hermitian matrix*. All matrices in quantum mechanics must be Hermitian because the eigenvalues of a Hermitian matrix are real, just as the eigenvalues of a Hermitian operator are real (see Section 4.6).

EXAMPLE H–2

Show that the eigenvectors of a Hermitian matrix are real and that the eigenvectors corresponding to distinct eigenvalues are orthogonal.

SOLUTION: Start with $\mathsf{A}\mathsf{c}_j = \lambda_j \mathsf{c}_j$. Multiply both sides from the left by c_i^* to obtain $\mathsf{c}_i^* \mathsf{A} \mathsf{c}_j = \lambda_j \mathsf{c}_i^* \mathsf{c}_j$, which we write in the notation

$$A_{ij} = \lambda_j \mathsf{c}_i^* \mathsf{c}_j \tag{H.11}$$

Now multiply $\mathsf{A}^*\mathsf{c}_i^* = \lambda_i^*\mathsf{c}_i^*$ from the left by c_j to obtain $\mathsf{c}_j\mathsf{A}^*\mathsf{c}_i^* = \lambda_i^*\mathsf{c}_j\mathsf{c}_i^*$, which we write as

$$A_{ji}^* = \lambda_i^*\mathsf{c}_j\mathsf{c}_i^* \tag{H.12}$$

But A is Hermitian, so $A_{ij} = A_{ji}^*$. Furthermore, $\mathsf{c}_j\mathsf{c}_i^* = \mathsf{c}_i^*\mathsf{c}_j$ because the dot product of two vectors is commutative. Comparing Equations H.11 and H.12 gives

$$(\lambda_i^* - \lambda_j)\mathsf{c}_i^*\mathsf{c}_j = 0 \tag{H.13}$$

If $i = j$, $\mathsf{c}_i^*\mathsf{c}_j \geq 0$, and so $\lambda_j = \lambda_j^*$, which says that the eigenvalues are real. If $i \neq j$, then $\lambda_i \neq \lambda_j$ if there is no degeneracy, and so $\mathsf{c}_i^*\mathsf{c}_j = 0$, which says that c_i and c_j are orthogonal.

Let's go back to Equation H.2, which we will write in the form

$$\mathsf{Ac}_k = \lambda_k\mathsf{c}_k \qquad k = 1, 2, \ldots, N \tag{H.14}$$

There are N eigenvalues λ_k and N corresponding eigenvectors, c_k. Now let's normalize the c_k and form a matrix

$$\mathsf{S} = (\mathsf{c}_1, \mathsf{c}_2, \ldots, \mathsf{c}_N) \tag{H.15}$$

where the notation means that the columns of S are the (normalized) eigenvectors of A. Because the columns of S consist of the eigenvectors of A, and because these eigenvectors form an orthonormal set if A is symmetric (which it usually is), S is an orthogonal matrix. In other words, $\mathsf{S}^{-1} = \mathsf{S}^{\mathsf{T}}$. Furthermore, the matrix S has a remarkable property that we can see by operating on S with A to obtain (Problem H–6)

$$\begin{aligned}\mathsf{AS} &= (\mathsf{Ac}_1, \mathsf{Ac}_2, \ldots, \mathsf{Ac}_N) \\ &= (\lambda_1\mathsf{c}_1, \lambda_2\mathsf{c}_2, \ldots, \lambda_N\mathsf{c}_N) \\ &= \mathsf{SD}\end{aligned} \tag{H.16}$$

where

$$\mathsf{D} = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_N \end{pmatrix} \tag{H.17}$$

is a diagonal matrix whose elements are the eigenvalues of A.

If we multiply Equation H.16 from the left by S^{-1}, then we obtain

$$\mathsf{D} = \mathsf{S}^{-1}\mathsf{AS} = \mathsf{S}^{\mathsf{T}}\mathsf{AS} \tag{H.18}$$

because S is orthogonal. Equation H.18 is called a *similarity transformation*. We say that the matrix A has been *diagonalized* by the similarity transformation in Equation H.18. Diagonalizing a matrix A is *completely equivalent* to solving the eigenvalue problem in Equation H.2, or, because Equations H.1 and H.2 are equivalent, diagonalizing the Hamiltonian matrix is completely equivalent to solving the Schrödinger equation. Physically, A and D represent the same operation (such as a rotation or a reflection through a plane). Their different forms result from the fact that D is expressed in an optimum, or natural, coordinate system. Because of the central importance of matrix diagonalization in quantum mechanics, there are many sophisticated and efficient algorithms for matrix diagonalization in the numerical analysis literature.

EXAMPLE H–3

Diagonalize the matrix A in Example H–1.

SOLUTION: The matrix S is given by

$$\mathsf{S} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

The inverse of S is

$$\mathsf{S}^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Using Equation H.18, we have

$$\begin{aligned} \mathsf{S}^{-1}\mathsf{A}\mathsf{S} &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} a & 1 \\ 1 & a \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \begin{pmatrix} a+1 & 0 \\ 0 & a-1 \end{pmatrix} = \mathsf{D} \end{aligned}$$

Notice that the elements of D are the eigenvalues of A. Notice also that the trace of A is equal to the trace of D, which equals $\lambda_1 + \lambda_2$ (Problem H–12).

Problems

H–1. Determine the eigenvalues and eigenvectors of $\mathsf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

H–2. Determine the eigenvalues and eigenvectors of $\mathsf{A} = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$.

H–3. Determine the eigenvalues and eigenvectors of $\mathsf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

H–4. Determine the eigenvalues and eigenvectors of $\mathsf{A} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$.

H–5. Show that the matrix $\mathsf{A} = \begin{pmatrix} 1 & i & 1-i \\ -i & 0 & -1+i \\ 1+i & -1-i & 3 \end{pmatrix}$ is Hermitian.

H–6. Verify that $(\lambda_1 \mathbf{c}_1, \lambda_2 \mathbf{c}_2, \ldots, \lambda_N \mathbf{c}_N) = \mathsf{SD}$ in Equation H.16.

H–7. The three eigenvectors of A in Problem H–4 are $c_1(-1, 0, 1)$, $c_2(0, 1, 0)$, and $c_3(1, 0, 1)$, where c_1, c_2, and c_3 are arbitrary. Choose them so that the three eigenvectors are normalized. Now form the matrix S whose columns consist of the three normalized eigenvectors. Find the inverse of S and then show explicitly that $\mathsf{S}^{-1} = \mathsf{S}^{\mathsf{T}}$, or that S is indeed orthogonal.

H–8. Diagonalize the matrix in Problem H–1.

H–9. Diagonalize the matrix in Problem H–2.

H–10. Diagonalize the matrix in Problem H–3.

H–11. Diagonalize the matrix in Problem H–4.

H–12. Show that $\mathrm{Tr}\,\mathsf{D} = \mathrm{Tr}\,\mathsf{A}$.

H–13. Programs such as *MathCad* and *Mathematica* can find the eigenvalues and corresponding eigenvectors of large matrices in seconds. Use one of these programs to find the eigenvalues and corresponding eigenvectors of

$$\mathsf{A} = \begin{pmatrix} a & 1 & 0 & 0 & 0 & 1 \\ 1 & a & 1 & 0 & 0 & 0 \\ 0 & 1 & a & 1 & 0 & 0 \\ 0 & 0 & 1 & a & 1 & 0 \\ 0 & 0 & 0 & 1 & a & 1 \\ 1 & 0 & 0 & 0 & 1 & a \end{pmatrix}$$