Eigensystems

We will discuss matrix diagonalization algorithms in *Numerical Recipes* in the context of the eigenvalue problem in quantum mechanics,

$$A|n\rangle = \lambda_n|n\rangle, \tag{1}$$

where A is a real, symmetric Hamiltonian operator and $|n\rangle$ is the n-th eigenvector with eigenvalue λ_n. In an N-dimensional vector space, Eq. (1) becomes

$$\sum_{j=1}^{N} A_{ij}x_j^{(n)} = \lambda_n x_i^{(n)}, \tag{2}$$

where A is an $N\times N$ matrix, and $x_i^{(n)}$ is the i-th element of the n-th eigenvector $x^{(n)} \in \mathbf{R}^N$.

ORTHONORMAL BASIS

(Orthogonality) The basis set $\{|n\rangle | n = 1, \dots, N\}$ can be made orthonormal, i.e.,

$$\langle m|n\rangle \equiv \sum_{i=1}^{N} x_i^{(m)}x_i^{(n)} = \delta_{mn}, \tag{3}$$

or, by defining the transformation matrix U as

$$U_{in} = x_i^{(n)} \tag{4}$$

(i.e., the n-th column of U is the n-th eigenvector), U is orthogonal,

$$U^T U = I \tag{5}$$

where I is the $N\times N$ identity matrix.

Proof of Eq. (3): First note that all eigenvalues λ_n are real. (∵ By multiplying eq. (1) by $\langle n|$ from the left, $\langle n|A|n\rangle = \lambda_n\langle n|n\rangle$. For a Hermitian matrix (and of course for a real, symmetric matrix), $\langle n|A|n\rangle$ is real, and $\langle n|n\rangle$ is also real since its complex conjugate is itself.) Next, by multiplying Eq. (1) by $\langle m|$ from the left, we obtain

$$\langle m|A|n\rangle = \lambda_n\langle m|n\rangle. \tag{6}$$

Similarly,

$$\langle n|A|m\rangle = \lambda_m\langle n|m\rangle. \tag{7}$$

By taking the complex conjugate of Eq. (7) and noting the reality of the eigenvalue,

$$\langle m|A|n\rangle = \lambda_m\langle m|n\rangle. \tag{8}$$

Subtracting Eq. (8) from Eq. (6),

$$0 = (\lambda_n - \lambda_m)\langle m|n\rangle. \tag{9}$$

If $\lambda_n \neq \lambda_m$, Eq. (9) requires that $\langle m|n\rangle = 0$. On the other hand, if $\lambda_n = \lambda_m$, we can still make them orthogonal without modifying the eigenvalue. For example, Gram-Schmidt orthogonalization procedure

$$|n'\rangle \leftarrow |n\rangle - |m\rangle\langle m|n\rangle \tag{10}$$

makes $\langle m|n'\rangle = \langle m|n\rangle - \langle m|m\rangle\langle m|n\rangle = \langle m|n\rangle - \langle m|n\rangle = 0$, followed by the normalization $|n'\rangle$ as $|n'\rangle \leftarrow |n'\rangle/\langle n'|n'\rangle^{1/2}$.

(Completeness) The orthonormal basis set $\{|n\rangle\}$ is also complete, i.e., in the N-dimensional vector space,

$$\sum_{n=1}^{N} |n\rangle\langle n| = 1 \tag{11}$$

is the identity operator. Equivalently,

$$\sum_{n=1}^{N} x_i^{(n)} x_j^{(n)} = \delta_{ij}, \tag{12}$$

or

$$UU^T = I. \tag{13}$$

Equation (11) states that any vector in the *N*-dimensional vector space $|\psi\rangle$ is a linear combination of the *N* basis functions,

$$|\psi\rangle = \sum_{n=1}^{N} |n\rangle\langle n|\psi\rangle, \tag{14}$$

since there are only *N* linearly independent vectors in this vector space.

The orthogonality and completeness together states that

$$U^T U = UU^T = I. \tag{15}$$

or

$$U^{-1} = U^T. \tag{16}$$

ORTHOGONAL TRANSFORMATION

Now, we use the orthogonal matrix *U* to restate the matrix eigenvalue problem. To do so, multiply Eq. (2) by $x_i^{(m)}$ and sum the resulting equation over *i*,

$$\sum_{i=1}^{N} \quad \sum_{j=1}^{N} x_i^{(m)} A_{ij} x_j^{(n)} = \lambda_n \sum_{i=1}^{N} x_i^{(m)} x_i^{(n)} = \lambda_n \delta_{mn}, \tag{17}$$

where we have used the orthonormality, Eq. (3). Using *U*, equation (17) can be rewritten as

$$U^T A U = \Lambda, \tag{18}$$

where

$$\Lambda_{mn} = \lambda_m \delta_{mn}, \tag{19}$$

is a diagonal matrix. Thus the matrix eigenvalue problem amounts to finding an orthogonal matrix, *U*, or the associated orthogonal transformation, Eq. (18), which eliminates all the off-diagonal matrix elements.

GRAND STRATEGY

The grand strategy of matrix diagonalization is to nudge the matrix *A* towards diagonal form by a sequence of orthogonal transformations,

$$A \to P_1^T \mathrm{AP}_1 \to P_2^T P_1^T \mathrm{AP}_1 P_2 \to \cdots, \tag{20}$$

so that its off-diagonal elements gradually disappear. At the end, the orthogonal matrix is

$$U = P_1 P_2 \cdots. \tag{21}$$

ORTHOGONAL TRANSFORMATION ~ ROTATION: JACOBI TRANSFORMATION

As an illustration, let us consider a two-state system, for which the most general Hamiltonian matrix is

$$H = \begin{bmatrix} \varepsilon_1 & \delta \\ \delta & \varepsilon_2 \end{bmatrix}. \tag{22}$$

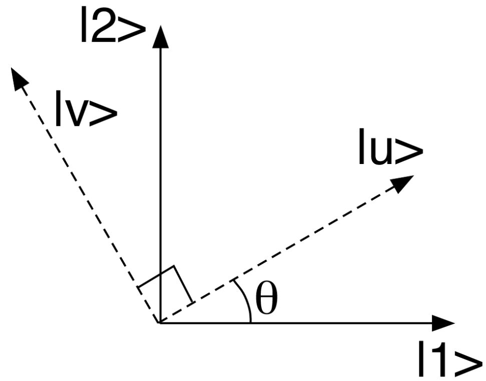
(We define the indices such that $\varepsilon_1 < \varepsilon_2$, i.e., the first state is the lower-energy state.) We express first eigenvector of this Hamiltonian as

$$|u\rangle = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \cos\theta|1\rangle + \sin\theta|2\rangle, \tag{22}$$

which is most general. (Because of the normalization condition, the any vector in the 2-dimensional vector space can be specified by one parameter.) Once we specify the first eigenvector, the second is readily determined from the orthonomality as

$$|v\rangle = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}, \tag{22}$$

see the figure below. The rotation angle θ specifies the deviation of the first eigenvector $|u\rangle$ from $|1\rangle$.



The orthogonal matrix is then

$$U = [u \quad v] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{23}$$

To find the specific rotation angle θ, let us return to the original eigenvalue problem,

$$\begin{bmatrix} \lambda - \varepsilon_1 & -\delta \\ -\delta & \lambda - \varepsilon_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{24}$$

The eigenvalues are obtained by solving the secular equation,

$$\det(\lambda I - H) = \begin{vmatrix} \lambda - \varepsilon_1 & -\delta \\ -\delta & \lambda - \varepsilon_2 \end{vmatrix} = (\lambda - \varepsilon_1)(\lambda - \varepsilon_2) - \delta^2 = 0, \tag{25}$$

and its two solutions are

$$\lambda_{\pm} = \frac{\varepsilon_1 + \varepsilon_2 \pm \sqrt{(\varepsilon_1 - \varepsilon_2)^2 + 4\delta^2}}{2}. \tag{26}$$

Now let us examine the lower eigenenergy λ_-. By substituting the eigenvalue and the corresponding eigenvector, Eq. (22), into Eq. (24), we obtain

$$\theta = \tan^{-1}\left(\frac{-\varepsilon_1 + \varepsilon_2 - \sqrt{(\varepsilon_1 - \varepsilon_2)^2 + 4\delta^2}}{2\delta}\right). \tag{27}$$

For example, if the off-diagonal element δ is small, we can expand Eq. (27) is its power series, the first term of which is (we have assumed $\varepsilon_1 < \varepsilon_2$)

$$\theta = \frac{\delta}{\varepsilon_1 - \varepsilon_2}. \tag{28}$$

Jacobi Transformation

In Jacobi transformation, each orthogonal transformation P_k in Eq. (20) is the two-dimensional rotation applied to a pair of rows, *i* and *j*, and the pair of columns of the same indices. One such rotation eliminates a pair—(*i*,*j*) and (*j*,*i*)—of off-diagonal elements. A sequence of two-dimensional rotations will eventually eliminate all the off-diagonal elements. (In fact, later rotations may partially restore off-diagonal elements eliminated earlier. Nevertheless, this procedure will converge, and the square sum of all the off-diagonal elements becomes smaller as we continue the procedure.)

HOUSEHOLDER TRANSFORMATIONS FOR TRIDIAGONALIZATION

Instead of eliminating a pair of off-diagonal elements at one time as in Jacobi transformation, Householder transformation eliminates an entire row but the first two elements at a time.

In Chapter 11 of *Numerical Recipes*, Householder transformations are used to reduce a real, symmetric matrix to a tridiagonal form, in which only the diagonal (A_{ii}), upper subdiagonal ($A_{i\,i+1}$), and lower subdiagonal ($A_{i+1\,i}$) elements may be nonzero. The function tred2() achieves this. The resulting tridiagonal matrix is then diagonalized (i.e., both subdiagonal elements are eliminated), using another set of orthogonal transformations in function `tqli()`.

The magical orthogonal matrix P is constructed from a vector in the N-dimensional vector space. First, let us prove a useful lemma.

(Lemma) Let $v \quad (\in R^N)$ and

$$P = I - \frac{2vv^T}{v^Tv}, \tag{29}$$

then P is symmetric and orthogonal, i.e.,

$$P^TP = PP = I. \tag{30}$$

∵First,

$$P_{ij} = \delta_{ij} - \frac{2v_iv_j}{\sum_{k=1}^{N} v_k^2}, \tag{31}$$

is symmetric with respect to the exchange of the indices i and j. Next,

$$\begin{aligned} P^TP &= \left(I - \frac{2vv^T}{v^Tv}\right)\left(I - \frac{2vv^T}{v^Tv}\right) \\ &= I - \frac{4vv^T}{v^Tv} + \frac{4vv^Tvv^T}{(v^Tv)^2} \qquad . \ // \\ &= I - \frac{4vv^T}{v^Tv} + \frac{4vv^T}{v^Tv} = I \end{aligned}$$

Now, given an arbitrary vector x in the N-dimensional vector space, we can device an orthogonal matrix that eliminates all the elements but the first one when multiplied to x.

(Theorem) For $\forall x \ (\in \mathbf{R}^N)$, let

$$v = x \pm \|x\|_2 e_1 \tag{32}$$

where

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{33}$$

and the vector 2-norm is defined as

$$\|x\|_2 = \sqrt{x^Tx} = \sqrt{\sum_{i=1}^{N} x_i^2}. \tag{34}$$

Then

$$Px = \left(I - \frac{2vv^T}{v^Tv}\right)x = \mp\|x\|_2 e_1, \tag{35}$$

i.e., the Householder matrix P, when multiplied, eliminates all the elements of x but the first one.

∵Note that,

$$
\begin{aligned}
v^T v &= (x^T \pm \|x\|_2 e_1^T)(x \pm \|x\|_2 e_1) \\
&= \|x\|_2^2 \pm 2\|x\|_2 x_1 + \|x\|_2^2 \\
&= 2\|x\|_2(\|x\|_2 \pm x_1)
\end{aligned}
.
$$

Then

$$
\begin{aligned}
Px &= x - \frac{2vv^T}{2\|x\|_2(\|x\|_2 \pm x_1)} x \\
&= x - \frac{(x \pm \|x\|_2 e_1)(x^T \pm \|x\|_2 e_1^T)x}{\|x\|_2(\|x\|_2 \pm x_1)} \\
&= x - \frac{(x \pm \|x\|_2 e_1)\|x\|_2(\|x\|_2 \pm x_1)}{\|x\|_2(\|x\|_2 \pm x_1)} \\
&= x - x \mp \|x\|_2 e_1 = \mp \|x\|_2 e_1
\end{aligned}
. \quad //
$$

The Householder matrix can be used for tridiagonalization as follows: Let us decompose a real, symmetric matrix A as

$$
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & & & \\ \vdots & & & \\ a_{N1} & & & \end{bmatrix} = \begin{bmatrix} a_{11} & A_{12} = A_{21}^T \\ & \\ A_{21} & A_{22} \end{bmatrix}. \tag{36}
$$

where A_{21}, A_{12}, and A_{22} are $(N\text{-}1)\times 1$, $1\times(N\text{-}1)$, and $(N\text{-}1)\times(N\text{-}1)$ matrices, respectively. Now let

$$
v \quad (\in \mathbf{R}^{N-1}) = A_{21} + \mathrm{sign}(a_{21})\|A_{21}\|_2 e_1. \tag{37}
$$

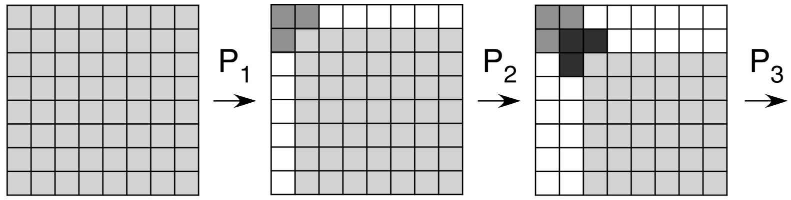
(The sign has been chosen to minimize the cancellation error.) Then

$$
\bar{P}A_{21} \equiv \left(I_{N-1} - \frac{2vv^T}{v^T v}\right) A_{21} = \text{-}\mathrm{sign}(a_{21})\|A_{21}\|_2 e_1 \equiv k e_1. \tag{38}
$$

Now

$$
\begin{aligned}
PAP &\equiv \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \bar{P} & \\ 0 & & & \end{bmatrix} \begin{bmatrix} a_{11} & A_{21}^T \\ & \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \bar{P} & \\ 0 & & & \end{bmatrix} \\
&\begin{bmatrix} a_{11} & A_{21}^T \\ k & \\ 0 & \bar{P}A_{22} \\ \vdots & \\ 0 & \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \bar{P} & \\ 0 & & & \end{bmatrix} \\
&\begin{bmatrix} a_{11} & k & 0 & \cdots & 0 \\ k & & & & \\ 0 & & \bar{P}A_{22}\bar{P} & & \\ \vdots & & & & \\ 0 & & & & \end{bmatrix}
\end{aligned}
, \tag{39}
$$

i.e., all the elements in the first row and first column but a_{11}, a_{12} and a_{21} have been eliminated by this transformation. Next, a similar Householder transformation is applied to the first column and first row of the $(N\text{-}1)\times(N\text{-}1)$ submatrix $\bar{P}A_{22}\bar{P}$, which eliminates all the elements in the second row and second column in the original $N\times N$ matrix but a_{22}, a_{23} and a_{32}, so on (see the figure below, in which white cells represent eliminated matrix elements).



After (N-2) such transformations, all the off-diagonal elements but the diagonal and upper/lower sub-diagonal elements are eliminated.

DIAGONALIZATION OF A TRIDIAGONAL MATRIX—QR DECOMPOSITION

QR Decomposition

The diagonalization of the tridiagonal matrix obtained above can use QR decomposition (or similar QL decomposition). That is, any square matrix A can be decomposed into

$$A = QR, \tag{40}$$

where Q is an orthogonal matrix and R is an upper-triangular matrix, i.e., $R_{ij} = 0$ for $i > j$.

For example, this can be achieved by using a Householder transformation as follows. First, we decompose the $N\times N$ matrix A into the first column A_1 and the rest A_2:

$$A = \begin{bmatrix} a_{11} & \\ \vdots & \\ a_{N1} & \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \end{bmatrix}. \tag{41}$$

Let

$$v \quad (\in R^N) = A_1 + \mathrm{sign}(a_{11})\|A_1\|_2 e_1, \tag{42}$$

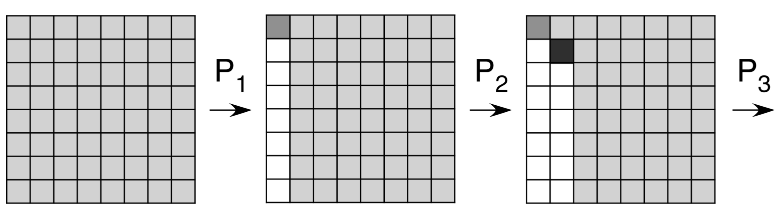
then

$$PA_1 \equiv \left(I_N - \frac{2vv^T}{v^Tv}\right)A_1 = -\mathrm{sign}(a_{11})\|A_1\|_2 e_1 \equiv ke_1, \tag{43}$$

and thus

$$PA = \begin{bmatrix} PA_1 & PA_2 \end{bmatrix} = \begin{bmatrix} k & \\ 0 & \\ \vdots & PA_2 \\ 0 & \end{bmatrix}, \tag{44}$$

i.e., all the elements in the first column but one have been eliminated. Next, we can apply a similar elimination to $A(2{:}N,2{:}N)$ submatrix to eliminate all the lower-triangular elements in the second column, see the figure below.



After (N-1) transformation, the resulting matrix is upper-triangular, i.e.,

$$P_{N-1} \cdots P_2 P_1 A = R, \tag{45}$$

or

$$A = P_1^{-1} P_2^{-1} \cdots P_{N-1}^{-1} R \equiv QR. \tag{46}$$

Orthogonal Transformation

Let Eq. (40) be the QR decomposition of matrix A. Then, define another matrix by

$$A' = RQ. \tag{48}$$

Since $R = Q^{-1}A = Q^TA$ from Eq. (40), Eq. (48) defines an orthogonal transformation,

$$A \to A' = Q^TAQ. \tag{49}$$

It can be proven that, if A is tridiagonal, then A' is also tridiagonal, i.e., the orthogonal transformation preserves the tridiagonality. The QR algorithm consists of successive applications of this orthogonal transformation.

(QR algorithm)

$$\begin{cases} 1. \quad Q_sR_s \leftarrow A_s \\ 2. \quad A_{s+1} \leftarrow R_sQ_s \end{cases} \quad s = 1,2, \dots . \tag{50}$$

The following theorems then guarantee that the eigenvalues can be obtained by the QR algorithm.

(Theorem)

1. $\lim_{s\to\infty} A_s$ is upper-triangular, and
2. The eigenvalues appear on its diagonal.

In Chapter 11 of *Numerical Recipes*, function `tqli()` uses QL algorithm, instead of the above QR algorithm, to achieve lower-triangularity, to minimize the cancellation error.

Once the eigenvalues are obtained, eigenvectors are determined based on inversion,

$$|y\rangle = \frac{1}{A - \tau I}|b\rangle, \tag{51}$$

which is a linear system of equations for $|y\rangle$ given a starting vector $|b\rangle$. When parameter τ is close to λ_n, $|y\rangle$ is dominated by $|n\rangle$. Equation (51) is used to iteratively obtain eigenvectors while refining eigenvalues.