Diagonalization of a general 2 × 2 hermitian matrix

Consider a general 2×2 hermitian matrix

$$A = \begin{pmatrix} a & c \\ c^* & b \end{pmatrix},$$

where a and b are real numbers and c is a complex number. The eigenvalues are the roots of the characteristic equation:

$$\begin{vmatrix} a-\lambda & c \\ c^* & b-\lambda \end{vmatrix} = (a-\lambda)(b-\lambda) - |c|^2 = \lambda^2 - \lambda(a+b) + (ab - |c|^2) = 0\,.$$

Noting that $(a+b)^2 - 4(ab - |c|^2) = (a-b)^2 + 4|c|^2$, the two roots can be written as:

$$\lambda_1 = \tfrac{1}{2}\left[a + b + \sqrt{(a-b)^2 + 4|c|^2}\right] \quad \text{and} \quad \lambda_2 = \tfrac{1}{2}\left[a + b - \sqrt{(a-b)^2 + 4|c|^2}\right], \qquad (1)$$

where by convention we take $\lambda_1 \geq \lambda_2$. As expected, the eigenvalues of an hermitian matrix are real.

An hermitian matrix can be diagonalized by a unitary matrix U,

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \qquad (2)$$

where λ_1 and λ_2 are the eigenvalues obtained in eq. (1). Note that one can always transform $U \to e^{i\zeta}U$ without modifying eq. (2), since the phase cancels out. Since $\det U$ is a pure phase, one can choose $\det U = 1$ in eq. (2) without loss of generality. The most general 2×2 unitary matrix of unit determinant can be written as:

$$U = \begin{pmatrix} e^{i\beta}\cos\theta & -e^{-i\chi}\sin\theta \\ e^{i\chi}\sin\theta & e^{-i\beta}\cos\theta \end{pmatrix}.$$

The columns of U are the normalized eigenvectors of A corresponding to the eigenvalues λ_1 and λ_2, respectively. But, we are always free to multiply any normalized eigenvector by an arbitrary complex phase. Thus, without loss of generality, we can choose $\beta = 0$ and $\cos\theta \geq 0$. Moreover, the sign of $\sin\theta$ can always be absorbed into the definition of χ. Hence, we will take

$$U = \begin{pmatrix} \cos\theta & -e^{-i\chi}\sin\theta \\ e^{i\chi}\sin\theta & \cos\theta \end{pmatrix}, \qquad (3)$$

where

$$0 \leq \theta \leq \tfrac{1}{2}\pi\,, \quad \text{and} \quad 0 \leq \chi < 2\pi\,. \qquad (4)$$

We now plug in eq. (3) into eq. (2). Since the off-diagonal terms must vanish, one obtains constraints on the angles θ and χ. It is convenient to define,

$$c = |c|e^{i\psi}\,, \qquad \text{where } 0 \leq \psi < 2\pi\,. \qquad (5)$$

Then,

$$
\begin{aligned}
U^{-1}AU &= \begin{pmatrix} \cos\theta & e^{-i\chi}\sin\theta \\ -e^{i\chi}\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a & |c|e^{i\psi} \\ |c|e^{-i\psi} & b \end{pmatrix} \begin{pmatrix} \cos\theta & -e^{-i\chi}\sin\theta \\ e^{i\chi}\sin\theta & \cos\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & e^{-i\chi}\sin\theta \\ -e^{i\chi}\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a\cos\theta + |c|e^{i(\psi+\chi)}\sin\theta & -ae^{-i\chi}\sin\theta + |c|e^{i\psi}\cos\theta \\ |c|e^{-i\psi}\cos\theta + be^{i\chi}\sin\theta & -|c|e^{-i(\psi+\chi)}\sin\theta + b\cos\theta \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & Z \\ Z^* & \lambda_2 \end{pmatrix},
\end{aligned}
$$

where

$$\lambda_1 = a\cos^2\theta + 2|c|\cos\theta\sin\theta\cos(\psi+\chi) + b\sin^2\theta\,, \tag{6}$$

$$\lambda_2 = a\sin^2\theta - 2|c|\cos\theta\sin\theta\cos(\psi+\chi) + b\cos^2\theta\,, \tag{7}$$

$$Z = e^{-i\chi}\left\{(b-a)\cos\theta\sin\theta + |c|\left[e^{i(\psi+\chi)}\cos^2\theta - e^{-i(\psi+\chi)}\sin^2\theta\right]\right\}. \tag{8}$$

The vanishing of the off-diagonal elements of $U^{-1}AU$ implies that:

$$(b-a)\cos\theta\sin\theta + |c|\left[e^{i(\psi+\chi)}\cos^2\theta - e^{-i(\psi+\chi)}\sin^2\theta\right] = 0\,.$$

This is a complex equation. Taking real and imaginary parts yields two real equations,

$$\tfrac{1}{2}(b-a)\sin 2\theta + |c|\cos 2\theta\cos(\psi+\chi) = 0\,, \tag{9}$$

$$|c|\sin(\psi+\chi) = 0\,. \tag{10}$$

Consider first the special case of $c = 0$. In light of the convention that $\lambda_1 \geq \lambda_2$,

$$
\begin{aligned}
c = 0 \quad &\text{and} \quad a > b \quad \Longrightarrow \quad \theta = 0 \quad \text{and } \chi \text{ is undefined}\,, \\
c = 0 \quad &\text{and} \quad a < b \quad \Longrightarrow \quad \theta = \tfrac{1}{2}\pi \ \text{ and } \chi \text{ is undefined}\,, \\
c = 0 \quad &\text{and} \quad a = b \quad \Longrightarrow \quad \theta \text{ and } \chi \text{ are undefined}\,.
\end{aligned}
$$

In particular, if $c = 0$ and $a = b$, then $A = a\mathbf{I}$ and it follows that $U^{-1}AU = U^{-1}U = a\mathbf{I}$, which is satisfied for any unitary matrix U. Consequently, in this limit θ and χ are arbitrary and hence undefined, as indicated above.

If $c \neq 0$ then eq. (10) yields

$$\sin(\psi+\chi) = 0 \quad \text{and} \quad \cos(\psi+\chi) = \varepsilon\,, \quad \text{where } \varepsilon = \pm 1. \tag{11}$$

We can determine the sign ε as follows. Since $\lambda_1 \geq \lambda_2$, we subtract eqs. (6) and (7) and make use of eq. (11) to obtain,

$$(a-b)\cos 2\theta + 2\varepsilon|c|\sin 2\theta \geq 0\,. \tag{12}$$

Likewise, we insert eq. (11) into eq. (9), which yields

$$(b-a)\sin 2\theta + 2\varepsilon|c|\cos 2\theta = 0\,. \tag{13}$$

Finally, we multiply eq. (12) by $\sin 2\theta$ and eq. (13) by $\cos 2\theta$ and add the two resulting equations. The end result is,

$$2\varepsilon|c| \geq 0\,. \tag{14}$$

By assumption, $c \neq 0$. Thus, it follows that $\varepsilon \geq 0$. Since $\varepsilon = \pm 1$, we can conclude that $\varepsilon = 1$. Hence,

$$\cos(\psi+\chi) = 1\,, \quad \text{for } c \neq 0. \tag{15}$$

Equivalently, $\psi + \chi = 2\pi n$ for some integer n. By the conventions established in eqs. (4) and (5), we take $0 \leq \psi\,,\ \chi < 2\pi$. Hence, it follows that

$$\chi = \begin{cases} 2\pi - \psi\,, & \text{for } c \neq 0 \text{ and } \psi \neq 0\,, \\ 0\,, & \text{for } c \neq 0 \text{ and } \psi = 0\,. \end{cases} \tag{16}$$

We can now determine θ. Inserting eq. (15) into eq. (9) yields

$$\tan 2\theta = \frac{2|c|}{a-b}\,, \quad \text{for } c \neq 0 \text{ and } a \neq b\,. \tag{17}$$

Note that if $a = b$, then eq. (13) yields $\cos 2\theta = 0$. In light of our convention stated in eq. (4),

$$c \neq 0 \quad \text{and} \quad a = b \quad \Longrightarrow \quad \theta = \tfrac{1}{4}\pi\,. \tag{18}$$

If $c \neq 0$ and $a \neq b$, we can use eq. (17) along with the convention that $\sin 2\theta \geq 0$ [cf. eq. (4)] to conclude that

$$\sin 2\theta = \frac{2|c|}{\sqrt{(a-b)^2 + 4|c|^2}}\,. \tag{19}$$

$$\cos 2\theta = \frac{a-b}{\sqrt{(a-b)^2 + 4|c|^2}}\,. \tag{20}$$

Eq. (20) implies that the sign of $a - b$ determines whether $0 < \theta < \frac{1}{4}\pi$ or $\frac{1}{4}\pi < \theta < \frac{1}{2}\pi$. The former corresponds to $a - b > 0$ while the latter corresponds to $a - b < 0$. The borderline case of $a = b$ has already been treated in eq. (18).

To summarize, if $c \neq 0$, then eqs. (16), (19) and (20) uniquely specify the diagonalizing matrix U [in the conventions stated in eqs. (4) and (5)]. When $c = 0$ and $a \neq b$, χ is arbitrary and $\theta = 0$ or $\frac{1}{2}\pi$ for the two cases of $a > b$ or $a < b$, respectively.* Finally, if $c = 0$ and $a = b$, then $A = a\mathbf{I}$, in which case U is arbitrary.

*Note that in the case of $c = 0$ and $a < b$, the matrix A is diagonal. Nevertheless, the "diagonalizing" matrix, $U \neq \mathbf{I}$. Indeed, in this case $\theta = \frac{1}{2}\pi$, and $U^{-1}AU$ simply interchanges the two diagonal elements of A to ensure that $\lambda_1 \geq \lambda_2$ in eq. (2), as required by the convention adopted below eq. (1).

Diagonalization of a real symmetric 2 × 2 matrix

Finally, we can easily treat the special case in which the matrix A is real. In this case, we can diagonalize a real symmetric 2×2 matrix by a real orthogonal matrix. The two eigenvalues are still given by eq. (1) in the convention that $\lambda_1 \geq \lambda_2$, although the absolute values signs are no longer needed since for real values of c, we have $|c|^2 = c^2$. Moreover, since c is real, eq. (5) implies that if $c \neq 0$ then $\psi = 0$ or $\psi = \pi$. Eq. (16) then yields

$$\chi = \begin{cases} 0\,, & \text{for } c \neq 0 \text{ and } \psi = 0\,, \\ \pi\,, & \text{for } c \neq 0 \text{ and } \psi = \pi\,, \end{cases} \tag{21}$$

which is equivalent to the statement that

$$e^{i\chi} = \operatorname{sgn} c\,, \quad \text{for real } c \neq 0. \tag{22}$$

It is convenient to redefine $\theta \to \theta \operatorname{sgn} c$ in eq. (3). With this modification, the range of θ can be taken as†

$$-\tfrac{1}{2}\pi < \theta \leq \tfrac{1}{2}\pi\,. \tag{23}$$

The diagonalizing matrix U is now a real orthogonal 2×2 matrix,

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \text{where} \quad \begin{cases} c > 0 & \Longrightarrow \qquad 0 < \theta < \tfrac{1}{2}\pi\,, \\ c < 0 & \Longrightarrow \qquad -\tfrac{1}{2}\pi < \theta < 0\,. \end{cases}$$

Hence, for real $c \neq 0$ with the range of θ given by eq. (23), eqs. (17) and (19) are modified by replacing $|c|$ with c in their numerators. That is,

$$\sin 2\theta = \frac{2c}{\sqrt{(a-b)^2 + 4c^2}}\,, \tag{24}$$

$$\cos 2\theta = \frac{a-b}{\sqrt{(a-b)^2 + 4c^2}}\,. \tag{25}$$

The sign of c determines the quadrant in which θ lives. Moreover, eq. (25) provides additional information. For $c > 0$, the sign of $a - b$ determines whether $0 < \theta < \frac{1}{4}\pi$ or $\frac{1}{4}\pi < \theta < \frac{1}{2}\pi$. The former corresponds to $a - b > 0$ while the latter corresponds to $a - b < 0$. Likewise, for $c < 0$, the sign of $a - b$ determines whether $-\frac{1}{2}\pi < \theta < -\frac{1}{4}\pi$ or $-\frac{1}{4}\pi < \theta < 0$. The former corresponds to $a - b < 0$ while the latter corresponds to $a - b > 0$. The borderline cases are likewise determined:

$$\begin{aligned} c = 0 \quad &\text{and} \quad a > b \quad \Longrightarrow \quad \theta = 0\,, \\ c = 0 \quad &\text{and} \quad a < b \quad \Longrightarrow \quad \theta = \tfrac{1}{2}\pi\,, \\ a = b \quad &\text{and} \quad c > 0 \quad \Longrightarrow \quad \theta = \tfrac{1}{4}\pi\,, \\ a = b \quad &\text{and} \quad c < 0 \quad \Longrightarrow \quad \theta = -\tfrac{1}{4}\pi\,. \end{aligned}$$

If $c = 0$ and $a = b$, then $A = a\mathbf{I}$ and it follows that $U^{-1}AU = U^{-1}U = \mathbf{I}$, which is satisfied for any invertible matrix U. Consequently, in this limit θ is arbitrary (and hence undefined).

† Using $\cos(\theta + \pi) = -\cos\theta$ and $\sin(\theta + \pi) = -\sin\theta$, it follows that shifting $\theta \to \theta + \pi$ simply multiplies U by an overall factor of -1. In particular, $U^{-1}AU$ is unchanged. Hence, the convention $-\frac{1}{2}\pi < \theta \leq \frac{1}{2}\pi$ may be chosen without loss of generality.