## Diagonalization of a $2 \times 2$ real symmetric matrix

Consider the most general real symmetric $2 \times 2$ matrix

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

where $a$, $b$ and $c$ are arbitrary *real* numbers. In these notes, we will compute the eigenvalues and eigenvectors of $A$, and then find the real orthogonal matrix that diagonalizes $A$.

The eigenvalues are the roots of the characteristic equation:

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

The two roots, $\lambda_1$ and $\lambda_2$, can be determined from the quadratic formula. Noting that $(a+b)^2 - 4(ab-c^2) = (a-b)^2 + 4c^2$, the two roots can be written as:

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

where by convention we take $\lambda_1 \geq \lambda_2$.

Since $(a-b)^2 + 4c^2 \geq 0$ (as the sum of two squares must be non-negative), eq. (1) implies that $\lambda_1$ and $\lambda_2$ are real. We next work out the two eigenvectors and demonstrate that they are orthogonal. It is convenient to define

$$D \equiv \sqrt{(a-b)^2+4c^2} \tag{2}$$

We first solve the eigenvalue equation,

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \tfrac{1}{2}(a+b+D)\begin{pmatrix} x \\ y \end{pmatrix},$$

This yields two equations:

$$\begin{aligned} ax + cy &= \tfrac{1}{2}(a+b+D)x\,, \\ cx + by &= \tfrac{1}{2}(a+b+D)y\,, \end{aligned}$$

which can be rewritten as:

$$\tfrac{1}{2}(a-b-D)x + cy = 0\,, \tag{3}$$

$$cx + \tfrac{1}{2}(b-a-D)y = 0\,. \tag{4}$$

One can show that eq. (4) is a multiple of eq. (3) [as it must be since the rank of the matrix $A - \lambda_1 \mathbf{I}$ is one]. Simply multiply eq. (4) by $(a-b-D)/(2c)$ to obtain

$$\tfrac{1}{2}(a-b-D)x + \frac{(a-b-D)(b-a-D)y}{4c} = \tfrac{1}{2}(a-b-D)x + \frac{[D^2-(a-b)^2]y}{4c} = 0\,.$$

Using eq. (2), $D^2 - (a-b)^2 = 4c^2$, and the above equation reduces to

$$\tfrac{1}{2}(a-b-D)x + cy = 0\,,$$

which is equivalent to eq. (3). Solving for $y$ yields

$$y = \frac{(b-a+D)x}{2c}\,,$$

which means that the eigenvector corresponding to eigenvalue $\lambda_1$ is given by

$$\begin{pmatrix} x \\ y \end{pmatrix}_1 = \frac{x}{2c}\begin{pmatrix} 2c \\ b-a+D \end{pmatrix}.$$

Since $\lambda_2$ differs from $\lambda_1$ by changing the sign of $D$, it follows without further computation that the eigenvector corresponding to eigenvalue $\lambda_2$ is given by

$$\begin{pmatrix} x \\ y \end{pmatrix}_2 = \frac{x}{2c}\begin{pmatrix} 2c \\ b-a-D \end{pmatrix}.$$

To show that the two eigenvectors are orthogonal, we evaluate the dot product of $(x \;\; y)_1$ and $(x \;\; y)_2$, which is equal to $x_1x_2 + y_1y_2$. Inserting the corresponding vector components, we end up with:

$$\frac{x^2}{4c^2}\left[4c^2 + (b-a+D)(b-a-D)\right] = \frac{x^2}{4c^2}\left[4c^2 + (a-b)^2 - D^2\right] = \frac{x^2}{c^2}\left[4c^2 - 4c^2\right] = 0\,,$$

after making use of $D^2 - (a-b)^2 = 4c^2$ [cf. eq. (2)].

We now propose to find the real orthogonal matrix that diagonalizes $A$. The most general $2 \times 2$ real orthogonal matrix $S$ with determinant equal to 1 must have the following form:

$$S = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Using this result, we shall determine $\theta$ in terms $a$, $b$ and $c$ such that

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

where $\lambda_1$ and $\lambda_2$ are the eigenvalues of $A$ obtained in eq. (1). The most straightforward approach is to compute $S^{-1}AS$ explicitly. Since the off-diagonal terms must vanish, one obtains a constraint on the angle $\theta$.

$$\begin{aligned}
S^{-1}AS &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a & c \\ c & b \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a\cos\theta + c\sin\theta & -a\sin\theta + c\cos\theta \\ c\cos\theta + b\sin\theta & -c\sin\theta + b\cos\theta \end{pmatrix} \\
&= \begin{pmatrix} a\cos^2\theta + 2c\cos\theta\sin\theta + b\sin^2\theta & (b-a)\cos\theta\sin\theta + c(\cos^2\theta - \sin^2\theta) \\ (b-a)\cos\theta\sin\theta + c(\cos^2\theta - \sin^2\theta) & a\sin^2\theta - 2c\cos\theta\sin\theta + b\cos^2\theta \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} . \qquad (5)
\end{aligned}$$

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

$$(b-a)\cos\theta\sin\theta + c(\cos^2\theta - \sin^2\theta) = 0 \,.$$

Using $\sin 2\theta = 2\sin\theta\cos\theta$ and $\cos 2\theta = \cos^2\theta - \sin^2\theta$, we can rewrite the above equation as

$$\tfrac{1}{2}(b-a)\sin 2\theta + c\cos 2\theta = 0 \,.$$

It follows that:

$$\boxed{\tan 2\theta = \frac{2c}{a-b}} \qquad (6)$$

after writing $\tan 2\theta = \sin 2\theta / \cos 2\theta$.

Let us now consider the range of the angle $\theta$. You might think that $0 \leq \theta < 2\pi$. However, since

$$\cos(\theta + \pi) = -\cos\theta \,, \qquad \text{and} \qquad \sin(\theta + \pi) = -\sin\theta \,,$$

it follows that shifting $\theta \to \theta + \pi$ simply multiplies $S$ by an overall factor of $-1$. Thus, $S^{-1}AS$ is unchanged. Hence, without loss of generality, we may assume that $0 \leq \theta < \pi$. Unfortunately, eq. (6) does not distinguish between the two intervals $0 \leq \theta \leq \pi/2$ or $\pi/2 \leq \theta < \pi$, since $\tan 2\theta = \tan(2\theta + \pi)$ is unchanged if $\theta \to \theta + \pi/2$.

However, we have not yet used all the available information. In particular, the diagonal elements of eq. (5) also provide some information on the possible values of $\theta$. Summing the diagonal terms of the matrices in eq. (5) yields:

$$\begin{aligned}
\lambda_1 + \lambda_2 &= (a\cos^2\theta + 2c\cos\theta\sin\theta + b\sin^2\theta) + (a\sin^2\theta - 2c\cos\theta\sin\theta + b\cos^2\theta) \\
&= (a+b)(\cos^2\theta + \sin^2\theta) = a + b \,,
\end{aligned}$$

which is independent of $\theta$. This is not surprising since we know that

$$\mathrm{Tr}\ A = \lambda_1 + \lambda_2 = a + b \,.$$

However, $\lambda_1 - \lambda_2$ does depend on $\theta$:

$$\begin{aligned}\lambda_1 - \lambda_2 &= (a\cos^2\theta + 2c\cos\theta\sin\theta + b\sin^2\theta) - (a\sin^2\theta - 2c\cos\theta\sin\theta + b\cos^2\theta) \\ &= (a-b)(\cos^2\theta - \sin^2\theta) + 4c\sin\theta\cos\theta = (a-b)\cos 2\theta + 2c\sin 2\theta\,. \qquad (7)\end{aligned}$$

From eqs. (1) and (7), we obtain

$$\lambda_1 - \lambda_2 = \sqrt{(a-b)^2 + 4c^2} = (a-b)\cos 2\theta + 2c\sin 2\theta\,. \qquad (8)$$

Using eq. (6) to write:

$$a - b = \frac{2c}{\tan 2\theta} = \frac{2c\cos 2\theta}{\sin 2\theta}\,,$$

and inserting this on the left hand side of eq. (8), the latter reduces to:

$$(a-b)\cos 2\theta + 2c\sin 2\theta = 2c\frac{\cos^2 2\theta}{\sin 2\theta} + 2c\sin 2\theta = \frac{2c}{\sin 2\theta}\left(\cos^2 2\theta + \sin^2 2\theta\right) = \frac{2c}{\sin 2\theta}\,.$$

Substituting this result back into eq. (8) and solving for $\sin 2\theta$, we find:

$$\boxed{\sin 2\theta = \frac{2c}{\sqrt{(a-b)^2 + 4c^2}}} \qquad (9)$$

We can also obtain $\cos 2\theta$ using eqs. (6) and (9):

$$\boxed{\cos 2\theta = \frac{a-b}{\sqrt{(a-b)^2 + 4c^2}}} \qquad (10)$$

Eq. (9) tells us in which quadrant $\theta$ lives. If $0 < \theta < \frac{1}{2}\pi$, then $\sin 2\theta > 0$, which implies that $c > 0$. If $\frac{1}{2}\pi < \theta < \pi$, then $\sin 2\theta < 0$, which implies that $c < 0$. Thus, the sign of $c$ determines the quadrant of $\theta$. Eq. (10) 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, if $c < 0$, the sign of $a - b$ determines whether $\frac{1}{2}\pi < \theta < \frac{3}{4}\pi$ or $\frac{3}{4}\pi < \theta < \pi$. The former corresponds to $a - b < 0$ while the latter corresponds to $a - b > 0$. The borderline cases are likewise determined:

$$\begin{aligned}
c = 0 \quad &\text{and} \quad a > b \quad \Longrightarrow \quad \theta = 0\,, \\
c = 0 \quad &\text{and} \quad a < b \quad \Longrightarrow \quad \theta = \tfrac{1}{2}\pi\,, \\
a = b \quad &\text{and} \quad c > 0 \quad \Longrightarrow \quad \theta = \tfrac{1}{4}\pi\,, \\
a = b \quad &\text{and} \quad c < 0 \quad \Longrightarrow \quad \theta = \tfrac{3}{4}\pi\,.
\end{aligned}$$

If $c = 0$ and $a = b$, then $A = \mathbf{I}$ and it follows that $S^{-1}AS = S^{-1}S = \mathbf{I}$, which is satisfied for any invertible matrix $S$. Consequently, in this limit $\theta$ is undefined.