Symmetry Primer
Laurette S. Tuckerman

Laurette S. Tuckerman

1 Complex eigenvalues

What does it mean for a real matrix to have complex eigenvalues and eigenvectors? Essentially, these serve as a convenient shorthand for rotation, i.e. trigonometry, just as $e^{i\theta}$ is taken to mean $\cos(\theta) + i\sin(\theta)$. When eigenvalues and eigenvectors are real, multiplication of the eigenvector by the matrix results in a stretched or contracted version of the same eigenvector. When eigenvalues and eigenvectors are complex, the eigenspace is two-dimensional. Any vector in this two-dimensional space is rotated as well as being lengthened or shortened. The complex equation

$$A(\psi_r + i\psi_i) = (\sigma + i\omega)(\psi_r + i\psi_i) \tag{1}$$

is equivalent to the pair of real equations

$$A\psi_r = \sigma\psi_r - \omega\psi_i \tag{2a}$$
$$A\psi_i = \omega\psi_r + \sigma\psi_i \tag{2b}$$

showing that the action of A is to mix ψ_r and ψ_i.

There is no significance to the designation of the real and imaginary components of an eigenvector. This is confirmed by the fact that an eigenvector is determined up to multiplication by a constant. Multiplication of $\psi_r + i\psi_i$ by i interchanges the real and imaginary components:

$$i(\psi_r + i\psi_i) = -\psi_i + i\psi_r \tag{3}$$

Any mixture can be obtained by multiplication by any complex number, e.g.

$$e^{i\theta}(\psi_r + i\psi_i) = [\cos(\theta)\psi_r - \sin(\theta)\psi_i] + i\,[\sin(\theta)\psi_r + \cos(\theta)\psi_i] \tag{4}$$

The action of A is similar on ψ_r or ψ_i. The value of $|\lambda|$ determines how much the vector is stretched or compressed, while the value of θ determines how much it rotates in the (ψ_r, ψ_i) plane.

The eigenvector-eigenvalue or Jordan decomposition of A is written

$$AE = E\Lambda \tag{5}$$

where E is a matrix whose columns are the eigenvectors and Λ is the diagonal matrix of eigenvalues. For pairs of complex eigenvalues, $\psi_{1,2} = \psi_r \pm i\psi_i$, the matrices E and Λ are of the form:

$$E = \begin{bmatrix} \psi_r & \psi_i & \psi_3 & \cdots & \psi_N \\ \downarrow & \downarrow & \downarrow & \cdots & \downarrow \\ \downarrow & \downarrow & \downarrow & \cdots & \downarrow \\ \downarrow & \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} \qquad \Lambda = \begin{bmatrix} \sigma & -\omega & 0 & \cdots & 0 \\ \omega & \sigma & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \lambda_N \end{bmatrix} \tag{6}$$

We have

$$f(A)E = Ef(\Lambda) \tag{7}$$

That is, the eigenvalues of the matrix $f(A)$ are f applied to the eigenvalues of A and the eigenvectors of $f(A)$ are the same as those of A. For an analytic function f, this can be shown by Taylor expanding f and using $A = E^{-1}\Lambda E$ to transform each power:

$$\begin{aligned} f(A) &= f_0 I + f_1 A + f_2 A^2 + f_3 A^3 + \dots \\ &= f_0 I + f_1 E^{-1}\Lambda E + f_2 (E^{-1}\Lambda E)\,(E^{-1}\Lambda E) + f_3 (E^{-1}\Lambda E)\,(E^{-1}\Lambda E)\,(E^{-1}\Lambda E) + \dots \\ &= f_0 I + f_1 E^{-1}\Lambda E + f_2 E^{-1}\Lambda^2 E + f_3 E^{-1}\Lambda^3 E + \dots \\ &= E^{-1}(f_0 I + f_1 \Lambda + f_2 \Lambda^2 + f_3 \Lambda^3) E + \dots \\ &= E^{-1} f(\Lambda) E \end{aligned} \tag{8}$$

where $f_0, f_1, \dots$ are real coefficients. Λ is block diagonal. Entries corresponding to real eigenvalues just squared, while the 2×2 blocks corresponding to complex eigenvalues become:

$$\begin{bmatrix} \sigma & -\omega \\ \omega & \sigma \end{bmatrix} \begin{bmatrix} \sigma & -\omega \\ \omega & \sigma \end{bmatrix} = \begin{bmatrix} \sigma^2 - \omega^2 & -2\sigma\omega \\ 2\sigma\omega & \sigma^2 - \omega^2 \end{bmatrix} \tag{9}$$

Note that $(\sigma + i\omega)^2 = \sigma^2 - \omega^2 + i\,2\sigma\omega$.

We turn our attention to the evolution equation

$$\frac{du}{dt} = Au \tag{10}$$

The solution of this equation is

$$u(t) = e^{At} u(0) \tag{11}$$

The matrix $\exp(\Lambda t)$ is:

$$\Lambda = \begin{bmatrix} e^{\sigma t}\cos(\omega t) & -e^{\sigma t}\sin(\omega t) & 0 & \cdots & 0 \\ e^{\sigma t}\sin(\omega t) & e^{\sigma t}\cos(\omega t) & 0 & \cdots & 0 \\ 0 & 0 & e^{\lambda_3 t} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & e^{\lambda_N t} \end{bmatrix} \tag{12}$$

Multiplication by A causes almost all vectors to increase in length if any eigenvalue satisfies $|\lambda| > 1$ (since almost all vectors will contain a component of the growing eigenvector). Integration of (10), or equivalently, multiplication by $\exp(At)$ causes almost all vectors to increase in length if any eigenvalue satisfies $Re(\lambda) > 0$.

If we begin with an initial condition belonging to one of the real eigenspaces, e.g.:

$$u(0) = \alpha\psi_3 \tag{13}$$

then the evolution will be

$$u(t) = e^{At}u(0) = e^{\sigma t}\alpha\psi_3 = e^{\sigma t}u(0) \tag{14}$$

If the initial condition belongs to one of the complex eigenspaces, e.g.:

$$u(0) = \alpha\psi_r + \beta\psi_i \tag{15}$$

then the evolution will be

$$\begin{aligned} u(t) = e^{At}u(0) &= e^{\sigma t}\left[\alpha(\cos(\omega t)\psi_r - \sin(\omega t)\psi_i) + \beta(\sin(\omega t)\psi_r + \cos(\omega t)\psi_i)\right] \\ &= e^{\sigma t}\left[(\alpha\cos(\omega t) + \beta\sin(\omega t))\psi_r + (-\alpha\sin(\omega t) + \beta\cos(\omega t))\psi_i\right] \\ &= ce^{\sigma(t-t_0)}\left[\cos(\omega(t-t_0)\psi_r + \sin(\omega(t-t_0))\psi_i\right] \end{aligned} \tag{16}$$

where

$$\tan(\omega t_0) = \beta/\alpha \qquad ce^{-\sigma t_0} = \sqrt{\alpha^2 + \beta^2} \tag{17}$$

This describes a combination of stretching (or shrinking) at a rate measured by σ and rotation at the rate ω.

Linear algebra texts usually state:

If a matrix is self-adjoint, then its eigenvalues are real. (I)

In order to define self-adjointness, we must first define an inner product.

An inner product $\langle x, y\rangle$ is a function of the vectors x, y satisfying:

$$\langle x, y\rangle = \langle y, x\rangle^* \tag{18a}$$

$$\langle x, \alpha y\rangle = \alpha\langle x, y\rangle \tag{18b}$$

$$\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle \tag{18c}$$

Given an inner product $\langle,\rangle$ and a matrix A, the adjoint A^* is defined to satisfy:

$$\langle x, Ay\rangle = \langle A^*x, y\rangle \tag{19}$$

The proof of (I) uses the properties of the adjoint and the inner product:

$$\begin{aligned} \lambda_k\psi_k &= A\psi_k \\ \langle\psi_k, \lambda_k\psi_k\rangle &= \langle\psi_k, A\psi_k\rangle \\ \lambda_k\langle\psi_k, \psi_k\rangle &= \langle A^*\psi_k, \psi_k\rangle = \langle A\psi_k, \psi_k\rangle \\ &= \langle\lambda_k\psi_k, \psi_k\rangle = \lambda_k^*\langle\psi_k, \psi_k\rangle \\ \lambda_k &= \lambda_k^* \end{aligned}$$

How should the oft-quoted statement (I) be interpreted, given that the hypothesis (self-adjointness) is inner-product dependent and the conclusion (real eigenvalues) is not? In addition, (I) gives sufficient but not necessary conditions for the eigenvalues to be real. (It is easy to find matrices which are not self-adjoint under some inner product, but which nevertheless have real eigenvalues.) What are the necessary conditions for eigenvalues to be real?

There can be only two possible logically correct interpretations of (I), either:

If a matrix is self-adjoint under one inner product, then its eigenvalues are real. (II)

or:

If a matrix is self-adjoint under all inner products, then its eigenvalues are real. (III)

Actually, a third interpretation, combining these two, is also possible:

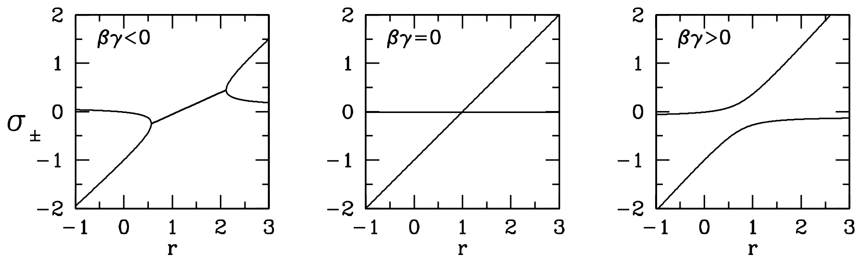

Figure 1: Behavior of real part of eigenvalues of 2×2 matrix depending on a parameter r. There are three cases, depending on the sign of the product $\beta\gamma$ of the off-diagonal terms.

If a matrix is self-adjoint under one inner product, then it is self-adjoint under all inner products and its eigenvalues are real. (IV)

The correct statement is (II), and it also provides the necessary conditions. We make the additional hypothesis that the eigenvectors form a complete set, i.e. that there are N linearly independent eigenvectors. Suppose a matrix A has real eigenvectors ψ_k and real eigenvalues λ_k. Define an inner product via $\langle \psi_j, \psi_k \rangle = \delta_{jk}$. (This defines an inner product only if the eigenvectors are real.) By construction, A is self-adjoint under this inner product.

A typical $N \times N$ matrix has N distinct eigenvalues and N linearly independent eigenvectors (counting the real and imaginary components as two vectors). However, for matrices which depend on a parameter r, it is common for a pair of real eigenvalues to approach one another as the parameter is varied. Consider the 2×2 matrix

$$\begin{pmatrix} \alpha r & \beta \\ \gamma & \delta r \end{pmatrix} \tag{20}$$

Its eigenvalues are:

$$\lambda_\pm = \frac{(\alpha+\delta)r}{2} \pm \sqrt{\left(\frac{(\alpha-\delta)r}{2}\right)^2 + \beta\gamma} \tag{21}$$

For r large, the eigenvalues (21) are real. As r approaches zero, there are three possibilities for the behavior of the eigenvalues and eigenvectors, illustrated in figure (1). If $\beta\gamma > 0$, then the curves $(r, \lambda_\pm)$ described by (21) form two hyperbolas whose asymptotes are $\lambda = \alpha r$ and $\lambda = \delta r$, the eigenvalues of matrix (20) with its off-diagonal entries set to zero. This situation is termed *avoided crossing*). If $\beta\gamma = 0$, then the eigenvalues are identical with these asymptotes. In either case, the eigenvalues and eigenvectors are real and distinct before and after the intersection. At the point of intersection, the eigenvalue is double, and the eigenspace is two-dimensional and spanned by any two linearly independent members of this eigenspace. If (20) is symmetric, then $\beta = \gamma$, so the behavior of the eigenvalues of symmetric matrices is described above.

The other possibility, if $\beta\gamma < 0$, is that the eigenvalues form a complex conjugate pair after the intersection of the two real eigenvalues. Away from $r = 0$, the real eigenvalues are again hyperbolas, with asymptotes $\lambda = \alpha r$ and $\lambda = \delta r$, but occupy different quadrants of the plane.

When real eigenvalues merge to form a complex conjugate pair, then at the point of intersection, the matrix is a Jordan block, which has a single eigenvector ψ, accompanied by a generalized eigenvector ϕ.

These obey:

$$(A - \lambda I)\psi = 0 \tag{22}$$
$$(A - \lambda I)\phi = \psi \tag{23}$$

Since

$$(A - \lambda I)(c\psi + d\phi) = d\psi \tag{24}$$

the generalized eigenvector is not uniquely defined; any multiple of the eigenvector can be added to it. An inner product can be specified to remove this degeneracy by imposing $\langle \phi, \psi \rangle$ = 0.

We use as an example the matrix

$$\begin{bmatrix} 0 & 1 \\ r & 0 \end{bmatrix} \tag{25}$$

Its eigenvalues are $\lambda = \pm\sqrt{r}$, so they are real for r positive and imaginary for r negative. For $r > 0$ we have:

$$\lambda_\pm = \pm\sqrt{r}, \quad \psi_\pm = \begin{pmatrix} 1 \\ \pm\sqrt{r} \end{pmatrix} \tag{26}$$

The two eigenvectors approach one another as r approaches zero. For $r = 0$, the matrix becomes the Jordan block

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{27}$$

We have a multiple eigenvalue $\lambda_+ = \lambda_-$ and a single eigenvalue $\psi_+ = \psi_- = (1, 0)$. We have

$$\lambda_+ = \lambda_- = 0, \quad \psi_+ = \psi_- = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \phi = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{28}$$

where ϕ is a generalized eigenvector. For $r < 0$, we have:

$$\lambda_\pm = \pm i\sqrt{-r}, \quad \psi_\pm = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \pm i \begin{pmatrix} 0 \\ \sqrt{-r} \end{pmatrix} \tag{29}$$

We emphasize again that there is no significance to which of the two vectors we call the real component and which the imaginary component. We could just as well write

$$\psi_\pm = \begin{pmatrix} 0 \\ \sqrt{-r} \end{pmatrix} \pm i \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{30}$$

or

$$\psi_\pm = \begin{pmatrix} \cos(\theta) \\ \mp\sqrt{-r}\sin(\theta) \end{pmatrix} \pm i \begin{pmatrix} \sin(\theta) \\ \pm\sqrt{-r}\cos(\theta) \end{pmatrix} \tag{31}$$

for any θ. The eigenvectors are thus located on an ellipse, whose major axis $|x| \leq 1$ corresponds to the single eigenvector that exists at $r = 0$ and whose minor axis is $|y| \leq \sqrt{-r}$. Note that, while real eigenvectors $\psi_\pm$ can be normalized separately, the real and imaginary parts of a complex eigenpair must be normalized by the same factor.