Overview

We've looked at eigenvalues and eigenvectors from several perspectives, studying how to find them and what they tell you about the linear transformation associated to a matrix.

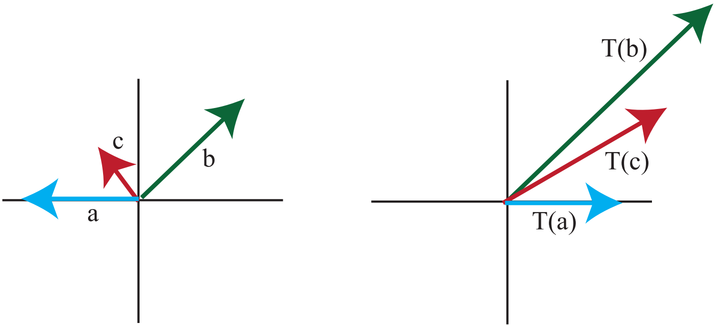
Question

What happens when the characteristic equation has complex roots?

From Lay, §5.5

Warm-up unquiz for review

Suppose that a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts as shown in the picture:



Write a matrix for T with respect to a basis of your choice.

Existence of Complex Eigenvalues

Since the characteristic equation of an $n \times n$ matrix involves a polynomial of degree n, there will be times when the roots of the characteristic equation will be complex. Thus, even if we start out considering matrices with real entries, we're naturally lead to consider complex numbers.

We'll focus on understanding what **complex** eigenvalues mean when **the entries of the matrix with which we are working are all real numbers**. For simplicity, we'll restrict to the case of 2×2 matrices.

Example 1

Let $A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$ for some real φ. The roots of the characteristic equation are $\cos\varphi \pm i\sin\varphi$.

Example 1

Let $A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$ for some real φ. The roots of the characteristic equation are $\cos\varphi \pm i\sin\varphi$.

What does the linear transformation $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T_A(\mathbf{x}) = A\mathbf{x}$ (for all $\mathbf{x} \in \mathbb{R}^2$) do to vectors in $\mathbb{R}^2$?

Example 1

Let $A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$ for some real φ. The roots of the characteristic equation are $\cos\varphi \pm i\sin\varphi$.

What does the linear transformation $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T_A(\mathbf{x}) = A\mathbf{x}$ (for all $\mathbf{x} \in \mathbb{R}^2$) do to vectors in $\mathbb{R}^2$?

Since the i^{th} column of the matrix is $T(\mathbf{e_i})$, we see that the linear transformation T_A is the transformation that rotates each point in $\mathbb{R}^2$ about the origin through an angle φ, with counterclockwise rotation for a positive angle.

A rotation in $\mathbb{R}^2$ cannot have a real eigenvector unless $\varphi = 2k\pi$ or $\varphi = \pi + 2k\pi$ for $k \in \mathbb{Z}$!

What about (complex) eigenvectors for such an A?

Let's take $\varphi = \pi/3$, so that multiplication by A corresponds to a rotation through $\pi/3$ (60^0). Then we get

$$A = \begin{bmatrix} \cos \pi/3 & -\sin \pi/3 \\ \sin \pi/3 & \cos \pi/3 \end{bmatrix} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$$

What happens when we try to find eigenvalues and eigenvectors?

Let's take $\varphi = \pi/3$, so that multiplication by A corresponds to a rotation through $\pi/3$ (60^0). Then we get

$$A = \begin{bmatrix} \cos\pi/3 & -\sin\pi/3 \\ \sin\pi/3 & \cos\pi/3 \end{bmatrix} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$$

What happens when we try to find eigenvalues and eigenvectors?

The characteristic polynomial of A is

$$(1/2 - \lambda)^2 + (\sqrt{3}/2)^2 = \lambda^2 - \lambda + 1$$

and the eigenvalues are

$$\lambda = \frac{1 \pm \sqrt{1-4}}{2} = \frac{1}{2} \pm \frac{\sqrt{3}}{2} i.$$

Take $\lambda_1 = \dfrac{1}{2} + \dfrac{\sqrt{3}}{2}i$. We find the eigenvectors in the usual way by solving $(A - \lambda_1 I)\mathbf{x} = \mathbf{0}$.

Take $\lambda_1 = \frac{1}{2} + \frac{\sqrt{3}}{2} i$. We find the eigenvectors in the usual way by solving $(A - \lambda_1 I)\mathbf{x} = \mathbf{0}$.

$$A - \lambda_1 I = \begin{bmatrix} -i\sqrt{3}/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -i\sqrt{3}/2 \end{bmatrix} \to \begin{bmatrix} i & 1 \\ 0 & 0 \end{bmatrix}.$$

Take $\lambda_1 = \frac{1}{2} + \frac{\sqrt{3}}{2}i$. We find the eigenvectors in the usual way by solving $(A - \lambda_1 I)\mathbf{x} = \mathbf{0}$.

$$A - \lambda_1 I = \begin{bmatrix} -i\sqrt{3}/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -i\sqrt{3}/2 \end{bmatrix} \rightarrow \begin{bmatrix} i & 1 \\ 0 & 0 \end{bmatrix}.$$

We solve the associated equation as usual, so we see that $ix + y = 0$.

Thus one possible eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$.

(All the other associated eigenvectors are of the form $\alpha\mathbf{x}_1 = \begin{bmatrix} \alpha \\ -i\alpha \end{bmatrix}$, where α is any non-zero number in $\mathbb{C}$.)

For $\lambda_2 = \frac{1}{2} - \frac{\sqrt{3}}{2}i$ we get $\mathbf{x}_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ as an associated complex eigenvector.

(All the other associated eigenvectors are of the form $\alpha\mathbf{x}_2 = \begin{bmatrix} \alpha \\ i\alpha \end{bmatrix}$, where α is any non-zero number in $\mathbb{C}$.)

We can check that these two vectors are in fact eigenvectors:

$$\begin{aligned} A\mathbf{x}_1 &= \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} \\ &= \begin{bmatrix} 1/2 + i\sqrt{3}/2 \\ \sqrt{3}/2 - i/2 \end{bmatrix} \\ &= \left(\frac{1}{2} + \frac{\sqrt{3}}{2} i \right) \begin{bmatrix} 1 \\ -i \end{bmatrix}. \end{aligned}$$

Similarly,

$$A\mathbf{x}_2 = \left(\frac{1}{2} - \frac{\sqrt{3}}{2} i \right) \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

Example 2

Find the eigenvectors associated to the matrix $\begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix}$.

Example 2

Find the eigenvectors associated to the matrix $\begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix}$.

The characteristic polynomial is

$$\det \begin{bmatrix} 5-\lambda & -2 \\ 1 & 3-\lambda \end{bmatrix} = (5-\lambda)(3-\lambda)+2 = \lambda^2 - 8\lambda + 17.$$

The roots are

$$\lambda = \frac{8 \pm \sqrt{64-68}}{2} = \frac{8 \pm \sqrt{-4}}{2} = \frac{8 \pm 2i}{2} = 4 \pm i.$$

Example 2

Find the eigenvectors associated to the matrix $\begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix}$.

The characteristic polynomial is

$$\det \begin{bmatrix} 5-\lambda & -2 \\ 1 & 3-\lambda \end{bmatrix} = (5-\lambda)(3-\lambda)+2 = \lambda^2 - 8\lambda + 17.$$

The roots are

$$\lambda = \frac{8 \pm \sqrt{64-68}}{2} = \frac{8 \pm \sqrt{-4}}{2} = \frac{8 \pm 2i}{2} = 4 \pm i.$$

Since complex roots always come in conjugate pairs, it follows that if $a+bi$ is an eigenvalue for A, then $a-bi$ will also be an eigenvalue for A.

Take $\lambda_1 = 4 + i$. We find a corresponding eigenvector:

$$A - \lambda_1 I = \begin{bmatrix} 5 - (4+i) & -2 \\ 1 & 3 - (4+i) \end{bmatrix} = \begin{bmatrix} 1 - i & -2 \\ 1 & -1 - i \end{bmatrix}$$

Row reduction of the usual augmented matrix is quite unpleasant by hand because of the complex numbers.

Take $\lambda_1 = 4 + i$. We find a corresponding eigenvector:

$$A - \lambda_1 I = \begin{bmatrix} 5-(4+i) & -2 \\ 1 & 3-(4+i) \end{bmatrix} = \begin{bmatrix} 1-i & -2 \\ 1 & -1-i \end{bmatrix}$$

Row reduction of the usual augmented matrix is quite unpleasant by hand because of the complex numbers.

However, there is an observation that simplifies matters: Since $4 + i$ is an eigenvalue, the system of equations

$$\begin{array}{rcrcl} (1-i)x_1 & - & 2x_2 & = & 0 \\ x_1 & - & (1+i)x_2 & = & 0 \end{array}$$

has a non trivial solution.

Therefore both equations determine the same relationship between x_1 and x_2, and either equation can be used to express one variable in terms of the other.

As these two equations both give the same information, we can use the second equation. It gives

$$x_1 = (1+i)x_2,$$

where x_2 is a free variable. If we take $x_2 = 1$, we get $x_1 = 1 + i$ and hence an eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1+i \\ 1 \end{bmatrix}$.

As these two equations both give the same information, we can use the second equation. It gives

$$x_1 = (1 + i)x_2,$$

where x_2 is a free variable. If we take $x_2 = 1$, we get $x_1 = 1 + i$ and hence an eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1+i \\ 1 \end{bmatrix}$.

If we take $\lambda_2 = 4 - i$, and proceed as for λ_1 we get that $\mathbf{x}_2 = \begin{bmatrix} 1-i \\ 1 \end{bmatrix}$ is a corresponding eigenvector.

Just as the eigenvalues come in a pair of complex conjugates, and so do the eigenvectors.

Normal form

When a matrix is diagonalisable, it's similar to a diagonal matrix: $A = PDP^{-1}$.
It's also similar to many other matrices, but we think of the diagonal matrix as the "best" representative of the class, in the sense that it expresses the associated linear transformation with respect to a most natural basis (i.e., a basis of eigenvectors.)
Of course, not all matrices are diagonalisable, so today we consider the following question:

Question

Given an arbitrary matrix, is there a "best" representative of its similarity class?

"Best" isn't a precise term, but let's interpret this as asking whether there's some basis for which the action of the associate linear transformation is most transparent.

Example 3

Consider the matrix $\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}$.

The characteristic polynomial is $1 - \lambda^3$, with roots $1, -1 \pm i\frac{\sqrt{3}}{2}$, the three cube roots of unity in $\mathbb{C}$.

A choice of corresponding eigenvectors is, for example,

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 + i\frac{\sqrt{3}}{2} \\ 1 + i\frac{\sqrt{3}}{2} \\ 1 \end{bmatrix}, \begin{bmatrix} -1 - i\frac{\sqrt{3}}{2} \\ 1 - i\frac{\sqrt{3}}{2} \\ 1 \end{bmatrix}.$$

Notice that we have one real eigenvector corresponding to the real eigenvalue 1, and two complex eigenvectors corresponding to the complex eigenvalues. Notice that also in this case the complex eigenvalues and eigenvectors come in pairs of conjugates.

Advantages of complex linear algebra

Doing computations by hand is messier when we work over $\mathbb{C}$, but much of the theory is cleaner! When the scalars are complex, rather than real

- matrices always have eigenvalues and eigenvectors; and
- every linear transformation $T : \mathbb{C}^n \to \mathbb{C}^n$ can be represented by an upper triangular matrix.

We don't have time to explore the implications fully, but we can take a quick look at some of the interesting structure that emerges immediately.

A real matrix acting on $\mathbb{C}$

Eigenvalues come in conjugate pairs.

If A is an $m \times n$ matrix with entries in $\mathbb{C}$, then $\bar{A}$ denotes the matrix whose entries are the complex conjugates of the entries in A.

Let A be an $n \times n$ matrix whose entries are real. Then $\overline{A} = A$.

A real matrix acting on $\mathbb{C}$

Eigenvalues come in conjugate pairs.

If A is an $m \times n$ matrix with entries in $\mathbb{C}$, then $\bar{A}$ denotes the matrix whose entries are the complex conjugates of the entries in A.

Let A be an $n \times n$ matrix whose entries are real. Then $\overline{A} = A$. So

$$\overline{A\mathbf{x}} = \overline{A}\overline{\mathbf{x}} = A\overline{\mathbf{x}}$$

for any vector $\mathbf{x} \in \mathbb{C}^n$.

A real matrix acting on $\mathbb{C}$

Eigenvalues come in conjugate pairs.

If A is an $m \times n$ matrix with entries in $\mathbb{C}$, then $\bar{A}$ denotes the matrix whose entries are the complex conjugates of the entries in A.

Let A be an $n \times n$ matrix whose entries are real. Then $\overline{A} = A$. So

$$\overline{A\mathbf{x}} = \overline{A}\overline{\mathbf{x}} = A\overline{\mathbf{x}}$$

for any vector $\mathbf{x} \in \mathbb{C}^n$.

If λ is an eigenvalue of A and $\mathbf{x}$ is a corresponding eigenvector in $\mathbb{C}^n$, then

$$A\overline{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}.$$

This shows that $\overline{\lambda}$ is also an eigenvalue of A with $\overline{\mathbf{x}}$ a corresponding eigenvector.

So...

A real matrix acting on $\mathbb{C}$

Eigenvalues come in conjugate pairs.
If A is an $m \times n$ matrix with entries in $\mathbb{C}$, then $\bar{A}$ denotes the matrix whose entries are the complex conjugates of the entries in A.
Let A be an $n \times n$ matrix whose entries are real. Then $\overline{A} = A$. So

$$\overline{A\mathbf{x}} = \overline{A}\overline{\mathbf{x}} = A\overline{\mathbf{x}}$$

for any vector $\mathbf{x} \in \mathbb{C}^n$.
If λ is an eigenvalue of A and $\mathbf{x}$ is a corresponding eigenvector in $\mathbb{C}^n$, then

$$A\overline{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}.$$

This shows that $\overline{\lambda}$ is also an eigenvalue of A with $\overline{\mathbf{x}}$ a corresponding eigenvector.
So...
...when A is a real matrix, its complex eigenvalues occur in conjugate pairs.

Some special 2×2 matrices

Consider the matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where a and b are real numbers and not both 0.

Some special 2×2 matrices

Consider the matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where a and b are real numbers and not both 0.

$$C - I\lambda = \begin{bmatrix} a-\lambda & -b \\ b & a-\lambda \end{bmatrix},$$

so the characteristic equation for C is

$$0 = (a-\lambda)^2 + b^2 = \lambda^2 - 2a\lambda + a^2 + b^2.$$

Using the quadratic formula, the eigenvalues of C are

$$\lambda = a \pm bi.$$

So if $b \neq 0$, the eigenvalues are not real.

Some special 2×2 matrices

Consider the matrix $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where a and b are real numbers and not both 0.

$$C - I\lambda = \begin{bmatrix} a-\lambda & -b \\ b & a-\lambda \end{bmatrix},$$

so the characteristic equation for C is

$$0 = (a-\lambda)^2 + b^2 = \lambda^2 - 2a\lambda + a^2 + b^2.$$

Using the quadratic formula, the eigenvalues of C are

$$\lambda = a \pm bi.$$

So if $b \neq 0$, the eigenvalues are not real.

Notice that this generalises our earlier observation about rotation matrices. In fact...

...apply some magic...

If we now take $r = |\lambda| = \sqrt{a^2 + b^2}$ then we can write

$$C = r \begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

where φ is the angle between the positive x-axis and the ray from $(0, 0)$ through (a, b).

...apply some magic...

If we now take $r = |\lambda| = \sqrt{a^2 + b^2}$ then we can write

$$C = r \begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

where φ is the angle between the positive x-axis and the ray from $(0,0)$ through (a,b). Here we used the fact that

$$\left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2 = \frac{a^2+b^2}{r^2} = \frac{r^2}{r^2} = 1\,.$$

...apply some magic...

If we now take $r = |\lambda| = \sqrt{a^2 + b^2}$ then we can write

$$C = r \begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

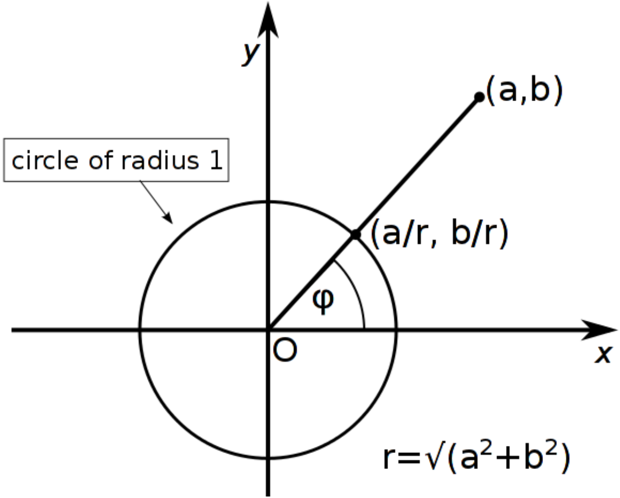
where φ is the angle between the positive x-axis and the ray from $(0, 0)$ through (a, b). Here we used the fact that

$$\left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2 = \frac{a^2 + b^2}{r^2} = \frac{r^2}{r^2} = 1\,.$$

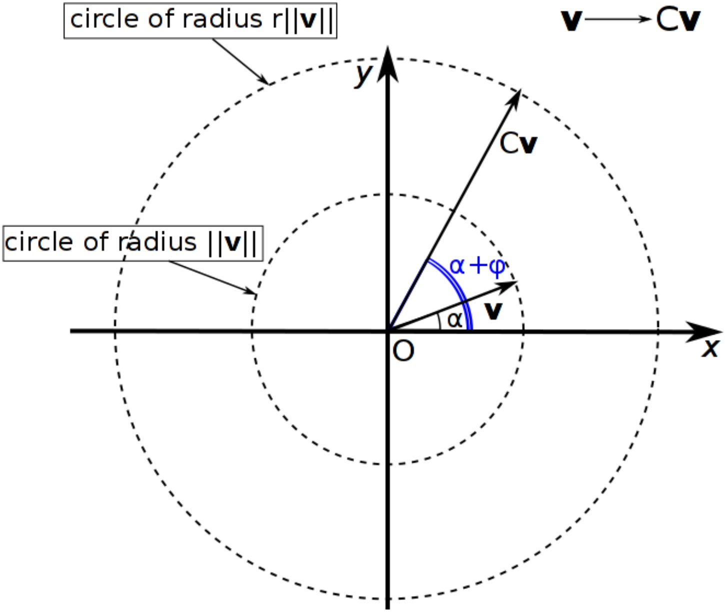
Thus the point $(a/r, b/r)$ lies on the circle of radius 1 with center at the origin and a/r, b/r can be seen as the cosine and sine of the angle between the positive x-axis and the ray from $(0, 0)$ through $(a/r, b/r)$ (which is the same as the angle between the positive x-axis and the ray from $(0, 0)$ through (a, b)).

The transformation $\mathbf{x} \mapsto C\mathbf{x}$ may be viewed as the composition of a rotation through the angle φ and a scaling by $r = |\lambda|$.

The angle φ



The action of C



Example 4

What is the geometric action of $C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ on $\mathbb{R}^2$?

From what we've just seen, C has eigenvalues $\lambda = 1 \pm i$, so $r = \sqrt{1^2 + 1^2} = \sqrt{2}$. We can therefore rewrite C as

$$C = \sqrt{2} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \sqrt{2} \begin{bmatrix} \cos \pi/4 & -\sin \pi/4 \\ \sin \pi/4 & \cos \pi/4 \end{bmatrix}.$$

So C acts as a rotation through $\pi/4$ together with a multiplication by $\sqrt{2}$.

To verify this, we look at the repeated action of C on a point $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
(Note $|\mathbf{x}_0| = 1$.)

$$\mathbf{x}_1 = C\mathbf{x}_0 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, ||\mathbf{x}_1|| = \sqrt{2},$$

$$\mathbf{x}_2 = C\mathbf{x}_1 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, ||\mathbf{x}_2|| = 2,$$

$$\mathbf{x}_3 = C\mathbf{x}_2 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}, ||\mathbf{x}_3|| = 2\sqrt{2}, \ldots$$

If we continue, we'll find a spiral of points each one further away from $(0,0)$ than the previous one.

Real and imaginary parts of vectors

The *complex conjugate* of a complex vector $\mathbf{x}$ in $\mathbb{C}^n$ is the vector $\bar{\mathbf{x}}$ in $\mathbb{C}^n$ whose entries are the complex conjugates of the entries in $\mathbf{x}$.

Real and imaginary parts of vectors

The *complex conjugate* of a complex vector $\mathbf{x}$ in $\mathbb{C}^n$ is the vector $\bar{\mathbf{x}}$ in $\mathbb{C}^n$ whose entries are the complex conjugates of the entries in $\mathbf{x}$.

The *real* and *imaginary parts* of a complex vector $\mathbf{x}$ are the vectors $\operatorname{Re} \mathbf{x}$ and $\operatorname{Im} \mathbf{x}$ formed from the real and imaginary parts of the entries of $\mathbf{x}$.

Real and imaginary parts of vectors

The *complex conjugate* of a complex vector $\mathbf{x}$ in $\mathbb{C}^n$ is the vector $\bar{\mathbf{x}}$ in $\mathbb{C}^n$ whose entries are the complex conjugates of the entries in $\mathbf{x}$.

The *real* and *imaginary parts* of a complex vector $\mathbf{x}$ are the vectors Re $\mathbf{x}$ and Im $\mathbf{x}$ formed from the real and imaginary parts of the entries of $\mathbf{x}$.

If $\mathbf{x} = \begin{bmatrix} 1+2i \\ -3i \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} + i \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}$, then

$$\text{Re } \mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}, \text{Im } \mathbf{x} = \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}, \text{ and}$$

$$\bar{\mathbf{x}} = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} - i \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1-2i \\ 3i \\ 5 \end{bmatrix}.$$

We'll use this idea in the next example.

The rotation hidden in a real matrix with a complex eigenvalue

Example 5

Show that $A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix}$ is similar to a matrix of the form $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

The characteristic polynomial of A is

$$\det \begin{bmatrix} 2-\lambda & 1 \\ -2 & -\lambda \end{bmatrix} = (2-\lambda)(-\lambda) + 2 = \lambda^2 - 2\lambda + 2.$$

So A has complex eigenvalues

$$\lambda = \frac{2 \pm \sqrt{4-8}}{2} = \frac{2 \pm 2i}{2} = 1 \pm i.$$

Take $\lambda_1 = 1 - i$. To find a corresponding eigenvector we find $A - \lambda_1 I$:

$$A - \lambda_1 I = \begin{bmatrix} 2-(1-i) & 1 \\ -2 & 0-(1-i) \end{bmatrix} = \begin{bmatrix} 1+i & 1 \\ -2 & -1+i \end{bmatrix}$$

We can use the first row of the matrix to solve $(A - \lambda_1 I)\mathbf{x} = \mathbf{0}$:

$$(1+i)x_i + x_2 = 0 \quad \text{or} \quad x_2 = -(1+i)x_1.$$

If we take $x_1 = 1$ we get an eigenvector

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1-i \end{bmatrix}$$

We now construct a real 2×2 matrix P:

$$P = \begin{bmatrix} \text{Re } \mathbf{v}_1 & \text{Im } \mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}.$$

We have not justified why we would try this!

We now construct a real 2×2 matrix P:

$$P = \begin{bmatrix} \text{Re } \mathbf{v}_1 & \text{Im } \mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}.$$

We have not justified why we would try this!

Note that $P^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}$.

Then calculate

$$\begin{aligned} C &= P^{-1}AP \\ &= \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

We recognise this matrix, from the previous example, as the composition of a counterclockwise rotation by $\pi/4$ and a scaling by $\sqrt{2}$. This is the rotation "inside" A.

We recognise this matrix, from the previous example, as the composition of a counterclockwise rotation by $\pi/4$ and a scaling by $\sqrt{2}$. This is the rotation "inside" A. We can write A:

$$A = PCP^{-1} = P \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} P^{-1}.$$

From the last lecture, we know that C is the matrix of the linear transformation $\mathbf{x} \to A\mathbf{x}$ relative to the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$ formed by the columns of P.

We recognise this matrix, from the previous example, as the composition of a counterclockwise rotation by $\pi/4$ and a scaling by $\sqrt{2}$. This is the rotation "inside" A. We can write A:

$$A = PCP^{-1} = P \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} P^{-1}.$$

From the last lecture, we know that C is the matrix of the linear transformation $\mathbf{x} \to A\mathbf{x}$ relative to the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$ formed by the columns of P.

This shows that when we represent the transformation in terms of the basis $\mathcal{B}$, the transformation $\mathbf{x} \to A\mathbf{x}$ "looks like" the composition of a scaling and a rotation. As promised, using a non-standard basis we can sometimes uncover the hidden geometric properties of a linear transformation!

Example 6

Consider the matrix $A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$.

The characteristic polynomial of A is given by

$$\det \begin{bmatrix} 1-\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = (1-\lambda)(-\lambda) + 1 = \lambda^2 - \lambda + 1.$$

This is the same polynomial as for the matrix in Example 1. So we know that A has complex eigenvalues and therefore complex eigenvectors.

Example 6

Consider the matrix $A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$.

The characteristic polynomial of A is given by

$$\det \begin{bmatrix} 1-\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = (1-\lambda)(-\lambda) + 1 = \lambda^2 - \lambda + 1.$$

This is the same polynomial as for the matrix in Example 1. So we know that A has complex eigenvalues and therefore complex eigenvectors.
To see how multiplication by A affects points, take an arbitrary point, say $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and then plot successive images of this point under repeated multiplication by A.

The first few points are

$$\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

$$\mathbf{x}_3 = A\mathbf{x}_2 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix},$$

$$\mathbf{x}_4 = A\mathbf{x}_3 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \ldots$$

You could try this also for matrices $\begin{bmatrix} 0.1 & -0.2 \\ 0.1 & 0.3 \end{bmatrix}$ and $\begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix}$.

The theorem (and why it's true)

Theorem

Let A be a 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ $(b \neq 0)$ and an associated eigenvector $\mathbf{v}$ in $\mathbb{C}^2$. Then

$$A = PCP^{-1}, \text{ where } P = \begin{bmatrix} Re\,\mathbf{v} & Im\,\mathbf{v} \end{bmatrix}$$

$$\text{and } C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Sketch of proof

Suppose that A is a real 2×2 matrix, with a complex eigenvalue $\lambda = a - ib$, $b \neq 0$, and a corresponding complex eigenvector $\mathbf{v} = \mathbf{v}_1 + i\mathbf{v}_2$ where $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$. Then

- $\mathbf{v}_2 \neq \mathbf{0}$ because otherwise $A\mathbf{v} = A\mathbf{v}_1$ would be real, whereas $\lambda\mathbf{v} = \lambda\mathbf{v}_1$ is not.
- If $\mathbf{v}_1 = \alpha\mathbf{v}_2$, for some (necessarily real) α,

  $$A(\mathbf{v}) = A((\alpha + i)\mathbf{v}_2) = (\alpha + i)A\mathbf{v}_2 = (\alpha + i)\lambda\mathbf{v}_2$$

  whence the real vector $A\mathbf{v}_2$ equals $\lambda\mathbf{v}_2$ which is not real.

Thus the real vectors $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent, and give a basis for $\mathbb{R}^2$.

Equate the real and imaginary parts in the two formulas

$$A\mathbf{v} = (a - ib)\mathbf{v} = (a - ib)(\mathbf{v}_1 + i\mathbf{v}_2) = (a\mathbf{v}_1 + b\mathbf{v}_2) + i(a\mathbf{v}_2 - b\mathbf{v}_1)\,,$$

and

$$A\mathbf{v} = A(\mathbf{v}_1 + i\mathbf{v}_2) = A\mathbf{v}_1 + iA\mathbf{v}_2.$$

Equate the real and imaginary parts in the two formulas

$$A\mathbf{v} = (a - ib)\mathbf{v} = (a - ib)(\mathbf{v}_1 + i\mathbf{v}_2) = (a\mathbf{v}_1 + b\mathbf{v}_2) + i(a\mathbf{v}_2 - b\mathbf{v}_1)\,,$$

and

$$A\mathbf{v} = A(\mathbf{v}_1 + i\mathbf{v}_2) = A\mathbf{v}_1 + iA\mathbf{v}_2.$$

This gives $A\mathbf{v}_1 = a\mathbf{v}_1 + b\mathbf{v}_2$ and $A\mathbf{v}_2 = a\mathbf{v}_2 - b\mathbf{v}_1$ so that

$$\begin{aligned} A\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} &= \begin{bmatrix}A\mathbf{v}_1 & A\mathbf{v}_2\end{bmatrix} \\ &= \begin{bmatrix}a\mathbf{v}_1 + b\mathbf{v}_2 & a\mathbf{v}_2 - b\mathbf{v}_1\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix}\begin{bmatrix}a & -b \\ b & a\end{bmatrix}. \end{aligned}$$

So with respect to the basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, the transformation T_A has matrix

$$\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix}^{-1} A \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{bmatrix} = \begin{bmatrix}a & -b \\ b & a\end{bmatrix}$$

Setting $\sin\varphi = \dfrac{b}{\sqrt{a^2+b^2}}, \cos\varphi = \dfrac{a}{\sqrt{a^2+b^2}}$,

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \sqrt{a^2+b^2}\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix},$$

which is a scaling and rotation. And all of this is determined by the complex eigenvalue $a - ib$.

Of course, if $a - ib$ is an eigenvalue with eigenvector $\mathbf{v}_1 + i\mathbf{v}_2$, $a + ib$ is an eigenvalue, with eigenvector $\mathbf{v}_1 - i\mathbf{v}_2$.

Example 7

What is the geometric action of $A = \begin{bmatrix} -5 & -5 \\ 5 & -5 \end{bmatrix}$ on $\mathbb{R}^2$?

As a first step we find the eigenvalues and eigenvectors associated with A.

$$\begin{aligned} \det(A - \lambda I) &= \begin{bmatrix} -5-\lambda & -5 \\ 5 & -5-\lambda \end{bmatrix} \\ &= (-5-\lambda)^2 + 25 \\ &= \lambda^2 + 10\lambda + 50 \end{aligned}$$

This gives

$$\lambda = \frac{-10 \pm \sqrt{100-200}}{2} = \frac{-10 \pm 10i}{2} = -5 \pm 5i.$$

Consider the eigenvalue $\lambda = -5 - 5i$. We will find the corresponding eigenspace:

$$\begin{aligned} E_\lambda &= \text{Nul}\,(A - \lambda I) \\ &= \text{Nul}\begin{bmatrix} 5i & -5 \\ 5 & 5i \end{bmatrix} \\ &= \text{Span}\left\{\begin{bmatrix} 1 \\ i \end{bmatrix}\right\} \end{aligned}$$

where Span here stands for *complex* span, that is the set of all scalar multiples $\alpha \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} \alpha \\ i\alpha \end{bmatrix}$ of $\begin{bmatrix} 1 \\ i \end{bmatrix}$, where α is in $\mathbb{C}$.

Choosing $\begin{bmatrix} 1 \\ i \end{bmatrix}$ as our eigenvector we find the associated matrices P and C:

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} -5 & -5 \\ 5 & -5 \end{bmatrix}.$$

It is easy to check that

$$A = PCP^{-1} \text{ or equivalently } AP = PC.$$

Further

$$C = \begin{bmatrix} -5 & -5 \\ 5 & -5 \end{bmatrix} = 5\sqrt{2} \begin{bmatrix} -1/\sqrt{2} & -1\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

The scaling factor is $5\sqrt{2}$. The angle of rotation is given by $\cos\varphi = -1/\sqrt{2}, \sin\varphi = 1/\sqrt{2}$, which gives $\phi = 3\pi/4$ (135°).