Linear Algebra – Lesson 9

Vectors and matrices over the complex numbers

Some motivation

This lesson is about vector and matrices over the complex numbers ($v_i \in \mathbb{C}, A_{ij} \in \mathbb{C}$).
We've seen complex eigenvalues before (do you remember which transformation it was?).
A strong motivation for studying complex matrices is related to the Fourier transform, which has numerous applications in science. It turns out that the Fourier transform can be calculated using matrix-vector product, defined over the complex numbers.

1. Complex numbers

A short reminder on complex numbers:

- A complex number $z \in \mathbb{C}$ is of the form $z = a + ib$, where a is the **real part** of z, b is the **imaginary part** of z, and i is the square root of -1 (i.e., $i^2 = -1$).
- The complex conjugate (denoted by z^* or $\bar{z}$) has the same real part, but its imaginary part has the opposite sign:
$$z = a + ib, \qquad z^* = a - ib$$
- A complex number can also be represented in polar representation:
$$z = |z|e^{i\theta}$$
    - $|z| = \sqrt{z^* z} = \sqrt{a^2 + b^2}$ is called the **norm** or **amplitude** of z
    - $\theta = \tan^{-1}\left(\frac{b}{a}\right)$ is called the **phase** of z

  Notice that $|z|$ is defined just as we define the norm of a vector in 2D.

2. Vectors over the complex numbers

2.1 Complex vectors

Complex vector are vectors whose entries are complex numbers. They are also called vectors over the complex field ($\mathbb{C}$), or vectors over the complex numbers. For example:

$$\bar{u} = \begin{pmatrix} 2 \\ 3i \\ 4 + 2i \end{pmatrix}, \quad \bar{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Notice that $\mathbb{R} \subset \mathbb{C}$, and therefore $\bar{v} \in \mathbb{C}^2$.

Many of the definitions from the real vectors generalize naturally to the complex numbers. For example: Vector addition, scalar multiplication ($k \in \mathbb{C}$) and, therefore, also linear combinations:

$$\bar{u} = \begin{pmatrix} 2 \\ 1 + i \end{pmatrix}, \bar{v} = \begin{pmatrix} i \\ 3 + 2i \end{pmatrix}, \quad \bar{w} = (1 - i)\bar{u} + 2\bar{v} = \begin{pmatrix} 2 \\ 8 + 4i \end{pmatrix}$$

2.2 The norm of a complex vector

Some definitions, however, require adjustments for complex vectors. For example, using our regular definition of the vector norm we get:

$$\bar{v} = \begin{pmatrix} 1 \\ i \\ i \end{pmatrix}, |\bar{v}|^2 = \bar{v}^T \bar{v} = -1$$

But this means that the norm is no longer positive as we want it to be, and it can be zero even for a non-zero vector.

The vector norm for a complex vector is therefore defined using the complex conjugate of the vector in the first entry:

$$|\bar{v}|^2 = \bar{v}^{*T}\bar{v} = \sum_{i=1}^{n} v_i^* v_i$$

The norm now behaves as expected: for each coordinate i we get $v_i^* v_i > 0$, and $v_i^* v_i = 0 \leftrightarrow v_i = 0$.

2. Matrices over the complex numbers

A complex matrix is a matrix whose entries are complex.

Here again, some new definitions are in order. For example, we know that a **real symmetric** matrix has only real eigenvalues. Now, let A be a **complex symmetric** matrix: $A = \begin{pmatrix} 1 & i \\ i & 0 \end{pmatrix}$. What are its eigenvalues?

$$\det(A - \lambda I) = (1-\lambda)(-\lambda) + 1 = \lambda^2 - \lambda + 1$$

$$\lambda_{1,2} = \frac{1 \pm \sqrt{1-4}}{2}$$

The eigenvalues are no longer real. In other words, $A^T = A$ is not a sufficient condition to make the matrix behave well. This leads us to the definition of the conjugate transpose.

2.1 Conjugate complex and conjugate transpose

The conjugate complex of a matrix A is denoted by A^* (and sometimes $\bar{A}$), for which: $(A^*)_{ij} = a_{ij}^*$.

We already have two hints that tell us it might be useful to define the **conjugate transpose** of a matrix ((1) The new definition for a vector norm; (2) The lack of good properties for a matrix that satisfies $A^T = A$).

The conjugate transpose of a matrix is obtained by taking the complex conjugate of each entry, and also transposing the matrix. It is called the **Hermitian transpose**, or the **adjoint matrix**, and it is usually denoted by[1]:

$$A^H = (A^*)^T = (\bar{A})^T$$

For example:

$$A = \begin{pmatrix} i & 3+2i \\ 2 & 4-i \end{pmatrix} \rightarrow A^H = \begin{pmatrix} -i & 2 \\ 3-2i & 4+i \end{pmatrix}$$

2.1 Hermitian matrices

A square complex matrix A is said to be **Hermitian** if:

$$A^H = A$$

If the matrix A is real, then an Hermitian matrix is simply a symmetric matrix.

[1] And some books use A^* to denote the conjugate transpose matrix.

Properties of an $n \times n$ Hermitian matrix:

- All its eigenvalues are real
- It has n independent eigenvectors
    - Hence, it is diagonalizable in $\mathbb{C}^n$
- Its eigenvectors are orthogonal to each other (so it is diagonalizable by a unitary matrix; see 2.2)

<u>2.2 Unitary matrices</u>

A **square** complex matrix A is said to be **unitary** if:

$$A^H = A^{-1}$$

A complex matrix is unitary if and only if its rows (and its columns) form an orthonormal set.
Note: we haven't yet defined the inner product for complex vectors, so we still don't know exactly what "orthogonal" means.
If the matrix A is real, then a unitary matrix is simply an orthogonal matrix.

<u>2.3 Normal matrices</u>

A complex matrix is said to be normal if it commutes with A^H:

$$AA^H = A^H A$$

A matrix is normal if and only if it is diagonalizable by a unitary matrix U:

$$UAU^{-1} = \Lambda$$

3. Diagonalization of complex matrices

<u>3.1 Diagonalizing a complex matrix</u>

Just like we learned for real matrices, an $n \times n$ complex matrix is diagonalizable if and only if it has n independent eigenvectors. And just like in the real case, eigenvectors associated with distinct eigenvalues are linearly independent.

In lesson 05, we found the eigenvalues of the rotation matrix by 90 degrees counterclockwise:

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \to \lambda_1 = i, \lambda_2 = -i$$

Since it has no real eigenvalues, A is not diagonalizable over the real field $\mathbb{R}$. But if we treat it as a complex matrix, it must be diagonalizable, since it has two distinct eigenvalues. Let's find the eigenvectors:

<u>For $\lambda = i$</u>

$$(A - iI)\bar{v} = 0$$

$$\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

We can solve this using Gaussian elimination, by $-iR_1 + R_2 \to R_2$:

$$\begin{pmatrix} -i & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$-iv_1 = v_2$$

We can set $v_1 = 1$:

$$\bar{v}_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

<u>For $\lambda = -i$</u>

$$(A + iI)\bar{v} = 0$$

$$\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

We can solve this using Gaussian elimination, by $iR_1 + R_2 \to R_2$:

$$\begin{pmatrix} i & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$iv_1 = v_2$$

We can set $v_1 = 1$:

$$\bar{v}_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

So $P = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$. Therefore:

$$P^{-1} = \frac{1}{\det(P)}\begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix} = \frac{1}{2i}\begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}$$

Let's verify that we got a diagonal matrix:

$$\Lambda = P^{-1}AP = \frac{1}{2i}\begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$$

$$= \frac{1}{2i}\begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} = \frac{1}{2i}\begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{i} & 0 \\ 0 & \frac{1}{i} \end{pmatrix}$$

And since $\frac{1}{i} = \frac{i}{ii} = -i$:

$$\Lambda = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

As expected (notice that the eigenvalues appear in the correct order).

3.2 Complex eigenvalues of real matrices

Theorem: Let A be a real matrix. If λ is a complex eigenvalue of A associated with eigenvector $\bar{v}$, then λ^* is also an eigenvalue of A associated with the eigenvector $\bar{v}^*$.

In other words, eigenvalues and eigenvectors of real matrices come in conjugate pairs.

Proof

Assume that:

$$(A - \lambda I)\bar{v} = 0$$

Take the conjugate from both sides:

$$(A - \lambda I)^*\bar{v}^* = 0$$
$$(A^* - \lambda^* I)\bar{v}^* = 0$$

But since A is real:

$$(A - \lambda^* I)\bar{v}^* = 0$$

And this equation means that λ^* is an eigenvalue of A with eigenvector $\bar{v}^*$.

Note: You can use this property when you solve real problems. You only need to find the eigenvectors of one eigenvalue, and easily obtain the eigenvectors of its complex conjugate.

3.3 A real symmetric matrix has real eigenvalues

A special case of a real matrix is a symmetric matrix. We've seen before that a real symmetric matrix has real eigenvalues. Now we prove it:

$$A\bar{v} = \lambda\bar{v}$$

Make sure you understand where we use the fact that A is real and when we use the fact that A is symmetric:

On the one hand	**On the other hand**

Multiply by $\bar{v}^{*T}$ from the left:

$$\bar{v}^{*T} A \bar{v} = \lambda \bar{v}^{*T} \bar{v}$$

Take the complex conjugate:

$$A\bar{v}^* = \lambda^* \bar{v}^*$$

Transpose:

$$\bar{v}^{*T} A = \lambda^* \bar{v}^{*T}$$

Multiply by $\bar{v}$ from the right:

$$\bar{v}^{*T} A \bar{v} = \lambda^* \bar{v}^{*T} \bar{v}$$

But this tells us that:

$$\lambda \bar{v}^{*T} \bar{v} = \lambda^* \bar{v}^{*T} \bar{v}$$

And therefore[2]:

$$\lambda = \lambda^* \quad \rightarrow \quad \lambda \in \mathbb{R}$$

4. Complex eigenvalues lead to spiraling dynamics

There is something surprising in getting complex eigenvalues for a real matrix. When we deal with real matrices, we want to be able to understand them as much as possible in the realm of real numbers.

We already saw that we can diagonalize a 2×2 matrix using complex eigenvalues and eigenvectors. Now we will see another way of thinking about the matrix, using only real matrices.

Let A be a 2×2 real matrix with a complex eigenvalue λ and an associated eigenvector $\bar{v}$. Then A is similar to a rotation-scaling matrix B:

$$A = CBC^{-1}$$

In other words, we can represent A in a new basis (by using the change of basis matrix C), to obtain a matrix B that is a rotation-scaling matrix. The forms of C and B are:

$$C = \begin{pmatrix} | & | \\ Re(\bar{v}) & Im(\bar{v}) \\ | & | \end{pmatrix}, B = \begin{pmatrix} Re(\lambda) & Im(\lambda) \\ -Im(\lambda) & Re(\lambda) \end{pmatrix}$$

where $Re(\bar{v}) = Re\begin{pmatrix} x + yi \\ z + wi \end{pmatrix} = \begin{pmatrix} x \\ z \end{pmatrix} \quad Im(\bar{v}) = Im\begin{pmatrix} x + yi \\ z + wi \end{pmatrix} = \begin{pmatrix} y \\ w \end{pmatrix}$

Notice that you can choose which of the two eigenvectors ($\bar{v}_1$ or $\bar{v}_2$) to use.

The scaling factor that B induces is: $\sqrt{\det(B)} = |\lambda|$.

Note: the scaling factor of B is exactly what we expect it to be. We know that the determinant $\det(B)$ is the change in volume induced by B. Since B is a 2×2 matrix, this means that each of the basis vectors is scaled by $\sqrt{\det(B)}$.

4.1 Rotation-scaling matrix

A rotation-scaling matrix is any matrix of the form:

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a, b \neq 0, a, b \in \mathbb{R}$$

[2] In fact, there is another simple proof. If A is real, then the characteristic polynomial $f(\lambda)$ has only real coefficients. It takes one line to show that if $f(\lambda_0) = 0$, then also $f(\lambda_0^*) = 0$.

Let's see why it is called a rotation-scaling matrix.

Any rotation-scaling matrix can be represented as a product of a rotation matrix and a scaling matrix:

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

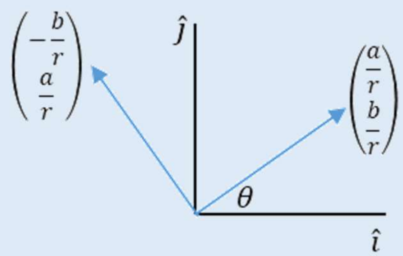
With a scaling factor $r = \sqrt{\det(A)} = \sqrt{a^2 + b^2}$.

Let's prove this:

First, since we know the scaling factor, we can already factor it out:

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = r \begin{pmatrix} \frac{a}{r} & -\frac{b}{r} \\ \frac{b}{r} & \frac{a}{r} \end{pmatrix} = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \begin{pmatrix} \frac{a}{r} & -\frac{b}{r} \\ \frac{b}{r} & \frac{a}{r} \end{pmatrix}$$

The first matrix is a scaling matrix by r. We need to convince ourselves that the second matrix is a rotation. Let's see where $\hat{\imath}$ and $\hat{\jmath}$ land:



The length of $A\hat{\imath}$ is $\sqrt{\left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2} = \sqrt{\frac{a^2}{a^2+b^2} + \frac{b^2}{a^2+b^2}} = 1$, and the same for $A\hat{\jmath}$, as expected from a matrix that only rotates. They also share the same angle θ. You can see that:

$$\cos(\theta) = \frac{a}{r}$$

$$\sin(\theta) = \frac{b}{r}$$

So we got:

$$A = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

We will not prove this theorem, but you can find a short proof here (press the "proof" link under "Rotation-Scaling Theorem").

Example

Let $A = \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}$

The eigenvalues of A are:

$$-\lambda(2 - \lambda) + 2 = 0$$

$$\lambda^2 - 2\lambda + 2 = 0$$

$$\lambda_{1,2} = \frac{2 \pm \sqrt{4 - 8}}{2} = 1 \pm i$$

Since the eigenvalues are complex, we know that A can be represented as a rotation-scaling matrix in some other basis. To find such a representation, we have to choose one of the eigenvalues, find its eigenvector, and follow the theorem above.

We will look at $\lambda_1 = 1 - i$. An eigenvector associated with λ_1 is:
$\bar{v} = \begin{pmatrix} 1 \\ 1+i \end{pmatrix}$. Now, using the theorem:

$$A = CBC^{-1}$$

Where:

$$C = \left(Re\begin{pmatrix} 1 \\ 1+i \end{pmatrix} \quad Im\begin{pmatrix} 1 \\ 1+i \end{pmatrix} \right) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

$$B = \begin{pmatrix} Re(1-i) & Im(1-i) \\ -Im(1-i) & Re(1-i) \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Question: What does this have to do with spirals?
Answer: You will answer this yourselves in the exercise.

Additional resources
A video example for diagonalization of a Hermitian matrix in this video.
A short review of complex numbers, complex eigenvalues and their geometric interpretation.