1. Quiz

Problem 1. Let $L$ be the linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ represented (in standard coordinates) by the matrix

$$
\begin{pmatrix}
1 & 2 \\
-3 & 0
\end{pmatrix}.
$$

Find the matrix that represents $L$ in the basis $\mathcal{B} = \{(4, -1), (-3, 1)\}$.

Problem 2. Let

$$
M = \begin{pmatrix} 8 & -2 \\ 24 & -6 \end{pmatrix}
$$

Find $M^{10}$ by diagonalizing $M$. Hint: The eigenvalues of $M$ are 2 and 0, with corresponding eigenvectors $(1, 3)$ and $(1, 4)$.

2. Complex vectors

The time has come to allow our vectors to have complex coordinates. One reason to do this has to do with eigenvalue and eigenvectors. You will recall that a good way to find the eigenvalues of a matrix $M$ is to calculate the characteristic polynomial of $M$, which is $\det(M - \lambda I)$ (or you can use $\det(\lambda I - M)$, and then the eigenvalues of $M$ will be the roots of this polynomial. In order to be sure that we have at least one root, we need complex numbers, even if the matrix $M$ has only real numbers as its entries. And once you have complex eigenvalues, you also need complex eigenvectors. Thus, even if we work exclusively with real matrices, we cannot avoid complex numbers and complex vectors. We’ll denote the complex number system by $\mathbb{C}$, and if $\vec{v}$ is a vector with $n$ coordinates that can be complex numbers, then we think of $\vec{v}$ as a vector in the vector space $\mathbb{C}^n$.

Here are two column vectors in $\mathbb{C}^3$:

$$
\vec{v}_1 = \begin{pmatrix} i \\ 3 - 2i \\ 6 \end{pmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 5 + i \\ 2 - i \end{pmatrix}.
$$

We can do the same things with these vectors that we do with vectors that have real coordinates. In particular, we can multiply $\vec{v}_1$ by a
scalar, such as $c_1 = 5 - 2i$:

$$c_1 \vec{v}_1 = (5 - 2i) \begin{pmatrix} i \\ 3 - 2i \\ 6 \end{pmatrix} = \begin{pmatrix} 2 + 5i \\ 11 - 16i \\ 30 - 12i \end{pmatrix}.$$  

Notice that we used ordinary complex arithmetic in this calculation. We can also add the two vectors:

$$\vec{v}_1 + \vec{v}_2 = \begin{pmatrix} i \\ 3 - 2i \\ 6 \end{pmatrix} + \begin{pmatrix} 0 \\ 5 + i \\ 2 - i \end{pmatrix} = \begin{pmatrix} i \\ 8 - i \\ 8 \end{pmatrix},$$

again using ordinary complex arithmetic. With a second scalar $c_2 = 1 + i$, we can form the following linear combination of $\vec{v}_1$ and $\vec{v}_2$:

$$c_1 \vec{v}_1 + c_2 \vec{v}_2 = (5-2i) \begin{pmatrix} i \\ 3 - 2i \\ 6 \end{pmatrix} + (1+i) \begin{pmatrix} 0 \\ 5 + i \\ 2 - i \end{pmatrix} = \begin{pmatrix} 2 + 5i \\ 11 - 16i \\ 30 - 12i \end{pmatrix} = \begin{pmatrix} 2 + 5i \\ 15 - 10i \\ 33 - 11i \end{pmatrix}.$$  

So far, everything is just as it was with real vectors, and all the usual rules of arithmetic apply, such as the associative, commutative, and distributive laws. The only real difference so far is that the geometry is no longer so easy to visualize. Nevertheless, we still find it useful to talk about the length of a complex vector, and also to ask whether two complex vectors are orthogonal. Since the geometric picture is lacking, all of this will be defined algebraically in terms of coordinates, by way of the inner product. (This is what we called the “dot product” with real vectors, but I’m going to use the fancier name, to emphasize that there is a difference.)

Let’s start with the length. If a complex vector $\vec{v} \in \mathbb{C}^n$ has complex coordinates $z_1, z_2, \ldots, z_n$, then we would like to define the inner product in such a way that

$$\vec{v} \cdot \vec{v} = \|\vec{v}\|^2 = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2.$$  

If one of the coordinates $z_i$ is a real number, then we know that $|z_i|^2$ is the same as $z_i^2$. But this is not typically true of a complex number. Instead, if $z$ is an arbitrary complex number, the correct formula is $|z|^2 = \overline{z}z$, where $\overline{z}$ is the complex conjugate of $z$. For example, if $z = 1 - 3i$,

$$z^2 = (1-3i)(1-3i) = -7 - 6i \neq |z|^2 = 1^2 + 3^2 = \overline{z}z = (1+3i)(1-3i) = 10.$$  

So in order to get things right using the inner product, we need to involve complex conjugates. Here is the definition. If we have two vectors $\vec{v}$ and $\vec{s}$ in $\mathbb{C}^n$,

$$\vec{r} = (r_1, r_2, \ldots, r_n) \quad \text{and} \quad \vec{s} = (s_1, s_2, \ldots, s_n),$$
or
\[ \vec{r} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \quad \text{and} \quad \vec{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}, \]

then
\[ \vec{r} \cdot \vec{s} = r_1 s_1 + r_2 s_2 + \cdots + r_n s_n. \]

This definition agrees perfectly with the definition of the dot product if \( \vec{r} \) and \( \vec{s} \) happen to be real vectors, and just as important, it ensures that
\[ \vec{v} \cdot \vec{v} = \|\vec{v}\|^2 \]
for all complex vectors \( \vec{v} \), which is an equality that we have used in many of the calculations and proofs that we have given in this course.

You need to be aware of the fact that not everyone uses our exact definition. Many mathematicians prefer that the complex conjugate be applied to the second vector rather than the first. This will not change the value of \( \vec{v} \cdot \vec{v} \) for any complex vector \( \vec{v} \), but it does affect the value of \( \vec{v} \cdot \vec{w} \) in general for complex vectors \( \vec{v} \) and \( \vec{w} \). Both choices would serve equally well. The choice I have made seems to work best when we work with column vectors. Since eigenvectors are important to us, and since we have chosen to use column vectors for eigenvectors, my choice for the definition of the inner product is the one I prefer.

You may have noticed that our inner product is not commutative. Instead, reversing the order changes the result by taking the conjugate:
\[ \vec{r} \cdot \vec{s} = \overline{\vec{s} \cdot \vec{r}}. \]

For example,
\[ \begin{pmatrix} i \\ 3 - 2i \\ 6 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 5 + i \\ 2 - i \end{pmatrix} = (-i) \times 0 + (3 + 2i) \times (5 + i) + 6 \times (2 - i) = 25 + 7i, \]
but
\[ \begin{pmatrix} 0 \\ 5 + i \\ 2 - i \end{pmatrix} \cdot \begin{pmatrix} i \\ 3 - 2i \\ 6 \end{pmatrix} = 0 \times i + (5 - i) \times (3 - 2i) + (2 + i) \times 6 = 25 - 7i. \]

Here is another aspect of arithmetic with the inner product that requires care. With the real dot product, we could move scalars around as follows:
\[ c(\vec{v} \cdot \vec{w}) = (c\vec{v}) \cdot \vec{w} = \vec{v} \cdot (c\vec{w}) \quad (\text{Real case}). \]
With complex numbers, the use of the conjugate changes this, and the correct rule is
\[ c(\bar{v} \cdot \bar{w}) = (\bar{c}\bar{v}) \cdot \bar{w} = \bar{v} \cdot (c\bar{w}) \] (Complex case).

So we need to be careful! (This is one place where choice of the definition of inner product makes a difference. With the other choice, \( c(\bar{v} \cdot \bar{w}) = (c\bar{v}) \cdot \bar{w} = \bar{v} \cdot (c\bar{w}). \))

Aside from these differences, the other rules for working with vectors and scalars are the same as they were in the real case. For example, the following distributive law holds:
\[ (\bar{v} + \bar{s}) \cdot \bar{v} = \bar{v} \cdot \bar{v} + \bar{s} \cdot \bar{v} \]

Just be careful whenever you want to change the order in an inner product (just as you must be careful with matrix multiplication), or when you move a scalar around in an inner product.

We say that two complex vectors \( \bar{v} \) and \( \bar{s} \) are orthogonal or perpendicular if \( \bar{v} \cdot \bar{s} = 0 \). A set of complex vectors \( \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \) is orthonormal if they are all orthogonal to each other and if they all have length 1:
\[ \vec{v}_i \cdot \vec{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases} \]

Again, since both 0 and 1 are real numbers, it does not matter which order we take the inner product in checking orthogonality or orthonormality.

What about orthogonal projections? It is important to get the formula exactly right. If we want to project a vector \( \bar{v} \) onto the linear span of a set of orthogonal vectors \( \bar{b}_1, \bar{b}_2, \ldots, \bar{b}_k \), the formula is
\[ \text{projection} = \frac{\bar{b} \cdot \bar{v}}{b \cdot b} \bar{b}_1 + \frac{\bar{b} \cdot \bar{v}}{b \cdot b} \bar{b}_2 + \cdots + \frac{\bar{b} \cdot \bar{v}}{b \cdot b} \bar{b}_k \]

and the order in which the inner product is taken in each of the numerators is critical. Let’s see why this is the case. For simplicity, let’s project a vector \( \vec{v} \) onto the line determined by the vector \( \vec{b} \):
\[ \text{projection} = \frac{\vec{b} \cdot \vec{v}}{\vec{b} \cdot \vec{b}} \vec{b}. \]

In earlier work, we saw that it was important that the difference between \( \vec{v} \) and this projection be perpendicular to \( \vec{b} \). Let’s do the calculation:
\[ \left( \vec{v} - \frac{\vec{b} \cdot \vec{v}}{\vec{b} \cdot \vec{b}} \vec{b} \right) \cdot \vec{b} = \vec{v} \cdot \vec{b} - \left( \frac{\vec{b} \cdot \vec{v}}{\vec{b} \cdot \vec{b}} \right) \cdot \vec{b} = \vec{v} \cdot \vec{b} - \frac{\vec{b} \cdot \vec{v}}{\vec{b} \cdot \vec{b}} (\vec{b} \cdot \vec{b}). \]
The denominator $\mathbf{b} \cdot \mathbf{b}$ in this expression is a positive number, so the complex conjugate only needs to be applied to the numerator. Using the fact that $\overline{\mathbf{b} \cdot \mathbf{v}} = \overline{\mathbf{v} \cdot \mathbf{b}}$ and canceling $\mathbf{b} \cdot \mathbf{b}$ in the numerator and denominator, we get 0.

Note that we were careful here in using the rules involving scalars and inner products. Just to emphasize this point, let’s do a similar calculation with the vector $\mathbf{b}$ on the left:

$$\mathbf{b} \cdot \left( \mathbf{v} - \frac{\mathbf{b} \cdot \mathbf{v}}{\mathbf{b} \cdot \mathbf{b}} \mathbf{b} \right) = \mathbf{b} \cdot \mathbf{v} - \left( \frac{\mathbf{b} \cdot \mathbf{v}}{\mathbf{b} \cdot \mathbf{b}} \right) (\mathbf{b} \cdot \mathbf{b}) = 0 .$$

So our previous formula for the orthogonal projection still works, provided you take the inner product in the correct order in the numerator of each of the coefficients. If you get this correct, then the Gram-Schmidt procedure will also still work out, just as before.

Exercise 1. Find the length of the vector $(2 - i, -6, 3i)$.

Exercise 2. Find the inner product of the two vectors $(2 - i, 6, -i)$ and $(3, 2 + i, 6 - i)$.

Exercise 3. Find the orthogonal projection of the vector $(i, 0, 1 - i)$ onto the linear span of the vectors $(1, 0, i)$ and $(-2i, 1, 0)$. Hint: First find an orthogonal basis for the plane determined by $(1, 0, i)$ and $(-2i, 1, 0)$.

3. Complex matrices

So far, we have seen that complex vectors behave like real vectors, except for some care that is needed with the inner product. The same is true of complex matrices. The only time we need to make a modification is when we connect the inner product to matrix multiplication. Recall that if $\mathbf{v}, \mathbf{w}$ are column vectors in $\mathbb{R}^n$, then the dot product $\mathbf{v} \cdot \mathbf{w}$ is the same as the matrix product $\mathbf{v}^T \mathbf{w}$, where $\mathbf{w}^T$ is the $1 \times n$ matrix (row vector) that you get by taking the transpose of the $n \times 1$ matrix (column vector) $\mathbf{w}$.

In order to make something like this work out for complex vectors, it is not enough to take the transpose of the first row vector $\mathbf{v}$. We also need to take the complex conjugate. So we define the Hermitian transpose of a complex matrix $M$ to be the matrix

$$M^H = \overline{M^T} ,$$
which is the matrix that you get from $M$ by first taking its transpose and then replacing all of the elements in $M^T$ by their complex conjugates. For example,

$$
\begin{pmatrix}
1 & i & 0 & 2 + i \\
1 - i & 3 + 2i & 5 & 6i \\
\end{pmatrix}^H = 
\begin{pmatrix}
1 & 1 + i \\
-i & 3 - 2i \\
0 & 5 \\
2 - i & -6i \\
\end{pmatrix}.
$$

Note that if $M$ is a real matrix, then $M^T = M^H$, because taking the complex conjugates doesn’t change anything. Also note that

$$(M^H)^H = M,$$

because the conjugate of the conjugate of a complex number $z$ is just the number $z$.

Now we get to the connection with inner products. If $\vec{v}$ and $\vec{w}$ are column vectors in $\mathbb{C}^n$, we have

$$
\vec{v} \cdot \vec{w} = 
\begin{pmatrix}
v_1 \\
v_2 \\
\vdots \\
v_n \\
\end{pmatrix}
\begin{pmatrix}
w_1 \\
w_2 \\
\vdots \\
w_n \\
\end{pmatrix} = v_1w_1 + v_2w_2 + \cdots + v_nw_n = 
\begin{pmatrix}
v_1 & v_2 & \cdots & v_n \\
\end{pmatrix}
\begin{pmatrix}
w_1 \\
w_2 \\
\vdots \\
w_n \\
\end{pmatrix} = \vec{v}^H \vec{w}.
$$

So the formula to remember is

$$\vec{v} \cdot \vec{w} = \vec{v}^H \vec{w}$$

for column vectors.

On the left side of this equation, $\vec{v}$ and $\vec{w}$ are being treated as vectors. On the right side, $\vec{v}^H$ is a matrix with one row and $\vec{w}$ is a matrix with one column.

Here is an important feature of the Hermitian transpose. If $M$ is a complex orthogonal matrix (which means that it has orthonormal rows), then instead of the transpose of $M$ being its inverse, as was the case with real orthogonal matrices, we have

$$M^{-1} = M^H.$$ 

We also note the following property of the Hermitian transpose, which can be proved in the same way as the corresponding property for transposes. If $A$ and $B$ are two complex matrices such that the product $AB$ makes sense, then

$$(AB)^H = B^HA^H.$$
Let’s relate this formula to the matrix calculation for the inner product of column vectors \( \vec{v}, \vec{w} \):

\[
\vec{v} \cdot \vec{w} = \vec{v}^H \vec{w} = (\vec{w}^H \vec{v})^H = \overline{\vec{w}^H \vec{v}} = \overline{\vec{w}} \cdot \overline{\vec{v}}.
\]

We used the fact that the Hermitian transpose of the \( 1 \times 1 \) matrix \( \overline{\vec{w}}^H \overline{\vec{v}} \) is simply its complex conjugate. So this calculation agrees with our earlier formula for what happens when you switch the order in the inner product.

4. The Spectral Theorem

This section is about an extremely important theorem in linear algebra. It concerns a certain type of complex square matrix. An \( n \times n \) complex matrix \( M \) is called a **Hermitian matrix** if \( M = M^H \). For example, the matrices

\[
\begin{pmatrix}
2 & 1 - i \\
1 + i & -4
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
2 & 1 & 0 \\
1 & -3 & -7 \\
0 & -7 & \sqrt{2}
\end{pmatrix}
\]

are both Hermitian matrices. Any real symmetric matrix is a Hermitian matrix and a real Hermitian matrix must be symmetric. Also note that the diagonal elements of a Hermitian matrix **must be real**, whether or not the matrix is real.

We now have all the tools we need to prove some remarkable properties of Hermitian matrices. I will state them as a series of propositions.

**Proposition 1.** Let \( M \) be an \( n \times n \) Hermitian matrix. Then for any column vectors \( \vec{v}, \vec{w} \) in \( \mathbb{C}^n \),

\[
(M\vec{v}) \cdot \vec{w} = \vec{v} \cdot (M\vec{w}).
\]

The proof is simply the following calculation involving some properties of inner product and matrix multiplication that we have learned:

\[
(M\vec{v})\cdot\vec{w} = (M\vec{v})^H \vec{w} = (\vec{v}^H M^H)\vec{w} = \vec{v}^H (M^H \vec{w}) = \vec{v}^H (M\vec{w}) = \vec{v} \cdot (M\vec{w}).
\]

Now suppose that \( \vec{v} \) is an eigenvector of a Hermitian matrix \( M \), with eigenvalue \( \lambda \). That is, suppose that \( M\vec{v} = \lambda \vec{v} \), where \( \vec{v} \) is a column vector. Then

\[
\lambda \|\vec{v}\|^2 = \lambda (\vec{v} \cdot \vec{v}) = \vec{v} \cdot (\lambda \vec{v}) = \vec{v} \cdot (M\vec{v}) = \lambda(\vec{v} \cdot \vec{v}) = \bar{\lambda} \vec{v} \cdot \vec{v} = \bar{\lambda} \|\vec{v}\|^2.
\]

Since \( \|\vec{v}\|^2 > 0 \), we see that \( \lambda = \bar{\lambda} \), which means that \( \lambda \) is a real number. We have proved

**Proposition 2.** Every eigenvalue of a Hermitian matrix is a real number.
Since a symmetric real matrix is a Hermitian matrix, and since an equation like $(\lambda I - M)\vec{v} = \vec{0}$ must have a real solution $\vec{v}$ if $\lambda$ is a real eigenvalue and $M$ is a real matrix, we have the following:

**Proposition 3.** Every eigenvalue of a real symmetric matrix is a real number, and the eigenvectors of such a matrix can always be chosen to be real eigenvectors.

This remarkable fact is not so easy to prove without the use of complex numbers, even though its statement does not involve anything but real numbers.

Suppose $\vec{v}$ and $\vec{w}$ are eigenvectors of a Hermitian matrix $M$, with distinct eigenvalues $\lambda$ and $\mu$. We know these eigenvalues are real. Let’s play with the inner product a bit, using this fact to avoid having to use the complex conjugate when we move the scalars $\lambda$ and $\mu$ around:

$$\lambda(\vec{v} \cdot \vec{w}) = (\lambda \vec{v}) \cdot \vec{w} = (M\vec{v}) \cdot \vec{w} = \vec{v} \cdot (M\vec{w}) = \vec{v} \cdot (\mu \vec{w}) = \mu(\vec{v} \cdot \vec{w}).$$

Since we have assumed that $\lambda \neq \mu$, it follows that $\vec{v} \cdot \vec{w} = 0$.

**Proposition 4.** If $\vec{v}$ and $\vec{w}$ are eigenvectors of a Hermitian matrix, and if the corresponding eigenvalues are distinct, then $\vec{v}$ and $\vec{w}$ are orthogonal.

From these propositions, we almost get the following result, which is almost the Spectral Theorem:

**Proposition 5** (Almost the Spectral Theorem). If $M$ is an $n \times n$ Hermitian matrix with $n$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then all of these eigenvalues are real numbers, and we can take the corresponding eigenvectors so that they are an orthonormal basis for $\mathbb{C}^n$. If $A$ is the matrix whose columns are those eigenvectors, then

$$A^H MA = D(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

If $A$ is real and symmetric, then the eigenvectors can be taken to be real, and they form an orthonormal basis for $\mathbb{R}^n$.

The only difference between this result and the actual Spectral Theorem is the assumption that there are $n$ distinct eigenvalues. In general, there may not be $n$ distinct eigenvalues. From the Fundamental Theorem of Algebra, we know that the characteristic polynomial of an $n \times n$ matrix $M$ has $n$ roots $\lambda_1, \lambda_2, \ldots, \lambda_n$, provided we count multiplicities. For example, the third degree polynomial

$$\lambda^3 - 2\lambda^2 + \lambda = (\lambda - 1)^2 \lambda.$$
has the roots 1 and 0, with the root 1 having multiplicity 2. So we could say that this polynomial has the three roots 1, 1, 0, listing the root 1 twice because the factor \((\lambda - 1)\) is repeated in the factorization.

In general, if we get repeated roots in the characteristic polynomial of an \(n \times n\) matrix, we may not be able to get \(n\) linearly independent eigenvectors. But with Hermitian matrices, we not only get \(n\) linearly independent eigenvectors, but we can make them orthonormal:

**Theorem 1** (Spectral Theorem). If \(M\) is an \(n \times n\) Hermitian matrix with eigenvalues \(\lambda_1, \lambda_2, \ldots, \lambda_n\) listed according to their multiplicities, then all of these eigenvalues are real numbers. Furthermore, if \(\lambda\) is one of these eigenvalues and \(\lambda\) has multiplicity \(k\), then there exist \(k\) orthonormal eigenvectors with eigenvalue \(\lambda\). Furthermore, there is an orthonormal basis for \(\mathbb{C}^n\) consisting of eigenvectors of \(M\), corresponding to the eigenvalues \(\lambda_1, \lambda_2, \ldots, \lambda_n\). If \(A\) is the matrix whose columns are those eigenvectors, then

\[
A^H M A = D(\lambda_1, \lambda_2, \ldots, \lambda_n).
\]

If \(A\) is real and symmetric, then the eigenvectors can be taken to be real, and they form an orthonormal basis for \(\mathbb{R}^n\).

Proving this full version of the Spectral Theorem is not too hard, but we will not spend the time to do it here.

**Exercise 4.** Explain why the diagonal elements of a Hermitian matrix must be real numbers.

**Exercise 5.** Let \(A\) be an \(n \times n\) complex orthogonal matrix. Show that for any two column vectors \(\vec{v}, \vec{w} \in \mathbb{C}^n\),

\[
\vec{v} \cdot \vec{w} = (M\vec{v}) \cdot (M\vec{w}).
\]

**Exercise 6.** Find the eigenvalues and eigenvectors of the matrix

\[
\begin{pmatrix}
1 & 1 + i \\
1 - i & 2
\end{pmatrix}
\]

**Exercise 7.** Let \(M\) be an \(n \times n\) complex matrix. Show that if \(M\) is invertible and \(\lambda\) is an eigenvalue of \(M\) with eigenvector \(\vec{v}\), then \(1/\lambda\) is an eigenvalue of \(M^{-1}\), with the same eigenvector \(\vec{v}\).
5. Topics for Exam

- Arithmetic with complex numbers: addition, subtraction, multiplication, division, conjugates, absolute value
- Real vectors: dot product, orthogonality, orthonormality, basis for a linear span, coordinates in different bases, orthogonal complement, orthogonal projection, Gram-Schmidt procedure
- Complex vectors: inner product, orthogonality, orthonormality, orthogonal projection, Gram-Schmidt procedure
- Real matrices: matrix form of dot product, rules for matrix multiplication involving transpose and inverse, determinant, eigenvalues, eigenvectors, matrix representation of linear transformation in different bases, orthogonal matrices, diagonalization
- Complex matrices: matrix form of inner product, rules for matrix multiplication involving Hermitian transpose, eigenvalues and eigenvectors (Spectral Theorem), orthogonal matrices, diagonalization
- Newton’s method
- Some common power series, power series of matrices using diagonalization.