(c)2023 Karen Smith, UM Math department
licensed under a Creative Commons
By-NC-SA 4.0 International License.

Worksheet 26: Diagonalization over Complex Numbers (§7.5)

In Math 217, the word "scalar" has been synonymous with "real number." However, it is natural to allow the scalars to be *complex numbers* instead of just real numbers.

INFORMAL DEFINITION: A **complex vector space** is a set V, equipped with operations of **vector addition** and **complex scalar multiplication**, which satisfy the eight axioms of a vector space from Worksheet 6, except that all scalars are taken from $\mathbb{C}$ (rather than $\mathbb{R}$).

Problem 1. Main Example of a Complex Vector Space. Let $\mathbb{C}^n$ denote the set of all $n \times 1$ matrices with *complex* entries. Its elements are **complex column vectors** $\vec{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$, where each $z_i \in \mathbb{C}$. Explain how to define a natural vector addition and a complex scalar multiplication on $\mathbb{C}^n$ in such a way that the axioms of a vector space hold. Thus the set $\mathbb{C}^n$ forms a **complex vector space** with this addition and scalar multiplication.

Solution: Addition of "complex vectors" is defined in the obvious way:

$$\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} z_1 + w_1 \\ \vdots \\ z_n + w_n \end{bmatrix},$$

as is "scalar multiplication" by any *complex scalar* $\lambda \in \mathbb{C}$:

$$\lambda \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} \lambda z_1 \\ \vdots \\ \lambda z_n \end{bmatrix}.$$

Problem 2. Let $\vec{e}_i \in \mathbb{C}^n$ denote the usual standard unit vector. Show that every column vector in $\mathbb{C}^n$ can be written as a *unique* linear combination of the standard unit vectors $\{\vec{e}_1, \ldots, \vec{e}_n\}$ if we allow the coefficients to be *complex scalars* from $\mathbb{C}$ (not just $\mathbb{R}$).

Discuss what it means for complex vectors to **span** a complex vector space, and what it means for a set of complex vectors to be **linearly independent.** Define **basis** for a complex vector space. Find two different bases for the complex vector space $\mathbb{C}^n$.

Solution: An arbitrary $\vec{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$ can be written as $z_1\vec{e}_1 + \cdots + z_n\vec{e}_n$. A set of vectors $\{v_1, \ldots, v_d\}$ spans a complex vector space V means that every $v \in V$ can be written as $z_1v_1 + \cdots + z_nv_d$ where the $z_i \in \mathbb{C}$. The complex vectors $\{v_1, \ldots, v_d\}$ are linearly independent if the only relation $z_1v_1 + \cdots + z_nv_d = 0$ with $z_i \in \mathbb{C}$ is the trivial relation (meaning that all $z_i = 0$). A basis is a set of complex vectors that span a complex vector space and are linearly independent. Two bases for $\mathbb{C}^n$ are $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ and, for example, $\{i\vec{e}_1, i\vec{e}_2, \ldots, i\vec{e}_n\}$.

Definition: A function $T : \mathbb{C}^n \to \mathbb{C}^m$ is a **complex linear transformation** if

(i) $T(\vec{z} + \vec{w}) = T(\vec{z}) + T(\vec{w})$ for all vectors $\vec{z}, \vec{w} \in \mathbb{C}^n$.

(ii) $T(\lambda \vec{z}) = \lambda T(\vec{z})$ for all vectors $\vec{z} \in \mathbb{C}^n$ and scalars $\lambda \in \mathbb{C}$.

Problem 3. Prove that if $T : \mathbb{C}^n \to \mathbb{C}^p$ is a complex linear transformation, then there exists a unique $p \times n$ matrix $A \in \mathbb{C}^{p\times n}$ (with complex entries) such that $T(\vec{z}) = A\vec{z}$.
[HINT: Guess what A might be, using your knowledge of the analogous case over $\mathbb{R}$. Then prove your guess.]

Solution: Let A be the $p \times n$ matrix in $\mathbb{C}^{p\times n}$ whose columns are $T(\vec{e}_1), T(\vec{e}_2), \ldots, T(\vec{e}_n)$. We claim that

$$T(\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}) = A \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}.$$

for all complex complex vectors $\vec{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{C}^n$. To check this, write $\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = z_1\vec{e}_1 + z_2\vec{e}_2 + \cdots + z_n\vec{e}_n$ and then apply T:

$$\begin{aligned} T(\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}) &= T(z_1\vec{e}_1 + z_2\vec{e}_2 + \cdots + z_n\vec{e}_n) \\ &= z_1T(\vec{e}_1) + z_2T(\vec{e}_2) + \cdots + z_nT(\vec{e}_n) \\ &= \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \ldots & T(\vec{e}_n) \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = A \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}. \end{aligned}$$

Note that A is the *unique* matrix with this property because if B also satisfies $T(\vec{z}) = B\vec{z}$, then we can compute that the j-th column of B is $B\vec{e}_j = T(\vec{e}_j) = A(\vec{e}_j)$, which is the j-th column of A. Since this holds for every column, we see $B = A$.

Definition: If A is an $n \times n$ matrix with complex (so possibly real) entries, then $\lambda \in \mathbb{C}$ is called a **complex eigenvalue** of A if there is a non-zero vector $\vec{z} \in \mathbb{C}^n$ such that $A\vec{z} = \lambda\vec{z}$. The vector $\vec{z}$ is called a **complex eigenvector** with eigenvalue λ.

Problem 4. Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Note that we can think of A as a matrix with real entries, or as a matrix with complex entries, since $A \in \mathbb{R}^{2\times 2} \subseteq \mathbb{C}^{2\times 2}$.

(a) Consider the linear transformation $T_{\mathbb{R}} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T_{\mathbb{R}}(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^2$. Prove that A (or $T_{\mathbb{R}}$) has no real eigenvalues in two different ways: by considering its characteristic polynomial and by consider what the transformation does geometrically to $\mathbb{R}^2$.

(b) Let $T_{\mathbb{C}} : \mathbb{C}^2 \to \mathbb{C}^2$ defined by $T_{\mathbb{C}}(\vec{z}) = A\vec{z}$ for all $\vec{z} \in \mathbb{C}^2$. Show that the complex numbers i and $-i$ are complex eigenvalues of A (or $T_{\mathbb{C}}$). Find corresponding eigenvectors $\vec{v}$ and $\vec{w}$ for each.

(c) Prove that there is a basis for $\mathbb{C}^2$ consisting of *eigenvectors* for $T_{\mathbb{C}}$.

(d) The matrix A is said to be **diagonalizable over** $\mathbb{C}$ but not diagonalizable over $\mathbb{R}$. Discuss and interpret what this means.

Solution:

(a) The characteristic polynomial of A is $x^2 + 1$, which has no (real) roots. Also, the matrix A is the standard matrix of rotation by $90°$ counterclockwise. Thinking geometrically, we see there is no vector taken to a scalar multiple of itself.

(b) Over $\mathbb{C}$, the polynomial $x^2 + 1$ has roots $\pm i$. To find eigenvectors, we can solve $A \begin{bmatrix} x \\ y \end{bmatrix} = \pm i \begin{bmatrix} x \\ y \end{bmatrix}$. This is

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ -x \end{bmatrix} = \pm i \begin{bmatrix} x \\ y \end{bmatrix}.$$

By inspection, we see that $\vec{v} = \begin{bmatrix} i \\ 1 \end{bmatrix}$ is an eigenvector for $-i$ and $\vec{w} = \begin{bmatrix} -i \\ 1 \end{bmatrix}$ is an eigenvector for i.

(c) The vectors $\{\vec{v}, \vec{w}\}$ from (b) are a basis (and consist of eigenvectors) since their span includes $\vec{e}_2 = \frac{1}{2}\vec{v} + \frac{1}{2}\vec{w}$ and $\vec{e}_1 = \frac{1}{2i}\vec{v} - \frac{1}{2i}\vec{w}$. They are obviously linearly independent, since if $c_1\vec{v} + c_2\vec{w} = 0$, then $c_1 + c_2 = 0$ but also $ic_1 - ic_2 = 0$. This means that $c_1 + c_2 = 0$ and $c_1 - c_2 = 0$, a contradiction unless $c_1 = c_2 = 0$.

(d) The matrix A is said to be **diagonalizable over** $\mathbb{C}$ because it has an eigenbasis consisting of complex vectors. It is not diagonalizable over $\mathbb{R}$ because it does not have an eigenbasis: there are no eigenvectors (and no eigenvalues) over $\mathbb{R}$! In this case, we can write

$$A = S \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} S^{-1}$$

where S is the change of basis matrix $\begin{bmatrix} -i & i \\ 1 & 1 \end{bmatrix}$ from the eigenbasis to the standard basis.

Let A be a matrix with *real* entries. Since $\mathbb{R} \subseteq \mathbb{C}$, A is also a matrix with complex entries.

Theorem: Let $A \in \mathbb{R}^{n \times n}$. If λ is a **complex eigenvalue** of A, then its complex conjugate $\overline{\lambda}$ is also a complex eigenvalue of A.

Similarly, if $\vec{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{C}^n$ is an eigenvector for A, then $\overline{\vec{z}} = \begin{bmatrix} \overline{z}_1 \\ \vdots \\ \overline{z}_n \end{bmatrix}$ is an eigenvector for A.

Problem 5.

(a) Verify the Theorem for the matrix in Problem 4.

(b) Determine which of the following are true:

   (i) For all $z, w \in \mathbb{C}$, $\overline{z + w} = \overline{z} + \overline{w}$.

   (ii) For all $z, w \in \mathbb{C}$, $\overline{zw} = \overline{z}\,\overline{w}$.

   (ii) For $z \in \mathbb{C}$, $\overline{z} = z$ if and only if $z \in \mathbb{R}$.

(c) Show that if $f(x)$ is a polynomial with *real* coefficients, and λ is a *complex* root, then $\overline{\lambda}$ is also a complex root of f. [HINT: Conjugate $f(\lambda)$, using (b).]

Solution: The eigenvalues in Problem 4 were the complex conjugate vectors i and $-i$. Similarly, we saw that $\begin{bmatrix} i \\ 1 \end{bmatrix}$ is an eigenvector for A, and so was the conjugate $\begin{bmatrix} -i \\ 1 \end{bmatrix}$.

All the statements are true.

(i) $\overline{(a+bi)+(c+di)} = \overline{(a+c)+(b+d)i} = (a+c)-(b+d)i = (a-bi)+(c-di) = \overline{a+bi} + \overline{c+di}$.

(ii) If $z_1 = a + ib$ and $z_2 = c + id$, then $\overline{z_1 z_2} = \overline{(a+ib)(c+id)} = (ac-bd) - i(ad+bc)$, and $\overline{z_1}\,\overline{z_2} = (a-ib)(c-id) = ac+bd - i(ad+bc)$. These are equal.

(ii) $z = a + bi \in \mathbb{R} \iff b = 0 \iff a + bi = a - bi \iff \overline{z} = z$.

For (c), suppose $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ where $a \in \mathbb{R}$. We have $f(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0$. Conjugating, and applying (i), (ii) and (iii) repeatedly, we have

$$\begin{aligned} \overline{a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0} &= \overline{a_n}\,\overline{\lambda}^n + \overline{a_{n-1}}\,\overline{\lambda}^{n-1} + \cdots + \overline{a_1}\,\overline{\lambda} + \overline{a}_0 \\ &= a_n\overline{\lambda}^n + a_{n-1}\overline{\lambda}^{n-1} + \cdots + a_1\overline{\lambda} + a_0, \end{aligned}$$

So $f(\lambda) = 0$ as well.

Problem 6. For complex column vectors $\vec{z}$ in $\mathbb{C}^n$ and complex matrices $A \in \mathbb{C}^{n\times p}$, we use the notation $\overline{\vec{z}}$ and $\overline{A}$, respectively, for the vector (respectively, matrix) whose entries are the conjugates of the entries in $\vec{z}$ (respectively, A). Determine which of the following are true:

(a) For any $c \in \mathbb{C}$ and $\vec{z} \in \mathbb{C}^n$, $\overline{c\vec{z}} = \overline{c}\,\overline{\vec{z}}$.

(b) For any $c \in \mathbb{C}$ and complex matrix A in $\mathbb{C}^{m\times n}$, $\overline{cA} = \overline{c}\overline{A}$.

(c) For any complex vector $\vec{z} \in \mathbb{C}^n$ and matrix $A \in \mathbb{C}^{m\times n}$, $\overline{A\vec{z}} = \overline{A}\,\overline{\vec{z}}$.

(d) For all $\vec{z} \in \mathbb{C}^n$, $\overline{\vec{z}} = \vec{z}$ if and only if $\vec{z} \in \mathbb{R}^n$.

(e) For all $A \in \mathbb{C}^{m\times n}$, $\overline{A} = A$ if and only if $A \in \mathbb{R}^{m\times n}$.

Solution: All the statements are true, as can be shown using the results from Problem 5. For instance, given $A \in \mathbb{C}^{m\times n}$ with (i,j)-entry a_{ij}, we have $\overline{cA} = (\overline{ca_{ij}}) = (\overline{c}\,\overline{a_{ij}}) = \overline{c}\overline{A}$.

Problem 7. Suppose that A is an $n \times n$ matrix with real entries and that $\vec{z} \in \mathbb{C}^n$ is a complex eigenvector of A with corresponding complex eigenvalue λ.

(a) Show that $\overline{\vec{z}}$ is also a complex eigenvector of A. What is the corresponding complex eigenvalue? Explain why this proves the Theorem.

(b) Give a second explanation why $\overline{\lambda}$ must be an eigenvalue using the characteristic polynomial of A. [HINT See 5c.]

Solution:

(a) If $A\vec{z} = \lambda\vec{z}$, then since A has real entries we have $A\overline{\vec{z}} = \overline{A\vec{z}} = \overline{\lambda\vec{z}} = \overline{\lambda}\,\overline{\vec{z}}$, showing that $\overline{\vec{z}}$ is also a complex eigenvalue of A, with corresponding eigenvalue $\overline{\lambda}$.

(b) If A is an $n \times n$ matrix with real entries, then the characteristic polynomial of A has real coefficients, so its complex roots (which are the eigenvalues of A) occur in conjugate pairs by Problem 5c. Therefore if λ is an eigenvalue of A, so is $\overline{\lambda}$.

Problem 8. Let a and b be real numbers, not both zero, and let $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

(a) Find the complex eigenvalues of A.

Solution: The eigenvalues are $\lambda = a \pm bi$.

(b) Factor A as a product of a scalar matrix rI_2 and a rotation matrix R_θ. How are r and θ related to the complex eigenvalues of A? [HINT: You can use QR factorization to factor A.]

Solution: $A = rI_2R_\theta$ where $r = \sqrt{a^2 + b^2}$ and $\tan\theta = b/a$. Thus $re^{i\theta}$ is the polar form of $a + bi$.

(c) Diagonalize A *over* $\mathbb{C}$; that is, find complex matrices P and D such that $A = PDP^{-1}$ where D is diagonal.

Solution: $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a + bi & 0 \\ 0 & a - bi \end{bmatrix} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}^{-1}$.

(d) Describe geometrically the effect of applying the transformation T_A repeatedly to a given point in $\mathbb{R}^2$. What is the difference between the cases $r > 1$, $r = 1$, and $0 < r < 1$?

Solution: Repeated applications of T_A move a point $\vec{x}$ in $\mathbb{R}^2$ around the origin in a spiral pattern, making a jump of angle θ with each iteration. If $r = 1$ then the spiral is a circle, whereas the path spirals inward toward the origin if $r < 1$ and outward to infinity if $r > 1$.

Problem 9. Let A be any 2×2 matrix with real entries that has a pair of (non-real) complex eigenvalues $a \pm bi$. Show that A is similar (over $\mathbb{R}$) to the matrix $B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

[HINT: Consider the fact that both A and B are similar *over* $\mathbb{C}$ to the complex diagonal matrix $\begin{bmatrix} a + bi & 0 \\ 0 & a - bi \end{bmatrix}$.]

Solution: Suppose $A(\vec{v}+i\vec{w}) = (a+bi)(\vec{v}+i\vec{w})$, so also $A(\vec{v}-i\vec{w}) = (a-bi)(\vec{v}-i\vec{w})$. Then, diagonalizing A over $\mathbb{C}$, we have

$$\begin{bmatrix} | & | \\ \vec{v}+i\vec{w} & \vec{v}-i\vec{w} \\ | & | \end{bmatrix}^{-1} A \begin{bmatrix} | & | \\ \vec{v}+i\vec{w} & \vec{v}-i\vec{w} \\ | & | \end{bmatrix} = \begin{bmatrix} a+bi & \\ & a-bi \end{bmatrix},$$

which by Problem 8(c) implies

$$\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} | & | \\ \vec{v}+i\vec{w} & \vec{v}-i\vec{w} \\ | & | \end{bmatrix}^{-1} A \begin{bmatrix} | & | \\ \vec{v}+i\vec{w} & \vec{v}-i\vec{w} \\ | & | \end{bmatrix} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

By direct computation we have $\begin{bmatrix} | & | \\ \vec{v}+i\vec{w} & \vec{v}-i\vec{w} \\ | & | \end{bmatrix} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix}$, so

$$\begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix}^{-1} A \begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Problem 10. Let $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$.

(a) Find the characteristic polynomial of A. Does A have any real eigenvalues?

Solution: $t^2 - 1.6t + 1$

(b) Find the complex eigenvalues of A.

Solution: $\lambda = \frac{4}{5} \pm \frac{3}{5}i$.

(c) Find complex eigenvectors corresponding to the complex eigenvalues you found in (b).

Solution: We have $A - (0.8+0.6i)I_2 = \begin{bmatrix} -0.3-0.6i & -0.6 \\ 0.75 & 0.3-0.6i \end{bmatrix}$. The second column of this matrix is $0.4-0.8i$ times the first, so an eigenvector corresponding to $0.8+0.6i$ is

$$\begin{bmatrix} 0.4-0.8i \\ -1 \end{bmatrix}, \qquad \text{or (after scaling),} \qquad \begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \end{bmatrix} + \begin{bmatrix} 4 \\ 0 \end{bmatrix} i.$$

So two linearly independent (complex) eigenvectors are

$$\begin{bmatrix} -2 \\ 5 \end{bmatrix} \pm \begin{bmatrix} 4 \\ 0 \end{bmatrix} i.$$

These are equivalent to $\begin{bmatrix} -2 \\ 1 \end{bmatrix} \pm \begin{bmatrix} 0 \\ 2 \end{bmatrix} i$, since $(1-2i)\begin{bmatrix} -2 \\ 1+2i \end{bmatrix} = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix}$.

(d) Choose one of the complex eigenvectors $\vec{z}$ that you found in (c), and write $\vec{z}$ as $\vec{z} = \vec{v} + i\vec{w}$ where $\vec{v}, \vec{w} \in \mathbb{R}^2$. Find $P^{-1}AP$ where $P = [\vec{v} \quad \vec{w}]$. What kind of matrix is $P^{-1}AP$?

Solution: $P^{-1}AP = \begin{bmatrix} -2 & 4 \\ 5 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix} \begin{bmatrix} -2 & 4 \\ 5 & 0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix}$, a rotation! So A is similar to a rotation, which means that A *is* a rotation "relative to a suitable basis," such as the basis of $\mathbb{R}^2$ given by the columns of P.

(e) Can you describe geometrically the action of the transformation T_A on $\mathbb{R}^2$?

Solution: It moves each point in $\mathbb{R}^2$ in an elliptical orbit around the origin.