1 Last time: methods to check diagonalizability

Let n be a positive integer and let A be an $n \times n$ matrix.

Remember that A is *diagonalizable* if $A = PDP^{-1}$ where P is an invertible $n \times n$ matrix and D is an $n \times n$ diagonal matrix. In other words, A is diagonalizable if A is similar to a diagonal matrix. When this holds and

$$P = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

then $Av_i = PDP^{-1}v_i = PDe_i = \lambda_i Pe_i = \lambda_i v_i$ for each $i = 1, 2, \dots, n$. In other words, if $A = PDP^{-1}$ is diagonalizable then the columns of P are a basis for $\mathbb{R}^n$ made up of eigenvectors of A.

Matrices which are not diagonalizable.

Proposition. $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not diagonalizable.

Proof. To check this directly, suppose $ad - bc \neq 0$ and compute

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} -ac & a^2 \\ -c^2 & ac \end{bmatrix}.$$

The only way the last matrix can be diagonal is if $a = c = 0$, but then we would have $ad - bc = 0$ so $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ would not be invertible. Therefore $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not similar to a diagonal matrix. □

Here is a second family of examples.

Let A be an $n \times n$ upper-triangular matrix with all entries on the diagonal equal to 1:

$$A = \begin{bmatrix} 1 & * & \dots & * \\ & 1 & \ddots & \vdots \\ & & \ddots & * \\ & & & 1 \end{bmatrix}$$

All entries in A below the diagonal are zero, and the entries above the diagonal can be anything.

Proposition. If $A \neq I$ is not the identity matrix then A is not diagonalizable.

Proof. Suppose $A = PDP^{-1}$ where D is diagonal.

Every diagonal entry of D is an eigenvalue for A. (Why?)

But A has characteristic polynomial $(1 - x)^n$ so its only eigenvalue is 1.

Therefore $D = I$ so $A = PIP^{-1} = PP^{-1} = I$. □

Theorem. Let A be an $n \times n$ matrix.

Suppose $\lambda_1, \lambda_2, \dots, \lambda_p$ are the distinct eigenvalues of A.

Let $d_i = \dim \mathrm{Nul}(A - \lambda_i I)$ for $i = 1, 2, \dots, p$ be the dimension of the corresponding eigenspace.

1. For each $i = 1, 2, \dots, p$ it holds that $d_i \geq 1$, and $p \leq d_1 + d_2 + \cdots + d_p \leq n$.
2. The matrix A is diagonalizable if and only if $d_1 + d_2 + \cdots + d_p = n$.
3. Suppose A is diagonalizable. Let $D_i = \lambda_i I_{d_i}$ and define D as the $n \times n$ diagonal matrix

$$D = \begin{bmatrix} D_1 & & & \\ & D_2 & & \\ & & \ddots & \\ & & & D_p \end{bmatrix}.$$

   Choose n vectors

$$a_1, a_2, \dots, a_{d_1}, b_1, b_2, \dots, b_{d_2}, \dots, z_1, z_2, \dots, z_{d_p}$$

   which are bases for $\mathrm{Nul}(A - \lambda_1 I)$, $\mathrm{Nul}(A - \lambda_2 I)$, $\dots$, $\mathrm{Nul}(A - \lambda_p I)$. Then $A = PDP^{-1}$ for

$$P = \begin{bmatrix} a_1 & a_2 & \dots & a_{d_1} & b_1 & b_2 & \dots & b_{d_2} & \dots & z_1 & z_2 & \dots & z_{d_p} \end{bmatrix}$$

2 Complex eigenvalues

We write $\mathbb{C}$ for the set of complex numbers $\{a + bi : a, b \in \mathbb{R}\}$.

Each complex number is a formal linear combination of two real numbers $a + bi$.

The symbol i is defined as the square root of -1, so $i^2 = -1$.

We add complex numbers like this:

$$(a + bi) + (c + di) = (a + c) + (b + d)i.$$

We multiply complex numbers just like polynomials, but substituting -1 for i^2:

$$(a + bi)(c + di) = ac + (ad + bc)i + bd(i^2) = (ac - bd) + (ad + bc)i.$$

The order of multiplication doesn't matter since $(a + bi)(c + di) = (c + di)(a + bi)$.

Example. The complex numbers $\mathbb{C}$ contain the real numbers $\mathbb{R}$ as a subset. Numbers of the form $bi \in \mathbb{C}$ with $b \in \mathbb{R}$ are called *imaginary*, though this is mostly just a historical convention.

Another way to think of the complex numbers is as the set of 2×2 matrices

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \qquad \text{for } a, b \in \mathbb{R}.$$

We identify this matrix with the number $a + bi \in \mathbb{C}$.

Addition and multiplication of complex numbers correspond, in terms as these matrices, to the usual notions of addition and multiplication:

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} a+c & -(b+d) \\ b+d & a+c \end{bmatrix}$$

and

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac-bd & -(ad+bc) \\ ad+bc & ac-bd \end{bmatrix}.$$

It can be helpful to draw the complex number $a+bi \in \mathbb{C}$ as the vector $\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$.

The number $i(a+bi) = -b+ai \in \mathbb{C}$ then corresponds to the vector $\begin{bmatrix} -b \\ a \end{bmatrix} \in \mathbb{R}^2$, which is given by *rotating* $\begin{bmatrix} a \\ b \end{bmatrix}$ *ninety degrees counterclockwise.* (Try drawing this yourself.)

The main reason it is helpful to work with complex numbers is the following theorem about polynomials.

Theorem (Fundamental theorem of algebra)**.** Suppose

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \dots a_1 x + a_0$$

is a polynomial of degree n (meaning $a_n \neq 0$) with coefficients $a_0, a_1, \dots, a_n \in \mathbb{C}$. There are n (not necessarily distinct) numbers $r_1, r_2, \dots, r_n \in \mathbb{C}$ such that

$$p(x) = a_n(x-r_1)(x-r_2)\cdots(x-r_n).$$

One calls the numbers $r_1, r_2, \dots, r_n$ the *roots* of $p(x)$.

A root r has *multiplicity* m if exactly m of the numbers $r_1, r_2, \dots, r_n$ are equal to r.

The characteristic equation of an $n \times n$ matrix A is a degree n polynomial with real coefficients.

Counting multiplicities, $\det(A - xI)$ has exactly n roots but some roots may be complex numbers.

Define $\mathbb{C}^n$ to be the set of vectors $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ with n rows and entries $v_1, v_2, \dots, v_n \in \mathbb{C}$.

Note that $\mathbb{R}^n \subset \mathbb{C}^n$ since $\mathbb{R} = \{a \in \mathbb{R}\} \subset \mathbb{C} = \{a+bi : a,b \in \mathbb{R}\}$.

The sum $u+v$ and scalar multiple cv for $u, v \in \mathbb{C}^n$ and $c \in \mathbb{C}$ are defined exactly as for vectors in $\mathbb{R}^n$, except we use the addition and multiplication operations from $\mathbb{C}$ instead of $\mathbb{R}$.

If A is an $n \times n$ matrix and $v \in \mathbb{C}^n$ then we define Av in the same way as when $v \in \mathbb{R}^n$.

Definition. Let A be an $n \times n$ matrix whose entries are all real numbers. Call $\lambda \in \mathbb{C}$ a *(complex) eigenvalue* of A if there exists a nonzero vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$.

Equivalently, $\lambda \in \mathbb{C}$ is an eigenvalue of A if λ is a root of the characteristic polynomial $\det(A - xI)$.

This is no different from our first definition of an eigenvalue, except that now we permit λ to be in $\mathbb{C}$.

Example. Let $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Then $\det(A - xI) = x^2 + 1 = (i-x)(-i-x)$.

The roots of this polynomial are the complex numbers i and $-i$. We have

$$A\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i\begin{bmatrix} 1 \\ -i \end{bmatrix} \quad \text{and} \quad A\begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i\begin{bmatrix} 1 \\ i \end{bmatrix}$$

so i and $-i$ are eigenvalues of A, with corresponding eigenvectors $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$.

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

$$\det(A - xI) = \det \begin{bmatrix} .5 - x & -.6 \\ .75 & 1.1 - x \end{bmatrix} = x^2 - 1.6x + 1.$$

Via the quadratic formula, we find that the roots of this characteristic polynomial are

$$x = \frac{1.6 \pm \sqrt{1.6^2 - 4}}{2} = .8 \pm .6i$$

since $i = \sqrt{-1}$. To find a basis for the $(.8 - .6i)$-eigenspace, we row reduce as usual

$$\begin{aligned} A - (.8 - .6i)I &= \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} - \begin{bmatrix} .8 - .6i & 0 \\ 0 & .8 - .6i \end{bmatrix} \\ &= \begin{bmatrix} -.3 + .6i & -.6 \\ .75 & .3 + .6i \end{bmatrix} \\ &\sim \begin{bmatrix} .5 - i & 1 \\ 1 & .8(.5 + i) \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & .8(.5 + i) \\ .5 - i & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & .8(.5 + i) \\ 0 & 1 - .8(.5 + i)(.5 - i) \end{bmatrix} = \begin{bmatrix} 1 & .8(.5 + i) \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

The last equality holds since $.8(.5 + i)(.5 - i) = .8(.25 - i^2) = .8(1.25) = 1$.

This implies that $Ax = (.8 - .6i)x$ if and only if $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ where $x_1 + .8(.5 + i)x_2 = 0$, i.e., where $5x_1 = -4(.5 + i)x_2 = -(2 + 4i)x_2$. Satisfying these conditions is the vector

$$v = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$$

which is therefore an eigenvector for A with eigenvalue $.8 - .6i$.

Similar calculations show that the vector

$$w = \begin{bmatrix} -2 + 4i \\ 5 \end{bmatrix}$$

is an eigenvector for A with eigenvalue $.8 + .6i$.

3 Complex conjugation

Given $a, b \in \mathbb{R}$, we define the *complex conjugate* of the complex number $a + bi \in \mathbb{C}$ to be

$$\overline{a + bi} = a - bi \in \mathbb{C}.$$

If A is a matrix and $v \in \mathbb{C}^n$ then we define $\overline{A}$ and $\overline{v}$ as the matrix and vector given by replacing all entries of A and v by their complex conjugates.

Lemma. Let $z \in \mathbb{C}$. Then $\overline{z} = z$ if and only if $z \in \mathbb{R}$.

Proof. Write $z = a + bi$ for $a, b \in \mathbb{R}$. Then $z - \overline{z} = (a + bi) - (a - bi) = 2bi$. This is zero if and only if $b = 0$, in which case $z = a \in \mathbb{R}$. □

Lemma. If $y, z \in \mathbb{C}$ then $\overline{y+z} = \overline{y} + \overline{z}$ and $\overline{yz} = \overline{y} \cdot \overline{z}$.

Hence if A is an $m \times n$ matrix (with real or complex entries) and $v \in \mathbb{C}^n$ then $\overline{Av} = \overline{A}\overline{v}$.

Proof. Write $y = a + bi$ and $z = c + di$ for $a, b, c, d \in \mathbb{R}$. Then

$$\overline{y+z} = \overline{(a+c)+(b+d)i} = (a+c) - (b+d)i = \overline{y} + \overline{z}$$

and

$$\overline{y \cdot z} = \overline{(ad-bc)+(ad+bc)i} = (ad-bc) - (ad+bc)i = (a-bi)(c-di) = \overline{y} \cdot \overline{z}.$$

Combining these properties shows that $\overline{Av} = \overline{A}\overline{v}$. □

If $z = a + bi \in \mathbb{C}$ then $z\overline{z} = (a+bi)(a-bi) = a^2 + b^2 \in \mathbb{R}$.

This indicates how to divide complex numbers:

$$\boxed{\frac{1}{a+bi} = \frac{a-bi}{(a+bi)(a-bi)} = \frac{a-bi}{a^2+b^2} = \tfrac{a}{a^2+b^2} - \tfrac{b}{a^2+b^2} i}$$

and more generally

$$\frac{c+di}{a+bi} = \tfrac{1}{a+bi} \cdot (c+di).$$

Example. We compute

$$\frac{3-4i}{2+i} = \frac{(3-4i)(2-i)}{(2+i)(2-i)} = \frac{6-3i-8i+4i^2}{4-i^2} = \frac{6-11i-4}{4+1} = \frac{2-11i}{5} = \tfrac{2}{5} - \tfrac{11}{5}i.$$

Proposition. Suppose A is an $n \times n$ matrix with real entries. If A has a complex eigenvalue $\lambda \in \mathbb{C}$ with eigenvector $v \in \mathbb{C}^n$ then $\overline{v} \in \mathbb{C}^n$ is an eigenvector for A with eigenvalue $\overline{\lambda}$.

Proof. Since A has real entries, it holds that $\overline{A} = A$.

Therefore $A\overline{v} = \overline{A}\overline{v} = \overline{Av} = \overline{\lambda v} = \overline{\lambda}\overline{v}$. □

The *real part* of a complex number $a + bi \in \mathbb{C}$ is $\Re(a+bi) = a \in \mathbb{R}$.

The *imaginary part* of a $a + bi \in \mathbb{C}$ is $\Im(a+bi) = b \in \mathbb{R}$.

Define $\Re(v)$ and $\Im(v)$ for $v \in \mathbb{C}^n$ by applying $\Re(\cdot)$ and $\Im(\cdot)$ to each entry in v.

Example. Let $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$ as in our earlier example.

Let $\lambda = .8 - .6i$ and $v = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$. Define

$$P = \begin{bmatrix} \Re(v) & \Im(v) \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} \qquad \text{so that} \qquad P^{-1} = \frac{1}{20}\begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}.$$

Let $C = P^{-1}AP$ so that $A = PCP^{-1}$. We compute

$$C = P^{-1}AP = \frac{1}{20}\begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}\begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} .8 & -.6 \\ .6 & .8 \end{bmatrix}.$$

Since $.8^2 + .6^2 = .64 + .36 = 1$, C is the *rotation matrix*

$$C = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

for $\phi \in [0, 2\pi)$ with $\cos\phi = .8$. Thus $A = PCP^{-1}$ is similar to a rotation matrix.

This phenomenon holds for all real 2×2 matrices.

Theorem. Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ ($b \neq 0$) and associated eigenvector $v \in \mathbb{C}^2$. Then $A = PCP^{-1}$ where

$$P = \begin{bmatrix} \Re(v) & \Im(v) \end{bmatrix} \qquad \text{and} \qquad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Moreover, $C = r \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$ where $r = \sqrt{a^2 + b^2}$ and $\phi \in [0, 2\pi)$ is such that $r\cos\phi = a$.

4 Vocabulary

Keywords from today's lecture:

1. **Complex number**.

   An expression of the form "$a + bi$" where $a, b \in \mathbb{R}$ and i is a formal symbol.

   The set of complex numbers is denoted $\mathbb{C}$.

   Complex numbers can be added and multiplied together. These operations are carried out by treating $a + bi$ as a polynomial in a variable i that satisfies $i^2 = -1$.

   Concretely, the notation $a + bi$ is just a shorthand for the 2×2 matrix $\left[\begin{array}{cc} a & -b \\ b & a \end{array}\right]$. Such matrices can be added and multiplied together, and the order of multiplication doesn't matter. The way we define addition and multiplication for complex numbers corresponds exactly to the way we define addition and multiplication for 2×2 matrices.

   Example:

   $1 + 2i$ corresponds to $\left[\begin{array}{cc} 1 & -2 \\ 2 & 1 \end{array}\right]$.

   $(1 + 2i) + (2 + 3i) = 3 + 5i$ corresponds to $\left[\begin{array}{cc} 1 & -2 \\ 2 & 1 \end{array}\right] + \left[\begin{array}{cc} 2 & -3 \\ 3 & 2 \end{array}\right] = \left[\begin{array}{cc} 3 & -5 \\ 5 & 3 \end{array}\right]$.

   $(1 + 2i)(2 + 3i) = -4 + 7i$ corresponds to $\left[\begin{array}{cc} 1 & -2 \\ 2 & 1 \end{array}\right]\left[\begin{array}{cc} 2 & -3 \\ 3 & 2 \end{array}\right] = \left[\begin{array}{cc} -4 & -6 \\ 7 & -4 \end{array}\right]$.

   $(1 + 2i)^{-1} = \frac{1}{5} - \frac{2}{5}i$ corresponds to $\left[\begin{array}{cc} 1 & -2 \\ 2 & 1 \end{array}\right]^{-1} = \frac{1}{5}\left[\begin{array}{cc} 1 & 2 \\ -2 & 1 \end{array}\right]$.

2. **Complex conjugation**.

   If $a, b \in \mathbb{R}$ then *complex conjugate* of $a + bi \in \mathbb{C}$ is $\overline{a + bi} = a - bi \in \mathbb{C}$.

   If $y, z \in \mathbb{C}$ then $\overline{y + z} = \overline{y} + \overline{z}$ and $\overline{yz} = \overline{y} \cdot \overline{z}$ and $\overline{y^{-1}} = \overline{y}^{-1}$.

3. **Fundamental theorem of algebra**.

   Any polynomial
   $$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$
   with coefficients $a_0, a_1, \ldots, a_n \in \mathbb{C}$ and $a_n \neq 0$ can be factored as
   $$f(x) = a_n(x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$$
   for some not necessarily distinct complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$.

4. **Complex eigenvalue**.

   Let $\mathbb{C}^n$ be the set of vectors with n rows with entries in $\mathbb{C}$. Since $\mathbb{R} \subset \mathbb{C}$, we have $\mathbb{R}^n \subset \mathbb{C}^n$.

   If A is an $n \times n$ matrix and there exists a nonzero vector $v \in \mathbb{C}^n$ with $Av = \lambda v$ for some $\lambda \in \mathbb{C}$, then λ is a *complex eigenvalue* for A. If $\lambda \in \mathbb{R} \subset \mathbb{C}$, then λ is a *real eigenvalue* for A.

   Example: The matrix $A = \left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]$ has complex eigenvalues i and $-i$.

   We have $\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]\left[\begin{array}{c} 1 \\ i \end{array}\right] = \left[\begin{array}{c} -i \\ 1 \end{array}\right] = -i\left[\begin{array}{c} 1 \\ i \end{array}\right]$ and $\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]\left[\begin{array}{c} 1 \\ -i \end{array}\right] = \left[\begin{array}{c} i \\ -1 \end{array}\right] = i\left[\begin{array}{c} 1 \\ -i \end{array}\right]$