Summary

Quick summary of today's notes. Lecture starts on next page.

- The characteristic equation of an $n \times n$ matrix A is a degree n polynomial in one variable.

  We can always factor this polynomial as

  $$\det(A - xI) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x)$$

  for some $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$. These complex numbers are the *(complex) eigenvalues* of A.

- 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 $v_1, v_2, \ldots, v_n \in \mathbb{C}$.

  The sum $u + v$ and scalar multiple cv for $u, v \in \mathbb{C}^n$, $c \in \mathbb{C}$ are defined just 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 $v \in \mathbb{R}^n$.

  A complex number $\lambda \in \mathbb{C}$ an *eigenvalue* of A if and only if there exists $0 \neq v \in \mathbb{C}^n$ with $Av = \lambda v$.

- The *trace* of a square matrix A, denoted $\mathrm{tr} A$, is the sum of the diagonal entries of A.

  If A and B are both $n \times n$ then $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ and $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

  But usually $\mathrm{tr}(AB) \neq \mathrm{tr}(A)\mathrm{tr}(B)$.

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

  Suppose the roots of the characteristic polynomial $\det(A - xI)$ are $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$.

  These are the eigenvalues of A, repeated accordingly to their multiplicity.

  Then $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$ and $\mathrm{tr} A = \lambda_1 + \lambda_2 + \ldots \lambda_n$.

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

  The matrices A and A^T have the same characteristic polynomial and same eigenvalues.

  If A is invertible, then A and A^{-1} have the same eigenvectors.

  However, λ is an eigenvalue of A if and only if λ^{-1} is an eigenvalue for A^{-1}.

  If A is diagonalizable then so is A^T and A^{-1} (when A is invertible).

1 Last time: complex numbers

Given $a, b \in \mathbb{R}$, we interpret $a + bi$ as the matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, so $1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $i = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

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

We view $\mathbb{R} = \{a + 0i : a \in \mathbb{R}\}$ as a subset of $\mathbb{C}$.

According to our definition, every complex number is a 2×2 matrix. It can also be helpful to think of a complex number $a + bi$ as a polynomial with real coefficient in a variable i that satisfies $i^2 = -1$.

We can add, subtract, multiply, and invert complex numbers. These operations correspond to the usual ways of adding, subtracting, multiplying, and inverting matrices.

Let $a, b, c, d \in \mathbb{R}$. We add complex numbers in the following way:

$$\boxed{(a + bi) + (c + di) = (a + c) = (b + d)i \in \mathbb{C}.}$$

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

$$\boxed{(a + bi)(c + di) = ac + (ad + bc)i + bd(i^2) = (ac - bd) + (ad + bc)i \in \mathbb{C}.}$$

The order of multiplication does not matter since $(a + bi)(c + di) = (c + di)(a + bi)$.

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

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

If $z = a + bi \in \mathbb{C}$. Then $\overline{z} = z$ if and only if $b = 0$ and $z \in \mathbb{R}$.

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

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 invert complex numbers $0 \neq a + bi$:

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}^{-1} = \boxed{(a + bi)^{-1} = \frac{a - bi}{(a + bi)(a - bi)} = \frac{a - bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i} = \frac{1}{a^2 + b^2} \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

Finally, complex division is defined by

$$\boxed{\frac{a + bi}{c + di} = (a + bi)(c + di)^{-1} = (c + di)^{-1}(a + bi).}$$

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

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.

Example. We have $9x^2 + 36 = 9(x - 2i)(x + 2i)$.

2 Complex eigenvalues

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 $v_1, v_2, \dots, v_n \in \mathbb{C}$.

We have $\mathbb{R}^n \subset \mathbb{C}^n$ 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 $v \in \mathbb{R}^n$. For example:

$$\begin{bmatrix} i & 1 \\ 3 & 2i \end{bmatrix} \begin{bmatrix} 1 \\ 1-i \end{bmatrix} = \begin{bmatrix} i + (1-i) \\ 3 + 2i(1-i) \end{bmatrix} = \begin{bmatrix} 1 \\ 3 + 2i - 2i^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 + 2i \end{bmatrix}.$$

Definition. Let A be an $n \times n$ matrix with entries in $\mathbb{R}$ or $\mathbb{C}$.

Let $\lambda \in \mathbb{C}$. The following statements are equivalent:

- λ is an *eigenvalue* of A.
- $Av = \lambda v$ for some nonzero vector $v \in \mathbb{C}^n$
- $\det(A - \lambda I) = 0$.

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

Example. The eigenvalues of $A = \begin{bmatrix} i & 1 \\ 3 & 2i \end{bmatrix}$ are the solutions to

$$0 = \det(A - xI) = \det \begin{bmatrix} i - x & 1 \\ 3 & 2i - x \end{bmatrix} = (i-x)(2i-x) - 3 = 2i^2 - 3ix + x^2 - 3 = -5 - 3ix + x^2.$$

By the quadratic formula these solutions are

$$\lambda = \frac{3i \pm \sqrt{(-3i)^2 - 4(-5)}}{2} = \frac{3i \pm \sqrt{-9+20}}{2} = \pm\tfrac{\sqrt{11}}{2} + \tfrac{3}{2}i.$$

The fundamental theorem of algebra implies the following essential property:

Fact. If A is an $n \times n$ matrix then A has n (not necessarily real or distinct) eigenvalues $\lambda \in \mathbb{C}$, counting repeated eigenvalues with their respective multiplicities.

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

Proposition. Suppose A is an $n \times n$ matrix with real entries, so that $A = \overline{A}$. 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 $\overline{\lambda}$.

This proposition does **not** apply to $A = \begin{bmatrix} i & 1 \\ 3 & 2i \end{bmatrix}$ from above since A does not have all real entries.

3 Some final properties of eigenvalues of eigenvectors

We discuss a few more properties of eigenvalues and eigenvectors.

Lemma. Suppose we can write a polynomial in x in two ways as

$$(\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

for some complex numbers $\lambda_1, \lambda_2, \dots, \lambda_n, a_0, a_1, \dots, a_n \in \mathbb{C}$. Then

$$a_n = (-1)^n \qquad \text{and} \qquad a_{n-1} = (-1)^{n-1}(\lambda_1 + \lambda_2 + \cdots + \lambda_n) \qquad \text{and} \qquad a_0 = \lambda_1\lambda_2\cdots\lambda_n.$$

Proof. The product $(\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x)$ is a sum of 2^n monomials corresponding to a choice of either λ_i or $-x$ for each of the n factors, multiplied together.

The only such monomial of degree n is $(-x)^n = (-1)^n x^n = a_n x^n$ so $a_n = (-1)^n$.

The only such monomial of degree 0 is $\lambda_1\lambda_2\cdots\lambda_n = a_0$.

Finally, there are n monomials of degree $n-1$ that arise:

$$\lambda_1(-x)^{n-1} + (-x)\lambda_2(-x)^{n-2} + (-x)^2\lambda_3(-x)^{n-3} + \cdots + (-x)^{n-1}\lambda_n = (-1)^{n-1}(\lambda_1 + \cdots + \lambda_n)x^{n-1}.$$

This sum must be equal to $a_{n-1}x^{n-1}$ so $a_{n-1} = (-1)^{n-1}(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$. □

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

Define $\mathrm{tr}(A)$ to be the sum of the diagonal entries of A. Call $\mathrm{tr}(A)$ the *trace* of A.

Example. $\mathrm{tr}\left(\begin{bmatrix} 1 & 0 & 7 \\ -1 & 2 & 8 \\ 2 & 4 & 3 \end{bmatrix}\right) = 1 + 2 + 3 = 6.$

Proposition. If A, B are $n \times n$ matrices then $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ and $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

However, usually $\mathrm{tr}(AB) \neq \mathrm{tr}(A)\mathrm{tr}(B)$, unlike for the determinant.

Proof. The diagonal entries of $A + B$ are given by adding together the diagonal entries of A with those of B in corresponding positions, so it follows that $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.

Let E_{ij} be the $n \times n$ matrix with 1 in position (i, j) and 0 in all other positions.

(In this proof, we use the symbol i to mean an integer index rather than a complex number.)

You can check that $E_{ij}E_{kl}$ is the zero matrix if $j \neq k$ and that $E_{ij}E_{jk} = E_{ik}$.

Moreover, $\mathrm{tr}(E_{ij}) = 0$ if $i \neq j$ and $\mathrm{tr}(E_{ii}) = 1$.

We conclude that $\mathrm{tr}(E_{ij}E_{kl})$ is 1 if $i = l$ and $j = k$ and is 0 otherwise.

This formula is symmetric so $\mathrm{tr}(E_{ij}E_{kl}) = \mathrm{tr}(E_{kl}E_{ij})$.

It follows that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ since if A_{ij} and B_{ij} are the entries of A and B in positions (i, j), then

$$A = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}E_{ij} \qquad \text{and} \qquad B = \sum_{k=1}^{n}\sum_{l=1}^{n} B_{kl}E_{kl}.$$

□

Theorem. Let A be an $n \times n$ matrix (with entries in $\mathbb{R}$ or $\mathbb{C}$).

Suppose the characteristic polynomial of A factors as

$$\det(A - xI) = (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x).$$

Then $\det A = \lambda_1\lambda_2\cdots\lambda_n$ and $\mathrm{tr}A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$. In other words:

(a) The product of the (complex) eigenvalues of A, counted with multiplicity, is $\det(A)$.

(b) The sum of the (complex) eigenvalues of A, counted with multiplicity if $\mathrm{tr}(A)$.

Remark. The theorem is true for all matrices, but is much easier to prove for diagonalizable matrices. If $A = PDP^{-1}$ where D is a diagonal matrix, then $\det(A) = \det(PDP^{-1}) = \det(D) = \lambda_1\lambda_2\cdots\lambda_n$ and

$$\mathrm{tr}(A) = \mathrm{tr}(PDP^{-1}) = \mathrm{tr}(DP^{-1}P) = \mathrm{tr}(D) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

Before proving the theorem let's see an example.

Example. If $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & i \end{bmatrix}$ then $\begin{bmatrix} -i \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} i \\ 1 \\ 0 \end{bmatrix}$ are eigenvectors of A.

The corresponding eigenvalues are i, i, and $-i$.

One can check that $\det(A - xI) = -x^3 + ix^2 - x + i = (i - x)^2(-i - x)$.

The theorem asserts that $(i)(i)(-i) = -i^3 = i = \det(A)$ and $i + i + (-i) = i = \mathrm{tr}(A)$.

Proof of the theorem. We can write $\det(A - xI) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ for some numbers $a_0, a_1, \ldots, a_n \in \mathbb{C}$. By the lemma it suffices to show that $a_0 = \det(A)$ and $a_{n-1} = (-1)^{n-1}\mathrm{tr}(A)$.

The first claim is easy. The value of a_0 is given by setting $x = 0$ in $\det(A - xI)$, so $a_0 = \det(A)$.

Showing that $a_{n-1} = (-1)^{n-1}\mathrm{tr}(A)$ takes a little more work.

Consider the coefficient a_{n-1} of x^{n-1} in the characteristic polynomial $\det(A - xI)$. Remember our formula

$$\det(A - xI) = \sum_{Z \in S_n} (-1)^{\mathsf{inv}(Z)} \mathsf{prod}(Z, A - xI) \qquad (*)$$

where $\mathsf{prod}(Z, A - xI)$ is the product of the entries of $A - xI$ in the nonzero positions of the permutation matrix Z. The key observation to make is that if $Z \in S_n$ is not the identity matrix then Z has at most $n - 2$ nonzero entries on the diagonal, so $\mathsf{prod}(Z, A - xI)$ is a polynomial in x degree at most $n - 2$.

Therefore the formula (*) implies that

$$\det(A - xI) = \mathsf{prod}(I, A - xI) + (\text{polynomial terms of degree} \le n - 2).$$

Let d_i be the diagonal entry of A in position (i, i). Then $\mathsf{prod}(I, A - xI) = (d_1 - x)(d_2 - x)\cdots(d_n - x)$ and the coefficient of x^{n-1} in this polynomial must be equal to the coefficient of x^{n-1} in $\det(A - xI)$.

By the lemma, the coefficient of x^{n-1} in $(d_1 - x)(d_2 - x)\cdots(d_n - x)$ is

$$(-1)^{n-1}(d_1 + d_2 + \cdots + d_n) = (-1)^{n-1}\mathrm{tr}(A),$$

and so $a_{n-1} = (-1)^{n-1}\mathrm{tr}(A)$. □

Corollary. Suppose A is a 2×2 matrix. Let $p = \det A$ and $q = \mathrm{tr} A$.

Then A has distinct eigenvalues if and only if $q^2 \neq 4p$.

Proof. Suppose $a, b \in \mathbb{C}$ are the eigenvalues of A (repeated with multiplicity).

Then $ab = p$ and $a + b = q$ so $a(q - a) = qa - a^2 = p$ and therefore $a^2 - qa + p = 0$.

The quadratic formula implies that $a = \frac{q \pm \sqrt{q^2 - 4p}}{2}$ and $b = \frac{q \mp \sqrt{q^2 - 4p}}{2}$ so $a \neq b$ if and only if $q^2 - 4p \neq 0$. □

Proposition. If A is a square matrix then A and A^T have the same eigenvalues.

Proof. This follows since $\det(A - xI) = \det((A - xI)^T) = \det(A^T - xI^T) = \det(A^T - xI)$. □

Proposition. Let A be a square matrix. Then A is invertible if and only if 0 is not one of its eigenvalues. Assume A is invertible. Then A and A^{-1} have the same eigenvectors, but v is an eigenvector of A with eigenvalue λ if and only if v is an eigenvector of A^{-1} with eigenvalue $1/\lambda$.

Proof. 0 is an eigenvalue of A if and only if $\det A = 0$ which occurs precisely when A is not invertible.

If A is invertible and $Av = \lambda v$ then $v = A^{-1}Av = A^{-1}\lambda v = \lambda A^{-1} v$ so $A^{-1}v = \lambda^{-1} v$. □

Corollary. If A is invertible and diagonalizable then A^{-1} is diagonalizable.

Proof. If A is invertible and diagonalizable, then $\mathbb{R}^n$ has a basis consisting of eigenvectors of A, but this basis is then also made up of eigenvectors of A^{-1}, so A^{-1} is diagonalizable. □

Corollary. If A is diagonalizable then A^T is diagonalizable.

Proof. If $A = PDP^{-1}$ then $A^T = (PDP^{-1})^T = (P^{-1})^T D^T P^T = QEQ^{-1}$ for the invertible matrix $Q = (P^{-1})^T = (P^T)^{-1}$ and the diagonal matrix $E = D^T$. □

4 Vocabulary

Keywords from today's lecture:

1. **(Complex) eigenvalues and eigenvectors**.

   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 an *eigenvalue* for A. The vector v is called an *eigenvector*.

   Example: The matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has eigenvalues i and $-i$.

   We have $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i \begin{bmatrix} 1 \\ i \end{bmatrix}$ and $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ -1 \end{bmatrix} = i \begin{bmatrix} 1 \\ -i \end{bmatrix}$.

2. **Trace** of a square matrix.

   The sum of the diagonal entries of a square matrix A, denote $\mathrm{tr}(A)$.

   The value of $\mathrm{tr}(A)$ is also the sum of the complex eigenvalues of A, counted with multiplicity.

   Example: $\mathrm{tr} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = 1 + 4 = 5$.