

# Linear Algebra Notes — Lecture 10/20

## Complex Eigenvalues, Multiplicity, Similarity and Diagonalization

Lecture pages 207–222 · Lay §5.2–5.5

### Contents

|                                                                                                 |    |
|-------------------------------------------------------------------------------------------------|----|
| Overview                                                                                        | 1  |
| 1 Recap: the characteristic polynomial (p. 207)                                                 | 2  |
| 2 Complex scalars: $\mathbb{C}^n$ , $\mathbb{R}$ -linear and $\mathbb{C}$ -linear (pp. 208–209) | 2  |
| 3 Complex eigenvalues and eigenvectors (pp. 209–210)                                            | 4  |
| 4 Real matrices: conjugate pairs (pp. 211–212)                                                  | 5  |
| 5 Multiplicity (pp. 213–215)                                                                    | 6  |
| 6 Determinant and trace from the eigenvalues (pp. 216–218)                                      | 8  |
| 7 Eigenvalues of a linear map $T : V \rightarrow V$ (pp. 219–220)                               | 11 |
| 8 Similar matrices (pp. 221–222)                                                                | 12 |
| 9 Diagonalizable matrices (p. 222)                                                              | 14 |
| Summary                                                                                         | 15 |
| Corrections to the handwritten notes                                                            | 15 |

---

### Overview

The 10/15 lecture found eigenvalues as the roots of the characteristic polynomial  $\det(A - \lambda I)$ . By the fundamental theorem of algebra those roots may be complex, so this lecture allows complex scalars and then asks what the eigenvalues tell us about  $A$ :

1.  $\mathbb{C}^n$  with complex scalars; complex eigenvalues and eigenvectors (the  $90^\circ$  rotation has  $\lambda = \pm i$ );
2. for a *real* matrix, non-real eigenvalues and eigenvectors come in conjugate pairs  $\lambda, \bar{\lambda}$  and  $v, \bar{v}$ , even counting multiplicity;

3.  $\det A = \lambda_1 \cdots \lambda_n$  and  $\text{trace } A = \lambda_1 + \cdots + \lambda_n$ ;
4. eigenvalues of a linear map  $T : V \rightarrow V$ ; the matrices of  $T$  in two bases are *similar*;
5.  $A$  is *diagonalizable* ( $A = PDP^{-1}$ ) exactly when  $\mathbb{R}^n$  has a basis of eigenvectors of  $A$ .

Every blank and question in the notes is answered in a blue box. Typos are listed at the end.

### Notation used throughout.

- $\bar{z} = a - ib$  is the complex conjugate of  $z = a + ib$ ; for a vector or matrix,  $\bar{v}$  and  $\bar{A}$  conjugate every entry (9/24 notes, p. 128).
- $\text{trace } A = a_{11} + \cdots + a_{nn}$ , the trace (10/15 notes, p. 207).
- $S(v) = [v]_\beta$  is the coordinate map of a basis  $\beta$  (9/22 notes, p. 115; 9/29 notes, p. 139).
- $P_{\gamma \leftarrow \beta}$  is the change-of-basis matrix,  $[v]_\gamma = P_{\gamma \leftarrow \beta} [v]_\beta$  (9/29 notes, p. 144).

The final notes end on p. 222, where the diagonalization theorem is stated but not yet proved. Its proof (p. 223) is in the 10/22 notes.

## 1 Recap: the characteristic polynomial (p. 207)

Page 207 is also the last page of the 10/15 notes, and it is covered in full there, together with pp. 205–206 (10/15 notes, pp. 205–207, Section 5). In brief:

- $\det(A - \lambda I)$  is the **characteristic polynomial** of the  $n \times n$  matrix  $A$ .
- It has degree  $n$ , and

$$\det(A - \lambda I) = (-\lambda)^n + (\text{trace } A)(-\lambda)^{n-1} + \cdots + \det A.$$

The constant term is  $f(0) = \det(A - 0I) = \det A$ .

- **Fundamental theorem of algebra:** every polynomial  $x^n + a_{n-1}x^{n-1} + \cdots + a_0$  with complex coefficients factors as  $(x - \lambda_1) \cdots (x - \lambda_n)$  for some complex  $\lambda_1, \dots, \lambda_n$ .
- For  $n \leq 4$  there are explicit root formulas; for  $n > 4$  there are none.

Because of the fundamental theorem, eigenvalues can be complex even when  $A$  is real. The rest of the lecture deals with that.

## 2 Complex scalars: $\mathbb{C}^n$ , $\mathbb{R}$ -linear and $\mathbb{C}$ -linear (pp. 208–209)

$\mathbb{C}^n$  can be a vector space with real scalars or with complex scalars, and whether a map is “linear” depends on which.

**Recall (p. 208).**  $\mathbb{C}^n$  is the set of  $n \times 1$  matrices with complex entries. It is a vector space in two ways:

1. when the scalars are real numbers (then  $\mathbb{C}^n$  behaves like  $\mathbb{R}^{2n}$ :  $z = a + ib$  is the pair  $(a, b)$ );
2. when the scalars are complex numbers.

Next to item 1 the notes write the product  $\begin{bmatrix} i+2 \\ -3 \end{bmatrix} (-i+2)$ , a vector of  $\mathbb{C}^2$  times a complex scalar. Each entry is multiplied by the scalar:

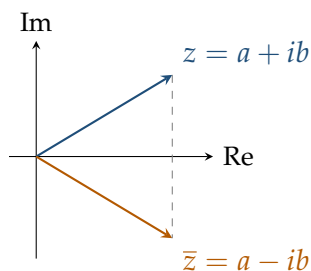
$$\begin{bmatrix} i+2 \\ -3 \end{bmatrix} (-i+2) = \begin{bmatrix} (2+i)(2-i) \\ -3(2-i) \end{bmatrix} = \begin{bmatrix} 4-i^2 \\ -6+3i \end{bmatrix} = \begin{bmatrix} 5 \\ -6+3i \end{bmatrix}.$$

This product is allowed in way 2 but not in way 1, because the scalar  $-i+2$  is not real.

The addition is the same both times. Only the set of allowed scalars changes, and so does the meaning of “linear”:  $T$  is  **$\mathbb{R}$ -linear** if  $T(c_1 z_1 + c_2 z_2) = c_1 T(z_1) + c_2 T(z_2)$  for real  $c_1, c_2$ , and  **$\mathbb{C}$ -linear** if this holds for all complex  $c_1, c_2$ .

**Example (pp. 208–209): complex conjugation**  $T : \mathbb{C} \rightarrow \mathbb{C}$ ,  $T(z) = \bar{z}$

For  $z = a + ib$  with  $a, b$  real,  $\bar{z} = a - ib$ . The notes draw  $z$  and  $\bar{z}$  in the plane (real part across, imaginary part up):  $\bar{z}$  is the mirror image of  $z$  in the real axis.



Since conjugation respects sums and products (9/24 notes, pp. 127–128),

$$T(c_1 z_1 + c_2 z_2) = \overline{c_1 z_1 + c_2 z_2} = \overline{c_1} \bar{z}_1 + \overline{c_2} \bar{z}_2.$$

**Is  $T(z) = \bar{z}$  linear with real scalars ( $\mathbb{R}$ -linear)? Yes**

The notes mark the key step under the display:  $\overline{c_1} = c_1$  and  $\overline{c_2} = c_2$ , because  $c_1, c_2$  are real. So the display gives  $T(c_1 z_1 + c_2 z_2) = c_1 \bar{z}_1 + c_2 \bar{z}_2 = c_1 T(z_1) + c_2 T(z_2)$ .

Meaning: in real coordinates  $z = (a, b)$ ,  $T$  is  $(a, b) \mapsto (a, -b)$ , the reflection across the real axis, with matrix  $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ .

**Is  $T(z) = \bar{z}$  linear with complex scalars ( $\mathbb{C}$ -linear)? No**

The notes (p. 209) test the scalar  $c = i$  on  $z = a + ib$ :

$$T(i(a + ib)) = \overline{ia - b} = -ia - b \neq i T(a + ib) = i(a - ib) = ia + b.$$

(Here  $i(a + ib) = ia + i^2 b = ia - b$ , whose real part is  $-b$  and imaginary part  $a$ , so its conjugate is  $-b - ia$ .) The two results are negatives of each other, so they differ whenever  $z \neq 0$ .

In general, for non-real  $c$  we get  $T(cz) = \bar{c}\bar{z}$ , while linearity needs  $cT(z) = c\bar{z}$ ; these differ whenever  $z \neq 0$ , because  $\bar{c} \neq c$ . Simplest instance:  $a = 1, b = 0$  gives  $T(i) = -i$  but  $iT(1) = i$ . So  $T$  is not  $\mathbb{C}$ -linear.

### 3 Complex eigenvalues and eigenvectors (pp. 209–210)

The definition of an eigenvalue is the same as before, except that  $\lambda$  and  $v$  may now be complex.

#### Definition (p. 209)

Let  $A$  be a complex (possibly real)  $n \times n$  matrix,  $v \neq 0$  in  $\mathbb{C}^n$  and  $\lambda \in \mathbb{C}$  with  $Av = \lambda v$ . Then  $\lambda$  is a **complex eigenvalue** of  $A$ , and  $v$  is a corresponding **eigenvector**.

As before,  $\lambda$  is an eigenvalue exactly when  $\det(A - \lambda I) = 0$ , and the eigenvectors are the nonzero vectors of  $\text{Nul}(A - \lambda I)$ , now computed with complex numbers (10/15 notes, pp. 193–196). Row reduction works the same way over  $\mathbb{C}$ .

**Example (pp. 209–210):**  $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

This is the matrix that already appeared on 9/29 (the recursion  $x_{k+2} + x_k = 0$ ). In the plane it is a rotation by  $90^\circ$  clockwise:  $A(1,0)^T = (0,-1)^T$ . No real vector keeps its direction, so there are no real eigenvectors. Over  $\mathbb{C}$  there are some:

$$A - \lambda I = \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix}, \quad 0 = \det(A - \lambda I) = (-\lambda)^2 + 1 = \lambda^2 + 1,$$

so  $\lambda^2 = -1$ , and the eigenvalues are  $\lambda_+ = i$  and  $\lambda_- = -i$ .

**Fill-in:**  $v_+ = (1, i)$  for  $\lambda_+ = i$

$$(A - iI)v_+ = \begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Row 1 says  $-iz_1 + z_2 = 0$ , i.e.  $z_2 = iz_1$ . Row 2,  $(-1, -i)$ , is  $-i$  times row 1,  $(-i, 1)$ , so it says the same thing. With  $z_1$  free,

$$\text{Nul}(A - iI) = \text{span} \left\{ \begin{bmatrix} 1 \\ i \end{bmatrix} \right\}, \quad v_+ = \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

The notes check this row by row:  $\begin{bmatrix} -i & 1 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = -i + i = 0$  and  $\begin{bmatrix} -1 & -i \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = -1 - i^2 = -1 + 1 = 0$ .

Check:  $Av_+ = \begin{bmatrix} 0 \cdot 1 + 1 \cdot i \\ -1 \cdot 1 + 0 \cdot i \end{bmatrix} = \begin{bmatrix} i \\ -1 \end{bmatrix} = i \begin{bmatrix} 1 \\ i \end{bmatrix}$ , since  $i \cdot i = -1$ . ✓

**Fill-in:**  $v_- = (1, -i) = \overline{v_+}$  for  $\lambda_- = -i$

Here  $A - \lambda_- I = A - (-i)I = A + iI$ , as the notes point out:

$$(A + iI) \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Row 1 says  $iz_1 + z_2 = 0$ , i.e.  $z_2 = -iz_1$ ; row 2 ( $-z_1 + iz_2 = 0$ ) is  $i$  times row 1. So

$$\text{Nul}(A + iI) = \text{span} \left\{ \begin{bmatrix} 1 \\ -i \end{bmatrix} \right\}, \quad v_- = \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

Check:  $Av_- = \begin{bmatrix} -i \\ -1 \end{bmatrix} = -i \begin{bmatrix} 1 \\ -i \end{bmatrix}$ , since  $(-i)(-i) = -1$ . ✓

The notes end the example with  $v_- = \begin{bmatrix} 1 \\ -i \end{bmatrix} = \overline{v_+}$ . Together with  $\lambda_- = \overline{\lambda_+}$ , this is no accident: see the theorem on p. 212.

## 4 Real matrices: conjugate pairs (pp. 211–212)

For a real matrix, a non-real eigenvalue needs a non-real eigenvector, and conjugating an eigenpair gives another eigenpair.

### Theorem (p. 211)

If  $A$  is a real  $n \times n$  matrix,  $Av = \lambda v$  with  $v \neq 0$ , and  $\lambda$  is not real, then  $v$  is not real.

*Proof (p. 211).* The notes: “Argue by contradiction. If  $v$  is real,  $Av$  is also real  $\neq \lambda v \leftarrow$  not real.” In full: suppose  $v$  were real. Then  $Av$  is real, because  $A$  and  $v$  are. But  $\lambda v$  is not real: since  $v \neq 0$ , some entry  $v_j$  is a nonzero real number, and entry  $j$  of  $\lambda v$  is  $\lambda v_j$ , which is not real (if it were,  $\lambda = (\lambda v_j)/v_j$  would be a quotient of real numbers, hence real). So  $Av \neq \lambda v$ , a contradiction. ■

**Question (p. 211):** if  $A$  and  $\lambda$  are real, must  $v$  be real? No, but a real one always exists

**No.** The notes put it as “Real  $v$  exist but so does nonreal  $v$ ”: multiply an eigenvector by  $i$ . Since

$$A(iv) = iAv = i\lambda v = \lambda(iv), \quad \text{equivalently} \quad (A - \lambda I)(iv) = i(A - \lambda I)v = 0 \cdot i = 0,$$

$iv$  is again an eigenvector, and if  $v$  is real and nonzero then  $iv$  is not real. For example  $A = I$  ( $2 \times 2$ ),  $\lambda = 1$ ,  $v = (1, 0)^T$ ,  $iv = (i, 0)^T$ : then  $A(iv) = iv = 1 \cdot (iv)$ .

**But** the eigenspace always contains real eigenvectors. Write  $v = a + ib$  with  $a, b \in \mathbb{R}^n$ . Then  $Aa + iAb = Av = \lambda v = \lambda a + i\lambda b$ . Here  $Aa$ ,  $Ab$ ,  $\lambda a$ ,  $\lambda b$  are all real, so comparing real and imaginary parts gives  $Aa = \lambda a$  and  $Ab = \lambda b$ . Since  $v \neq 0$ , at least one of  $a, b$  is nonzero, and it is a real eigenvector. (In the example,  $iv = (i, 0)$  has  $a = 0$ ,  $b = (1, 0)$ .) Equivalently:  $A - \lambda I$  is a real matrix, so row reducing it never creates an  $i$ , and the usual basis of its null space is real.

**Theorem (p. 212)**

If  $Av = \lambda v$ ,  $v \neq 0$ ,  $A$  is real and  $\lambda$  is not real, then  $A\bar{v} = \bar{\lambda}\bar{v}$ .

*Proof (p. 212).* The notes conjugate  $Av = \lambda v$ : “ $\overline{Av} = \overline{\lambda v} = \bar{\lambda}\bar{v}$ ” and “ $\overline{Av} = \overline{A}\bar{v} = A\bar{v}$ , as  $A$  is real.” The details: conjugation respects sums and products (9/24 notes, pp. 127–128), and each entry of  $Av$  is a sum of products  $a_{jk}v_k$ , so  $\overline{Av} = \overline{A}\bar{v}$ . Since  $A$  is real,  $\overline{A} = A$ . Hence

$$A\bar{v} = \overline{A}\bar{v} = \overline{Av} = \bar{\lambda}\bar{v} = \bar{\lambda}\bar{v}.$$

Also  $\bar{v} \neq 0$ , because  $v \neq 0$ . So  $\bar{v}$  is an eigenvector with eigenvalue  $\bar{\lambda}$ . ■

*Check on p. 210:*  $\overline{v_+} = \overline{(1, i)} = (1, -i) = v_-$  and  $\overline{\lambda_+} = -i = \lambda_-$ . ✓

**Corollary (p. 212)**

Non-real eigenvalues and eigenvectors of a real matrix come in pairs  $\lambda, \bar{\lambda}$  and  $v, \bar{v}$ .

(When  $\lambda$  is not real,  $\bar{\lambda} \neq \lambda$ , so the pair really consists of two different eigenvalues.)

**Ex (p. 212): is there a real  $3 \times 3$  matrix with eigenvalues  $1, 2, i$ ? No**

The notes: “No!  $-i$  must be an eigenvalue.  $3 \times 3$  has 3 eigenvalues.” In detail: if  $A$  is real and  $i$  is an eigenvalue, the corollary says  $\bar{i} = -i$  is an eigenvalue too. Then  $A$  has the four different eigenvalues  $1, 2, i, -i$ . But they are roots of  $\det(A - \lambda I)$ , a polynomial of degree 3 (10/15 notes, p. 205), and a nonzero polynomial of degree 3 has at most 3 roots. Contradiction.

(A complex  $3 \times 3$  matrix can have eigenvalues  $1, 2, i$ : take the diagonal matrix with entries  $1, 2, i$ . See the Caution on p. 215.)

## 5 Multiplicity (pp. 213–215)

A non-real eigenvalue of a real matrix and its conjugate are repeated the same number of times.

**Definition (multiplicity of a root, p. 213)**

If  $r$  is a root of  $f(x) = x^n + \cdots + a_1x + a_0$ , the **multiplicity** of  $r$  is the positive integer  $m$  with

$$f(x) = (x - r)^m g(x)$$

for some polynomial  $g$  with  $g(r) \neq 0$ .

So  $m$  counts how many factors  $(x - r)$  appear in the complete factorization given by the fundamental theorem of algebra. Example:  $f(x) = (x - 3)^2(x + 1)$  has the root 3 with multiplicity 2 (here  $g(x) = x + 1$  and  $g(3) = 4 \neq 0$ ) and  $-1$  with multiplicity 1.

**Definition (algebraic multiplicity, p. 213)**

If  $\det(A - \lambda I) = (\lambda_1 - \lambda)^{m_1} \cdots (\lambda_k - \lambda)^{m_k}$  with  $\lambda_1, \dots, \lambda_k$  distinct, then  $m_i$  is the **algebraic multiplicity** of  $\lambda_i$ .

Example: for  $A = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$  from the 10/15 notes (p. 194),  $\det(A - \lambda I) = (3 - \lambda)^2$ , so 3 has algebraic multiplicity 2. The multiplicities add up to  $n$ , because the polynomial has degree  $n$ . (The word “algebraic” distinguishes this from the *geometric* multiplicity  $\dim E_\lambda(A)$ , which for this  $A$  is only 1.)

**Theorem (p. 213)**

If  $A$  is a real  $n \times n$  matrix with a non-real eigenvalue  $\lambda$  of algebraic multiplicity  $m$ , then the algebraic multiplicity of  $\bar{\lambda}$  is also  $m$ .

Write  $f(x) = \det(A - xI)$ . It has real coefficients, because  $A$  is real.

**Proof (pp. 213–215), steps filled in**

Suppose the algebraic multiplicity of  $\bar{\lambda}$  is  $k$ . Arguing by contradiction, assume  $k \neq m$ . Say  $k < m$ . (If  $k > m$ , run the same argument with  $\lambda$  and  $\bar{\lambda}$  swapped:  $\bar{\lambda}$  is also non-real and its conjugate is  $\lambda$ .)

**Step 1: a real factor (p. 214).**  $\lambda$  and  $\bar{\lambda}$  are roots of  $f$ . Also

$$(\lambda - x)(\bar{\lambda} - x) = \lambda\bar{\lambda} - x(\lambda + \bar{\lambda}) + x^2$$

has real coefficients:  $\lambda\bar{\lambda} = a^2 + b^2$  and  $\lambda + \bar{\lambda} = 2a$  for  $\lambda = a + ib$ . Hence so does its  $k$ -th power  $h(x) = (\lambda - x)^k(\bar{\lambda} - x)^k$  (products of real polynomials are real).

**Step 2: factor it out.** Both  $\lambda$  (multiplicity  $m > k$ ) and  $\bar{\lambda}$  (multiplicity  $k$ ) occur at least  $k$  times among the linear factors of  $f$ , so

$$f(x) = g(x) (\lambda - x)^k (\bar{\lambda} - x)^k$$

for some polynomial  $g$  (namely the product of the remaining linear factors, up to sign).

**Step 3: key point,  $g$  is real.**  $g$  is the quotient of two real polynomials:

$$g(x) = \frac{f(x)}{(x - \lambda)^k (x - \bar{\lambda})^k}.$$

(The denominator equals  $h(x)$  because  $(\lambda - x)(\bar{\lambda} - x) = (x - \lambda)(x - \bar{\lambda})$ .) Long division of a real polynomial by a real polynomial only ever uses  $+$ ,  $-$ ,  $\times$ ,  $\div$  on real numbers, so the quotient is real; here the division is exact, and the quotient is  $g$ .

**Step 4:  $g(\lambda) = 0$ , because  $k < m$ .**  $f$  contains  $(x - \lambda)^m$  but  $h$  only  $(x - \lambda)^k$ , so  $m - k \geq 1$  factors  $(x - \lambda)$  are left in  $g$ .

**Step 5: so  $g(\bar{\lambda}) = 0$ , because non-real roots come in pairs.** The notes justify this (p. 215): “because by homework  $g$  is a real multiple of a characteristic polynomial of a real matrix.” Then the corollary on p. 212 applies:  $\lambda$  is an eigenvalue of that real matrix, so  $\bar{\lambda}$  is too. (The notes do not say which homework or which matrix. One real matrix that works is

the companion matrix of  $g$ , whose characteristic polynomial is  $\pm g$  divided by its leading coefficient.) A direct way to see it, without the homework: the coefficients  $c_j$  of  $g$  are real, so  $g(\bar{\lambda}) = \sum c_j \bar{\lambda}^j = \overline{\sum c_j \lambda^j} = \overline{g(\lambda)} = \bar{0} = 0$ .

**Step 6: contradiction.**  $g(\bar{\lambda}) = 0$  means  $g(x) = (x - \bar{\lambda})q(x)$  for some polynomial  $q$ , so  $f(x) = \pm(x - \bar{\lambda})^{k+1}(x - \lambda)^k q(x)$ . Thus  $\bar{\lambda}$  appears at least  $k + 1$  times in the factorization of  $f$ , so its multiplicity is at least  $k + 1 > k$ . This contradicts the choice of  $k$ . Hence  $k = m$ . ■

*Example.* The real  $4 \times 4$  matrix with two rotation blocks,  $\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$ , has  $\det(A - \lambda I) = (\lambda^2 + 1)^2 = (\lambda - i)^2(\lambda + i)^2$  (it is block diagonal, so the determinant is the product of the two  $2 \times 2$  blocks' determinants). Both  $i$  and  $-i$  have multiplicity 2, as the theorem predicts. ✓

**Q (p. 215):** is there a real  $4 \times 4$  matrix with eigenvalues  $i, i, -i, 1$ ? No

The notes: “No, since  $-i$  has to occur twice.” The list says  $i$  has algebraic multiplicity 2 and  $-i = \bar{i}$  has algebraic multiplicity 1. For a real matrix the theorem on p. 213 forces these to be equal. So no real matrix has these eigenvalues.

(The possible lists for a real  $4 \times 4$  matrix with  $i$  appearing twice are exactly  $i, i, -i, -i$ , as in the example above.)

**Caution (p. 215).** The theorem fails if the matrix is not real. Example:

$$A = \begin{bmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$A$  is diagonal, so  $\det(A - \lambda I) = (i - \lambda)^2(1 - \lambda)$  (10/15 notes, p. 195). As the notes say, it has eigenvalues  $i, i, 1$ :  $i$  has multiplicity 2, while  $\bar{i} = -i$  is not an eigenvalue at all (multiplicity 0). The proof breaks at Step 3:  $f$  no longer has real coefficients.

## 6 Determinant and trace from the eigenvalues (pp. 216–218)

The product of the eigenvalues is  $\det A$  and their sum is  $\text{trace } A$ ; in fact the eigenvalues determine the whole characteristic polynomial.

### Theorem (p. 216)

Let  $\lambda_1, \dots, \lambda_n$  be the (possibly non-distinct) eigenvalues of a real or complex  $n \times n$  matrix  $A$ , listed with algebraic multiplicity. Then

$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n.$$

**Proof (p. 216), completed**

By the fundamental theorem of algebra, the degree- $n$  polynomial  $\det(A - \lambda I)$ , whose leading term is  $(-\lambda)^n$  (p. 205), factors into linear factors. Its roots are the eigenvalues, so

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda).$$

(The sign is right: the right side also has leading term  $(-\lambda)^n$ .) The notes then say “Plug in  $\lambda = 0$ ”: this is an identity of polynomials, so we may put  $\lambda = 0$ :

$$\det A = \det(A - 0I) = (\lambda_1 - 0) \cdots (\lambda_n - 0) = \lambda_1 \cdots \lambda_n. \quad \blacksquare$$

*Check:* for  $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$  (p. 209),  $\lambda_+ \lambda_- = i(-i) = 1 = 0 \cdot 0 - 1 \cdot (-1) = \det A$ . ✓

**Consequence.**  $A$  is invertible  $\iff \det A \neq 0 \iff$  no eigenvalue is 0, matching the 10/15 notes (p. 196).

**Theorem (p. 216)**

If  $A = \begin{bmatrix} a_{11} & & * \\ & \ddots & \\ * & & a_{nn} \end{bmatrix}$  and  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of  $A$  (with multiplicity), then

$$\lambda_1 + \cdots + \lambda_n = a_{11} + \cdots + a_{nn} = \text{trace } A.$$

**The case  $n = 2$  (p. 216).** Before the proof, the notes multiply out two factors:

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda) = (-\lambda)^2 - \lambda(\lambda_1 + \lambda_2) + \lambda_1 \lambda_2.$$

The coefficient of  $-\lambda$  is the sum  $\lambda_1 + \lambda_2$ , and the constant term is the product  $\lambda_1 \lambda_2$ . Comparing with  $\det(A - \lambda I) = (-\lambda)^2 + (-\lambda)(a_{11} + a_{22}) + \det A$  (p. 207) gives both theorems for  $n = 2$ . *Check:* expanding the left side gives  $\lambda_1 \lambda_2 - \lambda_1 \lambda - \lambda_2 \lambda + \lambda^2$ , the same. ✓

*Proof (p. 217).* We saw before (p. 207) that

$$\det(A - \lambda I) = (-\lambda)^n + (-\lambda)^{n-1} \text{trace } A + \cdots + \det A,$$

and in the proof above that this equals  $(\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$ . The notes expand the product as

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda) = (-\lambda)^n + (-\lambda)^{n-1}(\lambda_1 + \lambda_2 + \cdots + \lambda_n) + \cdots,$$

underlining the  $\lambda_j$  and the  $-\lambda$ 's to show where the terms come from: a term with  $(-\lambda)^{n-1}$  comes from choosing  $-\lambda$  from all factors but one, and  $\lambda_j$  from the remaining factor  $j$ . So the coefficient of  $(-\lambda)^{n-1}$  is  $\lambda_1 + \cdots + \lambda_n$ . Comparing coefficients of  $(-\lambda)^{n-1}$  on both sides gives  $\lambda_1 + \cdots + \lambda_n = \text{trace } A$ . ■

*Check:* for the rotation,  $i + (-i) = 0 = 0 + 0 = \text{trace } A$  ✓; for  $\begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix}$  (10/15 notes, p. 198),  $-1 + 2 = 1 = 5 - 4$ . ✓

**Example (p. 217):  $n = 2$ , by the quadratic formula**

$$\det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - \lambda(a + d) + ad - bc.$$

By the theorems, the coefficient  $a + d$  should be  $\lambda_+ + \lambda_-$  and the constant  $ad - bc$  should be  $\lambda_+ \lambda_-$ . To confirm directly, the quadratic formula (roots of  $\lambda^2 + p\lambda + q$  are  $\frac{-p \pm \sqrt{p^2 - 4q}}{2}$ ) gives

$$\lambda_{\pm} = \frac{a + d \pm \sqrt{(a + d)^2 - 4(ad - bc)}}{2}.$$

Sum: the square roots cancel,  $\lambda_+ + \lambda_- = \frac{2(a+d)}{2} = a + d$ . ✓

**Fill-in (p. 217):**  $\lambda_+ \lambda_- = ad - bc$

Write  $\sqrt{\phantom{x}}$  for  $\sqrt{(a + d)^2 - 4(ad - bc)}$  and use  $(u + w)(u - w) = u^2 - w^2$ :

$$\begin{aligned} \lambda_+ \lambda_- &= \frac{(a + d + \sqrt{\phantom{x}})(a + d - \sqrt{\phantom{x}})}{2 \cdot 2} = \frac{(a + d)^2 - ((a + d)^2 - 4(ad - bc))}{4} \\ &= \frac{4(ad - bc)}{4} = ad - bc = \det A. \checkmark \end{aligned}$$

(This works even when the square root is of a negative number, i.e. for complex  $\lambda_{\pm}$ .)

**Q (p. 218):** do  $\lambda_1, \dots, \lambda_n$  determine all coefficients of  $\det(A - \lambda I)$ ? Yes

The notes answer “Yes”. By the proof on p. 216,  $\det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$ . Once the  $\lambda_i$  are known (with multiplicity), multiplying out this product gives every coefficient. Conversely the polynomial determines its roots. So knowing the eigenvalues with multiplicity is the same as knowing the characteristic polynomial.

The coefficient of  $(-\lambda)^{n-j}$  is the sum of all products of  $j$  different  $\lambda_i$ ’s:  $j = 1$  gives the sum (the trace),  $j = n$  the product (the determinant).

**Fill-in (p. 218,  $n = 3$ ):** the “?” is  $\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3$

Multiply out  $(\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda)$  by choosing one term from each factor. To get  $(-\lambda)^1$ , choose  $-\lambda$  from one factor and  $\lambda_i$  from the other two:

$$\det(A - \lambda I) = (-\lambda)^3 + (-\lambda)^2(\lambda_1 + \lambda_2 + \lambda_3) + (-\lambda)(\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3) + \lambda_1 \lambda_2 \lambda_3.$$

So “?” =  $\lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3$ , exactly as the notes fill it in.

Check:  $B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}$  is triangular with eigenvalues 1, 2, 3, and  $\det(B - \lambda I) = (1 - \lambda)(2 - \lambda)(3 - \lambda) = -\lambda^3 + 6\lambda^2 - 11\lambda + 6$ . The formula predicts  $(-\lambda)^3 + 6(-\lambda)^2 + 11(-\lambda) + 6$  with  $6 = 1 + 2 + 3$ ,  $11 = 2 + 3 + 6$ ,  $6 = 1 \cdot 2 \cdot 3$ . ✓

(Addition, not in the notes: in terms of the entries of  $A$ , this middle coefficient is the sum of the three  $2 \times 2$  determinants obtained by deleting row  $i$  and column  $i$ , for  $i = 1, 2, 3$ . For  $B$ :  $6 + 3 + 2 = 11$ . ✓)

## 7 Eigenvalues of a linear map $T : V \rightarrow V$ (pp. 219–220)

Eigenvalues make sense for any linear map from a space to itself, and the goal is to choose a basis in which the matrix of  $T$  is as simple as possible.

**Next objective (p. 219).** Let  $T : V \rightarrow V$  be linear and  $n = \dim V$ . Find a basis  $\beta$  of  $V$  in which the matrix  $A$  of  $T$  (9/29 notes, p. 140) is the simplest. (The answer, when it exists: a basis of eigenvectors, which makes  $A$  diagonal. See Section 9.)

### Definition (p. 219)

If  $T(v) = \lambda v$  with  $v \neq 0$ , then  $\lambda$  is an **eigenvalue** of  $T$  with **eigenvector**  $v$ .

**Ex (p. 219).**  $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $T(x) = Ax$  with  $A$  an  $n \times n$  matrix. Then  $T(v) = \lambda v$  says exactly  $Av = \lambda v$ , so the eigenvalues and eigenvectors of  $T$  and of  $A$  are the same.

**Ex (pp. 219–220): differentiation**  $T : \mathbb{R}_{\leq n}[x] \rightarrow \mathbb{R}_{\leq n}[x]$ ,  $T(f) = f'$

$T$  is linear, since  $(f + g)' = f' + g'$  and  $(cf)' = cf'$  (9/29 notes, p. 141). Here  $\mathbb{R}_{\leq n}[x]$  is the space of polynomials of degree at most  $n$ .

### Can you think of an eigenvalue of $T$ ? Yes: $\lambda = 0$

The notes: “ $T(1) = 0 = 0 \cdot 1$ .” More generally, every nonzero constant  $f = c$  has  $f' = 0 = 0 \cdot f$ . So 0 is an eigenvalue, and the nonzero constants are eigenvectors. The 0-eigenspace is  $\text{Ker } T = \{\text{constants}\}$ .

### Are there any eigenvectors $f$ with $\deg f > 0$ ? No

The notes (top of p. 220): “ $T(\text{degree } k > 0 \text{ polynomial } g) = \text{degree } k - 1 \text{ polynomial} \neq \lambda g$ .” In detail: differentiating lowers the degree, so if  $\deg g = k > 0$  then  $g'$  is a nonzero polynomial of degree exactly  $k - 1$ .

- If  $\lambda \neq 0$ , then  $\lambda g$  has degree  $k$ , but  $g'$  has degree  $k - 1$ . They cannot be equal.
- If  $\lambda = 0$ , then  $\lambda g = 0$ , but  $g' \neq 0$  (its leading coefficient is  $k$  times that of  $g$ ).

So the only eigenvectors are the nonzero constants, and the only eigenvalue is 0.

**The matrix of  $T$  in the basis  $\beta = \{1, x, \dots, x^n\}$  (p. 220).** The coordinate map is  $S(a_0 + a_1x + \dots + a_nx^n) = (a_0, a_1, \dots, a_n)^T \in \mathbb{R}^{n+1}$  (9/22 notes, p. 116), and the matrix  $A$  satisfies  $[T(f)]_\beta = A[f]_\beta$  (9/29 notes, p. 140).

**Fill-in:**  $T(a_0 + a_1x + \dots + a_kx^k + \dots + a_nx^n) = a_1 + 2a_2x + \dots + ka_kx^{k-1} + \dots + na_nx^{n-1}$

Differentiate term by term,  $(x^k)' = kx^{k-1}$ ; the notes write the result as  $0 + a_1 + \dots + ka_kx^{k-1} + \dots + na_nx^{n-1}$ , the 0 being the derivative of  $a_0$ . In coordinates, the coefficient of  $x^{k-1}$  in  $f'$  is

$ka_k$ , and the coefficient of  $x^n$  is 0:

$$A \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ \vdots \\ na_n \\ 0 \end{bmatrix}.$$

**Fill-in:**  $A$  is the  $(n+1) \times (n+1)$  matrix with  $1, 2, \dots, n$  just above the diagonal

Column  $k+1$  of  $A$  is  $[T(x^k)]_\beta = [kx^{k-1}]_\beta = k e_k$  (and column 1 is  $[T(1)]_\beta = 0$ ):

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & n \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

This is the matrix the notes fill in; they draw a red line along the main diagonal to show that it is all zeros, with  $1, 2, \dots, n$  just above it.

*Check* ( $n = 2$ , as in the 9/29 notes, p. 141):  $f = 1 + x + x^2$  has  $f' = 1 + 2x$ , and  $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} (1, 1, 1)^T = (1, 2, 0)^T$ , the coordinates of  $1 + 2x$ . ✓

*Consistency with the eigenvalue answers:*  $A$  is upper triangular with zero diagonal, so  $\det(A - \lambda I) = (-\lambda)^{n+1}$  and 0 is the only eigenvalue (10/15 notes, p. 196).  $A$  has  $n$  pivots, so  $\text{Nul } A$  is 1-dimensional, spanned by  $e_1 = [1]_\beta$ : the constants, as found above.

## 8 Similar matrices (pp. 221–222)

The matrices of the same linear map in two different bases are similar:  $A_1 = PA_2P^{-1}$ , where  $P$  is the change-of-basis matrix.

### Definition (p. 221)

Two  $n \times n$  matrices  $A, B$  are **similar** if  $A = PBP^{-1}$  for some invertible  $n \times n$  matrix  $P$ .

### Theorem (p. 221)

Let  $T : V \rightarrow V$  be linear and  $\dim V = n$ . Let  $\beta_1, \beta_2$  be two bases of  $V$ , and let  $S_1, S_2 : V \rightarrow \mathbb{R}^n$  be the coordinate maps,  $S_i(v) = [v]_{\beta_i}$ . Let  $A_i$  be the standard matrix of  $S_iTS_i^{-1}$  ( $i = 1, 2$ ). Then  $A_1$  and  $A_2$  are similar.

The two diagrams below are the 9/29 picture (p. 139) drawn once for each basis. Going right is “take coordinates”; the right column is the same map  $T$  seen in coordinates, so  $A_i$  is the matrix

of  $T$  relative to  $\beta_i$ .

$$\begin{array}{ccc} V & \xrightarrow{S_1} & \mathbb{R}^n \\ T \downarrow & & \downarrow S_1 T S_1^{-1} \\ V & \xrightarrow{S_1} & \mathbb{R}^n \end{array} \qquad \begin{array}{ccc} V & \xrightarrow{S_2} & \mathbb{R}^n \\ T \downarrow & & \downarrow S_2 T S_2^{-1} \\ V & \xrightarrow{S_2} & \mathbb{R}^n \end{array}$$

$S_i T S_i^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a composition of linear maps, hence linear, so it has a standard matrix  $A_i$ .

**Proof (p. 222).** We need  $A_1 = P A_2 P^{-1}$ , or in terms of maps,

$$S_1 T S_1^{-1} = P S_2 T S_2^{-1} P^{-1}.$$

**What is  $P$ ?**  $P =$  the standard matrix of  $S_1 S_2^{-1}$ , i.e.  $P = P_{\beta_1 \leftarrow \beta_2}$

The notes (p. 222) take “ $P = S_1 S_2^{-1}$ ,  $P^{-1} = S_2 S_1^{-1}$ ” and cancel:

$$S_1 S_2^{-1} S_2 T S_2^{-1} S_2 S_1^{-1} = S_1 T S_1^{-1} = A_1.$$

(In red they label the left side of the goal  $A_1 = S_1 T S_1^{-1}$  and the middle  $S_2 T S_2^{-1} = A_2$ , identifying each map with its standard matrix.) The idea: to turn the right side into the left side,  $P$  should undo  $S_2$  and apply  $S_1$  instead.  $P = S_1 S_2^{-1}$  is an invertible linear map  $\mathbb{R}^n \rightarrow \mathbb{R}^n$  with inverse  $S_2 S_1^{-1}$ , and the cancellation uses  $S_2^{-1} S_2 = \text{id}_V$  twice. Since the standard matrix of a composition is the product of the standard matrices,  $A_1 = P A_2 P^{-1}$ , with  $P$  invertible. ■

**What  $P$  looks like.** If  $\beta_2 = \{b_1, \dots, b_n\}$ , column  $j$  of  $P$  is  $S_1 S_2^{-1}(e_j) = S_1(b_j) = [b_j]_{\beta_1}$ . So  $P = [[b_1]_{\beta_1} \cdots [b_n]_{\beta_1}] = P_{\beta_1 \leftarrow \beta_2}$ , the change-of-basis matrix from the 9/29 notes (p. 144): it converts  $\beta_2$ -coordinates into  $\beta_1$ -coordinates.

*Check with  $V = \mathbb{R}^2$ ,  $T(x) = Ax$ ,  $A = \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix}$  (10/15 notes, p. 198). Take  $\beta_1$  the standard basis, so  $S_1 = \text{id}$  and  $A_1 = A$ . Take  $\beta_2 = \{v_1, v_2\}$  with the eigenvectors  $v_1 = (1, 1)$ ,  $v_2 = (2, 1)$ . Since  $T(v_1) = -v_1$  and  $T(v_2) = 2v_2$ , the matrix of  $T$  in  $\beta_2$  is  $A_2 = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$ . Here  $P = [[v_1]_{\beta_1} \ [v_2]_{\beta_1}] = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$  and  $P^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$  ( $2 \times 2$  inverse formula, 9/10 notes). Then*

$$P A_2 P^{-1} = \begin{bmatrix} -1 & 4 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix} = A_1. \checkmark$$

**Why this matters (addition).** Similar matrices have the same characteristic polynomial: since  $P I P^{-1} = I$ ,

$$\det(P B P^{-1} - \lambda I) = \det(P (B - \lambda I) P^{-1}) = \det(B - \lambda I)$$

by  $|P C P^{-1}| = |C|$  (10/8 notes, p. 174). So all matrices of  $T$  have the same eigenvalues, determinant and trace, and it makes sense to speak of “the characteristic polynomial of  $T$ ”. The next objective (p. 219) is then to pick, among all these similar matrices, the simplest one.

## 9 Diagonalizable matrices (p. 222)

$A$  is similar to a diagonal matrix exactly when there are enough eigenvectors to form a basis; the eigenvectors are the columns of  $P$  and the eigenvalues are the diagonal of  $D$ .

### Definition (p. 222)

A matrix  $A$  is **diagonalizable** if  $A = PDP^{-1}$  (with  $P$  invertible) where  $D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix}$  is a diagonal matrix.

In words:  $A$  is similar to a diagonal matrix. By Section 8, this says that in some basis the map  $x \mapsto Ax$  just scales each basis vector.

### Theorem (p. 222)

An  $n \times n$  matrix  $A$  is diagonalizable over  $\mathbb{R} \iff \mathbb{R}^n$  has a basis of eigenvectors of  $A$ .

Under “diagonalizable over  $\mathbb{R}$ ” the notes write “by def:  $A = PDP^{-1}$ ,  $P, D$  real”. So “over  $\mathbb{R}$ ” means that  $P$  and  $D$  are real; the eigenvectors in the basis are then real, and so are the eigenvalues.

The final notes stop here, after the statement of the theorem.

**The proof is in the 10/22 notes.** The proof of this theorem (p. 223), its recipe and its examples, including diagonalizing the rotation over  $\mathbb{C}$ , were covered in the next lecture. See the 10/22 notes, pp. 223–232.

## Summary

| Fact                                                     | Formula                                                                                       |
|----------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| Complex eigenpair                                        | $Av = \lambda v, v \in \mathbb{C}^n, v \neq 0, \lambda \in \mathbb{C}$                        |
| Rotation $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ | $\lambda = \pm i, v_{\pm} = (1, \pm i)$                                                       |
| Real $A, \lambda$ not real                               | $v$ not real; $A\bar{v} = \bar{\lambda}v$                                                     |
| Real $A$ : multiplicities                                | alg. mult. of $\lambda$ = alg. mult. of $\bar{\lambda}$                                       |
| Determinant                                              | $\det A = \lambda_1 \cdots \lambda_n$                                                         |
| Trace                                                    | $\text{trace } A = \lambda_1 + \cdots + \lambda_n$                                            |
| Characteristic polynomial                                | $\det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$                    |
| $d/dx$ on $\mathbb{R}_{\leq n}[x]$                       | only eigenvalue 0; eigenvectors = nonzero constants                                           |
| Similar                                                  | $A = PBP^{-1}, P$ invertible                                                                  |
| Matrices of $T$ in two bases                             | $A_1 = PA_2P^{-1}, P = P_{\beta_1 \leftarrow \beta_2}$                                        |
| Diagonalizable                                           | $A = PDP^{-1} \iff$ basis of eigenvectors; $P = [v_1 \cdots v_n], D = \text{diag}(\lambda_i)$ |

## Corrections to the handwritten notes

| Page | As written                                                                        | Should be                                                                         | Why                                      |
|------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|------------------------------------------|
| 210  | $\begin{bmatrix} z_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 1 \\ i \end{bmatrix}$ | $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 \\ i \end{bmatrix}$ | the unknown vector is $v_+ = (z_1, z_2)$ |
| 221  | "for some $n \times n$ matrix $P$ "                                               | some <i>invertible</i> $n \times n$ matrix $P$                                    | $P^{-1}$ must exist                      |

(The typo on p. 207,  $a_{n-1}x^n$  for  $a_{n-1}x^{n-1}$ , is listed in the 10/15 notes.)