Linear Algebra Notes — Lecture 10/27

Algebraic vs. Geometric Multiplicity, the Diagonalizability Test, and Eigenvalues of Maps

Lecture pages 237–256 · Lay §5.3 (Theorem 7), §5.4

Contents

Overview 1

1 Algebraic and geometric multiplicity (pp. 237–238) 2

2 Schur's theorem (pp. 238–239) 3

3 Proof that $1 \le g_i \le m_i$ (pp. 240–241) 5

4 The diagonalizability test (pp. 242–243) 6

5 Proof of the test (pp. 243–246) 7

6 Algorithm and a 4×4 example (pp. 246–248) 9

7 Eigenvalues of $f \mapsto f''$ on polynomials (pp. 249–250) 10

8 Eigenvalues of $A \mapsto AB$ on matrices (pp. 251–253) 11

9 $A \mapsto AB$ for diagonal B (pp. 254–256) 13

Summary 15

Corrections to the handwritten notes 15

Overview

The 10/22 lecture showed that A is diagonalizable exactly when there is a basis of eigenvectors. This lecture answers the practical question "*how do I decide* whether A is diagonalizable?" by counting:

1. each eigenvalue λ_i has an *algebraic* multiplicity m_i (how often it is a root of $\det(A - \lambda I)$) and a *geometric* multiplicity $g_i = \dim \mathrm{Nul}(A - \lambda_i I)$ (how many independent eigenvectors it has);

2. always $1 \le g_i \le m_i$, proved using Schur's theorem (every matrix is similar to an upper triangular one);
3. A is diagonalizable $\iff$ $g_i = m_i$ for every i $\iff$ $\sum g_i = n$;
4. an algorithm, a 4×4 example, and then eigenvalues of linear *maps*: $f \mapsto f''$ on polynomials and $A \mapsto AB$ on matrices.

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

Notation used throughout.

- $\lambda_1, \ldots, \lambda_k$ are the *distinct* eigenvalues of A; m_i and g_i are the algebraic and geometric multiplicities of λ_i.
- $E_\lambda(A) = \mathrm{Nul}(A - \lambda I)$ is the λ-eigenspace (10/15 notes), so $g_i = \dim E_{\lambda_i}(A)$.
- $M_{n\times n}(\mathbb{R})$ is the vector space of real $n \times n$ matrices (dimension n^2); E_{ij} is the matrix with a 1 in row i, column j and zeros elsewhere.
- $\overline{P}$ conjugates every entry of P (10/20 notes, p. 211).

Recipe (pp. 246–247). To decide whether A is diagonalizable and find P, D:

1. factor $\det(A - \lambda I) = (\lambda_1 - \lambda)^{m_1} \cdots (\lambda_k - \lambda)^{m_k}$;
2. for each i, compute $g_i = \dim \mathrm{Nul}(A - \lambda_i I)$ = number of free variables of $A - \lambda_i I$;
3. if $g_i = m_i$ for all i, a basis of each eigenspace, put together, is a basis of eigenvectors; otherwise (some $g_i < m_i$) A is not diagonalizable.

1 Algebraic and geometric multiplicity (pp. 237–238)

The algebraic multiplicity counts roots of the characteristic polynomial; the geometric multiplicity counts independent eigenvectors. They can differ.

Definition (recall, p. 237; 10/20 notes, p. 213)

The **algebraic multiplicity** of the eigenvalue λ_i of A is the integer $m_i \ge 1$ such that

$$g(\lambda) = \frac{\det(A - \lambda I)}{(\lambda - \lambda_i)^{m_i}} \text{ is a polynomial and } g(\lambda_i) \neq 0.$$

In words: $(\lambda - \lambda_i)$ divides $\det(A - \lambda I)$ exactly m_i times. (The g here is just the leftover polynomial; it has nothing to do with the g_i below.)

Definition (p. 237)

The **geometric multiplicity** of the eigenvalue λ_i of A is $g_i = \dim \mathrm{Nul}(A - \lambda_i I) = \dim E_{\lambda_i}(A)$.

By the Rank Theorem (Lay §2.9), $g_i = n - \mathrm{rank}(A - \lambda_i I)$, the number of free variables when you row reduce $A - \lambda_i I$.

Note (p. 237): $g_i \geq 1$. Why? λ_i has an eigenvector

λ_i is an eigenvalue, so by definition there is a $v \neq 0$ with $Av = \lambda_i v$, i.e. $(A - \lambda_i I)v = 0$. So $\text{Nul}(A - \lambda_i I)$ contains a nonzero vector, and a subspace containing a nonzero vector has dimension at least 1.

Theorem (p. 237)

$m_i \geq g_i$ for every eigenvalue λ_i.

The proof is on pp. 240–241 (Section 3).

Ex (p. 237): $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$, $\lambda_1 = 2$

Fill-in: $m_1 = 3$, $g_1 = 1$

The notes: $m_1 = 3$, $g_1 = 1$, and they write out $A - 2I$. A is upper triangular, so $\det(A - \lambda I) = (2 - \lambda)^3$ (the determinant of a triangular matrix is the product of its diagonal entries, 10/15 notes, p. 195). So 2 is the only eigenvalue and $m_1 = 3$.
$A - 2I = \left[\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{smallmatrix}\right]$ is already in echelon form, with pivots in columns 2 and 3. Only x_1 is free, so $E_2(A) = \text{span}\{e_1\}$ and $g_1 = 1$.
Check: $g_1 = 1 \leq 3 = m_1$, as the theorem says. ✓ **Meaning:** A has only one independent eigenvector, so there is no basis of eigenvectors and A is not diagonalizable (this is the case "$A - \lambda_1 I \neq 0$" on p. 243).

Ex (p. 238): A is $n \times n$ with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$

Fill-in: $m_i = 1$ and $g_i = 1$ for every i

The notes fill in $m_i = 1$, $g_i = 1$. Why: The multiplicities add up to $\deg \det(A - \lambda I) = n$ (10/20 notes, p. 213), and there are n eigenvalues each with $m_i \geq 1$, so every $m_i = 1$. Then $1 \leq g_i \leq m_i = 1$ (Note and Theorem on p. 237) forces $g_i = 1$.
This matches the 10/15 result that n distinct eigenvalues give a basis of eigenvectors (10/15 notes, p. 203): one eigenvector from each eigenspace.

2 Schur's theorem (pp. 238–239)

Every square matrix is similar to an upper triangular matrix R, and the diagonal of R lists the eigenvalues of A with multiplicity.

Theorem (Schur, p. 238; to be proved later)

Any real $n \times n$ matrix A is of the form $A = PRP^{-1}$ where $\overline{P}^T = P^{-1}$, R is upper triangular, and P, R are complex matrices. Also, if A has real eigenvalues, P and R can be chosen real (and then $P^T = P^{-1}$).

What the extra condition says. $\overline{P}^T = P^{-1}$ means P is invertible *and* its inverse is easy: conjugate and transpose. (The notes write $P^T = P^{-1}$; for a complex P the conjugate is needed, see the corrections.) For the multiplicity argument below, only "P invertible" is used, so A and R are similar.

Note (p. 239): $R = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ **where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. Why? Same characteristic polynomial**

The notes: "similar matrices A, R have the same eigenvalues." In detail, with $\lambda_1, \ldots, \lambda_n$ the diagonal entries of R (repeats allowed):

1. Similar matrices have the same characteristic polynomial (10/22 notes, p. 228): $\det(A - \lambda I) = \det\left(P(R - \lambda I)P^{-1}\right) = \det P \det(R - \lambda I) \det P^{-1} = \det(R - \lambda I)$.
2. $R - \lambda I$ is upper triangular with diagonal entries $\lambda_j - \lambda$, so $\det(R - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$ (10/15 notes, p. 195).

So $\det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$: the diagonal entries of R are exactly the eigenvalues of A, and each eigenvalue appears on the diagonal as many times as its algebraic multiplicity.

Bunching (p. 239). The notes go further: one can choose R so that equal eigenvalues are bunched together on the diagonal,

$$R = \begin{bmatrix} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{matrix}} & & & * \\ & \boxed{\begin{matrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{matrix}} & & \\ & & \ddots & \\ 0 & & & \boxed{\begin{matrix} \lambda_k & & * \\ & \ddots & \\ 0 & & \lambda_k \end{matrix}} \end{bmatrix},$$

where $\lambda_1, \ldots, \lambda_k$ are distinct and the block $\begin{bmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{bmatrix}$ has size $m_i \times m_i$. (The size is m_i by the answer box above: λ_i appears m_i times on the diagonal.) Everything below the diagonal is 0; the $*$ entries above it can be anything.

3 Proof that $1 \le g_i \le m_i$ (pp. 240–241)

In $R - \lambda_i I$ only the λ_i-block has zeros on its diagonal, so only that block can produce free variables, and it has m_i columns.

Proof (Schur $\Rightarrow m_i \ge g_i$), steps filled in

Step 1 (p. 240): move to R. With $A = PRP^{-1}$, and writing $\lambda_i I = P\lambda_i I P^{-1}$ as the notes do,

$$A - \lambda_i I = PRP^{-1} - \lambda_i PP^{-1} = P(R - \lambda_i I)P^{-1}.$$

Step 2: rank does not change. By homework, $\operatorname{rank}(A - \lambda_i I) = \operatorname{rank}(R - \lambda_i I)$. (Reason: multiplying by an invertible matrix does not change rank. On the right, $\operatorname{Col}(MP^{-1}) = \operatorname{Col} M$ because P^{-1} is onto. On the left, P is one-to-one, so it maps $\operatorname{Col} M$ isomorphically onto $\operatorname{Col}(PM)$.) By the Rank Theorem, $\operatorname{rank}(A - \lambda_i I) = n - \dim \operatorname{Nul}(A - \lambda_i I) = n - g_i$. So

$$g_i = n - \operatorname{rank}(R - \lambda_i I) = \dim \operatorname{Nul}(R - \lambda_i I).$$

Step 3: the shape of $R - \lambda_i I$. Subtracting λ_i from the diagonal of the bunched R gives an upper triangular matrix whose diagonal blocks have diagonal entries $\lambda_1 - \lambda_i, \ \dots, \ \lambda_i - \lambda_i = 0, \ \dots, \ \lambda_k - \lambda_i$. Since the λ_j are distinct, $\lambda_j - \lambda_i \neq 0$ for $j \neq i$. So every diagonal entry outside the λ_i-block is nonzero, and the λ_i-block $N = \begin{bmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{bmatrix}$ (size $m_i \times m_i$, p. 241) has zeros on its diagonal.

Step 4 (p. 241): g_i = number of free variables of this block. Write $R - \lambda_i I = \begin{bmatrix} B_1 & X & Y \\ 0 & N & Z \\ 0 & 0 & B_3 \end{bmatrix}$, where B_1 and B_3 are the parts before and after the λ_i-block. They are upper triangular with nonzero diagonals, hence invertible. Solve $(R - \lambda_i I)(x, y, z) = 0$ from the bottom up:

- $B_3 z = 0$ forces $z = 0$;
- then $Ny = 0$: the solutions y form $\operatorname{Nul} N$;
- then $B_1 x + Xy = 0$ forces $x = -B_1^{-1} Xy$, so x is determined by y.

So the solutions correspond exactly to $y \in \operatorname{Nul} N$, and $g_i = \dim \operatorname{Nul} N$ = the number of free variables of the block.

Step 5: the bounds. N has m_i columns, so it has at most m_i free variables: $g_i \le m_i$. Its first column is zero (the diagonal entry is 0 and everything below is 0), so x_1 is always free: $g_i \ge 1$. Hence

$$1 \le g_i \le m_i. \qquad \blacksquare$$

Ex (p. 241): $R = \begin{bmatrix} 1 & * & * & 2 & 3 \\ 0 & 1 & * & -1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$

The green lines in the notes split R into the 3×3 block for $\lambda = 1$ ($m = 3$) and the 2×2 block for $\lambda = 2$ ($m = 2$).

$R - 1I$ can have 1, 2 or 3 free variables, depending on the $*$'s

In $R - I$ the $\lambda = 2$ block becomes $\left[\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right]$, which is invertible, so by Step 4 only the block $N = \left[\begin{smallmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{smallmatrix}\right]$ (the three $*$'s called a, b, c) matters:

$*$ entries	$\operatorname{rank} N$	$g = \dim E_1(R)$
$a = b = c = 0$	0	3
exactly one of a, c nonzero, or $a = c = 0 \neq b$	1	2
$a \neq 0$ and $c \neq 0$	2	1

Check (NumPy, $\dim \operatorname{Nul}(R - I)$): $(a, b, c) = (0, 0, 0) \mapsto 3$, $(1, 0, 0) \mapsto 2$, $(0, 1, 0) \mapsto 2$, $(1, 0, 1) \mapsto 1$, $(1, 5, 1) \mapsto 1$. ✓

Side remark (addition). For $\lambda = 2$: $R - 2I$ has the block $\left[\begin{smallmatrix} 0 & -1 \\ 0 & 0 \end{smallmatrix}\right]$, with one free variable, so $g = 1 < 2 = m$ no matter what the $*$'s are. By the theorem on p. 242, this R is never diagonalizable.

4 The diagonalizability test (pp. 242–243)

A real matrix whose eigenvalues are all real is diagonalizable exactly when no eigenvalue is "short" of eigenvectors.

Theorem (p. 242; Lay §5.3, Theorem 7)

Let A be a real $n \times n$ matrix with $\det(A - \lambda I) = (\lambda_1 - \lambda)^{m_1} \cdots (\lambda_k - \lambda)^{m_k}$, where $\lambda_1, \ldots, \lambda_k$ are real and distinct. Then the following are equivalent:

① $\mathbb{R}^n$ has a basis of eigenvectors of A;

② $g_i = m_i$ for each λ_i;

③ $\sum_{i=1}^{k} g_i = n$.

Complex version (p. 242). The same theorem holds with "real" replaced by "complex" and $\mathbb{R}^n$ by $\mathbb{C}^n$. Over $\mathbb{C}$ the hypothesis is automatic: by the fundamental theorem of algebra every characteristic polynomial factors into linear factors (10/20 notes, p. 213).

Ex (pp. 242–243): $k = 1$, $n = m_1$, $A = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{bmatrix}$

Here A is upper triangular with the single eigenvalue λ_1 on its diagonal, so $m_1 = n$, and

$$g_1 = \dim \mathrm{Nul}(A - \lambda_1 I) = \dim \mathrm{Nul} \begin{bmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{bmatrix}.$$

Fill-in (p. 243): A is diagonalizable $\iff$ $A - \lambda_1 I = 0$, i.e. $A = \lambda_1 I$

The notes: $A - \lambda_1 I = 0$ (with a check mark at the $*$ of $A - \lambda_1 I$: the $*$'s must all be 0). By the theorem, A is diagonalizable $\iff$ $g_1 = m_1 = n$ $\iff$ $\mathrm{Nul}(A - \lambda_1 I)$ is all of $\mathbb{R}^n$ $\iff$ $(A - \lambda_1 I)x = 0$ for every x $\iff$ $A - \lambda_1 I = 0$ (apply it to $e_1, \ldots, e_n$: every column is 0).
Meaning: a triangular matrix with one repeated eigenvalue is diagonalizable only if it is already diagonal, and then it is $\lambda_1 I$. Any nonzero $*$ ruins it. Example: the 3×3 matrix on p. 237 has $*$'s equal to 1 and is not diagonalizable ($g_1 = 1 < 3$). ✓

5 Proof of the test (pp. 243–246)

The proof goes around the circle ②⇒③⇒①⇒②.

②⇒③ (p. 243), the "??": the m_i add up to n

The notes answer with $\sum_{i=1}^{k} m_i = n$. In detail: if $g_i = m_i$ for all i, then $\sum_{i=1}^{k} g_i = \sum_{i=1}^{k} m_i$. The right side is n: by hypothesis $\det(A - \lambda I) = (\lambda_1 - \lambda)^{m_1} \cdots (\lambda_k - \lambda)^{m_k}$, and comparing degrees, $m_1 + \cdots + m_k = \deg \det(A - \lambda I) = n$ (10/20 notes, p. 213). So $\sum g_i = n$.

③⇒① (p. 243). Suppose $\sum_{i=1}^{k} g_i = n$. Let $\mathcal{B}_i$ be a basis of $E_{\lambda_i}(A)$; it has g_i vectors. Then $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ is a set of $\sum g_i = n$ eigenvectors. (The $\mathcal{B}_i$ don't overlap: a vector in two eigenspaces would satisfy $\lambda_i v = Av = \lambda_j v$, so $v = 0$, but basis vectors are nonzero.)

Why is $\mathcal{B}$ a basis? It is linearly independent (proof by example, p. 244)

$\mathcal{B}$ has n vectors in $\mathbb{R}^n$, so it is enough to show it is linearly independent (a set of n independent vectors in $\mathbb{R}^n$ is a basis, Lay §4.5). The notes prove this in a typical case.
Ex (p. 244): $k = 2$, $n = 4$, $g_1 = g_2 = 2$, $\mathcal{B}_1 = \{v_1, v_2\}$, $\mathcal{B}_2 = \{v_3, v_4\}$. Suppose

$$\underbrace{c_1 v_1 + c_2 v_2}_{w_1 \in E_{\lambda_1}(A)} + \underbrace{c_3 v_3 + c_4 v_4}_{w_2 \in E_{\lambda_2}(A)} = 0.$$

- w_1 and w_2 lie in eigenspaces (eigenspaces are subspaces, so linear combinations stay inside).
- Since $\lambda_1 \neq \lambda_2$, the vectors w_1 and w_2 are independent unless one of them is zero. (If both are nonzero they are eigenvectors for distinct eigenvalues, and those are independent, 10/15 notes, p. 203.)

- But $w_1 + w_2 = 0$ says $w_1 = -w_2$: they are dependent. So one of them is zero, and then the other is too. In either case $c_1v_1 + c_2v_2 = 0 = c_3v_3 + c_4v_4$.

So $c_i = 0$ for all i (the notes' last line on p. 244): $c_1 = c_2 = 0$ because $\{v_1, v_2\}$ is a basis of $E_{\lambda_1}(A)$, hence independent, and $c_3 = c_4 = 0$ because $\{v_3, v_4\}$ is independent. All the c's are 0, so $\{v_1, v_2, v_3, v_4\}$ is independent, hence a basis of $\mathbb{R}^4$. ✓
In general: group the combination as $w_1 + \cdots + w_k = 0$ with $w_i \in E_{\lambda_i}(A)$. The nonzero w_i's would be eigenvectors for distinct eigenvalues satisfying a dependence relation, which is impossible; so every $w_i = 0$, and then every coefficient is 0 because each $\mathcal{B}_i$ is independent.

①⇒② **(p. 245).** Suppose A has a basis of eigenvectors $v_1, \ldots, v_n$, $Av_i = \lambda_i v_i$ (here the λ's are listed with repeats). By the 10/22 diagonalization theorem, $A = PDP^{-1}$ with $P = [v_1 \ \cdots \ v_n]$ and D diagonal with the matching eigenvalues on its diagonal. Group equal eigenvalues together, so $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_1, \ \ldots, \ \lambda_k, \ldots, \lambda_k)$.

$\dim E_{\lambda_i}(D) = m_i$ (p. 245)

How many times does λ_i appear on D's diagonal? A and D are similar, so $\det(A - \lambda I) = \det(D - \lambda I) = \prod_j (d_{jj} - \lambda)$ (10/22 notes, p. 228). So λ_i appears exactly m_i times.
Then: $D - \lambda_i I$ is diagonal with exactly m_i zeros on the diagonal and nonzero entries elsewhere. The zero diagonal entries give zero columns, i.e. m_i free variables, so $\dim E_{\lambda_i}(D) = \dim \operatorname{Nul}(D - \lambda_i I) = m_i$. (Concretely, $E_{\lambda_i}(D)$ is spanned by the e_j with $d_{jj} = \lambda_i$.)

The notes then show that P is a linear isomorphism from $E_{\lambda_i}(D)$ to $E_{\lambda_i}(A)$. If true, then $g_i = \dim E_{\lambda_i}(A) = \dim E_{\lambda_i}(D) = m_i$, as desired (isomorphic spaces have the same dimension).

p. 246: $x \in E_{\lambda_i}(D) \Rightarrow Px \in E_{\lambda_i}(A)$. Why? $(A - \lambda_i I)Px = 0$

Using $AP = PD$ (from $A = PDP^{-1}$):

$$(A - \lambda_i I)Px = APx - \lambda_i Px = PDx - P\lambda_i x = P(D - \lambda_i I)x = P\,0 = 0.$$

(The notes fill in the last two steps, $= P0 = 0$.) The last step uses $x \in E_{\lambda_i}(D)$, i.e. $(D - \lambda_i I)x = 0$.
"Similarly": if $y \in E_{\lambda_i}(A)$ then, using $DP^{-1} = P^{-1}A$,

$$(D - \lambda_i I)P^{-1}y = P^{-1}Ay - \lambda_i P^{-1}y = P^{-1}(A - \lambda_i I)y = 0,$$

so $P^{-1}y \in E_{\lambda_i}(D)$.
Why this makes P an isomorphism between the eigenspaces: $x \mapsto Px$ is linear and maps $E_{\lambda_i}(D)$ into $E_{\lambda_i}(A)$; $y \mapsto P^{-1}y$ maps back; the two are inverse to each other. A linear map with an inverse is an isomorphism. ■

Thus ① ⟺ ② ⟺ ③ (p. 246).

6 Algorithm and a 4×4 example (pp. 246–248)

The algorithm is the theorem read as a procedure: compute every m_i and g_i, compare, and if they match, collect eigenspace bases.

Algorithm (finding a basis of eigenvectors if possible, pp. 246–247)

Step 1. Find the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ and their algebraic multiplicities m_i.
Step 2. Compute $g_i = \dim \operatorname{Nul}(A - \lambda_i I)$.
Step 3. If $m_i = g_i$ for each i, find a basis $\mathcal{B}_i$ of each $E_{\lambda_i}(A)$. Then $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ is a basis of eigenvectors. (If some $g_i < m_i$, stop: there is no basis of eigenvectors.)

Ex (pp. 247–248): $A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

Fill-in: $\lambda_1 = 2$, $m_1 = 2$; $\lambda_2 = 1$, $m_2 = 2$

A is lower triangular, so $\det(A - \lambda I) = (2-\lambda)^2(1-\lambda)^2$ (product of the diagonal entries, 10/15 notes, p. 195). *Check:* NumPy's characteristic polynomial is $\lambda^4 - 6\lambda^3 + 13\lambda^2 - 12\lambda + 4 = (\lambda-2)^2(\lambda-1)^2$. ✓

Fill-in: $g_1 = 2$, $g_2 = 2$

$\lambda_1 = 2$: $A - 2I = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}$ has two nonzero rows, and they are independent (pivots in columns 1 and 3 after swapping rows), so its rank is 2 and $g_1 = 4 - 2 = 2$.
$\lambda_2 = 1$: $A - I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$. Row 4 equals row 1, so the rank is 2 (pivots in columns 1, 2) and $g_2 = 4 - 2 = 2$.
Since $g_1 = m_1$ and $g_2 = m_2$, A is diagonalizable (Step 3).

Fill-in: bases $E_2(A) = \operatorname{span}\{(0,1,0,0), (1,0,0,1)\}$, $E_1(A) = \operatorname{span}\{e_3, e_4\}$

$E_2(A)$: $(A - 2I)x = 0$ says $x_1 = x_4$ and $-x_3 = 0$ (as in the notes). So $x_3 = 0$, $x_1 = x_4$, with x_2, x_4 free:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_4 \\ x_2 \\ 0 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

$E_1(A)$ **(p. 248):** $(A - I)x = 0$ says $x_1 = 0$, $x_2 = 0$ and (row 4) $x_1 = 0$ again, with x_3, x_4 free: $x = x_3 e_3 + x_4 e_4$.
Check: $A(1,0,0,1)^T = (2,0,0,2)^T$, $Ae_2 = 2e_2$, $Ae_3 = e_3$, $Ae_4 = e_4$. ✓

Fill-in (p. 248): $P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$ **with** $D = \operatorname{diag}(2, 2, 1, 1)$

Put the basis vectors in as columns, in the same order as the eigenvalues on D's diagonal: the two $\lambda = 2$ eigenvectors first, then the two $\lambda = 1$ eigenvectors. Then $A = PDP^{-1}$, with

$$P^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}.$$

(This is the notes' P. Any order of the basis vectors works, as long as D uses the same order.) *Check:* $AP = PD$ column by column is exactly the four eigenvector equations above, and NumPy confirms $PDP^{-1} = A$. ✓

7 Eigenvalues of $f \mapsto f''$ on polynomials (pp. 249–250)

Differentiating twice lowers the degree, so the only eigenvalue is 0, and its eigenvectors are the polynomials of degree at most 1.

Let $T : \mathbb{R}_{\leq n}[x] \to \mathbb{R}_{\leq n}[x]$, $T(f) = f''$. **Recall:** $T(f) = \lambda f$ with $f \neq 0$ means f is an eigenvector of T for the eigenvalue λ (Lay §5.4).

What are the eigenvalues of T? Only $\lambda = 0$

The notes: $\lambda = 0$, $T(f) = 0$. Why no other λ: suppose $f'' = \lambda f$ with $f \neq 0$, and say $\deg f = d$.

- If $\lambda \neq 0$: the right side λf has degree d, but f'' has degree $d - 2$ (or is 0 if $d \leq 1$). A polynomial can't equal one of a different degree, so this is impossible.
- If $\lambda = 0$: we need $f'' = 0$, and $f = 1$ works. So 0 is an eigenvalue.

Check: $T(1) = 0 = 0 \cdot 1$. ✓

What are the eigenvectors of T? The nonzero $f = a + bx$

The notes: $f(x) = a_0 + a_1 x$, and $E_0(T)$ has basis $\{1, x\}$. The eigenvectors are the nonzero solutions of $f'' = 0$. Write $f = a_0 + a_1 x + \cdots + a_n x^n$. Then $f'' = 2a_2 + 6a_3 x + \cdots + n(n-1)a_n x^{n-2}$, which is 0 exactly when $a_2 = \cdots = a_n = 0$. So the eigenvectors are the nonzero polynomials $a + bx$ (for $n \geq 1$), and $E_0(T) = \mathbb{R}_{\leq 1}[x] = \operatorname{span}\{1, x\}$ has dimension 2.

Does T have a basis of eigenvectors? No for $n \geq 2$; yes for $n \leq 1$

The notes: "Yes if $n \leq 1$, No if $n > 1$." Every eigenvector lies in $E_0(T)$, which has dimension 2. A basis of $\mathbb{R}_{\leq n}[x]$ needs $n + 1$ vectors, so for $n \geq 2$ there is no basis of eigenvectors. (For $n \leq 1$, T is the zero map and every basis works, e.g. $\{1, x\}$.)

In the language of p. 242 (using the matrix below): $m = n+1$ but $g = 2$, so $g < m$ when $n \geq 2$.

Fill-in (p. 250): the matrix of T in $\mathcal{B} = \{1, x, \ldots, x^n\}$

The notes write the matrix and the rule $(x^k)'' = k(k-1)x^{k-2}$ (e.g. $(x^3)'' = 6x$). Column j is $[T(x^j)]_{\mathcal{B}}$ (the matrix of a linear map in a basis, Lay §5.4), and $T(x^j) = j(j-1)x^{j-2}$:

f	1	x	x^2	x^3	$\cdots$	x^n
$T(f)$	0	0	2	$6x$	$\cdots$	$n(n-1)x^{n-2}$

So A has the single nonzero entry $j(j-1)$ in each column x^j ($j \geq 2$), two rows above the diagonal:

$$A = \begin{bmatrix} 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 6 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n(n-1) \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{(n+1)\times(n+1)}.$$

For $n = 3$: $A = \left[\begin{smallmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix}\right]$.
A is upper triangular with zero diagonal, so $\det(A - \lambda I) = (-\lambda)^{n+1}$: the only eigenvalue is 0, with $m = n+1$. The first two columns are zero and the rest have pivots, so $g = \dim \mathrm{Nul}\, A = 2$. *Check* ($n = 3$, NumPy): $\dim \mathrm{Nul}\, A = 2$. ✓

8 Eigenvalues of $A \mapsto AB$ on matrices (pp. 251–253)

Right-multiplication by B has the same eigenvalues as B.

Set-up (p. 251). Fix a real $n \times n$ matrix B and let $T : M_{n\times n}(\mathbb{R}) \to M_{n\times n}(\mathbb{R})$, $T(A) = AB$. It is linear (bottom of p. 250): $(c_1A_1 + c_2A_2)B = c_1A_1B + c_2A_2B$. Its eigenvectors are nonzero *matrices* A with $AB = \lambda A$.

Fill-in: $B = 0 \Rightarrow T = 0$ (the zero map); $B = I \Rightarrow T(A) = A$ (the identity map)

$T(A) = A \cdot 0 = 0$ for every A, and $T(A) = AI = A$ for every A.

What are the eigenvalues of T? For $B = 0$: only 0; for $B = I$: only 1; in general: those of B

For $B = 0$, $T(A) = 0 = 0 \cdot A$, so every nonzero A is an eigenvector with eigenvalue 0, and no other λ works ($0 = \lambda A$ with $A \neq 0$ forces $\lambda = 0$). For $B = I$, similarly, only 1. In both cases these are the eigenvalues of B, which is the theorem below.

Theorem (p. 251)

λ is an eigenvalue of $B \iff \lambda$ is an eigenvalue of T.

Proof of "⇐" (p. 251), including the "Why?"

Let λ be an eigenvalue of T: there is an $A \neq 0$ with

$$AB = T(A) = \lambda A = A(\lambda I).$$

So $A(B - \lambda I) = 0$, with $A \neq 0$.
Then $B - \lambda I$ is singular. Why? The notes: "If not, $A(B - \lambda I)(B - \lambda I)^{-1} = 0(B - \lambda I)^{-1} = 0$, but $A \neq 0$." That is: if $B - \lambda I$ were invertible, multiply on the right by its inverse: $A = A(B - \lambda I)(B - \lambda I)^{-1} = 0 \cdot (B - \lambda I)^{-1} = 0$, contradicting $A \neq 0$. So $B - \lambda I$ is singular, i.e. $\det(B - \lambda I) = 0$, and λ is an eigenvalue of B. ■

Proof of "⇒" (p. 252). Suppose λ is an eigenvalue of B. So $B - \lambda I$ is singular, so $\mathrm{Col}(B - \lambda I) \neq \mathbb{R}^n$ (Invertible Matrix Theorem, Lay §2.3). Then (by the Lemma below) there is a nonzero $n \times n$ matrix A with $\mathrm{Nul}\, A \supseteq \mathrm{Col}(B - \lambda I)$. If true, then for any x, the vector $(B - \lambda I)x$ lies in $\mathrm{Col}(B - \lambda I) \subseteq \mathrm{Nul}\, A$, so

$$A(B - \lambda I)x = 0 \text{ for any } x.$$

So $A(B - \lambda I) = 0$, i.e. $AB = \lambda A$ with $A \neq 0$: λ is an eigenvalue of T with eigenvector A. ■

Lemma (p. 252)

For any subspace $W \subsetneq \mathbb{R}^n$ (a *proper* subspace, not all of $\mathbb{R}^n$) there is a nonzero $n \times n$ matrix A with $W \subseteq \mathrm{Nul}\, A$.

Proof (pp. 252–253). Let $v_1, \ldots, v_k$ be a basis of W; $k < n$ because $W \neq \mathbb{R}^n$. Extend it to a basis $v_1, \ldots, v_k, v_{k+1}, \ldots, v_n$ of $\mathbb{R}^n$ (Lay §4.5). Define $S : \mathbb{R}^n \to \mathbb{R}^n$ on the basis by

$$S(v_i) = \begin{cases} 0, & i \leq k, \\ v_i, & i > k, \end{cases}$$

and extend linearly (a linear map is determined by its values on a basis).

Fill-in (p. 253): S sends $c_1v_1 + \cdots + c_nv_n$ to $c_{k+1}v_{k+1} + \cdots + c_nv_n$

By linearity, $S(c_1v_1 + \cdots + c_nv_n) = c_1S(v_1) + \cdots + c_nS(v_n) = 0 + \cdots + 0 + c_{k+1}v_{k+1} + \cdots + c_nv_n$. In words, S keeps the part outside W and deletes the part in W.

Then take A = the standard matrix of S, that is, $Ax = S(x)$ for all x (Lay §1.9). Then $A \neq 0$ since $Av_n = v_n \neq 0$ (here $n > k$), and A sends W to 0 because it sends each basis vector $v_1, \ldots, v_k$ of W to 0. ■

Q (p. 253): is A diagonalizable? Yes: $v_1, \ldots, v_n$ is a basis of eigenvectors

The notes: "Yes, $\{v_i\}$ basis of eigenvectors." Indeed $Av_i = 0 = 0 \cdot v_i$ for $i \leq k$ and $Av_i = v_i =$

$1 \cdot v_i$ for $i > k$. So $v_1, \ldots, v_n$ is a basis of eigenvectors, and

$$A = PDP^{-1}, \qquad P = [v_1 \ \cdots \ v_n], \quad D = \operatorname{diag}(\underbrace{0, \ldots, 0}_{k}, \underbrace{1, \ldots, 1}_{n-k}).$$

Meaning: A is a *projection* onto $\operatorname{span}\{v_{k+1}, \ldots, v_n\}$ along W: it satisfies $A^2 = A$. *Check:* $D^2 = D$, so $A^2 = PD^2P^{-1} = A$. ✓

Ex (p. 253): $T(A) = AB$ with $B^k = 0$ for some $k > 0$, but $B \neq 0$

Fill-in: $T^k = 0$ (the zero map)

The notes: $T^k(A) = AB^k = A0 = 0$. (Each application of T multiplies by one more B: $T^2(A) = T(AB) = (AB)B = AB^2$, and so on.)

Does T have a basis of eigenvectors? No

The notes: "no, since $B \neq 0$," with the argument at the top of p. 254. **Only eigenvalue 0 (p. 254).** If $T(A) = \lambda A$ with $A \neq 0$, then $T^k(A) = \lambda^k A$; but $T^k(A) = 0$ and $A \neq 0$, so $\lambda^k = 0$ and $\lambda = 0$. So every eigenvector satisfies $T(A) = 0A = 0$. (Equivalently, by the theorem on p. 251: if $Bv = \lambda v$, $v \neq 0$, then $0 = B^k v = \lambda^k v$, so $\lambda = 0$.)
So a basis of eigenvectors would make $T = 0$. If $A_1, \ldots, A_{n^2}$ were a basis of eigenvectors, each would satisfy $T(A_j) = 0$, so by linearity $T(A) = 0$ for every A. Take $A = I$: $T(I) = B = 0$, contradicting $B \neq 0$.
Check ($n = 2$, $B = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$, NumPy on the 4×4 matrix of T): all eigenvalues 0, $\dim E_0(T) = 2 < 4$, and $T^2 = 0$. ✓

9 $A \mapsto AB$ for diagonal B (pp. 254–256)

If B is diagonal with distinct entries, $T(A) = AB$ just scales column j of A by d_j, so the matrix units E_{ij} form a basis of eigenvectors.

Ex (p. 254). B is $n \times n$ diagonal with distinct eigenvalues $d_1, \ldots, d_n$, that is,

$$B = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix} = [d_1 e_1 \ \cdots \ d_n e_n],$$

and $T(A) = AB$. Is T diagonalizable?

The d_j-eigenspace of T (pp. 254–255). The notes: $E_{d_j}(T) = \{A \ (n \times n) \mid AB = d_j A\}$. It consists of the $A = [a_1 \ \cdots \ a_n]$ (a_i = column i) solving

$$\begin{aligned} d_j[a_1 \ \cdots \ a_n] = d_j A = T(A) = AB &= A[d_1 e_1 \ \cdots \ d_n e_n] \\ &= [d_1 A e_1 \ \cdots \ d_n A e_n] = [d_1 a_1 \ \cdots \ d_n a_n]. \end{aligned}$$

(Column-by-column multiplication; $Ae_i = a_i$.) Comparing column i: $d_i a_i = d_j a_i$. If $i \neq j$ then $d_i \neq d_j$, so $a_i = 0$. Thus $A = [a_1 \ \cdots \ a_n]$ solving $T(A) = d_j A$ satisfies $a_{i \neq j} = 0$, with a_j arbitrary: the eigenspace is the set of matrices whose only nonzero column is column j.

What is the dimension of the d_j-eigenspace? n; the geometric multiplicity of each d_j is n

A matrix whose only nonzero column is column j is determined by that column, a vector in $\mathbb{R}^n$. A basis is $E_{1j}, E_{2j}, \ldots, E_{nj}$ (a single 1 in column j, row i). So $\dim E_{d_j}(T) = n$ for each j.

Does T have a basis of eigenvectors? Yes. Sum of geometric multiplicities: $n \cdot n = n^2$

The notes: "Yes," and $n + \cdots + n$ (n terms) $= n^2 =$ size of the domain of T. That is, there are n eigenvalues $d_1, \ldots, d_n$, each with geometric multiplicity n, so $\sum g_j = n^2 = \dim M_{n\times n}(\mathbb{R})$. By ③⇒① (the proof works in any vector space; here $\mathbb{R}^n$ is replaced by $M_{n\times n}(\mathbb{R})$), the union of the eigenspace bases, i.e. all n^2 matrix units E_{ij}, is a basis of eigenvectors. So T is diagonalizable, with $T(E_{ij}) = d_j E_{ij}$.
(Directly: $\{E_{ij}\}$ is the standard basis of $M_{n\times n}(\mathbb{R})$, so no theorem is even needed.)

Ex (p. 256): $B = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$, $T(A) = AB$, A is 2×2

The notes give the basis of eigenvectors

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \text{ for } \lambda_1 = -1, \qquad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ for } \lambda_2 = 3.$$

These are E_{11}, E_{21} (column 1) and E_{12}, E_{22} (column 2), as the general example predicts: T multiplies column 1 by -1 and column 2 by 3.

Check: $\left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} -1 & 0 \\ 0 & 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 & 0 \\ -1 & 0 \end{smallmatrix}\right] = -1 \cdot \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} -1 & 0 \\ 0 & 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 & 3 \\ 0 & 0 \end{smallmatrix}\right] = 3\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$. NumPy on the 4×4 matrix of T: eigenvalues $-1, -1, 3, 3$. ✓

Summary

Fact	Formula
Algebraic multiplicity m_i	power of $(\lambda_i - \lambda)$ in $\det(A - \lambda I)$; $\sum m_i = n$ (if it factors)
Geometric multiplicity g_i	$\dim \mathrm{Nul}(A - \lambda_i I)$ = free variables of $A - \lambda_i I$
Bounds	$1 \leq g_i \leq m_i$ (proof via Schur)
Schur	$A = PRP^{-1}$, R upper triangular, $\mathrm{diag}(R)$ = eigenvalues
Diagonalizability test	basis of eigenvectors $\iff$ all $g_i = m_i$ $\iff$ $\sum g_i = n$
Triangular, one eigenvalue	diagonalizable $\iff$ $A = \lambda_1 I$
$f \mapsto f''$ on $\mathbb{R}_{\leq n}[x]$	only eigenvalue 0, $E_0 = \mathbb{R}_{\leq 1}[x]$; not diagonalizable for $n \geq 2$
$A \mapsto AB$	same eigenvalues as B; B diagonal, distinct: basis $\{E_{ij}\}$

Corrections to the handwritten notes

Page	As written	Should be	Why
238	$P^T = P^{-1}$ (P complex)	$\overline{P}^T = P^{-1}$	for complex P the inverse is the conjugate transpose; $P^T = P^{-1}$ is right when P is real
243	$\sum_{n=1}^{k} g_i = n$	$\sum_{i=1}^{k} g_i = n$	the summation index is i