

Linear Algebra Notes — Lecture 10/15

Eigenvalues, Eigenvectors and the Characteristic Polynomial

Lecture pages 193–207 · Lay §5.1–5.2

Contents

Overview	1
1 Eigenvectors and eigenspaces (pp. 193–194, 196)	2
2 Examples (pp. 195–200)	3
3 Distinct eigenvalues give independent eigenvectors (pp. 201–202)	6
4 n distinct eigenvalues: a basis of eigenvectors (pp. 203–204)	7
5 The characteristic polynomial (pp. 205–207)	8
Corrections to the handwritten notes	9

Overview

Eigenvalues first appeared on 9/29 (solving $X_{k+1} = AX_k$) and 10/8 (finding them with $\det(A - \lambda I) = 0$). This lecture makes them systematic:

1. eigenvectors for λ are the nonzero vectors of $\text{Nul}(A - \lambda I)$, the *eigenspace*;
2. a recipe: find λ from $\det(A - \lambda I) = 0$, then find a basis of each eigenspace;
3. eigenvectors with *different* eigenvalues are independent, so n distinct eigenvalues give a basis of $\mathbb{R}^n$, and then $A^k X_0$ is easy;
4. the characteristic polynomial $\det(A - \lambda I)$: its degree, its top and bottom coefficients, and the fundamental theorem of algebra.

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

The recipe in one place.

Step 1. Compute $\det(A - \lambda I)$ and solve $\det(A - \lambda I) = 0$. The roots are the eigenvalues.

Step 2. For each eigenvalue λ , row reduce $A - \lambda I$ and find a basis of $\text{Nul}(A - \lambda I)$. Its nonzero vectors are the eigenvectors.

Step 3. Check: $Av = \lambda v$ for each basis vector.

1 Eigenvectors and eigenspaces (pp. 193–194, 196)

Definition (recall)

If A is an $n \times n$ real matrix and $Av = \lambda v$ for some **nonzero** $v \in \mathbb{R}^n$ and some real number λ , then λ is an **eigenvalue** of A and v is an **eigenvector** for A, λ .

Moving $\lambda v = \lambda Iv$ to the left side:

$$Av = \lambda v \iff Av - \lambda Iv = 0 \iff (A - \lambda I)v = 0.$$

Theorem

The eigenvectors for λ, A are precisely the *nonzero* vectors in $\text{Nul}(A - \lambda I)$.

Why exclude $v = 0$?

$A \cdot 0 = 0 = \lambda \cdot 0$ for *every* number λ . If $v = 0$ were allowed, every number would be an eigenvalue of every matrix, and the definition would say nothing. Requiring $v \neq 0$ means λ is an eigenvalue only when A really acts as “multiply by λ ” on some actual direction.

The notes put it in one line: $A0 = \lambda 0$ has no information about A .

Theorem (10/8)

λ is an eigenvalue of $A \iff \det(A - \lambda I) = 0$.

(Reason: $\text{Nul}(A - \lambda I)$ contains a nonzero vector exactly when $A - \lambda I$ is singular, i.e. has determinant 0.)

Definition (p. 196)

The λ -**eigenspace** of A is $\text{Nul}(A - \lambda I)$. Notation: $E_\lambda(A)$.

For example, the 0-eigenspace is $E_0(A) = \text{Nul}(A - 0I) = \text{Nul } A$.

Is the λ -eigenspace a subspace? Yes

It is the null space of the matrix $A - \lambda I$, and every null space is a subspace (Lay §4.2). Directly: if $(A - \lambda I)u = 0$ and $(A - \lambda I)v = 0$, then $(A - \lambda I)(cu + dv) = c \cdot 0 + d \cdot 0 = 0$.

Note the eigenspace contains 0, which is *not* an eigenvector: the eigenspace is “all the eigenvectors, plus 0”.

Example (p. 194): $A = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$

$$\det \begin{bmatrix} 2-\lambda & 1 \\ -1 & 4-\lambda \end{bmatrix} = (2-\lambda)(4-\lambda) + 1 = \lambda^2 - 6\lambda + 9 = (\lambda - 3)^2.$$

So $\lambda = 3$ is the only eigenvalue (a *repeated* root).

Fill-in: $\text{Nul}(A - 3I) = \text{span}\{(1, 1)\}$

$A - 3I = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$. Both rows say $-x + y = 0$, i.e. $y = x$, with x free. So

$$\text{Nul}(A - 3I) = \text{span} \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

The eigenvectors are the nonzero multiples of $(1, 1)$; for example $25 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is also an eigenvector.

Check: $A(1, 1)^T = (3, 3)^T = 3(1, 1)^T$. ✓

$A - 3I$ row reduces to $\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$: **1 pivot**, so 1 free variable and $\dim E_3(A) = 1$.

Note. The eigenspace is only 1-dimensional, so this A does *not* have two independent eigenvectors: there is no basis of $\mathbb{R}^2$ made of eigenvectors of A . A repeated eigenvalue can cause this (compare Section 4).

2 Examples (pp. 195–200)

$A = cI$ (p. 195): eigenvalue c , every nonzero vector is an eigenvector

$Av = cIv = cv$ for every v . So c is an eigenvalue, the c -eigenspace is all of $\mathbb{R}^n$, and every nonzero vector is an eigenvector. There are no other eigenvalues: $\det(cI - \lambda I) = (c - \lambda)^n = 0$ only for $\lambda = c$.

Diagonal $A = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix}$ (p. 195): eigenvalues $d_1, \dots, d_n$, eigenvectors $e_1, \dots, e_n$

$Ae_i = d_i e_i$ (column i of A is $d_i e_i$), so e_i is an eigenvector with eigenvalue d_i . Since $\det(A - \lambda I) = (d_1 - \lambda) \cdots (d_n - \lambda)$, these are all the eigenvalues. The λ -eigenspace is spanned by the e_i with

$d_i = \lambda$. For example, $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ has 2-eigenspace $\text{span}\{e_1, e_3\}$ and 5-eigenspace $\text{span}\{e_2\}$.

Upper triangular A (p. 196): the eigenvalues are $a_{11}, \dots, a_{nn}$

$A - \lambda I$ is also triangular, with diagonal entries $a_{ii} - \lambda$, so $\det(A - \lambda I) = (a_{11} - \lambda) \cdots (a_{nn} - \lambda)$, which is 0 exactly when λ is a diagonal entry. (The *eigenvectors* are not just the e_i any more; you still have to compute null spaces, as in the example on p. 199.)

If $v \neq 0$ is in $\text{Nul } A$, is v an eigenvector? Yes, with eigenvalue 0

$Av = 0 = 0 \cdot v$. So 0 is an eigenvalue of A exactly when $\text{Nul } A \neq \{0\}$, i.e. when A is not invertible.

Combinations of eigenvectors (p. 197)

If $Av = 2v$ and $Au = 2u$ ($u, v \neq 0$), is $5u - v$ an eigenvector? Yes, unless $5u - v = 0$

$A(5u - v) = 5Au - Av = 10u - 2v = 2(5u - v)$. So $5u - v$ lies in the 2-eigenspace (a subspace!). It is an eigenvector with eigenvalue 2 as long as it is nonzero, i.e. unless $v = 5u$.

The exception really happens (p. 197): take $A = 2I = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $u = \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}$. Then $Av = 2v$ and $Au = 2u$, but $5u - v = 0$, which is not an eigenvector.

If $Av = 2v$ and $Au = -u$ ($u, v \neq 0$), is $5u - v$ an eigenvector? No

$A(5u - v) = 5Au - Av = -5u - 2v$. If this equaled $\mu(5u - v) = 5\mu u - \mu v$ for some number μ , then

$$(-5 - 5\mu)u + (-2 + \mu)v = 0.$$

Eigenvectors with different eigenvalues ($2 \neq -1$) are independent (Theorem, Section 3), so both coefficients are 0: $\mu = -1$ and $\mu = 2$. Impossible. So $5u - v$ is not an eigenvector.

The notes' version of the argument: the equation above says u and v are dependent, i.e. they span the same line, say $u = cv$. But then $Au = cAv = 2cv = 2u$, contradicting $Au = -u$ (as $u \neq 0$). So u and v cannot span the same line.

Moral: adding eigenvectors with the *same* eigenvalue gives an eigenvector (or 0); mixing *different* eigenvalues never does.

Example (p. 198): find the “eigenthings” of $A = \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix}$

Eigenvalues $\lambda_1 = -1$, $\lambda_2 = 2$; eigenvectors $v_1 = (1, 1)$, $v_2 = (2, 1)$

$$\begin{aligned} \det \begin{bmatrix} 5 - \lambda & -6 \\ 3 & -4 - \lambda \end{bmatrix} &= (5 - \lambda)(-4 - \lambda) + 18 = \lambda^2 - \lambda - 20 + 18 \\ &= \lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1). \end{aligned}$$

- $\lambda_1 = -1$: $A + I = \begin{bmatrix} 6 & -6 \\ 3 & -3 \end{bmatrix}$, so $x = y$: $\text{Nul}(A + I)$ is 1-dimensional, spanned by $v_1 = (1, 1)$.
Check: $A(1, 1)^T = (-1, -1)^T$. ✓
- $\lambda_2 = 2$: $A - 2I = \begin{bmatrix} 3 & -6 \\ 3 & -6 \end{bmatrix}$, so $x = 2y$: $\text{Nul}(A - 2I)$ is 1-dimensional, spanned by $v_2 = (2, 1)$.
Check: $A(2, 1)^T = (10 - 6, 6 - 4)^T = (4, 2)^T = 2(2, 1)^T$. ✓

Does $\mathbb{R}^2$ have a basis of eigenvectors of A ? Yes

$\{v_1, v_2\}$: two eigenvectors with different eigenvalues are independent (Section 3), and two independent vectors in $\mathbb{R}^2$ form a basis. (Directly: $(1, 1)$ and $(2, 1)$ are not multiples of each other.) Compare p. 194, where there was no such basis.

Example (p. 199): a basis for each eigenspace of $A = \begin{bmatrix} 0 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$

Eigenvalues $\lambda_1 = 0, \lambda_2 = 2$

A is upper triangular with diagonal $0, 2, 0$, so $\det(A - \lambda I) = (-\lambda)^2(2 - \lambda)$ and the eigenvalues are 0 (repeated) and 2.

$\lambda = 0$: $\text{Nul } A$. Both nonzero rows say $2x_2 + 3x_3 = 0$, with x_1 and x_3 free. So $x = x_1(1, 0, 0) + x_3(0, -\frac{3}{2}, 1)$. Basis of $E_0(A)$:

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -\frac{3}{2} \\ 1 \end{bmatrix} \right\} \quad (\text{dimension 2; scaling the second by 2 gives } (0, -3, 2), \text{ also fine}).$$

$\lambda = 2$: $A - 2I = \begin{bmatrix} -2 & 2 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & -2 \end{bmatrix}$. Rows 2 and 3 give $x_3 = 0$; then row 1 gives $x_1 = x_2$. Basis:

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\} \quad (\text{dimension 1}).$$

Checks: $A(0, -3, 2)^T = (-6 + 6, -6 + 6, 0)^T = 0$; $A(1, 1, 0)^T = (2, 2, 0)^T$. ✓

Together the three vectors form a basis of $\mathbb{R}^3$, even though 0 is a repeated eigenvalue: here the 0-eigenspace is big enough (dimension 2).

If A is invertible and $Av = \lambda v, v \neq 0$, then $A^{-1}v = \frac{1}{\lambda}v$ (p. 200)

First, $\lambda \neq 0$: otherwise $Av = 0$ with $v \neq 0$, impossible for invertible A . Apply A^{-1} to $Av = \lambda v$: $v = \lambda A^{-1}v$, so $A^{-1}v = \frac{1}{\lambda}v$. So A^{-1} has the same eigenvectors, with eigenvalues $1/\lambda$.

Example: $A = \begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix}$ from p. 198 has $A^{-1}(2,1)^T = \frac{1}{2}(2,1)^T$ and $A^{-1}(1,1)^T = -(1,1)^T$.

3 Distinct eigenvalues give independent eigenvectors (pp. 201–202)

Theorem

Let $v_1, \dots, v_k$ be eigenvectors of an $n \times n$ matrix A with **distinct** eigenvalues $\lambda_1, \dots, \lambda_k$. Then $\{v_1, \dots, v_k\}$ is independent.

Proof by induction on m (the number of vectors).

Base $m = 1$: $\{v_1\}$ is independent because $v_1 \neq 0$ (eigenvectors are nonzero).

Induction step $m \rightarrow m+1$: Assume any m eigenvectors with distinct eigenvalues are independent. Arguing by contradiction, assume $v_1, \dots, v_{m+1}$ is dependent. Since $v_1, \dots, v_m$ is independent, the dependence must involve v_{m+1} , so

$$v_{m+1} = c_1 v_1 + \dots + c_m v_m.$$

(Details: a dependence $d_1 v_1 + \dots + d_{m+1} v_{m+1} = 0$ with not all $d_i = 0$ must have $d_{m+1} \neq 0$, otherwise it would be a dependence among $v_1, \dots, v_m$. Divide by d_{m+1} .)

Apply A to both sides and use $Av_i = \lambda_i v_i$:

$$\lambda_{m+1} v_{m+1} = c_1 \lambda_1 v_1 + \dots + c_m \lambda_m v_m.$$

Also multiply the first equation by λ_{m+1} : $\lambda_{m+1} v_{m+1} = c_1 \lambda_{m+1} v_1 + \dots + c_m \lambda_{m+1} v_m$. Subtract and collect terms:

$$c_1 (\lambda_1 - \lambda_{m+1}) v_1 + \dots + c_m (\lambda_m - \lambda_{m+1}) v_m = 0.$$

By the induction hypothesis $v_1, \dots, v_m$ are independent, so every coefficient $c_i (\lambda_i - \lambda_{m+1}) = 0$. Since $\lambda_i \neq \lambda_{m+1}$ for $i \leq m$, we get $c_1 = \dots = c_m = 0$. So $v_{m+1} = 0$, which is not an eigenvector: contradiction. $\square$

The idea in one line: applying A multiplies each v_i by a *different* number, which a genuine dependence can't survive.

4 n distinct eigenvalues: a basis of eigenvectors (pp. 203–204)

Theorem

If an $n \times n$ matrix A has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$, then their eigenvectors $v_1, \dots, v_n$ form a basis of $\mathbb{R}^n$.

Proof (p. 203)

By the theorem of Section 3, $v_1, \dots, v_n$ are independent. n independent vectors in the n -dimensional space $\mathbb{R}^n$ form a basis (Basis Theorem, Lay §4.5).

Caution. The converse fails: I has only one eigenvalue but a basis of eigenvectors, and the 3×3 example on p. 199 has two eigenvalues but still a basis. Repeated eigenvalues *may* or *may not* cause trouble (compare p. 194).

If $Av = \lambda v$, then $A^k v = \lambda^k v$ for every integer k (pp. 200, 204)

$k > 0$ (p. 200): $A^k v = A \cdots A Av$, and each A multiplies by λ . Formally, by induction: $A^{k+1} v = A(\lambda^k v) = \lambda^k A v = \lambda^{k+1} v$.

$k < 0$ (p. 204, A invertible): $A^k = (A^{-1})^{|k|}$, and $A^{-1} v = \frac{1}{\lambda} v$ (p. 200), so

$$A^k v = \left(\frac{1}{\lambda}\right)^{|k|} v = \frac{1}{\lambda^{|k|}} v = \lambda^k v.$$

(The notes write $\frac{1}{\lambda^k} v$, meaning λ to the power $|k|$ in the denominator.)

$k = 0$: $A^0 v = I v = v = \lambda^0 v$.

Example (p. 203): a 4×4 matrix with eigenvalues $1, -\frac{1}{2}, \frac{1}{3}, 0$

Is $\{v_1, v_2, v_3, v_4\}$ a basis of $\mathbb{R}^4$? Yes

The four eigenvalues are distinct, so by the theorem the eigenvectors form a basis. (The eigenvalue 0 is allowed: v_4 is a nonzero vector in $\text{Nul } A$.)

Fill-in: $A^k X_0$

Write $X_0 = c_1 v_1 + c_2 v_2 + c_3 v_3 + c_4 v_4$ (possible because it's a basis). By linearity and the aside,

$$A^k X_0 = c_1 \cdot 1^k v_1 + c_2 \left(-\frac{1}{2}\right)^k v_2 + c_3 \left(\frac{1}{3}\right)^k v_3 + c_4 \cdot 0^k v_4.$$

For $k \geq 1$ the last term is 0. As $k \rightarrow \infty$, $\left(-\frac{1}{2}\right)^k \rightarrow 0$ and $\left(\frac{1}{3}\right)^k \rightarrow 0$, so

$$A^k X_0 \longrightarrow c_1 v_1.$$

In words: the long-run behaviour is controlled by the eigenvalue 1; the other components die out. (This is exactly what happened with the steady state of a Markov chain on 10/1.)

5 The characteristic polynomial (pp. 205–207)

Definition

For an $n \times n$ matrix A , $\det(A - \lambda I)$ is the **characteristic polynomial** of A .

Theorem

$\det(A - \lambda I)$ is a polynomial in λ of degree n . In fact,

$$\det(A - \lambda I) = (-\lambda)^n + (a_{11} + \cdots + a_{nn})(-\lambda)^{n-1} + \cdots + \det A.$$

Proof of “degree n , leading term $(-\lambda)^n$ ” (pp. 205–206). Induct on n .

Base $n = 1$: $\det[a_{11} - \lambda] = a_{11} - \lambda = (-\lambda) + a_{11}$. ✓

General n : expand along row 1:

$$\det(A - \lambda I) = (a_{11} - \lambda)C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n},$$

where $C_{1j} = (-1)^{1+j} \det(A - \lambda I \text{ with row 1, column } j \text{ removed})$.

- C_{11} is the characteristic polynomial of the $(n-1) \times (n-1)$ matrix left after removing row 1 and column 1. By induction it is $(-\lambda)^{n-1} + (\text{lower})$, so $(a_{11} - \lambda)C_{11} = (-\lambda)^n + (\text{lower})$.
- For $j \geq 2$, removing row 1 and column j removes *two* of the n entries containing λ (the one in row 1 and the one in column j). The smaller determinant is a sum of products of $n-1$ entries, at most $n-2$ of which contain λ , so C_{1j} has degree $\leq n-2$. (The notes say “ $\leq n-1$ ”, which is also true and is all we need.)

So the only λ^n term comes from $(a_{11} - \lambda)C_{11}$, and it is $(-\lambda)^n$. □

Example ($n = 3$, p. 206). Removing row 1 and column 2 leaves

$$C_{12} = (-1)^{1+2} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} - \lambda \end{bmatrix},$$

which contains only one λ : degree 1 = $n - 2$.

The constant term. Set $\lambda = 0$: $f(0) = \det(A - 0I) = \det A$.

Definition

The **trace** of A is the sum of its diagonal entries: $\text{trace } A = a_{11} + a_{22} + \cdots + a_{nn}$.

So the characteristic polynomial is $(-\lambda)^n + (\text{trace } A)(-\lambda)^{n-1} + \cdots + \det A$. (The $(-\lambda)^{n-1}$ terms also come only from $(a_{11} - \lambda)C_{11}$, and by induction they add up to the trace; for 2×2 see below.)

Check for 2×2 . $\det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = \lambda^2 - (a + d)\lambda + (ad - bc) = (-\lambda)^2 + (\text{trace } A)(-\lambda) + \det A$. ✓ For p. 198: $\text{trace } A = 5 - 4 = 1$, $\det A = -20 + 18 = -2$, giving $\lambda^2 - \lambda - 2$. ✓

Fundamental theorem of algebra

Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ with $a_{n-1}, \dots, a_0$ complex numbers. Then

$$f(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$$

for some complex numbers $\lambda_1, \dots, \lambda_n$ (not necessarily distinct). (Proof omitted.)

Consequence for matrices. Every $n \times n$ matrix has exactly n eigenvalues counted with repetition, if complex eigenvalues are allowed. Real matrices can have complex ones, like the 90° rotation with $\lambda = \pm i$ (9/29).

Sum and product of the eigenvalues. Writing $\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda)$ and comparing with the theorem:

- constant terms ($\lambda = 0$): $\lambda_1 \lambda_2 \cdots \lambda_n = \det A$;
- $(-\lambda)^{n-1}$ terms: $\lambda_1 + \lambda_2 + \cdots + \lambda_n = \text{trace } A$.

Quick check: on p. 198, $2 \cdot (-1) = -2 = \det A$ and $2 + (-1) = 1 = \text{trace } A$; on p. 194, $3 \cdot 3 = 9 = 8 + 1$ and $3 + 3 = 6 = 2 + 4$.

Note (p. 207). For $n \leq 4$ there are explicit formulas for the roots $\lambda_1, \dots, \lambda_n$: the quadratic formula, and the (long) cubic and quartic formulas. For $n > 4$ no such formula in terms of $+, -, \times, \div$ and roots exists (the Abel–Ruffini theorem). That is why computers find eigenvalues by iterative numerical methods, as mentioned on 10/8.

Corrections to the handwritten notes

Page	As written	Should be	Why
204	$A^k v = (A^{-1})^k v = \frac{1}{\lambda^k} v$ for $k < 0$	$(A^{-1})^{ k } v = \lambda^k v$	with $k < 0$, the exponent on A^{-1} must be $ k = -k$
206	“Ex $n = 3, j = 3$ ”	$j = 2$	the computation is C_{12}
207	$f(x) = x^n + a_{n-1}x^n + \cdots$	$a_{n-1}x^{n-1}$	exponent