

Linear Algebra Notes — Lecture 10/8

Determinants

Lecture pages 159–179 · plus pp. 180–185 from the draft · Lay Chapter 3

Contents

Overview	2
1 What we want from a determinant (p. 159)	2
2 Definition: cofactor expansion along row one (pp. 160–161)	3
3 Row operations and determinants (pp. 162–164)	4
4 Computing determinants by row reduction (pp. 164–165)	5
5 Invertibility and eigenvalues (pp. 166–168)	6
6 Computational note: roots of polynomials (pp. 169–171)	7
7 $\det(AB) = \det A \det B$ (pp. 171–174)	8
8 Digression: row space and rank (pp. 175–177)	9
9 $\det A^T = \det A$ (pp. 177–179)	11
10 Cofactor expansion along any row or column (p. 179; draft pp. 180–181)	12
11 Cramer’s rule (draft pp. 181–183)	13
12 A formula for A^{-1} (draft pp. 184–185)	14
Summary of determinant facts	15
Corrections to the handwritten notes	16

Overview

The determinant is a single number attached to a square matrix. It is zero exactly when the matrix is not invertible, it multiplies when matrices multiply, and it gives a way to find eigenvalues. The lecture:

1. defines det by cofactor expansion;
2. shows how row operations change it, which gives the fast way to compute it;
3. proves A invertible $\iff \det A \neq 0$, and uses it to find eigenvalues;
4. proves $\det(AB) = \det A \det B$ and $\det A^T = \det A$;
5. states cofactor expansion along any row or column.

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

About pp. 180–185. The final 10/8 notes stop at p. 179. The draft continued with the proof of cofactor expansion, linearity in a column, Cramer’s rule and a formula for A^{-1} . Those pages are kept at the end (Sections 10–12, marked “from the draft”), since they will probably be covered next lecture.

Words used throughout. *Singular* means not invertible. $|A|$ is another notation for $\det A$ (it is not an absolute value). REF = row echelon form, RREF = reduced row echelon form.

1 What we want from a determinant (p. 159)

Small cases.

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc, \quad \det[a] = a.$$

Wish list. We want a function with these properties:

1. $\det : M_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R}$ (it eats a square matrix and returns a number);
2. $\det$ is easily computable;
3. $\det(AB) = \det A \det B$;
4. $\det A = 0 \iff A$ is singular;
5. $|\det[v_1 \cdots v_n]|$ = volume of the parallelepiped spanned by $v_1, \dots, v_n$.

About item 5. For $n = 2$ this is the area of the parallelogram with sides v_1, v_2 ; for $n = 3$ it is the volume of the slanted box in the notes’ picture. Example: $v_1 = (2, 0)$, $v_2 = (1, 3)$ span a parallelogram with base 2 and height 3, area 6, and $\det \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} = 6$. This fits item 4: the volume is 0 exactly when the vectors are squashed into a lower dimension, i.e. dependent.

2 Definition: cofactor expansion along row one (pp. 160–161)

Definition

For an $n \times n$ matrix $A = [a_{ij}]$,

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n},$$

where the **cofactor** is

$$C_{ij} = (-1)^{i+j} \det(\text{the matrix obtained from } A \text{ by removing row } i \text{ and column } j).$$

The definition is recursive: an $n \times n$ determinant is built from $(n-1) \times (n-1)$ determinants, down to $\det[a] = a$. The signs $(-1)^{i+j}$ form a checkerboard starting with $+$ in the top-left corner.

Example (2×2). $C_{11} = (-1)^{1+1} \det[d] = d$ and $C_{12} = (-1)^{1+2} \det[c] = -c$, so

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = aC_{11} + bC_{12} = ad - bc.$$

This matches the formula from Section 1.

Example (3×3 , p. 161), finished

For each entry of row 1, cross out its row and column (the colored lines in the notes) and take the 2×2 determinant that is left:

$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix} &= 1 \cdot (-1)^{1+1} \det \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} + 2 \cdot (-1)^{1+2} \det \begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} \\ &\quad + (-1) \cdot (-1)^{1+3} \det \begin{bmatrix} 3 & 1 \\ -1 & 0 \end{bmatrix} \\ &= 1 \cdot (1) - 2 \cdot (3 + 2) + (-1) \cdot (0 + 1) = 1 - 10 - 1 = -10. \end{aligned}$$

Example (lower triangular). If every entry above the diagonal is 0, row 1 is $(a_{11}, 0, \dots, 0)$, so only the first term survives:

$$\det \begin{bmatrix} a_{11} & & 0 \\ & a_{22} & \\ * & & a_{nn} \end{bmatrix} = a_{11}(-1)^{1+1} \det \begin{bmatrix} a_{22} & & 0 \\ & \ddots & \\ * & & a_{nn} \end{bmatrix} + 0 + \cdots + 0 = a_{11}a_{22} \cdots a_{nn}.$$

(The smaller matrix is again lower triangular, so repeat.) In particular $\det I = 1$.

How does $\det A$ change under row operations? That is the next section.

3 Row operations and determinants (pp. 162–164)

Rules

1. **Swap.** If B is obtained from A by swapping two rows, then $\det B = -\det A$.
2. **Scale.** If B is obtained from A by scaling a row by a number t , then $\det B = t \det A$.
3. **Replace.** Row replacement (adding a multiple of one row to another) does not change $\det A$.

The notes check each rule for 2×2 matrices; the general proofs are by induction on n (Lay, Theorem 3 of §3.2).

Rule 1 for 2×2 . $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$ and $\det \begin{bmatrix} c & d \\ a & b \end{bmatrix} = cb - ad = -(ad - bc)$. ✓

Ex: if A has two identical rows, then $\det A = 0$

Swap the two identical rows. The matrix doesn't change, but by Rule 1 the determinant changes sign: $\det A = -\det A$. So $2 \det A = 0$ and $\det A = 0$.

Rule 2 for 2×2 . Scaling row 1 by t multiplies the row-one expansion by t , straight from the definition. To scale row 2, swap it into row 1 first (Rule 1), use the definition, and swap back:

$$\det \begin{bmatrix} a & b \\ ct & dt \end{bmatrix} \stackrel{(1)}{=} -\det \begin{bmatrix} ct & dt \\ a & b \end{bmatrix} \stackrel{\text{def}}{=} -t \det \begin{bmatrix} c & d \\ a & b \end{bmatrix} \stackrel{(1)}{=} t \det \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Ex: if A has a zero row, then $\det A = 0$

A zero row is 0 times itself, so scaling that row by $t = 0$ leaves A unchanged. By Rule 2, $\det A = 0 \cdot \det A = 0$.

Rule 3 for 2×2 . Add 3 times row 1 to row 2:

$$\det \begin{bmatrix} a & b \\ 3a+c & 3b+d \end{bmatrix} = a(3b+d) - b(3a+c) = ad - bc + \underbrace{3ab - 3ab}_0 = ad - bc. \checkmark$$

Why it works in general: the determinant is linear in each row, so adding $t \cdot (\text{row } j)$ to row i adds $t \cdot \det(\text{a matrix with row } j \text{ appearing twice}) = t \cdot 0$.

Ex: upper triangular 3×3 (pp. 163–164), finished

The definition expands along row 1, which is not the row with the zeros. So swap rows 1 and 3 (one sign change) to bring the zeros to the top:

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} = -\det \begin{bmatrix} 0 & 0 & a_{33} \\ 0 & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{bmatrix} = -a_{33} (-1)^{1+3} \det \begin{bmatrix} 0 & a_{22} \\ a_{11} & a_{12} \end{bmatrix}.$$

The 2×2 determinant is $0 \cdot a_{12} - a_{22}a_{11} = -a_{22}a_{11}$. Altogether

$$(-a_{33})(-a_{22}) \det[a_{11}] = a_{11}a_{22}a_{33}.$$

Theorem (triangular matrices)

If A is triangular (upper or lower), then $\det A = a_{11}a_{22} \cdots a_{nn}$, the product of the diagonal entries.

4 Computing determinants by row reduction (pp. 164–165)

Theorem

Suppose the $n \times n$ matrix A is row reduced to an echelon form

$$U = \begin{bmatrix} u_1 & & * \\ & \ddots & \\ 0 & & u_n \end{bmatrix}$$

using r row swaps and *no* row scaling (which is always possible). Then

$$\det A = (-1)^r \det U = (-1)^r u_1 u_2 \cdots u_n.$$

Proof. Only swaps and replacements are used. Each row swap changes the sign of the determinant (Rule 1); row replacement does not change it (Rule 3). So $\det A = (-1)^r \det U$, and U is triangular, so $\det U = u_1 \cdots u_n$. $\square$

Why scaling is never needed: to clear the entries below a pivot you only add multiples of the pivot row, and to get a nonzero pivot you only swap. Scaling is only for making pivots equal to 1, which REF doesn't require.

Example (p. 165): $\det A = -6$

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix} \xrightarrow[\substack{R_3+R_1}]{R_2-R_1} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -2 & 2 \\ 0 & 3 & 0 \end{bmatrix} \xrightarrow[\substack{R_2+\frac{2}{3}R_3}]{R_1-\frac{2}{3}R_3} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 2 \\ 0 & 3 & 0 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

The first two steps are row replacements (no change). The last is one swap, so $r = 1$:

$$\det A = (-1)^1 \cdot 1 \cdot 3 \cdot 2 = -6.$$

Check by cofactor expansion: $1(0 - 1) - 2(1 + 1) + (-1)(1 - 0) = -1 - 4 - 1 = -6$. $\checkmark$

This is what makes “wish 2” come true. Cofactor expansion of an $n \times n$ matrix has about $n!$ terms ($20! \approx 2.4 \times 10^{18}$). Row reduction takes about n^3 steps ($20^3 = 8000$). Computers always

use row reduction.

5 Invertibility and eigenvalues (pp. 166–168)

Theorem

An $n \times n$ matrix A is invertible $\iff \det A \neq 0$.

Proof. Row reduce A to an echelon form U as above. Then

$$\begin{aligned} A \text{ is invertible} &\iff A \text{ has } n \text{ pivots (Invertible Matrix Theorem)} \\ &\iff \text{every diagonal entry } u_1, \dots, u_n \text{ of } U \text{ is a pivot, i.e. nonzero} \\ &\iff u_1 u_2 \cdots u_n \neq 0. \end{aligned}$$

But $\det A = (-1)^r u_1 \cdots u_n$, so this happens exactly when $\det A \neq 0$. $\square$

(“Each u_i is a scaled pivot” in the notes means: u_i is what the pivot looks like before you scale it to 1.)

Recall. If $Av = \lambda v$ with $v \neq 0$ in $\mathbb{R}^n$ and $\lambda \in \mathbb{R}$, then λ is an eigenvalue and v an eigenvector of A (9/29 notes).

Note (p. 166). If we already know λ , finding v is easy: solve $(A - \lambda I)v = 0$, an ordinary null-space problem. The hard part is finding λ .

Q: How to find eigenvalues?

Theorem

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

Proof.

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A &\iff Av = \lambda v \text{ for some } v \neq 0 \\ &\iff (A - \lambda I)v = 0 \text{ for some } v \neq 0 && \text{(move } \lambda v = \lambda Iv \text{ to the left)} \\ &\iff A - \lambda I \text{ is singular} && \text{(its null space is not } \{0\}) \\ &\iff \det(A - \lambda I) = 0. && \square \end{aligned}$$

$\det(A - \lambda I)$ is a polynomial in λ of degree n , called the *characteristic polynomial*. Its roots are the eigenvalues.

Example: eigenvalues of $A = \begin{bmatrix} 2 & 1 \\ 6 & 1 \end{bmatrix}$ are $\lambda_1 = -1, \lambda_2 = 4$

Subtract λ from the diagonal and take the determinant:

$$\begin{aligned} \det \begin{bmatrix} 2-\lambda & 1 \\ 6 & 1-\lambda \end{bmatrix} &= (2-\lambda)(1-\lambda) - 6 = \lambda^2 - 3\lambda + 2 - 6 \\ &= \lambda^2 - 3\lambda - 4 = (\lambda - 4)(\lambda + 1) = 0. \end{aligned}$$

So $\lambda_1 = -1$ and $\lambda_2 = 4$.

Eigenvectors (solve $(A - \lambda I)v = 0$ for each):

- $\lambda = 4$: $A - 4I = \begin{bmatrix} -2 & 1 \\ 6 & -3 \end{bmatrix}$, so $-2x + y = 0$ and $v = (1, 2)$. Check: $A(1, 2)^T = (4, 8)^T = 4(1, 2)^T$. ✓
- $\lambda = -1$: $A + I = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix}$, so $3x + y = 0$ and $v = (1, -3)$. Check: $A(1, -3)^T = (-1, 3)^T = -1(1, -3)^T$. ✓

Example (p. 168): the eigenvalues of a triangular matrix are its diagonal entries

If A is upper or lower triangular, so is $A - \lambda I$, with diagonal entries $a_{11} - \lambda, \dots, a_{nn} - \lambda$. By the triangular theorem,

$$\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda),$$

which is 0 exactly when λ equals one of the a_{ii} . So the eigenvalues are $a_{11}, a_{22}, \dots, a_{nn}$. *Example:*

$\begin{bmatrix} 2 & 7 \\ 0 & 5 \end{bmatrix}$ has eigenvalues 2 and 5.

6 Computational note: roots of polynomials (pp. 169–171)

MATLAB knows many methods for finding roots of polynomials, but the most general one uses linear algebra: it computes the eigenvalues of a matrix whose characteristic polynomial is the given polynomial. (That matrix is the companion matrix from the 9/29 notes.)

Why can computing roots of polynomials be numerically hard? Because a tiny change in a coefficient can cause a big change in the roots.

Example: $x^n = r^n$ with $0 < r < 1$, say $r = \frac{1}{2}$

If n is large, r^n is too small for the computer to see, so it thinks the equation is $x^n = 0$, and answers $x = 0$. (In standard double precision, $(\frac{1}{2})^n$ is stored as exactly 0 once n is larger than about 1075.)

What are the roots of $x^n = 0$?

Only $x = 0$ (repeated n times).

What are the roots of $x^n = r^n$?

There are n complex roots, equally spaced on the circle of radius r in the complex plane:

$$x = r e^{2\pi i k/n} = r \left(\cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n} \right), \quad k = 0, 1, \dots, n-1.$$

Here $e^{i\theta}$ is shorthand for $\cos \theta + i \sin \theta$ (Euler's formula), the point on the unit circle at angle θ .

Multiplying such numbers adds their angles, so $(r e^{2\pi i k/n})^n = r^n e^{2\pi i k} = r^n \cdot 1$. ✓

For $n = 4$ the angles are $0^\circ, 90^\circ, 180^\circ, 270^\circ$, giving $x = r, ir, -r, -ir$ (the four red dots in the notes' picture).

The point. The true roots all have size $r = \frac{1}{2}$, no matter how large n is. But the computer's version of the equation, $x^n = 0$, has only the root 0. An invisible change in one coefficient moved every root by $\frac{1}{2}$.

Example: Wilkinson's polynomial

$$P(x) = (x-1)(x-2)\cdots(x-20) = x^{20} + a_{19}x^{19} + \cdots + a_1x + 20!$$

Its roots are $1, 2, \dots, 20$, nicely separated. Here $a_{19} = -(1+2+\cdots+20) = -210$.

Fact. A small change in a_{19} leads to a large change in the roots. Wilkinson's classic computation: changing a_{19} from -210 to $-210 - 2^{-23}$ (a change of about 0.0000001) moves the root 20 to about 20.85 and turns several of the larger roots into complex pairs, such as $16.73 \pm 2.81i$.

7 $\det(AB) = \det A \det B$ (pp. 171–174)

Theorem

For $n \times n$ matrices, $\det(AB) = \det A \det B$. In the notation $|C| = \det C$: $|AB| = |A| |B|$.

A surprise (p. 171). Determinants are numbers, so $|A| |B| = |B| |A|$. Hence

$$|AB| = |A| |B| = |B| |A| = |BA|,$$

even though $AB \neq BA$ in general. For example $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ give $AB = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ and

$BA = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$: different matrices, both with determinant 1.

Proof, Case 1: A is singular. Then AB is singular too. (If AB were invertible, then $x \mapsto ABx$ would be onto. Every output $A(Bx)$ is an output of A , so A would be onto, hence invertible.) Thus $\det(AB) = 0 = \det A$, and both sides of the theorem are 0.

Case 2: A is invertible. Then A row reduces to I . Reversing the steps, I row reduces to A . Each row operation is the same as multiplying on the left by an *elementary matrix* E , which is the matrix you get by doing that row operation to I (Lay §2.2). So

$$I \rightarrow E_1 I \rightarrow E_2 E_1 I \rightarrow \cdots \rightarrow E_p \cdots E_1 I = E_p \cdots E_1 = A.$$

Compare $|C|$ with $|EC|$ for any matrix C . Since EC is C after one row operation, the three rules give:

Fill-in (p. 173)

$$|EC| = \begin{cases} |C| & E \text{ encodes a row replacement} \\ -|C| & E \text{ encodes a row swap} \\ t|C| & E \text{ scales a row by } t \end{cases} \xrightarrow{C=I} |E| = \begin{cases} 1 & E \text{ row replacement} \\ -1 & E \text{ row swap} \\ t & E \text{ scales by } t \end{cases}$$

(For the second, put $C = I$ and use $|I| = 1$.) Comparing the two: in every case $|EC| = |E| |C|$.

Now peel off the elementary matrices one at a time:

$$|AB| = |E_p \cdots E_1 B| = |E_p| |E_{p-1} \cdots E_1 B| = \cdots = |E_p| \cdots |E_1| |B|.$$

Using the same fact in reverse (with $B = I$) gives $|E_p| \cdots |E_1| = |E_p \cdots E_1| = |A|$. So $|AB| = |A| |B|$. $\square$

Examples (p. 174)

Inverse. $|AA^{-1}| = |A| |A^{-1}|$, and $AA^{-1} = I$ with $|I| = 1$. So $|A| |A^{-1}| = 1$ and

$$|A^{-1}| = \frac{1}{|A|}.$$

Is $|PAP^{-1}| = |A|$? Yes. $|PAP^{-1}| = |P| |A| |P^{-1}| = |P| |A| \frac{1}{|P|} = |A|$. (The determinants are numbers, so they commute even though the matrices don't.)

Powers. For $k > 0$: $|A^k| = |A \cdot A \cdots A| = |A|^k$.

For $k < 0$ (A invertible): $A^k = (A^{-1})^{-k}$, so $|A^k| = \left(\frac{1}{|A|}\right)^{-k} = |A|^k$.

$A^0 = I$ by convention, and $|A^0| = 1 = |A|^0$. So $|A^k| = |A|^k$ for every integer k .

8 Digression: row space and rank (pp. 175–177)

The next goal is $|A^T| = |A|$. It needs one fact about transposes, which comes from the row space.

Definition

Row A = the span of the rows of A .

Ex: Row $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$

$\text{span}\{(1,2), (0,0)\} = \text{span}\{(1,2)\}$: the line through $(1,2)$. The zero row adds nothing.

Note. Row $A = \text{Col } A^T$, because the rows of A are the columns of A^T .

Theorem

$\dim \text{Row } A = \dim \text{Col } A$.

Proof. Row reduce $A \rightarrow \cdots \rightarrow \text{RREF}(A)$. At each step the row space is preserved:

- after a replacement, swap, or scaling by $c \neq 0$, every new row is a combination of the old rows, so the new row space is inside the old one;
- each operation can be undone by an operation of the same kind, so the old row space is inside the new one too.

So Row $A = \text{Row RREF}(A)$.

What is $\dim \text{Row RREF}$? The number of pivots

The nonzero rows of an RREF are the pivot rows, and they are independent. Look at the notes' example

$$\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Row 1 has a 1 in column 1 where row 2 has 0; row 2 has a 1 in column 3 where row 1 has 0. So $c_1(\text{row } 1) + c_2(\text{row } 2) = 0$ forces $c_1 = 0$ (look at column 1) and $c_2 = 0$ (column 3). In general each pivot row has a 1 in its pivot column, where every other row has 0.

So $\dim \text{Row } A = \# \text{pivots of } A = \text{rank } A = \dim \text{Col } A$. □

Consequence. $\text{rank } A^T = \dim \text{Col } A^T = \dim \text{Row } A = \text{rank } A$.

Corollary

If A is $n \times n$ (so $m = n$), then $\dim \text{Nul } A = \dim \text{Nul } A^T$.

Proof. For an $m \times n$ matrix A : $\dim \text{Nul } A = \text{number of non-pivot columns} = n - \text{rank } A$. The transpose is $n \times m$, so $\dim \text{Nul } A^T = m - \text{rank } A^T$. But $\text{rank } A = \text{rank } A^T$, so when $m = n$ the two are equal. □

Ex (p. 177): $A = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ shows $m = n$ is needed

A is 1×3 with rank 1: $\dim \text{Nul } A = 3 - 1 = 2$.

$A^T = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is 3×1 with rank 1: $\dim \text{Nul } A^T = 1 - 1 = 0$.

They differ because $m \neq n$.

Ex. For $n \times n$ matrices, A is invertible $\iff A^T$ is invertible, because $\dim \text{Nul } A^T = 0 \iff \dim \text{Nul } A = 0$. Equivalently, A is singular $\iff A^T$ is singular.

9 $\det A^T = \det A$ (pp. 177–179)

Theorem

$$|A| = |A^T|.$$

Proof, Case 1: A is singular. Then A^T is singular (previous example). So $|A| = 0 = |A^T|$.

Case 2: A is invertible. Then $A = E_1 \cdots E_p$ with each E_i elementary. Transposing reverses the order: $A^T = E_p^T \cdots E_1^T$.

Note. If E is elementary, then $|E^T| = |E|$:

(a) **Swap:** $E = E^T$. For example $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ are symmetric.

(b) **Scaling:** $E = E^T$, since E is diagonal, e.g. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix}$.

(c) **Row replacement:** E and E^T are both triangular with the same diagonal (all 1's), e.g. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Both have determinant 1.

So in any case $|E^T| = |E|$, and using the product rule twice,

$$|A^T| = |E_p^T \cdots E_1^T| = |E_p^T| \cdots |E_1^T| = |E_p| \cdots |E_1| = |E_1| \cdots |E_p| = |E_1 \cdots E_p| = |A|. \quad \square$$

(The middle step reorders *numbers*, which is allowed.)

Why this matters. Every row fact about determinants is now also a column fact: swapping two columns flips the sign, a zero column gives 0, and so on.

Ex (p. 179): $|A| = 2$, $|B| = 3$. Find $|A^T B^2 A B^{-1}|$

$$|A^T B^2 A B^{-1}| = |A^T| \cdot |B|^2 \cdot |A| \cdot |B^{-1}| = 2 \cdot 9 \cdot 2 \cdot \frac{1}{3} = 12.$$

10 Cofactor expansion along any row or column (p. 179; draft pp. 180–181)

Theorem

For an $n \times n$ matrix A , any row i and any column j :

$$\begin{aligned} |A| &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} && \text{(along row } i) \\ &= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} && \text{(along column } j) \end{aligned}$$

where $C_{ij} = (-1)^{i+j} \det(\text{matrix obtained from } A \text{ by removing row } i \text{ and column } j)$.

Proof, with the signs worked out

Rows. The notes say: swap row i to the top and take signs into account. To keep the other rows in their original order, move row i up one step at a time: that takes $i - 1$ swaps of neighbouring rows, so the new matrix A' has $|A| = (-1)^{i-1}|A'|$. Expand A' along its first row (which is row i of A). Removing row 1 and column j of A' leaves exactly A with row i and column j removed, so

$$|A| = (-1)^{i-1} \sum_j a_{ij} (-1)^{1+j} \det(A \text{ without row } i, \text{ column } j) = \sum_j a_{ij} \underbrace{(-1)^{i+j}}_{C_{ij}} \det(\cdots).$$

Columns. Use $|A| = |A^T|$: column j of A is row j of A^T .

Tip. Expand along the row or column with the most zeros.

From here on: pages 180–185 of the draft. The final 10/8 notes end with the theorem above (p. 179). The proof above and Sections 10–12 come from the draft version, and will likely be covered next lecture.

Corollary (the determinant is linear in each column)

Let $A = [a_1 \cdots a_i \cdots a_n]$ with columns a_k . Replace column i by a variable vector x and set

$$T(x) = \det[a_1 \cdots a_{i-1} \ x \ a_{i+1} \cdots a_n].$$

Then $T : \mathbb{R}^n \rightarrow \mathbb{R}$ is linear.

Proof. Expand along column i :

$$T(x) = x_1 C_{1i} + x_2 C_{2i} + \cdots + x_n C_{ni}.$$

The cofactors C_{ki} delete column i , so they don't involve x at all: they are fixed numbers. So $T(x)$ is (fixed numbers) $\cdot$ (entries of x), which gives $T(x + y) = T(x) + T(y)$ and $T(cx) = cT(x)$. The notes also get $T(cx) = cT(x)$ another way: transpose, so column i becomes row i , and scale that row (Rule 2). $\square$

Example (p. 181). $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \det \begin{bmatrix} a & x_1 \\ b & x_2 \end{bmatrix} = ax_2 - bx_1$, a linear function of (x_1, x_2) .

11 Cramer's rule (draft pp. 181–183)

Notation. For an $n \times n$ matrix $A = [a_1 \cdots a_n]$ and $b \in \mathbb{R}^n$, let $A_i(b)$ be the matrix A with column i replaced by b :

$$A_i(b) = [a_1 \cdots a_{i-1} \ b \ a_{i+1} \cdots a_n].$$

Theorem (Cramer's rule)

If A is invertible and $x = A^{-1}b$ (the solution of $Ax = b$), then

$$x_i = \frac{\det A_i(b)}{\det A} \quad \text{for } i = 1, \dots, n.$$

Proof. Let $I = [e_1 \cdots e_n]$ and $I_i(x) = [e_1 \cdots e_{i-1} \ x \ e_{i+1} \cdots e_n]$. Multiplying a matrix by A multiplies each column by A :

$$A I_i(x) = [Ae_1 \cdots Ae_{i-1} \ Ax \ Ae_{i+1} \cdots Ae_n] = [a_1 \cdots a_{i-1} \ b \ a_{i+1} \cdots a_n] = A_i(b),$$

using $Ae_k = a_k$ and $Ax = b$. Take determinants of both sides:

$$\det A_i(b) = \det (A I_i(x)) = \det A \cdot \det I_i(x).$$

It remains to show $x_i = |\det I_i(x)|$.

Example (p. 183): $n = 3, i = 2$

By linearity in column 2, with $x = x_1e_1 + x_2e_2 + x_3e_3$:

$$\det \begin{bmatrix} 1 & x_1 & 0 \\ 0 & x_2 & 0 \\ 0 & x_3 & 1 \end{bmatrix} = x_1 \underbrace{\det \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{=0 \text{ (zero row)}} + x_2 \underbrace{\det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{=1} + x_3 \underbrace{\det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}}_{=0 \text{ (zero row)}} = x_2.$$

The general case, with the last blank filled in

By linearity in column i ,

$$|I_i(x)| = x_1 |I_i(e_1)| + x_2 |I_i(e_2)| + \cdots + x_n |I_i(e_n)|.$$

- If $j \neq i$: $I_i(e_j)$ has e_j in column i and in column j , and no column e_i . So its row i is zero, and its determinant is 0.
- If $j = i$: $I_i(e_i) = [e_1 \cdots e_n] = I$, with determinant 1.

$$\det I_i(e_j) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \implies |I_i(x)| = x_i.$$

Therefore $\det A_i(b) = \det A \cdot x_i$, and dividing by $\det A \neq 0$ gives Cramer's rule. $\square$

Worked example. Solve $2x_1 + x_2 = 5$, $6x_1 + x_2 = 9$. Here $A = \begin{bmatrix} 2 & 1 \\ 6 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 5 \\ 9 \end{bmatrix}$, $\det A = 2 - 6 = -4$.

$$x_1 = \frac{\det \begin{bmatrix} 5 & 1 \\ 9 & 1 \end{bmatrix}}{-4} = \frac{5 - 9}{-4} = 1, \quad x_2 = \frac{\det \begin{bmatrix} 2 & 5 \\ 6 & 9 \end{bmatrix}}{-4} = \frac{18 - 30}{-4} = 3.$$

Check: $2(1) + 3 = 5$ and $6(1) + 3 = 9$. $\checkmark$

12 A formula for A^{-1} (draft pp. 184–185)

Theorem

If A is invertible, the (i, j) entry of A^{-1} is $\frac{C_{ji}}{\det A}$. (Note the switched indices: C_{ji} , not C_{ij} .)

Proof. Column j of A^{-1} is $A^{-1}e_j$. Call it x ; it is the solution of $Ax = e_j$. The (i, j) entry of A^{-1} is x_i , and by Cramer's rule

$$x_i = \frac{\det A_i(e_j)}{\det A}.$$

Why $\det A_i(e_j) = C_{ji}$

$A_i(e_j)$ is A with column i replaced by e_j . Expand along that column i : its only nonzero entry is the 1 in row j , so the determinant is $1 \cdot (\text{the } (j, i) \text{ cofactor of } A_i(e_j))$. That cofactor deletes row j and column i . Column i is the only place where $A_i(e_j)$ differs from A , so this is the (j, i) cofactor C_{ji} of A . $\square$

Example (p. 184), finished

This matrix is $A_2(e_3)$ for a 3×3 matrix A . Expand along column 2, whose only nonzero entry is the 1 in row 3:

$$\det \begin{bmatrix} a_{11} & 0 & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & 1 & a_{33} \end{bmatrix} = 1 \cdot (-1)^{3+2} \det \begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix} = -(a_{11}a_{23} - a_{13}a_{21}) = C_{32}.$$

So the $(2, 3)$ entry of A^{-1} is $C_{32} / \det A$.

Check with 2×2 . For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$: $C_{11} = d$, $C_{12} = -c$, $C_{21} = -b$, $C_{22} = a$. The theorem gives

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

the familiar formula. (The matrix of cofactors with indices switched is called the *adjugate* of A .)

Conclusion (p. 185)

- Cramer's rule gives a formula for $A^{-1}b$ in terms of determinants.
- We also get a formula for A^{-1} in terms of determinants.

The formulas look horrible but can be useful: not for computing (row reduction is far faster), but for theory. For instance, they show that the entries of A^{-1} depend continuously on the entries of A , and that if A has integer entries and $\det A = \pm 1$, then A^{-1} has integer entries too.

Summary of determinant facts

Fact	Formula
Swap two rows (or columns)	$\det$ changes sign
Scale a row (or column) by t	$\det$ is multiplied by t
Row replacement	$\det$ unchanged
Two equal rows, or a zero row	$\det = 0$
Triangular matrix	$\det =$ product of diagonal entries
Row reduce to U with r swaps, no scaling	$\det A = (-1)^r u_1 \cdots u_n$
Invertibility	A invertible $\iff \det A \neq 0$
Eigenvalues	λ eigenvalue $\iff \det(A - \lambda I) = 0$
Products, inverses, powers	$ AB = A B $, $ A^{-1} = 1/ A $, $ A^k = A ^k$
Similar matrices, transpose	$ PAP^{-1} = A $, $ A^T = A $
Cramer's rule	$x_i = \det A_i(b) / \det A$
Inverse	$(A^{-1})_{ij} = C_{ji} / \det A$

Corrections to the handwritten notes

Page	As written	Should be	Why
163	$\det \begin{bmatrix} 0 & a_{22} \\ a_{11} & a_2 \end{bmatrix}$	bottom-right entry a_{12}	it comes from row (a_{11}, a_{12}, a_{13})
179	column expansion ends $c_{nj}C_{nj}$	$a_{nj}C_{nj}$	entries are a 's
180– 181 (draft)	$A_i(x)$ is a determinant, then $A_i(b)$ is a matrix	(same symbol, two meanings)	these notes use $T(x)$ for the determinant
183 (draft)	last term $x_n I_n(e_n) $	$x_n I_i(e_n) $	column i is the one replaced
184 (draft)	“cofactor expansion along column j ”	along column i	e_j sits in column i of $A_i(e_j)$
172, 178	$A = E_p \cdots E_1$, later $A = E_1 \cdots E_p$	(either is fine)	only the labelling differs