

Linear Algebra Notes — Lecture 10/13

Cramer's Rule, a Formula for A^{-1} , and Determinants as Area/Volume

Lecture pages 180–192 · Lay §3.3

Contents

Overview	1
1 Goals and motivation: noise (pp. 180–181)	2
2 The determinant is linear in each column (pp. 182–184)	3
3 Cramer's rule (pp. 184–185)	4
4 A formula for A^{-1} (p. 186)	5
5 Area and determinant in $\mathbb{R}^2$ (pp. 187–189)	6
6 Linear maps scale every area by $ \det A $ (pp. 189–191)	8
7 Area-preserving linear maps (p. 192)	9
8 Main facts about determinants (p. 192)	10
Summary	11
Corrections to the handwritten notes	11

Overview

This lecture finishes determinants. It has two parts:

1. **Explicit formulas.** Using the fact that $\det$ is linear in each column, we get Cramer's rule (a formula for the solution $A^{-1}b$ of $Ax = b$) and a formula for A^{-1} itself. The motivation is *noise*: the formulas show exactly how sensitive the answers are when $\det A$ is tiny.
2. **Geometry.** $|\det A|$ is the factor by which $x \mapsto Ax$ scales areas (in $\mathbb{R}^2$) or volumes (in $\mathbb{R}^n$). This is item 5 of the “wish list” from 10/8.

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

Notation from 10/8. $|A| = \det A$. $C_{ij} = (-1)^{i+j} \det(A \text{ with row } i \text{ and column } j \text{ removed})$ is the (i, j) cofactor. Cofactor expansion works along *any* row or column (p. 179).

1 Goals and motivation: noise (pp. 180–181)

Goals. Get an explicit formula for A^{-1} for an $n \times n$ matrix A . In the process we get a formula for $A^{-1}b$.

Why explicit formulas? For 2×2 we already have one:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

How does noise affect A^{-1} ?

The formula divides by $\det A = ad - bc$. If $\det A$ is close to 0, then a tiny change (“noise”) in the entries can change $\det A$ by a large *fraction* of itself, and so change A^{-1} enormously. If $\det A$ is not small, small noise causes only a small change in A^{-1} .

If A is invertible, how does noise in A, b affect the unique solution of $Ax = b$?

The same way: the solution is $x = A^{-1}b$, and Cramer’s rule (Section 3) writes each x_i as a ratio with $\det A$ in the denominator. When $\det A$ is tiny, small changes in A or b can move x a lot. Section 3 shows this in an example.

Example (p. 181)

Let ε be small and positive, say $\varepsilon = 10^{-20}$.

Fill-ins: $\det A, A^{-1}, B^{-1}$

$$A = \begin{bmatrix} 1 + \varepsilon & 1 \\ 1 & 1 \end{bmatrix} : \quad \det A = (1 + \varepsilon) - 1 = \varepsilon, \quad A^{-1} = \frac{1}{\varepsilon} \begin{bmatrix} 1 & -1 \\ -1 & 1 + \varepsilon \end{bmatrix}.$$

$$B = \begin{bmatrix} 1 + \varepsilon^2 & 1 \\ 1 & 1 \end{bmatrix} : \quad \det B = \varepsilon^2, \quad B^{-1} = \frac{1}{\varepsilon^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 + \varepsilon^2 \end{bmatrix}.$$

With $\varepsilon = 10^{-20}$:

- A and B differ in one entry, by $\varepsilon - \varepsilon^2 \approx 10^{-20}$. They are practically the same matrix.
- But A^{-1} has entries of size about 10^{20} , while B^{-1} has entries of size about 10^{40} . They differ by a factor of 10^{20} .

Having an explicit formula helps. A computer working in double precision stores numbers with about 16 significant digits, so it stores $1 + 10^{-20}$ as exactly 1. To the computer, both A and B look

like $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, which is not invertible at all. Row reduction on the computer can't tell the three matrices apart. The formulas can: they show precisely that the inverse blows up like $1/\det$.

2 The determinant is linear in each column (pp. 182–184)

Because $\det$ is linear in each column separately (when the other columns are held fixed), $\det : M_{n \times n}(\mathbb{R}) \rightarrow \mathbb{R}$ is called a **multilinear** map. (It is *not* linear as a function of the whole matrix: $\det(2A) = 2^n \det A$, not $2 \det A$.)

Theorem (p. 182)

Let $A = [a_1 \cdots a_i \cdots a_n]$ (columns a_k), and for $x \in \mathbb{R}^n$ let

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

be A with column i replaced by x . Let $T(x) = \det A_i(x)$. Then $T : \mathbb{R}^n \rightarrow \mathbb{R}$ is linear.

Example ($i = 2$). $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Replacing column 2 by x :

$$T(x) = \det \begin{bmatrix} 1 & x_1 \\ 3 & x_2 \end{bmatrix} = x_2 - 3x_1 = \begin{bmatrix} -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

So T is multiplication by the 1×2 matrix $\begin{bmatrix} -3 & 1 \end{bmatrix}$: linear. (The original column 2, $(2, 4)$, played no role.)

Proof (pp. 182–183). We need $T(cx) = cT(x)$ and $T(x + y) = T(x) + T(y)$.

Since $|A| = |A^T|$, we have $T(x) = \det A_i(x) = \det A_i(x)^T$. In $A_i(x)^T$ the columns of $A_i(x)$ become rows, so row i of $A_i(x)^T$ is $x^T = (x_1, \dots, x_n)$. Expand along row i :

$$\det A_i(x)^T = x_1 C_{i1} + x_2 C_{i2} + \cdots + x_n C_{in}.$$

Key point: each cofactor C_{ik} deletes row i , the only row that contains x . So the C_{ik} are fixed numbers that do not depend on x . Therefore

$$\begin{aligned} \det A_i(cx) &= cx_1 C_{i1} + \cdots + cx_n C_{in} = c(x_1 C_{i1} + \cdots + x_n C_{in}) = c \det A_i(x), \\ \det A_i(x + y) &= (x_1 + y_1) C_{i1} + \cdots + (x_n + y_n) C_{in} \\ &= (x_1 C_{i1} + \cdots + x_n C_{in}) + (y_1 C_{i1} + \cdots + y_n C_{in}) = \det A_i(x) + \det A_i(y). \end{aligned} \quad \square$$

Ex (pp. 183–184): $\det[v \ u \ 2v - u] = 0$ for any $u, v \in \mathbb{R}^3$

Use linearity in column 3:

$$\det[v \ u \ 2v - u] = 2 \det[v \ u \ v] - \det[v \ u \ u] = 2 \cdot 0 - 0 = 0.$$

Each of the two matrices on the right has two equal columns, so its determinant is 0. (Two

equal *columns* give 0 because $|A| = |A^T|$ turns them into two equal rows.)

Why it makes sense: the third column $2v - u$ is a combination of the first two, so the columns are dependent and the matrix is not invertible.

3 Cramer's rule (pp. 184–185)

Theorem (Cramer's rule)

If $A = [a_1 \cdots a_n]$ is an invertible $n \times n$ matrix and $b \in \mathbb{R}^n$, then the i -th coordinate of $A^{-1}b$ (the solution of $Ax = b$) is

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

Example: the noisy matrix again (pp. 184–185)

Fill-ins: $\det A$, $\det A_1(b)$, $\det A_2(b)$, $A^{-1}b$

$$A = \begin{bmatrix} 1 + \varepsilon & 1 \\ 1 & 1 \end{bmatrix}, \det A = \varepsilon.$$

$$\det A_1(b) = \det \begin{bmatrix} b_1 & 1 \\ b_2 & 1 \end{bmatrix} = b_1 - b_2,$$

$$\det A_2(b) = \det \begin{bmatrix} 1 + \varepsilon & b_1 \\ 1 & b_2 \end{bmatrix} = (1 + \varepsilon)b_2 - b_1 = b_2 - b_1 + \varepsilon b_2.$$

So by Cramer's rule

$$A^{-1}b = \frac{1}{\varepsilon} \begin{bmatrix} b_1 - b_2 \\ b_2 - b_1 + \varepsilon b_2 \end{bmatrix}.$$

Check against the A^{-1} from Section 1: $\frac{1}{\varepsilon} \begin{bmatrix} 1 & -1 \\ -1 & 1 + \varepsilon \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \frac{1}{\varepsilon} \begin{bmatrix} b_1 - b_2 \\ -b_1 + (1 + \varepsilon)b_2 \end{bmatrix}. \checkmark$

If $b_1 - b_2 = \varepsilon$, then

$$A^{-1}b = \begin{bmatrix} \varepsilon/\varepsilon \\ (-\varepsilon + \varepsilon b_2)/\varepsilon \end{bmatrix} = \begin{bmatrix} 1 \\ -1 + b_2 \end{bmatrix},$$

and ε no longer appears in the answer. The answer looks perfectly tame.

The catch. The solution depends on the ratio $(b_1 - b_2)/\varepsilon$. If noise of size $\varepsilon = 10^{-20}$ changes $b_1 - b_2$ from ε to 2ε , then x_1 jumps from 1 to 2. If noise makes $b_1 - b_2 = 0$, then $x_1 = 0$. An invisible change in b produces a completely different answer. That is the danger of a tiny $\det A$.

Proof of Cramer's rule (p. 185). Let $x = A^{-1}b$, so $Ax = b$. Let $I = [e_1 \cdots e_n]$ and let $I_i(x)$ be I with column i replaced by x . Multiplying by A multiplies each column by A , and $Ae_k = a_k$ (the k -th column of A):

$$\begin{aligned} A I_i(x) &= A[e_1 \cdots e_{i-1} \ x \ e_{i+1} \cdots e_n] = [Ae_1 \cdots Ae_{i-1} \ Ax \ Ae_{i+1} \cdots Ae_n] \\ &= [a_1 \cdots a_{i-1} \ b \ a_{i+1} \cdots a_n] = A_i(b). \end{aligned}$$

Taking determinants and using $|AB| = |A||B|$: $\det A_i(b) = \det A \cdot \det I_i(x)$.

Now compute $\det I_i(x)$. $I_i(x)$ looks like the identity except for column i :

$$I_i(x) = \begin{bmatrix} 1 & & x_1 & & \\ & \ddots & \vdots & & \\ & & x_i & & \\ & & \vdots & \ddots & \\ & & x_n & & 1 \end{bmatrix}.$$

Row i is all zeros except the entry x_i in column i (every e_k with $k \neq i$ has a 0 in slot i). Expand along row i :

$$\det I_i(x) = x_i (-1)^{i+i} \det I_{n-1} = x_i,$$

because deleting row i and column i leaves the $(n-1) \times (n-1)$ identity. So $\det A_i(b) = \det A \cdot x_i$, and dividing by $\det A \neq 0$ gives $x_i = \det A_i(b) / \det A$. $\square$

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

$$x_1 = \frac{\det \begin{bmatrix} 5 & 1 \\ 9 & 1 \end{bmatrix}}{-4} = \frac{-4}{-4} = 1, \quad x_2 = \frac{\det \begin{bmatrix} 2 & 5 \\ 6 & 9 \end{bmatrix}}{-4} = \frac{-12}{-4} = 3.$$

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

4 A formula for A^{-1} (p. 186)

Theorem

If A is invertible, the (i, j) entry of A^{-1} is

$$\frac{C_{ji}}{\det A}, \quad C_{ji} = (-1)^{j+i} \det(A \text{ with row } j \text{ and column } i \text{ removed}).$$

Note the switched indices: the (i, j) entry uses the (j, i) cofactor.

Proof. Column j of A^{-1} is $A^{-1}e_j$, the solution of $Ax = e_j$. By Cramer's rule (with $b = e_j$), its i -th entry is

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

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

$A_i(e_j) = [a_1 \cdots a_{i-1} \ e_j \ a_{i+1} \cdots a_n]$. Expand along column i (the notes say column j ; it is column i that holds e_j). That column has a single nonzero entry, the 1 in row j . So

$$\det A_i(e_j) = 1 \cdot (-1)^{j+i} \det(A_i(e_j) \text{ with row } j, \text{ column } i \text{ removed}).$$

Removing column i removes the only column where $A_i(e_j)$ differs from A , so the remaining matrix is A with row j and column i removed. That is exactly C_{ji} . $\square$

Check with 2×2 . $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has $C_{11} = d$, $C_{12} = -c$, $C_{21} = -b$, $C_{22} = a$. The theorem puts C_{ji} in position (i, j) :

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

the formula from p. 180. $\checkmark$ (The matrix $[C_{ji}]$ is called the *adjugate* of A .)

Check with the noisy example. For $A = \begin{bmatrix} 1 + \varepsilon & 1 \\ 1 & 1 \end{bmatrix}$: $C_{11} = 1$, $C_{12} = -1$, $C_{21} = -1$, $C_{22} = 1 + \varepsilon$,

so $A^{-1} = \frac{1}{\varepsilon} \begin{bmatrix} 1 & -1 \\ -1 & 1 + \varepsilon \end{bmatrix}$, as on p. 181. $\checkmark$

Conclusion (p. 186)

If $\det A$ is small but nonzero, then computing A^{-1} and solving $Ax = b$ is delicate: adding noise to A (or b) can change the results significantly, exactly as the formulas show, since $\det A$ sits in every denominator.

5 Area and determinant in $\mathbb{R}^2$ (pp. 187–189)

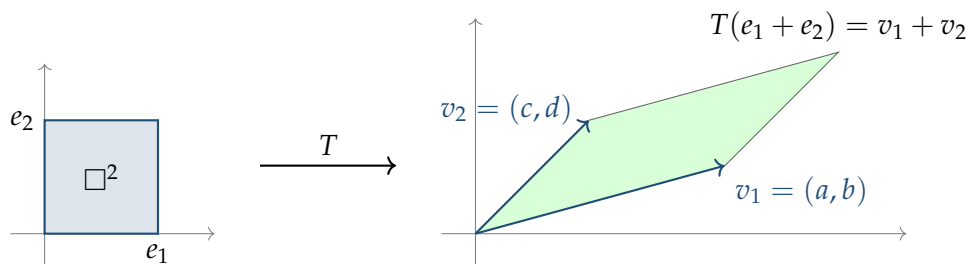
Key example. $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $T(x) = Ax$ with $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ (columns $v_1 = (a, b)$ and $v_2 = (c, d)$).

Let

$$\square^2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} = xe_1 + ye_2 : 0 \leq x \leq 1, 0 \leq y \leq 1 \right\}$$

be the unit square. What is the area of $T(\square^2)$?

By linearity, $T(xe_1 + ye_2) = xT(e_1) + yT(e_2) = x \begin{bmatrix} a \\ b \end{bmatrix} + y \begin{bmatrix} c \\ d \end{bmatrix}$ with $0 \leq x, y \leq 1$. So $T(\square^2)$ is the **parallelogram spanned by v_1 and v_2** : corners 0 , v_1 , v_2 and $v_1 + v_2 = T(e_1 + e_2)$.



So the question becomes: *what is the area of the parallelogram spanned by v_1, v_2 in $\mathbb{R}^2$?*

Special case (p. 188): $v_1 = (x, 0)$ with $x > 0$, $v_2 = (y, w)$

The base is v_1 , along the horizontal axis, with length x . The height is the vertical distance from v_2 to that axis, $|w|$. (If $w < 0$ the parallelogram hangs below the axis, as in the notes' second picture; the height is still $|w|$.) So

$$\text{area} = \text{base} \times \text{height} = x |w| = |xw|.$$

And $\det \begin{bmatrix} x & y \\ 0 & w \end{bmatrix} = xw$, so $\text{area} = |\det[v_1 \ v_2]|$ in this case.

Note. Any parallelogram with a corner at 0 can be rotated into the special case (turn it until v_1 points along the positive horizontal axis), and rotations preserve area.

Theorem (p. 188)

The area of the parallelogram spanned by $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$ is $\left| \det \begin{bmatrix} a & c \\ b & d \end{bmatrix} \right| = |ad - bc|$.

Proof (pp. 188–189). If $v_1 = 0$ the area and the determinant are both 0. Otherwise let θ be the angle of v_1 from the positive horizontal axis, and let

$$R = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

be the rotation by θ clockwise. It moves $v_1 = \begin{bmatrix} a \\ b \end{bmatrix}$ to $\begin{bmatrix} x \\ 0 \end{bmatrix}$ with $x = \sqrt{a^2 + b^2} > 0$, and moves

$v_2 = \begin{bmatrix} c \\ d \end{bmatrix}$ to some $\begin{bmatrix} y \\ w \end{bmatrix}$. Since R acts on each column,

$$R \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} x & y \\ 0 & w \end{bmatrix}.$$

The rotated parallelogram is the special case, with the same area. So

$$\text{area} = |xw| = \left| \det \left(R \begin{bmatrix} a & c \\ b & d \end{bmatrix} \right) \right| = \left| \det R \cdot \det \begin{bmatrix} a & c \\ b & d \end{bmatrix} \right| = |ad - bc|,$$

because $\det R = \cos^2 \theta + \sin^2 \theta = 1$. □

Example. $v_1 = (3, 1)$, $v_2 = (1, 2)$: $\text{area} = |3 \cdot 2 - 1 \cdot 1| = 5$.

6 Linear maps scale every area by $|\det A|$ (pp. 189–191)

Theorem (p. 189)

If $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is linear with $T(x) = Ax$ for all x , and R is a “nice” region in $\mathbb{R}^2$ of finite area, then

$$\text{area } T(R) = |\det A| \cdot \text{area } R, \quad \text{where } T(R) = \{T(x) : x \in R\}.$$

(“Nice” means the region has a sensible area, e.g. its boundary is a smooth curve; this rules out pathological sets.)

Example: the unit disk and ellipses (pp. 189–190)

Let $R = \{(x, y) : x^2 + y^2 \leq 1\}$ be the unit disk (area π) and $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ with $a, b \neq 0$. Then

$$T(R) = \left\{ \begin{bmatrix} ax \\ by \end{bmatrix} : x^2 + y^2 \leq 1 \right\}.$$

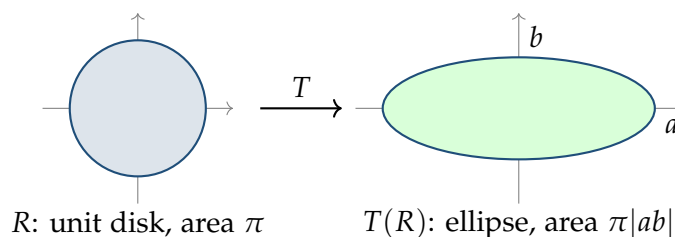
Rename the output $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} ax \\ by \end{bmatrix}$, so $x = x'/a$ and $y = y'/b$. The condition $x^2 + y^2 \leq 1$ becomes $\left(\frac{x'}{a}\right)^2 + \left(\frac{y'}{b}\right)^2 \leq 1$.

What is $T(R)$? What is its area?

$T(R) = \left\{ (x', y') : \left(\frac{x'}{a}\right)^2 + \left(\frac{y'}{b}\right)^2 \leq 1 \right\}$ is the **ellipse** with semi-axes $|a|$ (horizontal) and $|b|$ (vertical): the disk stretched by a sideways and b vertically. Its area is

$$\text{area } T(R) = |\det A| \cdot \text{area } R = |ab| \pi.$$

This is the familiar formula “area of an ellipse = πab ”, obtained with no calculus.



Proof sketch (p. 190). Put a grid of $\varepsilon \times \varepsilon$ squares on R .

- **Area of R .** Each small square has area ε^2 , so

$$\text{area } R = \lim_{\varepsilon \rightarrow 0} \varepsilon^2 \cdot \#(\varepsilon\text{-squares inside } R).$$

As $\varepsilon \rightarrow 0$ the squares fill R better and better, and the partial squares along the edge stop mattering.

- **Area of $T(R)$.** T sends each small square to a small parallelogram, and these parallelograms fill $T(R)$. So

$$\text{area } T(R) = \lim_{\varepsilon \rightarrow 0} \text{area}(T(\varepsilon\text{-square})) \cdot \#(\varepsilon\text{-squares inside } R).$$

- **Each small parallelogram.** An ε -square is spanned by $\varepsilon e_1, \varepsilon e_2$, so its image is spanned by $\varepsilon v_1, \varepsilon v_2$. By the area theorem, its area is $|\det[\varepsilon v_1 \ \varepsilon v_2]| = \varepsilon^2 |\det A|$ (pull ε out of each column). A shifted square, starting at a point p , maps to the same parallelogram shifted to $T(p)$, with the same area.

Substituting, $\text{area } T(R) = |\det A| \cdot \lim \varepsilon^2 \cdot \# = |\det A| \cdot \text{area } R$. □

(If $\det A = 0$, every small parallelogram is flat, and $T(R)$ lies on a line or point: area 0, which still matches.)

Theorem (p. 191, any dimension)

Let R be a nice region in $\mathbb{R}^n$ of finite n -dimensional volume, and $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear with $T(x) = Ax$ for all x . Then

$$\text{volume } T(R) = |\det A| \cdot \text{volume}(R).$$

Proof idea. As above, the grid argument reduces everything to $R =$ the unit cube. T maps the unit cube to the parallelepiped spanned by the columns $v_1, \dots, v_n$ of A . Its volume is

$$\text{volume} = \text{base} \times \text{height},$$

where the base is the $(n-1)$ -dimensional parallelepiped spanned by $v_1, \dots, v_{n-1}$ (red in the notes' picture) and the height is the distance from v_n to that base (green). So one can induct on n , with the $n=2$ case above as the model. In particular, with R the unit cube (volume 1), the volume of the parallelepiped spanned by $v_1, \dots, v_n$ is $|\det[v_1 \ \cdots \ v_n]|$: wish 5 from p. 159.

Example ($n=3$). $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ stretches the unit cube into a $2 \times 3 \times 4$ box: volume $24 = \det A$.

✓

7 Area-preserving linear maps (p. 192)

Q: Describe all area-preserving linear maps $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. What are their standard matrices?

$T(x) = Ax$ preserves area (area $T(R) = \text{area } R$ for every region R) exactly when

$$|\det A| = 1, \quad \text{i.e.} \quad ad - bc = \pm 1 \quad \text{for } A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

Why: if $|\det A| = 1$, the theorem gives area $T(R) = 1 \cdot \text{area } R$. Conversely, if T preserves area, apply it to the unit square: $|\det A| = \text{area } T(\square^2) = \text{area } \square^2 = 1$.

Examples. The professor's example on p. 190, $A = \begin{bmatrix} 25 & 1 \\ 26 & 1 \end{bmatrix}$, looks nothing like a rotation: it stretches e_1 to the long vector $(25, 26)$. But $\det A = 25 \cdot 1 - 1 \cdot 26 = -1$, so it preserves area (and flips orientation). The parallelogram spanned by $(25, 26)$ and $(1, 1)$ is very long and very thin, with area exactly 1.

Map	Standard matrix	det
Rotation by θ	$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$	1
Reflection across the x -axis	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	-1
Shear	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$	1
Stretch one way, squeeze the other	$\begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$	1
Professor's example (p. 190)	$\begin{bmatrix} 25 & 1 \\ 26 & 1 \end{bmatrix}$	$25 - 26 = -1$

Remarks.

- Area-preserving does *not* mean shape-preserving: a shear turns the unit square into a slanted parallelogram, and the last example turns a circle into a long thin ellipse, all with the same area.
- The sign of $\det A$ also means something: $\det A = +1$ keeps orientation (counterclockwise stays counterclockwise), while $\det A = -1$ flips it, like a mirror.
- The area-preserving maps are closed under composition and inverses, since $|\det(AB)| = |\det A| |\det B| = 1$ and $|\det A^{-1}| = 1/|\det A| = 1$.

8 Main facts about determinants (p. 192)

The lecture ends with a list of the main facts about $\det$ from 10/8 and 10/13. Here they are, with the blank in item 5 filled in.

1. **Row operations.** Swapping two rows changes the sign; scaling a row by t multiplies $\det$ by t ; row replacement does not change $\det$.
2. **Cofactor expansion** works along any row or column.
3. **Products and transposes.** $\det(AB) = \det A \det B$, and $\det A^T = \det A$.
4. **Formulas.** Cramer's rule and $(A^{-1})_{ij} = C_{ji} / \det A$; $|\det A|$ is the factor by which $x \mapsto Ax$ scales volume.
- 5.

Fill-in: the determinant of a triangular matrix

$$\det \begin{bmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix} = a_{11}a_{22} \cdots a_{nn}, \text{ the product of the diagonal entries (10/8, p. 164).}$$

6. **Invertibility.** A^{-1} exists $\iff \det A \neq 0$.

Summary

Fact	Formula
$\det$ is linear in column i	$\det A_i(cx + dy) = c \det A_i(x) + d \det A_i(y)$
Cramer's rule	$x_i = \det A_i(b) / \det A$
Inverse	$(A^{-1})_{ij} = C_{ji} / \det A$
Area of parallelogram on v_1, v_2	$ \det[v_1 \ v_2] $
Area/volume scaling	$\text{vol } T(R) = \det A \text{vol } R$
Area-preserving maps of $\mathbb{R}^2$	$ \det A = 1$
Small $\det A$	inverse and solutions are very sensitive to noise

Corrections to the handwritten notes

Page	As written	Should be	Why
181	$B^{-1} = \frac{1}{\varepsilon^2} \begin{bmatrix} 1 + \varepsilon & -1 \\ -1 & 1 \end{bmatrix}$	$\frac{1}{\varepsilon^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 + \varepsilon^2 \end{bmatrix}$	check $BB^{-1} = I$; the $1 + \varepsilon^2$ goes bottom right
186	$\det A_i(e_j) = C_{ji}$ "by co-factor expansion along column j "	along column i	e_j sits in column i of $A_i(e_j)$
189	area = $ \det R \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} $	(same value)	$\det A^T = \det A$, so either way gives $ad - bc$