10/28

3.1 Motivation

- You have most likely seen and computed the determinant of a matrix before. Our main interest in it will be as a computational tool to check whether or not a matrix is invertible.

- Consider $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Recall A is invertible iff the RREF of A is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$:

  $\boxed{\text{Case 1}}$ $a \neq 0$. Then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow{\frac{1}{a}R1} \begin{pmatrix} 1 & b/a \\ c & d \end{pmatrix} \xrightarrow{R2-cR1} \begin{pmatrix} 1 & b/a \\ 0 & d-\frac{cb}{a} \end{pmatrix} \xrightarrow[\frac{1}{d-cb/a}R_2]{} \begin{pmatrix} 1 & b/a \\ 0 & 1 \end{pmatrix} \xrightarrow{R1-\frac{b}{a}R2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

  Thus the RREF of A is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ iff $d - \frac{cb}{a} \neq 0 \iff ad - bc \neq 0$.

  $\boxed{\text{Case 2}}$ $a = 0$ Then

$$\begin{pmatrix} 0 & b \\ c & d \end{pmatrix} \xrightarrow{R1 \leftrightarrow R2} \begin{pmatrix} c & d \\ 0 & b \end{pmatrix} \xrightarrow[\frac{1}{b}R2]{\frac{1}{c}R1} \begin{pmatrix} 1 & d/c \\ 0 & 1 \end{pmatrix} \xrightarrow{R1-\frac{d}{c}R2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

  Here the RREF of A is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ iff $c \neq 0$ and $b \neq 0$. Since $a = 0$ that is equivalent to

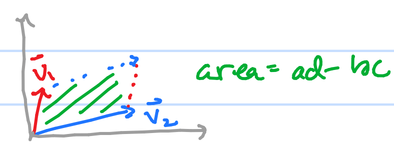
$$ad - bc = -bc \neq 0.$$

So in either case, it comes down to $ad - bc \neq 0$. We set

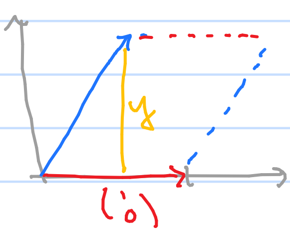
$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc.$$

- There is a geometric interpretation of this quantity. Let $\vec{v}_1 = \begin{pmatrix} a \\ c \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} b \\ d \end{pmatrix}$ so that $A = (\vec{v}_1, \vec{v}_2)$. These columns determine a parallelogram consisting of all vectors:

$$\vec{v} = t_1\vec{v}_1 + t_2\vec{v}_2 \qquad 0 \le t_1, t_2 \le 1.$$



$\boxed{\text{Ex}}$ Let $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\vec{v}_2 = \begin{pmatrix} x \\ y \end{pmatrix}$.
Note $\det\begin{pmatrix} 1 & x \\ 0 & y \end{pmatrix} = 1 \cdot y - x \cdot 0 = y$.



$\left.\begin{matrix} \text{height} = y \\ \text{base} = 1 \end{matrix}\right\} \Rightarrow \text{area} = y \cdot 1 = y$

- Thus we shall motivate our definition of the determinant by letting it be the area of the parallelogram determined by the columns of the matrix.

$\boxed{\text{Def}}$ For $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{F}^n$ the parallelepiped determined by these vectors is the set:

$$\{\vec{v} \in \mathbb{F}^n : \vec{v} = t_1\vec{v}_1 + \cdots + t_n\vec{v}_n \text{ for } 0 \le t_1, \dots, t_n \le 1\}.$$

- Notation: the determinant of $\vec{v}_1, \dots, \vec{v}_n$ is $D(\vec{v}_1, \dots, \vec{v}_n) := \det(A) = |A|$ where $A = (\vec{v}_1 \cdots \vec{v}_n)$.

3.2 Properties the determinant should have

Using our geometric interpretation of $D(\vec{v}_1, \dots, \vec{v}_n)$, we shall explore what properties this quantity ought to have. These properties will be sufficient to actually define $D(\vec{v}_1, \dots, \vec{v}_n)$ in Section 3.3.

In the parallelepiped determined by $\vec{v}_1, \dots, \vec{v}_n$, we can allow any of the vectors to determine the "height". The remaining $n-1$ vectors will then determine the "base", which together with the height determines the volume.

Linearity in each argument

- For any $\alpha > 0$ and any $k = 1, \dots, n$

$$D(\vec{v}_1, \dots, \alpha \vec{v}_k, \dots, \vec{v}_n) = \alpha D(\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{v}_n)$$

This is just saying when we let $\vec{v}_k$ be the height, scaling the height by α causes the volume to be scaled by α. However $\alpha \cdot \vec{v}_k$ makes sense for any scalar α, so we should allow arbitrary scalars. This means $D(\cdot)$ is not always positive but we can accept that (its sign will tell us about the orientation of the parallelepiped)

- For any $k = 1, \dots, n$

$$D(\vec{v}_1, \dots, \vec{v}_k + \vec{w}_k, \dots, \vec{v}_n) = D(\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{v}_n) + D(\vec{v}_1, \dots, \vec{w}_k, \dots, \vec{v}_n)$$

That is, a parallelepiped with combined height $\vec{v}_k + \vec{w}_k$ has volume equal to the sum of the volumes of the parallelepipeds with heights $\vec{v}_k$ and $\vec{w}_k$.

- Combining these two observations yields:

① $$D(\vec{v}_1, \dots, \alpha \vec{v}_k + \beta \vec{w}_k, \dots, \vec{v}_n) = \alpha D(\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{v}_n) + \beta D(\vec{v}_1, \dots, \vec{w}_k, \dots, \vec{v}_n)$$

for all $k = 1, \dots, n$ and all scalars α, β. That is, D is <u>linear</u> in each entry. In particular, scaling a column (column operation of the first type) scales D.

Preservation under column replacement

- If $\vec{v}_j = \vec{v}_k$ for any $j \neq k$, then the parallelepiped is "flat" and has volume zero. So we should expect

$$D(\vec{v}_1, \dots, \underbrace{\vec{v}_k}_{j}, \dots, \underbrace{\vec{v}_k}_{k}, \dots, \vec{v}_n) = 0$$

Combining this with ① gives:

② $$D(\vec{v}_1, \dots, \underbrace{\vec{v}_j + \alpha \vec{v}_k}_{j}, \dots, \underbrace{\vec{v}_k}_{k}, \dots, \vec{v}_n) = D(\vec{v}_1, \dots, \underbrace{\vec{v}_j}_{j}, \dots, \underbrace{\vec{v}_k}_{k}, \dots, \vec{v}_n)$$

for any $j, k = 1, \dots, n$ and any scalar α. That is, D is unchanged by <u>column operations</u> of the third type.

Antisymmetry

The next property is:

3 $\quad D(\vec{v}_1, \dots, \underbrace{\vec{v}_k}_{j}, \dots, \underbrace{\vec{v}_j}_{k}, \dots, \vec{v}_n) = -D(\vec{v}_1, \dots, \underbrace{\vec{v}_j}_{j}, \dots, \underbrace{\vec{v}_k}_{k}, \dots, \vec{v}_n)$

for any distinct $j, k = 1, \dots, n$. That is, swapping columns (column op. of the second type) yields a negative sign.

This follows from 1 and 2:

$$\begin{aligned}
D(\vec{v}_1, \dots, \vec{v}_j, \dots, \vec{v}_k, \dots, \vec{v}_n) &\overset{2}{=} D(\vec{v}_1, \dots, \vec{v}_j, \dots, \vec{v}_k - \vec{v}_j, \dots, \vec{v}_n) \\
&\overset{2}{=} D(\vec{v}_1, \dots, \vec{v}_j + (\vec{v}_k - \vec{v}_j), \dots, \vec{v}_k - \vec{v}_j, \dots, \vec{v}_n) \\
&= D(\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{v}_k - \vec{v}_j, \dots, \vec{v}_n) \\
&\overset{2}{=} D(\vec{v}_1, \dots, \vec{v}_k, \vec{v}_k - \vec{v}_j - \vec{v}_k, \dots, \vec{v}_n) \\
&= D(\vec{v}_1, \dots, \vec{v}_k, -\vec{v}_j, \dots, \vec{v}_n) \\
&\overset{1}{=} -D(\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{v}_j, \dots, \vec{v}_n)
\end{aligned}$$

- At this point, we know how D behaves under all 3 column operations.

Normalization

The parallelepiped determined by $\vec{e}_1, \dots, \vec{e}_n$ is the n-dimensional unit cube, which has volume 1.

4 $\quad D(\vec{e}_1, \dots, \vec{e}_n) = 1$

3.3 Constructing the determinant

Recall the four basic properties of the determinant we obtained last time:

1. Linear in each entry. In particular:
$$D(\vec{v}_1,\dots,\alpha\vec{v}_k,\dots,\vec{v}_n) = \alpha D(\vec{v}_1,\dots,\vec{v}_k,\dots,\vec{v}_n)$$
2. Invariant under column replacement
3. Antisymmetric
4. $D(\vec{e}_1,\dots,\vec{e}_n)=1$

These properties are enough for us to compute the determinant.

EX 1. ($a\neq 0$)
$$D\left(\begin{pmatrix}a\\b\end{pmatrix},\begin{pmatrix}c\\d\end{pmatrix}\right) = \left|\begin{pmatrix}a & b\\ c & d\end{pmatrix}\right| = a\left|\begin{pmatrix}1 & b\\ c/a & d\end{pmatrix}\right| = a\left|\begin{pmatrix}1 & 0\\ c/a & d-\frac{bc}{a}\end{pmatrix}\right| = a\left(d-\frac{bc}{a}\right)\left|\begin{pmatrix}1 & 0\\ c/a & 1\end{pmatrix}\right|$$
$$= a\left(d-\frac{bc}{a}\right)\left|\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right| = a\left(d-\frac{bc}{a}\right)\cdot 1 = ab-cd.$$

2.
$$\left|\begin{pmatrix}1 & 2 & 0\\ 0 & -2 & 3\\ 1 & 0 & 2\end{pmatrix}\right| = \left|\begin{pmatrix}1 & 0 & 0\\ 0 & -2 & 3\\ 1 & -2 & 2\end{pmatrix}\right| = -2\cdot\left|\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 3\\ 1 & 1 & 2\end{pmatrix}\right| = -2\cdot\left|\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 1 & -1\end{pmatrix}\right| = 2\left|\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 1 & 1\end{pmatrix}\right|$$
$$= 2\left|\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\right| = 2\cdot 1 = 2.$$
□

• From the above properties, we can deduce more:

Prop 3.1 For $A\in M_{n\times n}$ a square matrix, the following hold:

1. If A has a column of all zeros, then $\det(A)=0$.
2. If A has two equal columns, then $\det(A)=0$.
3. If one column of A is a multiple of another, then $\det(A)=0$.
4. If the columns of A are linearly dependent (equivalently, if A is not invertible) then $\det(A)=0$.

Proof Let $\vec{v}_1,\dots,\vec{v}_n$ be the columns of A, so that $\det(A)=D(\vec{v}_1,\dots,\vec{v}_n)$.

1: Suppose $\vec{v}_k=\vec{0}$. Then using 1 above we have:
$$\det(A)=D(\vec{v}_1,\dots,\vec{v}_k,\dots,\vec{v}_n) = D(\vec{v}_1,\dots,0\cdot\vec{v}_k,\dots,\vec{v}_n) = 0\cdot D(\vec{v}_1,\dots,\vec{v}_k,\dots,\vec{v}_n)=0.$$

2: Suppose $\vec{v}_j=\vec{v}_k$ for $j\neq k$. Then using 3:
$$\det(A)=D(\vec{v}_1,\dots,\vec{v}_j,\dots,\vec{v}_k,\dots,\vec{v}_n) = -D(\vec{v}_1,\dots,\vec{v}_k,\dots,\vec{v}_j,\dots,\vec{v}_n) = -\det(A).$$
So $\det(A)=-\det(A)$, which implies $\det(A)=0$.

3: Suppose $\vec{v}_j=\alpha\vec{v}_k$ for some $j\neq k$ and some scalar α. Then by 1 and 2
$$\det(A)= D(\vec{v}_1,\dots,\alpha\vec{v}_k,\dots,\vec{v}_k,\dots,\vec{v}_n) \overset{1}{=} \alpha D(\vec{v}_1,\dots,\vec{v}_k,\dots,\vec{v}_k,\dots,\vec{v}_n) \overset{2}{=} \alpha\cdot 0=0.$$

4: Suppose $\vec{v}_1,\dots,\vec{v}_n$ are linearly dependent. Then one of the vectors can be written as a linear comb. of the others, say
$$\vec{v}_k=\sum_{j\neq k}\alpha_j\vec{v}_j$$

So using ① and ②

$$\det(A) = D(\vec{v}_1,\dots,\vec{v}_k,\dots,\vec{v}_n) = D(\vec{v}_1,\dots,\sum_{j\neq k}\alpha_j\vec{v}_j,\dots,\vec{v}_n)$$

$$\overset{①}{=} \sum_{j\neq k}\alpha_j\, D(\vec{v}_1,\dots,\vec{v}_j,\dots,\vec{v}_n)$$

$$\overset{②}{=} \sum_{j\neq k}\alpha_j\cdot 0 = 0.$$

□

- Next, we generalize ② slightly:

Prop 3.2 For each $k=1,\dots,n$,

$$D(\vec{v}_1,\dots,\vec{v}_k+\sum_{j\neq k}\alpha_j\vec{v}_j,\dots,\vec{v}_n) = D(\vec{v}_1,\dots,\vec{v}_k,\dots,\vec{v}_n)$$

for any scalars $\alpha_1,\dots,\alpha_{j-1},\alpha_{j+1},\dots,\alpha_n$.

Proof Denote

$$\vec{u} = \sum_{j\neq k}\alpha_j\vec{v}_j$$

and observe that $\vec{v}_1,\dots,\vec{v}_{j-1},\vec{u},\vec{v}_{j+1},\dots,\vec{v}_n$ are lin. dep. (since $\vec{u}$ is a lin. comb. of the other vectors). Now, by ①

$$D(\vec{v}_1,\dots,\vec{v}_k+\vec{u},\dots,\vec{v}_n) = D(\vec{v}_1,\dots,\vec{v}_k,\dots,\vec{v}_n) + D(\vec{v}_1,\dots,\vec{u},\dots,\vec{v}_n)$$

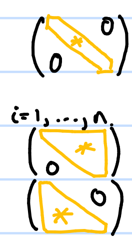
and the second term on the right vanishes by ④ in the previous proposition. □

11/2

Determinants of triangular matrices

We consider a special class of matrices for which we can easily compute the determinant.

Def A square matrix $A=(a_{ij})\in M_{n\times n}$ is called:



- *diagonal* if $a_{ij}=0$ for all $i\neq j$.
  In this case we write $A=\mathrm{diag}(a_1,a_2,\dots,a_n)$ where $a_i := a_{i,i}$ for $i=1,\dots,n$.
- *upper triangular* if $a_{ij}=0$ for all $i>j$.
- *lower triangular* if $a_{ij}=0$ for all $i<j$.
- *triangular* if it is upper or lower triangular (or both).

- Observe that a diagonal matrix is upper and lower triangular.

Prop 3.3 Let $A=(a_{ij})\in M_{n\times n}$ be a triangular matrix. Then

$$\det(A) = a_{1,1}a_{2,2}\cdots a_{n,n}.$$

Proof First suppose at least one of the diagonal entries of A is zero. Then the right-hand side is zero, so it remains in this case to show $\det(A)=0$. Observe that, because A is triangular with at least one diagonal entry zero, its REF does not have a pivot in every column. Hence A is not invertible, and so by ④ from Proposition 3.1, we have $\det(A)=0$.

Next, assume all diagonal entries of A are non-zero. Let $\tilde{A} \in M_{n\times n}$ be the matrix obtained from A after dividing column j by a_{jj} for each $j=1,\dots,n$. Then the diagonal entries of $\tilde{A}$ are all 1's and by ①

$$\det(A) = a_{1,1}a_{2,2}\cdots a_{n,n}\det(\tilde{A})$$

Now, using only column replacement one can turn $\tilde{A}$ into I_n. So by ② and ④ we have:

$$\det(\tilde{A}) \overset{②}{=} \det(I_n) \overset{④}{=} 1$$

Therefore $\det(A) = a_{1,1}a_{2,2}\cdots a_{n,n}$ as claimed. $\square$

Prop 3.4 For a square matrix $A \in M_{n\times n}$, A is invertible if and only if $\det(A) \neq 0$. Equivalently, is A is not invertible if and only if $\det(A) = 0$.

Proof ($\Longrightarrow$) Assume A is invertible. Then A^T is invertible with inverse $(A^T)^{-1} = (A^{-1})^T$. Thus the RREF of A^T is I_n. The row operations that turn A^T into I_n correspond to column operations that turn A into $I_n^T = I_n$. By ①–③, these operations don't change whether or not the determinant is zero, and by ④ it is not zero. Thus $\det(A) \neq 0$.

($\Longleftarrow$) This direction is equivalent to its contrapositive: if A is not invertible, then $\det(A)=0$. But this is precisely ④ from the first proposition in this section. $\square$

- It might seem like the determinant is fully defined at this point: do column operations on A until it is triangular, keeping track of how it affects the determinant through ①–③, then we then take the product of the diagonal entries. The problem is that there are multiple ways to go from A to a triangular matrix via column operations, and we do not know if the all give the same answer. So we do not know yet if $\det(A)$ is "well-defined". To see that it is in fact well-defined will require breaking things down to elementary matrices.

Determinants and transposes, products, and elementary matrices

In this section we will see how the determinant interacts with transposes, and—more importantly—products. We first require a lemma.

Lemma 3.5 For $A \in M_{n\times n}$ and an elementary matrix $E \in M_{n\times n}$

$$\det(AE) = \det(A)\det(E).$$

Proof Recall that AE is the matrix obtained by A after doing a column operation. Also recall elementary matrices come in three types:

$$\begin{array}{cc} & \begin{array}{ccccc} & j & & k & \end{array} \\ \begin{array}{c} \\ j \\ \\ k \\ \\ \end{array} & \begin{pmatrix} 1 & & & & \\ & 0 & \cdots & 1 & \\ & \vdots & & \vdots & \\ & 1 & \cdots & 0 & \\ & & & & 1 \end{pmatrix} \end{array} \quad j \neq k \qquad \begin{array}{cc} & \begin{array}{ccc} & j & \end{array} \\ \begin{array}{c} \\ j \\ \\ \end{array} & \begin{pmatrix} 1 & & \\ & \alpha & \\ & & 1 \end{pmatrix} \end{array} \quad \alpha \neq 0 \qquad \begin{array}{cc} & \begin{array}{ccccc} & j & & k & \end{array} \\ \begin{array}{c} \\ j \\ \\ k \\ \\ \end{array} & \begin{pmatrix} 1 & & & & \\ & 1 & \cdots & 0 & \\ & \vdots & & \vdots & \\ & \alpha & \cdots & 1 & \\ & & & & 1 \end{pmatrix} \end{array} \quad \alpha \text{ any scalar}$$

type 1   type 2   type 3

If E is type 1, it is the identity matrix with two columns swapped. So by ③ and ④

$\det(E) = -\det(I_n) = -1$. At the same time, AE is the matrix obtained from A after swapping two columns. So by ③ again

$$\det(AE) = -\det(A) = \det(A)\cdot(-1) = \det(A)\det(E).$$

If E is type 2, it is diagonal and so $\det(E) = \alpha$. Also AE is the matrix obtained from A after multiplying a column by α. So by ① we have

$$\det(AE) = \alpha\det(A) = \det(A)\cdot\alpha = \det(A)\det(E).$$

Finally, if E is type 3, it is triangular and so $\det(E) = 1$. AE is the matrix obtained from A after replacing column j with its sum with $\alpha\cdot$ column k. So by ② we have

$$\det(AE) = \det(A) = \det(A)\cdot 1 = \det(A)\det(E). \quad \square$$

• By iterating the previous lemma d times we obtain:

Cor 3.6 For $A \in M_{n\times n}$ and any elementary matrices $E_1, \dots, E_d \in M_{n\times n}$ we have

$$\det(AE_1\cdots E_d) = \det(A)\det(E_1\cdots E_d).$$

Thm 3.7 For any $A, B \in M_{n\times n}$, $\det(AB) = \det(A)\det(B)$.

Proof We first claim that AB is invertible iff A and B are invertible. Indeed, in Section 1.6 we showed if A and B are invertible, then so is AB with $(AB)^{-1} = B^{-1}A^{-1}$. Conversely, if AB is invertible, then by Exercise 3 on Homework 4, A is right invertible and B is left invertible. But A and B are square matrices and we showed in Section 2.3 that this implies they are both invertible.

Now, our claim implies that if either A or B is not invertible, then neither is AB. But in this case

$$\det(AB) = \det(A)\det(B)$$

holds because both sides are zero.

So it remains to consider the case when both A and B are invertible. Since B is invertible, then its RREF is I_n. It follows that $B = E_1\cdots E_d\cdot I_n = E_1\cdots E_d$ where $E_1, \dots, E_d$ are elementary matrices. So by the Corollary (applied twice) we have:

$$\det(AB) = \det(AE_1\cdots E_d) = \det(A)\det(E_1)\cdots\det(E_d) = \det(A)\det(E_1\cdots E_d) = \det(A)\det(B). \quad \square$$

Thm 3.8 For $A \in M_{n\times n}$, $\det(A) = \det(A^T)$

Proof We know A is invertible iff A^T is, so if A is not invertible then neither is A^T and hence:

$$\det(A) = 0 = \det(A^T).$$

Now suppose A is invertible, then as in the proof of the previous theorem $A = E_1\cdots E_d$ for $E_1, \dots, E_d$ elementary matrices. For any elementary matrix E, E^T is also elementary and one can directly check that $\det(E) = \det(E^T)$. So using the corollary twice we have:

$$\det(A) = \det(E_1\cdots E_d) = \det(E_1)\cdots\det(E_d) = \det(E_1^T)\cdots\det(E_d^T) = \det(E_d^T)\cdots\det(E_1^T)$$
$$= \det(E_d^T\cdots E_1^T) = \det((E_1\cdots E_d)^T) = \det(A^T). \quad \square$$

• One benefit of these theorems is that we now know how the determinant behaves under row operations:

$$\det(EA) = \det(E)\det(A)$$

Namely, the same way as under column operations.

Also $\det(A^T) = \det(A)$ means any result for the determinant relating to columns now also holds for rows.

• We conclude with a summary of all the properties we obtained in this section:

$\boxed{\text{Summary}}$ Let $A \in M_{n\times n}$ be a square matrix. Then

1. The determinant is linear in each row (column) so long as the other rows (columns) are fixed.
2. Interchanging two rows (columns) changes the sign of the determinant.
3. The determinant of a triangular (in particular, diagonal) matrix is the product of its diagonal entries.
4. If A has a zero row (column), then $\det(A)=0$.
5. If A has two identical rows (columns), then $\det(A)=0$.
6. If the rows (columns) are linearly dependent, then $\det(A)=0$.
7. If A is invertible, then $\det(A)\neq 0$.
8. If A is not invertible, then $\det(A)=0$.
9. The determinant is invariant under row (column) replacement operations.
10. $\det(A^T) = \det(A)$
11. $\det(AB) = \det(A)\det(B)$ for any square matrix $B \in M_{n\times n}$
12. $\det(\alpha A) = \alpha^n \det(A)$ for any scalar α.

3.4 Formal definition of the determinant.

We will show that a function satisfying the properties from last section not only exists but is unique. In particular, we will only need properties 1, 3, and 4 from the beginning of Section 3.3.

- Fix a matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \in M_{n\times n}$$

Denote its columns by $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \in F^n$. Observe that for each $j = 1, \dots, n$

$$\vec{v}_j = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix} = a_{1,j}\vec{e}_1 + a_{2,j}\vec{e}_2 + \dots + a_{n,j}\vec{e}_n = \sum_{i=1}^{n} a_{i,j}\vec{e}_i$$

Now, using linearity of D in each coordinate we get:

$$\begin{aligned} \det(A) = D(\vec{v}_1, \dots, \vec{v}_n) &= D\Big(\sum_{i_1=1}^{n} a_{i_1,1}\vec{e}_{i_1}, \vec{v}_2, \dots, \vec{v}_n\Big) \\ &= \sum_{i_1=1}^{n} a_{i_1,1}\, D(\vec{e}_{i_1}, \vec{v}_2, \dots, \vec{v}_n) \\ &= \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} a_{i_1,1}\, a_{i_2,2}\, D(\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{v}_n) \\ &\ \ \vdots \\ &= \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_n=1}^{n} a_{i_1,1}\, a_{i_2,2}\cdots a_{i_n,n}\, D(\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{e}_{i_n}) \end{aligned}$$

However, as $i_1, i_2, \dots, i_n$ vary, we will obtain a lot of repeated indices. But we know $D=0$ when two entries are the same. So in this giant sum, only terms corresponding to distinct $i_1, i_2, \dots, i_n$ will give us a non-zero value. We can think of such a choice as a function:

$$\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$$
$$\sigma(k) = i_k$$

Then the condition that $i_k \neq i_\ell$ for $k \neq \ell$ says precisely that σ is injective. Moreover, since the domain and codomain are the same size, σ is necessarily surjective and consequently bijective.

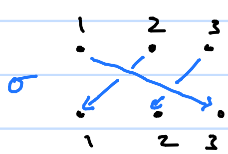
Def A bijective map

$$\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$$

is called a <u>permutation</u>. The collection of such maps is denoted $\mathrm{Perm}(n)$

Ex $\sigma : \{1,2,3\} \to \{1,2,3\}$

$\sigma(1) = 3$, $\sigma(2) = 1$, $\sigma(3) = 2$



$\sigma : 3, 1, 2$

Continuing our computation above with $\sigma : i_1, i_2, \dots, i_n$ we can write:

$$\det(A) = D(\vec{v}_1, \dots, \vec{v}_n) = \sum_{\sigma \in \mathrm{Perm}(n)} a_{\sigma(1),1}\, a_{\sigma(2),2} \cdots a_{\sigma(n),n}\, D(\vec{e}_{\sigma(1)}, \vec{e}_{\sigma(2)}, \dots, \vec{e}_{\sigma(n)})$$

Now the matrix $(\vec{e}_{\sigma(1)}\, \vec{e}_{\sigma(2)} \cdots \vec{e}_{\sigma(n)})$ can be turned into I_n using only column swaps. Thus $D(\vec{e}_{\sigma(1)}, \vec{e}_{\sigma(2)}, \dots, \vec{e}_{\sigma(n)}) = \pm 1$, depending on the number of swaps.

11/4

Def For $\sigma \in \mathrm{Perm}(n)$, a <u>disorder</u> of σ is a pair (i,j) with $i,j \in \{1,2,\dots,n\}$ such that $i<j$ while $\sigma(i) > \sigma(j)$. Let K be the number of disorders of σ. Then the <u>sign</u> of σ is

$$\mathrm{sign}(\sigma) := (-1)^K.$$

We say σ is <u>even</u> if $\mathrm{sign}(\sigma)=1$, and say σ is <u>odd</u> if $\mathrm{sign}(\sigma)=-1$.

Lemma 3.9 For $\sigma \in \mathrm{Perm}(n)$, let s be the number of swaps needed to rearrange $\sigma(1), \sigma(2), \dots, \sigma(n)$ into $1, 2, \dots, n$. Then $\mathrm{sign}(\sigma) = (-1)^s$.

Proof We first show that if $\tau \in \mathrm{Perm}(n)$, then the number of swaps needed to change from σ to τ is even if they have the same sign, and odd otherwise. First notice that swapping two <u>adjacent</u> entries changes the number of disorders by exactly one, and so changes the sign. More generally, swapping any two entries requires an <u>odd</u> number of adjacent swaps:

$$i, i+1, i+2, \dots, j-1, j \to i+1, i, i+2, \dots, j-1, j \to i+1, i+2, i, \dots, j-1, j \dashrightarrow i+1, i+2, \dots, j-1, i, j \to i+1, i+2, \dots, j-1, j, i \qquad i-j \text{ swaps}$$

$$\to i+1, i+2, \dots, j, j-1, i \dashrightarrow i+1, i+2, j, \dots, j-1, i \to i+1, j, i+2, \dots, j-1, i \to j, i+1, i+2, \dots, j-1, i \qquad i-j-1 \text{ swaps}$$

$2(i-j)-1$ swaps total

Thus swapping any two indices changes the sign. So if $\mathrm{sign}(\sigma) = \mathrm{sign}(\tau)$, then the number of swaps needed must be even. If $\mathrm{sign}(\sigma) \neq \mathrm{sign}(\tau)$, then the number of swaps needed is odd.
Now, observe that the trivial permutation $e: 1, 2, \dots, n$ is even. So if σ is even then s must be even and therefore $(-1)^s = 1 = \mathrm{sign}(\sigma)$. If σ is odd, then s is odd and therefore $(-1)^s = -1 = \mathrm{sign}(\sigma)$. $\square$

This lemma implies

$$D(\vec{e}_{\sigma(1)}, \dots, \vec{e}_{\sigma(n)}) = (-1)^s = \mathrm{sign}(\sigma).$$

So we can conclude our computation, the result of which we state as a theorem

Thm 3.10 Let $A = (a_{ij}) \in M_{n \times n}$. Then

$$\det(A) = \sum_{\sigma \in \mathrm{Perm}(n)} a_{\sigma(1),1}\, a_{\sigma(2),2} \cdots a_{\sigma(n),n}\, \mathrm{sign}(\sigma).$$

In particular, the determinant is well defined.

Ex For $A = I_n$, $a_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}$. Thus $a_{\sigma(1),1}\, a_{\sigma(2),2} \cdots a_{\sigma(n),n} = \begin{cases} 1 & \text{if } \sigma = e \\ 0 & \text{otherwise.} \end{cases}$
Therefore

$$\det(I_n) = \sum_{\sigma \in \mathrm{Perm}(n)} a_{\sigma(1),1}\, a_{\sigma(2),2} \cdots a_{\sigma(n),n}\, \mathrm{sign}(\sigma) = 1 \cdot \mathrm{sign}(e) = 1.$$

Exercise: Verify the other properties of the determinant directly from the formula in the theorem.

3.5 Cofactor Expansion

- In this section, we derive some additional formulas for computing the determinant.

- Given a matrix $A \in M_{n\times n}$, for each $i,j=1,\dots,n$ let $A_{i,j}$ be the $(n-1)\times(n-1)$ matrix obtained from A after removing row i and column j.

Thm 3.11 Let $A \in M_{n\times n}$. Then for each $i=1,\dots,n$ the determinant of A can be expanded along the ith row:

$$\det(A) = (-1)^{i+1} a_{i,1}\det(A_{i,1}) + (-1)^{i+2} a_{i,2}\det(A_{i,2}) + \cdots + (-1)^{i+n} a_{i,n}\det(A_{i,n})$$
$$= \sum_{j=1}^{n} (-1)^{i+j} a_{i,j}\det(A_{i,j}).$$

Similarly, for each $j=1,\dots,n$ the determinant of A can be expanded along the jth column:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j}\det(A_{i,j}).$$

- Before we prove the theorem, let's do a few examples:

Ex $A = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 2 & 1 \\ 1 & -3 & 1 \end{pmatrix}$

Expanding along first row:

$$\det(A) = (-1)^{1+1}\, 1 \det\begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix} + (-1)^{1+2}\cdot 0 \cdot \det\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} + (-1)^{1+3}\cdot(-1)\det\begin{pmatrix} 2 & 2 \\ 1 & -3 \end{pmatrix}$$
$$= 1\cdot(2-(-3)) + 0 + (-1)\cdot(-6-2) = 5+8 = 13$$

Expanding along first column:

$$\det(A) = (-1)^{1+1}\, 1 \det\begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix} + (-1)^{2+1}\, 2 \det\begin{pmatrix} 0 & -1 \\ -3 & 1 \end{pmatrix} + (-1)^{3+1}\cdot 1 \cdot \det\begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix}$$
$$= 1\cdot(2-(-3)) + (-2)(0-3) + 1\cdot(0-(-2)) = 5+6+2 = 13$$

Expanding along the second column:

$$\det(A) = (-1)^{1+2}\cdot 0\cdot\det\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} + (-1)^{2+2}\cdot 2\cdot\det\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} + (-1)^{3+2}(-3)\det\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$$
$$= 0 + 2\cdot(1-(-1)) + 3\cdot(1-(-2)) = 4+9 = 13. \quad \square$$

Proof (Thm 3.11) We first claim it suffices check for expanding along the first column:

$$* \quad \det(A) \overset{?}{=} \sum_{i=1}^{n} (-1)^{i+1} a_{i,1}\det(A_{i,1}).$$

Indeed, once we establish this formula then we can get the expansion along the second column by swapping the 1st and 2nd columns of A. The additional factor of (-1) in the formula comes from this column operation. By swapping the 2nd and 3rd columns, we get the expansion along the 3rd column and so on. To get the expansion along the rows, we replace A with A^T and use $\det(A) = \det(A^T)$. So it indeed suffices to verify $(*)$.

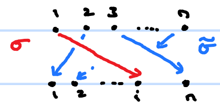
Recall the formula we derived last section:

$$\det(A) = \sum_{\sigma \in \mathrm{Perm}(n)} a_{\sigma(1),1}\, a_{\sigma(2),2} \cdots a_{\sigma(n),n}\, \mathrm{sign}(\sigma).$$

We will break this sum up by partitioning the $\sigma \in \mathrm{Perm}(n)$ according to the value of $\sigma(1)$:

$$\det(A) = \sum_{i=1}^{n} \sum_{\substack{\sigma \in \mathrm{Perm}(n) \\ \sigma(1)=i}} a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} \operatorname{sign}(\sigma)$$

$$= \sum_{i=1}^{n} a_{i,1} \sum_{\substack{\sigma \in \mathrm{Perm}(n) \\ \sigma(1)=i}} a_{\sigma(2),2} \cdots a_{\sigma(n),n} \operatorname{sign}(\sigma)$$



Fix $\sigma \in \mathrm{Perm}(n)$ with $\sigma(1)=i$. We want to consider the $\tilde{\sigma} \in \mathrm{Perm}(n-1)$ "induced" by σ. To get the indices correct, we need to do some shifting. For $j=1,\dots,n$ define

$$\tilde{\sigma}(j) := \begin{cases} \sigma(j+1) & \text{if } \sigma(j+1) < i \\ \sigma(j+1)-1 & \text{if } \sigma(j+1) > i \end{cases}$$

Then $\tilde{\sigma} \in \mathrm{Perm}(n-1)$ and by counting disorders we see

$$\operatorname{sign}(\sigma) = (-1)^{i-1} \operatorname{sign}(\tilde{\sigma}) = (-1)^{i+1} \operatorname{sign}(\tilde{\sigma})$$

As we vary over $\sigma \in \mathrm{Perm}(n)$ with $\sigma(1)=i$, $\tilde{\sigma}$ varies over all $\mathrm{Perm}(n-1)$. Also, if $A_{i,1}$ has entries (b_{jk}), then 11/9

$$a_{\sigma(2),2} \cdots a_{\sigma(n),n} = b_{\tilde{\sigma}(1),1} \cdots b_{\tilde{\sigma}(n-1),n-1}$$

Thus

$$\sum_{\substack{\sigma \in \mathrm{Perm}(n) \\ \sigma(1)=i}} a_{\sigma(2),2} \cdots a_{\sigma(n),n} \operatorname{sign}(\sigma) = \sum_{\tilde{\sigma} \in \mathrm{Perm}(n-1)} b_{\tilde{\sigma}(1),1} \cdots b_{\tilde{\sigma}(n-1),n-1} (-1)^{i+1} \operatorname{sign}(\tilde{\sigma})$$

$$\text{(Thm 3.10)} = (-1)^{i+1} \det(A_{i,1}).$$

Since this holds for each $i=1,\dots,n$, $(*)$ follows. $\square$

Def For $A \in M_{n,n}$, the numbers

$$C_{ij} = (-1)^{i+j} \det(A_{ij}) \qquad i,j = 1,\dots,n.$$

are called the <u>cofactors</u> of A. The matrix $C = (C_{ij}) \in M_{n \times n}$ is called the <u>cofactor matrix</u> of A.

Exercise Verify the properties of the determinant using any cofactor expansion:

$$\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij} \quad \text{for any } j=1,\dots,n \text{ or}$$

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij} \quad \text{for any } i=1,\dots,n.$$

Remark The cofactor expansion works great for 3×3 or 4×4 matrices, but as the matrices get larger this method becomes slower than using row operations to obtain a triangular matrix. For example, cofactor expansion on a 20×20 matrix would take a modern computer (i.e. with an Intel Core i7 7500 u) 8.47×10^7 seconds $\approx$ 2.686 years, whereas row reduction would take a fraction of a second.

Cofactor inversion formula

Thm 3.12 Let $A \in M_{n\times n}$ and let $C \in M_{n\times n}$ be its cofactor matrix. If A is invertible, then
$$A^{-1} = \frac{1}{\det(A)} C^T.$$

Proof Let us compute the entries of AC^T:
$$[AC^T]_{ij} = \sum_{k=1}^{n} (A)_{i,k} (C^T)_{kj} = \sum_{k=1}^{n} a_{i,k} C_{j,k}$$
If $i=j$, then by the cofactor expansion formula it equals $\det(A)$. If $i \neq j$, then let $\tilde{A}$ be the matrix obtained from A after changing row j to equal row i. Since all other rows of $\tilde{A}$ agree with A, we have $\tilde{A}_{j,k} = A_{j,k}$ for each $k=1,\ldots,n$. Consequently
$$C_{j,k} = (-1)^{j+k}\det(A_{j,k}) = (-1)^{j+k}\det(\tilde{A}_{j,k}).$$
Also $a_{i,k} = [\tilde{A}]_{j,k}$ for each $k=1,\ldots,n$. Thus
$$(AC^T)_{i,j} = \sum_{k=1}^{n} a_{i,k} C_{j,k} = \sum_{k=1}^{n} (\tilde{A})_{j,k} (-1)^{j+k}\det(\tilde{A}_{j,k}) = \det(\tilde{A})$$
by the cofactor expansion formula. But $\tilde{A}$ has two identical rows (j and i), and therefore $\det(\tilde{A})=0$. Thus AC^T is a diagonal matrix, namely
$$AC^T = \det(A)\cdot I_n$$
Since A is invertible, $\det(A)\neq 0$ and so $A\left(\frac{1}{\det(A)}C^T\right) = I_n$. Since A is square, this implies $A^{-1} = \frac{1}{\det(A)}C^T$. $\square$

- This formula also gives us a new way to solve linear systems:

Cor 3.13 (Cramer's rule) For an invertible matrix A, the unique solution to the linear system $A\vec{x} = \vec{b}$ has entries
$$x_j = \frac{\det(B_j)}{\det(A)} \qquad j=1,\ldots,n$$
where B_j is the matrix obtained from A after replacing column j with $\vec{b}$.

Proof Using the cofactor inversion formula, the unique solution is $\vec{x} = \frac{1}{\det(A)} C^T\vec{b}$. Its jth coordinate is
$$x_j = \frac{1}{\det(A)} \sum_{i=1}^{n} (C^T)_{j,i}\, b_i = \frac{1}{\det(A)} \sum_{i=1}^{n} b_i\, C_{i,j}$$
Now, because the columns of B_j and A agree (aside from column j), $(B_j)_{i,j} = A_{i,j}$ for each $i=1,\ldots,n$ and therefore
$$C_{i,j} = (-1)^{i+j}\det(A_{i,j}) = (-1)^{i+j}\det((B_j)_{i,j}).$$
Also, $b_i = (B_j)_{i,j}$ for each $i=1,\ldots,n$ by definition of B_j. Thus
$$x_j = \frac{1}{\det(A)} \sum_{i=1}^{n} (B_j)_{i,j} (-1)^{i+j}\det((B_j)_{i,j}) = \frac{\det(B_j)}{\det(A)}$$
by the cofactor inversion formula. $\square$

Applications

- Consider $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2}$. Then $\det(A) = ad - bc$. Suppose A is invertible so that $ad - bc \neq 0$. Now
$$A_{1,1} = (d) \qquad A_{1,2} = (c) \qquad A_{2,1} = (b) \qquad A_{2,2} = (a)$$
which all have determinants equal to their only entry. Thus the cofactor matrix is:
$$C = \begin{pmatrix} (-1)^{1+1}\cdot d & (-1)^{1+2}\cdot c \\ (-1)^{2+1}\cdot b & (-1)^{2+2}\cdot a \end{pmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$
Then the cofactor inversion formula implies:
$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

- Suppose $A \in M_{n\times n}$ is invertible with all entries integers. Then the entries of its cofactor matrix will be integers. If we also know $\det(A) = 1$, then the cofactor inversion formula says A^{-1} has integer entries. Otherwise, they will be fractions with a common denominator of $\det(A)$.