Linear Algebra
Week 11

G-07

05 XII 2024

1 The Determinant

In one sentence, the determinant of a matrix is a real number that corresponds to the volume of the unit cube after applying the linear transformation represented by the matrix. If the matrix flips the unit cube then the sign is negative. Similar to the script, this section will be less proof based and I am going to try to provide the intuition behind the properties of the determinant.

In 2×2 case, we calculate the determinant as follows:

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

This becomes obvious when you calculate the area of the image of the unit square by multiplication with the matrix. You can see the image in the lecture notes or watch the whole 3Blue1Brown video. We directly move to the $n \times n$ case before we discuss the properties of the determinant.

1.1 Permutations

To calculate the determinant in $n \times n$ case, we need to know what the sign of a permutation means. Here is the formal definiton:

Definition 6.0.4. Given a permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ of n elements, its sign sgn(σ) can be 1 or -1. The sign counts the parity of the number of pairs of elements that are out of order (sometimes called *inversions*) after applying the permutation. In other words,

$sgn(\sigma) =$
$$\begin{cases} 1 \text{ if } |\{(i,j) \in \{1, \ldots, n\} \times \{1, \ldots, n\} \mid i < j \text{ and } \sigma(i) > \sigma(j)\}| \text{ is even} \\ \text{-1 if } |\{(i,j) \in \{1, \ldots, n\} \times \{1, \ldots, n\} \mid i < j \text{ and } \sigma(i) > \sigma(j)\}| \text{ is odd} \end{cases}$$

What does this mean? We better see an example. Let the permutation π be defined as
$$\pi(1) = 2, \ \pi(2) = 3, \ \pi(3) = 1, \ \pi(4) = 4$$
Now forget about the specific permutation. Since $\pi : \{1,2,3,4\} \mapsto \{1,2,3,4\}$ we list all tuples $(i,j) \in \{1,2,3,4\} \times \{1,2,3,4\}$ such that $i < j$. These are:

$$(1,2), \ (1,3), \ (1,4), \ (2,3), \ (2,4), \ (3,4)$$
$$\downarrow \text{apply } \pi \text{ to every number you see} \downarrow$$
$$(2,3), \ (2,1), \ (2,4), \ (3,1), \ (3,4), \ (1,4)$$

We simply went from (i,j) to $(\pi(i), \pi(j))$. Now count for how many pairs $\pi(i) > \pi(j)$. This is the case for $(2,1)$ and $(3,1)$. The number of *out-of-order pairs* is 2 which is even. So $sgn(\pi) = 1$.

1.2 $\ n \times n$ **Case**

Definition 6.0.6 Given a square matrix $A \in \mathbb{R}^{n \times n}$ the determinant $det(A)$ is defined as
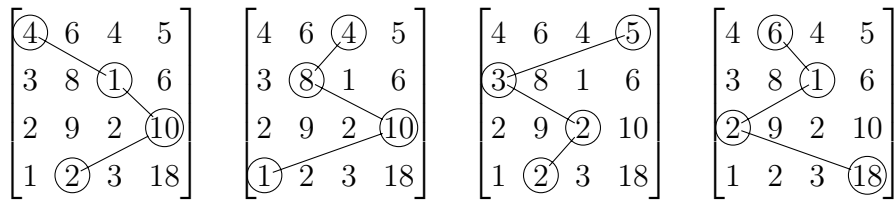$$det(A) = \sum_{\sigma \in \Pi_n} sgn(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)}$$
where Π_n is the set of all permutations of n elements.

This definition might seem confusing at first sight. Let's break it down. If you write this in java, this is nothing but just two for loops. You add some

product for each possible permutation of n elements. If you consider the product for one specific permutation, we have

$$det(A) = \cdots + sgn(\sigma) \cdot A_{1,\sigma(1)} \cdot A_{2,\sigma(2)} \cdots A_{n,\sigma(n)} + \ldots$$

This means for a certain permutation $\sigma \in \Pi_n$ you begin from the first row, get the element on the $\sigma(1)$-th column, go to second row, get the element on the $\sigma(2)$-th column etc. and multiply these terms with each other and the sign of σ. In other words you begin from the first row and traverse the matrix down to the last row by choosing each time one element per row and per column. No row or no column is allowed to have 2 selected elements. It must be exactly 1 per column and per row. What do I mean with this? Here is an example:

$$\begin{bmatrix} 4 & 6 & 4 & 5 \\ 3 & 8 & 1 & 6 \\ 2 & 9 & 2 & 10 \\ 1 & 2 & 3 & 18 \end{bmatrix} \quad \begin{bmatrix} 4 & 6 & 4 & 5 \\ 3 & 8 & 1 & 6 \\ 2 & 9 & 2 & 10 \\ 1 & 2 & 3 & 18 \end{bmatrix} \quad \begin{bmatrix} 4 & 6 & 4 & 5 \\ 3 & 8 & 1 & 6 \\ 2 & 9 & 2 & 10 \\ 1 & 2 & 3 & 18 \end{bmatrix} \quad \begin{bmatrix} 4 & 6 & 4 & 5 \\ 3 & 8 & 1 & 6 \\ 2 & 9 & 2 & 10 \\ 1 & 2 & 3 & 18 \end{bmatrix}$$

$sgn(\sigma_1) \cdot 4 \cdot 1 \cdot 10 \cdot 2 \quad sgn(\sigma_2) \cdot 4 \cdot 8 \cdot 10 \cdot 1 \quad sgn(\sigma_3) \cdot 5 \cdot 3 \cdot 2 \cdot 2 \quad sgn(\sigma_4) \cdot 6 \cdot 1 \cdot 2 \cdot 18$

The permutations $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ are some arbitrary examples of permutations in Π_4 and the products under each matrix is the term $sgn(\sigma) \prod_{i=1}^{4} A_{i,\sigma(i)}$. If you have all 24 possible permutations of 4 elements and add the corresponding terms $sgn(\sigma) \prod_{i=1}^{4} A_{i,\sigma(i)}$ together, you will have the determinant of this matrix.

As you might have already realized, it is some burdensome calculation that we have here. For any $n \times n$ matrix you have to consider all $n!$ permutations! This is too much, even for a computer when n gets large. Luckily we have other ways of calculating the determinant.

1.3   The Cofactor Expansion

The main idea is to reduce the calculation of the determinant of a larger matrix to the calculation of a bunch of 2×2 matrix determinants. The 3×3 case is the usual example to give the intuition.

Example 6.0.14. For 3×3 matrices there are $3! = 6$ permutations, so there will be 6 terms. For A a 3×3 matrix, we can write its determinant as (where an empty entry corresponds to a zero entry)

$$\det(A) = \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix}$$

$$= \begin{vmatrix} A_{11} & & \\ & A_{22} & \\ & & A_{33} \end{vmatrix} + \begin{vmatrix} A_{11} & & \\ & & A_{23} \\ & A_{32} & \end{vmatrix} + \begin{vmatrix} & A_{12} & \\ A_{21} & & \\ & & A_{33} \end{vmatrix}$$

$$+ \begin{vmatrix} & A_{12} & \\ & & A_{23} \\ A_{31} & & \end{vmatrix} + \begin{vmatrix} & & A_{13} \\ A_{21} & & \\ & A_{32} & \end{vmatrix} + \begin{vmatrix} & & A_{13} \\ & A_{22} & \\ A_{31} & & \end{vmatrix}$$

$$= A_{11}A_{22}A_{33} - A_{11}A_{23}A_{32} - A_{12}A_{21}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32} - A_{13}A_{22}A_{31}.$$

$$= A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{12}(A_{21}A_{33} - A_{23}A_{31}) + A_{13}(A_{21}A_{32} - A_{22}A_{31})$$

$$= A_{11} \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} - A_{12} \begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} + A_{13} \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix}.$$

In general, these terms in the last line are called the co-factors of A. Here is the formal definition:

Definition 6.0.15 Given $A \in \mathbb{R}^{n \times n}$, for each $1 \leq i, j \leq n$ let $\mathscr{A}_{ij}$ denote the $(n-1) \times (n-1)$ matrix obtained by removing row i and column j from A. Then we define the co-factors of A as

$$C_{ij} = (-1)^{i+j} \det(\mathscr{A}_{ij})$$

And if you multiply each co-factor with the corresponding matrix entry A_{ij} and take the sum over these products, you have the determinant of the matrix. This is the **Proposition 6.0.16.** Note that it corresponds to the last line in the calculation above for 3×3 case.

Proposition 6.0.16. Let $A \in \mathbb{R}^{n \times n}$ for any $1 \leq i \leq n$,

$$det(A) = \sum_{j=1}^{n} A_{ij} C_{ij}$$

In words, this proposition tells you that you can choose a row i arbitrarily, then go over the elements of that row to calculate the co-factors. Let's see an example:

$$\text{Let } A = \begin{bmatrix} 2 & 2 & 4 & 5 \\ 3 & 1 & 9 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 3 & 1 \end{bmatrix}$$

It makes sense to choose the row with the most 0's so that we have less work. So choose $i = 3$. Then we have

$$det(A) = \begin{vmatrix} 2 & 2 & 4 & 5 \\ 3 & 1 & 9 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 3 & 1 \end{vmatrix} = 0 \cdot C_{31} + 0 \cdot C_{32} + 2 \cdot C_{33} + 0 \cdot C_{34}$$

As you can see, the only relevant term is $2 \cdot C_{33}$ since other terms are all 0.

$$det(A) = 2 \cdot C_{33} = 2 \cdot (-1)^{3+3} \cdot \begin{vmatrix} 2 & 2 & 5 \\ 3 & 1 & 0 \\ 0 & 2 & 1 \end{vmatrix}$$

Now we can apply co-factor expansion once more to calculate the determinant of the remaining 3×3 matrix. If we choose our $i = 2$ we can write:

$$det(A) = 2 \cdot C_{33} = 2 \cdot (-1)^{3+3} \cdot \begin{vmatrix} 2 & 2 & 5 \\ 3 & 1 & 0 \\ 0 & 2 & 1 \end{vmatrix} = 2 \cdot (-1)^{3+3} \cdot (3 \cdot C_{21} + 1 \cdot C_{22} + 0 \cdot C_{23})$$

These co-factors require us to calculate 2×2 determinants which we can do. Remember $ad - bc$. Note that $0 \cdot C_{23} = 0$ anyways so we have:

$$\begin{aligned} det(A) =& 2 \cdot (-1)^{3,3} \cdot \left(3 \cdot (-1)^{2+1} \cdot \begin{vmatrix} 2 & 5 \\ 2 & 1 \end{vmatrix} + 1 \cdot (-1)^{2+2} \cdot \begin{vmatrix} 2 & 5 \\ 0 & 1 \end{vmatrix} \right) \\ =& 2 \cdot 1 \cdot (3 \cdot (-1) \cdot (2 \cdot 1 - 2 \cdot 5) + 1 \cdot 1 \cdot (2 \cdot 1 - 0 \cdot 5)) \\ =& 52 \end{aligned}$$

Note that at any given step we could have chosen any row. We chose the 3rd row the first time and the 2nd row the second time because then we have relatively less computations. However, doing this for any other row would work and yield the same result. Try it!

1.4 Properties of the Determinant

Most properties of the determinant are intuitive when you think of it as the volume of the unit cube under the associated linear transformation.[1]

Proposition 6.0.12 Given matrices $A, B \in \mathbb{R}^{n \times n}$ we have

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

Intuition: If you multiply the unit cube with a matrix A and then a matrix B, it is equally inflated compared to the case where you multiply A and B first with each other and then apply the product transformation to the unite cube.

Proposition 6.0.11. A matrix $A \in \mathbb{R}^{n \times n}$ is invertible if and only if

$$det(A) \neq 0$$

Intuition: Again think in terms of the volume of the unit cube. When is the volume 0? The volume of a 3 dimensional shape is 0, if in one of the dimensions the height is 0. This means our "cube" determined by the columns of A just exists in 2 dimensions. 3 vectors in 2 dimensions must be linearly dependent.

---

[1]The following intuitions are neither guaranteed to be correct nor generalizable to multiple dimensions. They are there to help you imagine the statements of the propositions in 3 dimensions but if they confuse you more than they help you, simply ignore them.

Proposition 6.0.13 Given a matrix $A \in \mathbb{R}^{n \times n}$ such that $det(A) \neq 0$, then A is invertible and

$$det(A^{-1}) = \frac{1}{det(A)}$$

Intuition: If A shrinks the unit cube, then A^{-1} inflates it back, if A inflates the unit cube, then A^{-1} shrinks it back. So if A multiplies every edge of the unit cube by 5 thereby inflating the volume from 1 up to 125, then the inverse must divide every edge by 5 such that the volume is 1 again. The matrix A that multiplies every edge of the unit cube -and hence the whole space- by 5 is the matrix $5 \cdot I$.

Theorem 6.0.9. Given a matrix $A \in \mathbb{R}^{n \times n}$ we have

$$det(A^{\top}) = det(A)$$

Intuition: This lemma has its proof in the lecture notes which argue over the permutations. Remember how we associated the determinant with "*traversing the matrix beginning from the first row and going until the last row in all possible ways*"? This lemma implies that you can also do the same thing with columns. So you can start from the leftmost column and traverse the matrix until you reach the rightmost column or vice versa. The sum of products will still be the same. This lemma has one more important implication: **You can expand through columns when you use the co-factor expansion!** So instead of choosing one row and going through its elements according to **Proposition 6.0.16.** You can choose a column and do the same thing for the elements of the column.

We go on by introducing some special matrices with special determinants.

Proposition 6.0.7. Given a permutation matrix $P \in \mathbb{R}^{n \times n}$ corresponding to a permutation σ, the $det(P) = sgn(P)$. It is also possible to write $sgn(P)$.

Proposition 6.0.10[2] If $Q \in \mathbb{R}^{n \times n}$ is an orthogonal matrix then $det(Q) = 1$, or $det(Q) = -1$.

Proposition 6.0.8. Given a triangular (either upper or lower[3]) matrix $T \in \mathbb{R}^{n \times n}$ we have

$$det(T) = \prod_{k=1}^{n} T_{kk}$$

in particular, $det(I) = 1$.

For the triangular matrices, try to find a traversal as we did above on page 3, that do not step to the side of the diagonal where all entries are 0. Because if you step on some 0 entry then the product corresponding your permutation becomes 0. You will realize your only option is going along the diagonal. In other words, the only permutation that contributes to the determinant is the one that chooses the diagonal entries.

Proposition 6.0.17. **(extended)** Given $A \in \mathbb{R}^{n \times n}$ with $det(A) \neq 0$ we have

$$A^{-1} = \frac{1}{det(A)} C^{\top}$$

where C is the $n \times n$ matrix with the co-factors of A as entries. Which can be rewritten as

$$A C^{\top} = det(A) I$$

In 2 dimensions for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ this corresponds to

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

---

[3]Proof in the lecture notes.

[3]Remember, diagonal matrices are upper and lower triangular at the same time. So the determinant of a diagonal matrix is the product of all diagonal elements.

1.5 Cramer's Rule

Using the determinant, we can also write a formula for the solution of the linear system $Ax = b$. This is what we call **Cramer's Rule**. You can -and you should- read the section 6.0.3. OR ***Proposition 6.0.19*** in lecture notes.

As you might encounter on several pages in the lecture notes, the formulas using the determinant repeatedly -like the inverse formula for $n \times n$ matrices above or Cramer's rule- are considered as a high workload in computational sense and therefore not preferred in practice.

1.6 PALU, Gauss and the Determinant

Remember the $PA = LU$ decomposition which is the product of Gaussian Elimination. We have

- $det(P) = sgn(P)$, since P is a permuation matrix.

- $det(U) = \prod_{i=1}^{n} U_{ii}$, in words determinant of U is equal to the product of the diagonal entries since it is an upper triangular matrix.

- $det(L) = 1$, since L is a lower triangular matrix that has only $1's$ on its diagonal per definition.

We have
$$PA = LU$$
$$det(PA) = det(LU)$$
$$det(P)det(A) = det(L)det(U)$$
$$det(A) = \frac{1}{sgn(P)} \cdot 1 \cdot det(U)$$

$$\boxed{det(A) = sgn(P)det(U)}$$

The last step holds because we know the sign of a permutation matrix is either 1 or -1 so we can say $\frac{1}{sgn(P)} = sgn(P)$.

To understand better how Gaussian Elimination can be associated with the determinant, we have 2 important propositions.

Proposition 6.0.21. If A is an $n \times n$ matrix and P is a permutation that swaps two elements, meaning that PA corresponds to swapping two rows of A then $det(PA) = -det(A)$.

Proposition 6.0.22. The determinant is linear in each row (or each column). In other words, for any $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$ and $\alpha_0, \alpha_1 \in \mathbb{R}$ we have

$$
\begin{vmatrix} \text{---} & \alpha_0 a_0^\top + \alpha_1 a_1^\top & \text{---} \\ \text{---} & a_2^\top & \text{---} \\ & \vdots & \\ \text{---} & a_n^\top & \text{---} \end{vmatrix} = \alpha_0 \begin{vmatrix} \text{---} & a_0^\top & \text{---} \\ \text{---} & a_2^\top & \text{---} \\ & \vdots & \\ \text{---} & a_n^\top & \text{---} \end{vmatrix} + \alpha_1 \begin{vmatrix} \text{---} & a_1^\top & \text{---} \\ \text{---} & a_2^\top & \text{---} \\ & \vdots & \\ \text{---} & a_n^\top & \text{---} \end{vmatrix}
$$

and

$$
\begin{vmatrix} | & | & & | \\ \alpha_0 a_0 + \alpha_1 a_1 & a_2 & \ldots & a_n \\ | & | & & | \end{vmatrix} = \alpha_0 \begin{vmatrix} | & | & & | \\ a_0 & a_2 & \ldots & a_n \\ | & | & & | \end{vmatrix} + \alpha_1 \begin{vmatrix} | & | & & | \\ a_1 & a_2 & \ldots & a_n \\ | & | & & | \end{vmatrix}
$$

This property is useful to prove the following fact:

(Not in lecture notes) For any $A \in \mathbb{R}^{n \times n}$ we have $det(k \cdot A) = k^n \cdot det(A)$.

2 Complex Numbers

In lecture notes you have the fundamental tools to work with complex numbers. Familiarizing yourself with these operations is not only important for linear algebra but also for your future courses. Here I am going to mention a part of this section.

Theorem 7.0.3. (Fundamental Theorem of Algebra) Any degree n non-constant polynomial $P(z) = \alpha_n z^n + \alpha_{n-1} z^{n-1} + \cdots + \alpha_1 z + \alpha_0$ with $a_n \neq 0$ has a zero $\lambda \in \mathbb{C}$ such that $P(\lambda) = 0$.

Corollary 7.0.4. Any degree n non-constant ($n \geq 1$) polynomial $P(z) = \alpha_n z^n + \alpha_{n-1} z^{n-1} + \cdots + \alpha_1 z + \alpha_0$ (with $\alpha_n \neq 0$) has n zeros: $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, perhaps with repetitions, such that

$$P(z) = \alpha_n (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n).$$

The number of times $\lambda \in \mathbb{C}$ appears in this expansion is called the *algebraic multiplicity* of the zero.

The fundamental theorem of algebra will later help us when we calculate the eigenvalues of a matrix and it will be the certificate of the fact that all (square) matrices have an eigenvalue.

In $\mathbb{C}^n$ we usually define linear independence, span and similar concepts just as in $\mathbb{R}^n$. An important difference is the transpose of a matrix. We usually speak of *hermitian transpose* or the conjugate transpose of a matrix. The notation is usually A^*, sometimes A^H. You take the complex conjugate of each entry and then transpose the matrix as usual.

$$A^* = \bar{A}^\top$$

Then we have

$$\|v\|^2 = v^* v = \bar{v}^\top v = \sum_{i=1}^n \bar{v}_i v_i = \sum_{i=1}^n |v_i|^2$$

where $\bar{v}_i$ is the complex conjugate of v_i.

A useful trick to prove some given $x \in \mathbb{C}$ is in $\mathbb{R}$, is to show that the complex conjugate of x is equal to itself as in

$$x = \bar{x} \implies x \in \mathbb{R}.$$

This fact is is not a lemma from the notes but it is used in the proof of **Proposition 7.3.7.** It is just a cool trick, good to know.

3 Hints

1. Solved in class. Select a column / row and iterate through it using proposition 6.0.16. You can also select a column because $det(A) = det(A^\top)$. Go for the one with most 0's and hence with least work. For b see section 6.0.4.

2. Use the definition of the determinant over permutations. If your permutation touches a 0 entry then that permutation does not contribute to the determinant. What are all the permutations that avoid the 0 entries, *i.e.* that only choose entries that are non zero?

3. No hints.

4. Write E as a linear combination of A,B,C, and D. Remember, the determinant is linear in each row / column. $\rightarrow$ Prop 6.0.22.

5. To show that λ and $\mathbf{v} \neq 0$ are an eigenvalue-eigenvector pair associated with a matrix M, all you have to show is $M\mathbf{v} = \lambda\mathbf{v}$. Do this for the matrix $M + cI$. For b you can guess the two eigenvalues and verify by direct computation.

6. Write $\mathbf{p}_j = \frac{1}{2}(\mathbf{p}_{j-1} + \mathbf{p}_{j+1})$ as $\mathbf{p}_j - \frac{1}{2}(\mathbf{p}_{j-1} + \mathbf{p}_{j+1}) = \mathbf{0}$ and formulate equations (3) and (4) similarly. Does it look more like a linear system of equations now? You should also include the vectors $\mathbf{c}_s$ into your matrix. You can do so by defining a big matrix $S = \left[\begin{array}{c} A \\ \hline B \end{array}\right]$ where A represents the LSE for $\mathbf{p}'s$ and B represents the LSE for $\mathbf{c}'s$. B will look like a permutation matrix but you only choose some elements so its shorter.

<div align="center">

mkilic

♠

</div>