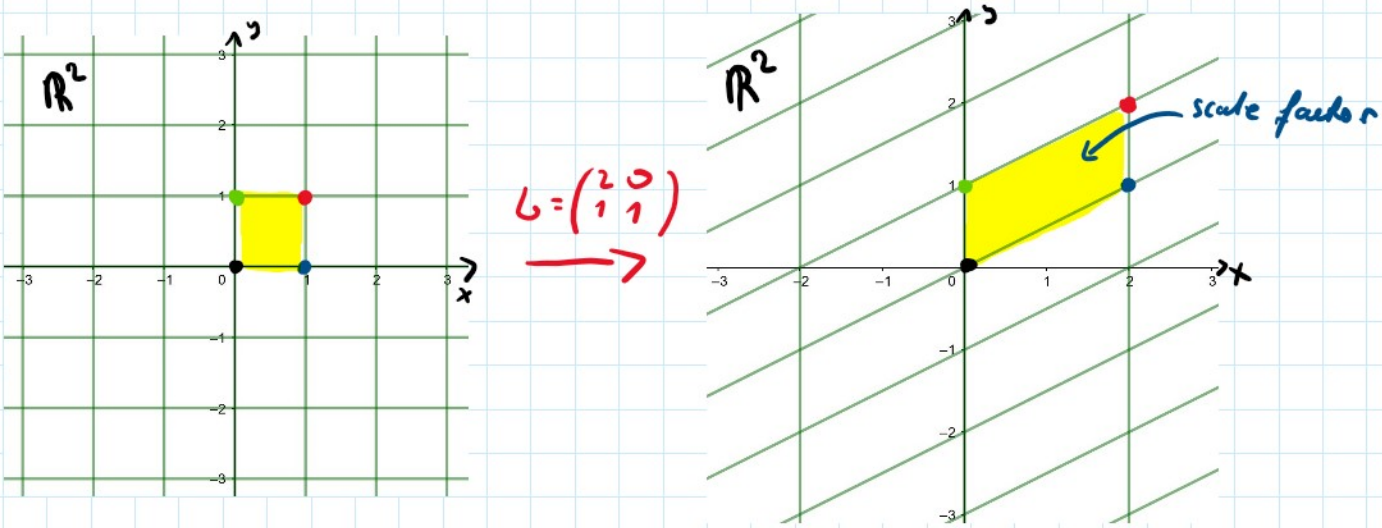
Lecture 11 - Determinant

Monday, December 6, 2021 20:20

Recall that a linearity of a transformation $L: \mathbb{R}^n \to \mathbb{R}^n$ precisely means that it maps parallel lines to parallel lines preserves the origin (maps zero to zero)



Therefore, a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$ uniformly scales n-volumes (area in $\mathbb{R}^2$) in $\mathbb{R}^n$. Intuitively, this scaling factor should carry certain information about the linear transformation in question. For example, if this scaling factor is 0, that means $\dim(\operatorname{Im} L) < n$, so L is not bijective, thus matrix of L is not invertible (and vice-versa).

Specifically, if $n \times n$ real matrix A is given by its column vectors $A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}$, then recall that $Ae_1 = v_1$, $Ae_2 = v_2, \ldots, Ae_n = v_n$. This means that A maps a unit n-cube to the n-dimensional parallelotope defined by vectors $v_1, v_2, \ldots, v_n$, that is, the region

$$P = \{ c_1 v_1 + \cdots + c_n v_n \mid 0 \leq c_i \leq 1 \quad \forall i \}.$$

$$P = \{ c_1 v_1 + \dots + c_n v_n \mid 0 \le c_i \le 1 \quad \forall i \}.$$

The determinant gives the signed n-dimensional volume of this parallelotope $\det(A) = \pm \operatorname{vol}(P)$, and hence describes more generally the n-dimensional volume scaling factor of the linear transformation given by A (the sign actually shows weather transformation preserves or reverses the orientation).

Definition: (Recall) For $n=2$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} := a_{11} a_{22} - a_{12} a_{21}$$

For $n=3$

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} := a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{31} a_{22} a_{13} - a_{32} a_{23} a_{11} - a_{33} a_{21} a_{12}$$

The case $n=3$ can be remembered by the rule of Sarrus:

$$\left(\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix}\right) \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

first two columns of matrix are written to the right, grey diagonals are with positive sign, yellow diagonals with negative sigh.

To define the determinant for $n \times n$ matrix, note that for $n=3$ each of the six terms in the expression is a product of three factors involving exactly one entry from each row and each column of the matrix.

row and each column of the matrix.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

We will proceed with same rational for general n. That is, each factor of the determinant will envolve exactly one entry from each row and each column, and the sign of the factor will be decided by signature of the corresponding permutation of entries.

Recall: (DS) Permutation of the set $\{1,2,3,\dots,n\}$ is a bijective function $\sigma \in S_n$ from $\{1,2,3,\dots n\}$ to itself. That is σ reorders this set of integers. Here S_n is a symmetric group, that is the group of all integers.

Recall (again from DS), that signature of σ is defined to be $+1$ whenever the reordering given by σ can be achived by succesively interchanging two entries even number of times, and -1 when it can be achived by an odd number of such interchanges.

even number of times, and -1 when it can be derived by an odd number of such interchanges.

Given $n \times n$ matrix A and a permutation $\sigma \in S_n$, consider a product $a_{1,\sigma(1)} a_{2,\sigma(2)} a_{3,\sigma(3)} \cdots a_{n,\sigma(n)} = \prod_{i=1}^{n} a_{i,\sigma(i)}$ and denote the signature of σ by $\operatorname{sgn}(\sigma)$.

Definition: For $n \times n$ matrix A,

$$\det(A) = \sum_{\sigma \in S_n} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \right)$$

This is sometimes called the Leibnitz formula.

Example: $n=3$.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Now there are $3! = 6$ permutations of $\{1,2,3\}$ (elements in S_3).

These permutations are:

$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ – zero exchanges, so $+1$

$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ – 2 exchanges, so $+1$

$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ – 2 exchanges, so $+1$

$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ – 1 exchange, so -1

$3 \quad 2 \quad 1$

$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ - 1 exchange, so -1

$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ - 1 exchange, so -1.

So, we get
$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$
which is what we already knew.

Lemma: The determinant of the upper (or lower) triangular matrix is the product of the diagonal entries of the matrix.

Proof: Exercise in CE.

Corollary: The determinant of the diagonal matrix is the product of the entries on a diagonal.

Computing determinants efficiently

How many operations does computing 30×30 matrix determinant need? Well, S_{30} has $30!$ elements, so determinant will have have $30!$ terms each with 29 multiplications, so, in total $29 \cdot 30! \approx 7 \cdot 10^{33}$ multiplications are needed. Even if our supercomputer performs 10 trillion (10^{13}) multiplications per second, performs 10 trillion (10^{13}) multiplications per second, it will take over trillion years to conclude the computation. We need more efficent ways to compute the determinant, that is, we need an algorithm in polinomial time rather than exponential.

We will again use Gauss-Jordan elimination.

Q: What happens to a determinant of A, when we perform elementary row operations on A?

1) **Row swapping**: First, say we get matrix B by swapping two **adjecent** rows in matrix A.

$$A = \begin{pmatrix} & \vdots & & \\ & \boxed{a_{k,1}} & a_{k,2} & \cdots \\ & a_{k+1,1} & \boxed{a_{k+1,2}} & \cdots \\ & \vdots & & \end{pmatrix} \rightarrow B = \begin{pmatrix} & \cdots & & \\ a_{k+1,1} & \boxed{a_{k+1,2}} & \cdots \\ \boxed{a_{k,1}} & a_{k,2} & \cdots \\ & \cdots & & \end{pmatrix}$$

Then each sum term in $\det(A)$ has a corresponding sum term in $\det(B)$ (for example, if $a_{k,1} \cdot a_{k+1,2} \cdots$ is a term in $\det(A)$, $a_{k+1,2}\, a_{k,1}$ is a term in $\det(B)$) with the opposite signature. So, we have $\det(A) = -\det(B)$.

Now, if B is obtained from A by swaping any arbitrary rows, then $\det(A) = -\det(B)$ because swaping arbitrary rows can be obtained by swaping adjacent rows odd number of times (Exercise in CE).

adjacent rows odd number of times (Exercise in CE).

Cor: If matrix has two equal rows, then $\det(A)=0$.

proof: Say B is a matrix where we swap two equal rows in A. So, $A=B$. But $\det(A) = -\det(B)$ $= -\det(A) \Rightarrow \det(A)=0$.

2) Determinant is linear in the rows and columns:

> **Linearity of the determinant in the rows and columns**
>
> Consider fixed row vectors $\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n$ with n components. Then the function
>
> $$T(\vec{x}) = \det \begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_{i-1} & - \\ - & \vec{x} & - \\ - & \vec{v}_{i+1} & - \\ & \vdots & \\ - & \vec{v}_n & - \end{bmatrix} \qquad \text{from } \mathbb{R}^{1\times n} \text{ to } \mathbb{R}$$
>
> is a linear transformation. This property is referred to as *linearity of the determinant in the* ith *row*. Likewise, the determinant is *linear in all the columns*.

Proof: Clearly every sum term is linear in all the rows and columns, since this product contains one factor from each row and one factor from each column. Determinant, being a linear combination of the sum terms, is itself then linear.

Corollary: If

terms is itself then linear.

Corollary 1: If

$$A = \begin{pmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_i & - \\ & \vdots & \\ - & \vec{v}_n & - \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \frac{\vec{v}_i}{k} & - \\ - & \vec{v}_n & - \end{pmatrix}$$

then $\det(B) = \frac{1}{k}\det(A)$.

Proof: by linearity in i^{th} row.

Corollary 2: If matrix B is given by adding k times j^{th} row to the i^{th} row of A, then $\det(A) = \det(B)$.

Proof:

$$\det B = \det \begin{pmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_i & - \\ & \vdots & \\ - & \vec{v}_j + k\vec{v}_i & - \\ - & \vec{v}_n & - \end{pmatrix} = \underbrace{\det \begin{pmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_i & - \\ & \vdots & \\ - & \vec{v}_j & - \\ & \vdots & \\ - & \vec{v}_n & - \end{pmatrix}}_{= \det(A)} +$$

$$k \underbrace{\det \begin{pmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_i & - \\ & \vdots & \\ - & \vec{v}_i & - \\ & \vdots & \\ - & \vec{v}_n & - \end{pmatrix}}_{= 0} = \det A$$

So, summing up

Elementary row operations and determinants

a. If B is obtained from A by dividing a row of A by a scalar k, then

Elementary row operations and determinants

a. If B is obtained from A by dividing a row of A by a scalar k, then

$$\det B = (1/k)\det A.$$

b. If B is obtained from A by a row swap, then

$$\det B = -\det A.$$

We say that the determinant is *alternating* on the rows.

c. If B is obtained from A by adding a multiple of a row of A to another row, then

$$\det B = \det A.$$

Analogous results hold for elementary column operations.

Now, lets analyze the relationship between $\det A$ and $\det(\operatorname{rref} A)$. Suppose that in the course of Gauss-Jordan elimination we swap rows s times and divide various rows by the scalars $k_1, k_2, \dots, k_r$. Then

$$\det(\operatorname{rref} A) = (-1)^s \frac{1}{k_1 \cdot k_2 \cdot \dots \cdot k_r} \det(A), \quad \text{or}$$

$$\det(A) = (-1)^s \cdot k_1 \cdot k_2 \cdots k_r \cdot \det(\operatorname{rref} A).$$

Theorem: A square matrix A is invertible if and only if $\det A \neq 0$.

Proof: A invertible $\Rightarrow \operatorname{rref} A = I_n$, so that

$$\det(\operatorname{rref} A) = \det(I_n) = 1 \quad \text{and}$$

$$\det A = (-1)^s k_1 \cdots k_r \neq 0$$

since all scalars $k_1, \dots, k_r$ are non zero.

if A is noninvertible then $\operatorname{rref} A$ contains last row since all scalars $k_1, \dots, k_r$ are non zero.

if A is noninvertible, then rref A contains last row of all zeros, and thus every sum term in $\det(A)$ is zero, therefore $\det(A) = 0$. $\square$

Discussion above also gives us a good method for computing determinants.

Using Gauss–Jordan elimination to compute the determinant

a. Consider an invertible $n \times n$ matrix A. Suppose you swap rows s times as you compute rref $A = I_n$, and you divide various rows by the scalars $k_1, k_2, \dots, k_r$. Then

$$\det A = (-1)^s k_1 k_2 \cdots k_r.$$

b. In fact, it is not always sensible to reduce A all the way to rref A. Suppose you can use elementary row operations to transform A into some matrix B whose determinant is easy to compute (B might be a triangular matrix, for example). Suppose you swap rows s times as you transform A into B, and you divide various rows by the scalars $k_1, k_2, \dots, k_r$. Then

$$\det A = (-1)^s k_1 k_2 \cdots k_r \det B.$$

Even though not hard, we give the following fact without proof:

Theorem: If A and B are $n \times n$ matrices then

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

Corollary: Similar matrices have the same determinant

Proof: $A \sim_{sim} B \iff AS = SB \Rightarrow$

$$\det(AS) = \det A \cdot \det S = \det(SB) =$$

$$= \det S \cdot \det B \Rightarrow \det A = \det B$$

since S is invertible and $\det S \neq 0$.

since S is invertible and $\det S \neq 0$.

Corollary: $\det(A^{-1}) = \dfrac{1}{\det A}$ for an invertible $n \times n$ matrix A.

Proof: $1 = \det I_n = \det A \cdot A^{-1} = \det A \cdot \det A^{-1}$.

Laplace expension:

We will give another way to compute determinants without proof.

Minors

For an $n \times n$ matrix A, let A_{ij} be the matrix obtained by omitting the ith row and the jth column of A. The determinant of the $(n-1) \times (n-1)$ matrix A_{ij} is called a *minor* of A.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix},$$

$$A_{ij} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix}$$

Laplace expansion (or cofactor expansion)

We can compute the determinant of an $n \times n$ matrix A by Laplace expansion down any column or along any row.

Expansion down the jth column:

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

Expansion along the ith row:

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