

# **Unit 3(b) Lecture Notes for MAT224**

Xuanqi Wei 1009353209

14-16 February 2023

## §1 3.1 Determinant as Area

Let  $\vec{a} = (a_1, a_2)$  and  $\vec{b} = (b_1, b_2)$  be two linearly independent vectors. What is the area of a 2-dimensional parallelogram that has  $\vec{a}$  and  $\vec{b}$  as adjacent sides?

The area of a parallelogram with adjacent sides  $\vec{a}$  and  $\vec{b}$  is given by the magnitude of their cross product:  $\|\vec{a} \times \vec{b}\|$ . Since we are in 2-dimensional space, we can think of  $\vec{a}$  and  $\vec{b}$  as living in 3-dimensional space, with a  $\vec{0}$  third component. Then we can compute their cross product. Then we can compute their cross product:

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

The magnitude of this vector is  $|a_1 b_2 - a_2 b_1|$ , which is the area of the parallelogram formed by  $\vec{a}$  and  $\vec{b}$

What happens to the above formula if we use vectors  $\vec{a}$  and  $\vec{b}$  that are linearly dependent?

If  $\vec{a}$  and  $\vec{b}$  are linearly dependent, then one of them can be written as a multiple of the other. Without loss of generality, suppose  $\vec{b} = k\vec{a}$  for some scalar  $k$ . Then the parallelogram formed by  $\vec{a}$  and  $\vec{b}$  is actually a parallelogram with two identical adjacent sides, and the height of the parallelogram is the length of the perpendicular from one of the sides to the line containing the other side. Since the sides are identical, any perpendicular from one side to the other will have length equal to the length of  $\vec{a}$  (or equivalently, the length of  $\vec{b}$ ). Therefore, the area of the parallelogram is 0.

The “area” function  $A(\vec{x}, \vec{y})$  is the function that takes in two vectors  $\vec{x}, \vec{y} \in \mathbb{R}^2$  and outputs the area of a parallelogram with adjacent sides equal to  $\vec{x}$  and  $\vec{y}$ . What are three of its properties?

1. Non-negativity: The area of a parallelogram is always non-negative, so  $A(\vec{x}, \vec{y}) \geq 0$  for any vectors  $\vec{x}$  and  $\vec{y}$  in  $\mathbb{R}^2$ .
2. Linearity: If we fix one of the vectors and scale the other, then the area of the parallelogram is scaled by the same factor. That is, for any  $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^2$  and  $c \in \mathbb{R}$ , we have  $A(\vec{x}, c\vec{y} + \vec{z}) = |c|A(\vec{x}, \vec{y}) + A(\vec{x}, \vec{z})$ . This property is often referred to as the ‘parallelogram law.’
3. Antisymmetry: If we switch the order of the inputs, the sign of the output changes. That is, for any  $\vec{x}, \vec{y} \in \mathbb{R}^2$ , we have  $A(\vec{x}, \vec{y}) = -A(\vec{y}, \vec{x})$ . This

property reflects the fact that the area of a parallelogram does not depend on the order in which we choose its adjacent sides.

Do any other functions have all three properties?

(Answer this question after reading up to and including Theorem 3.2.8.)

Yes, there are other functions that have all three properties. In fact, Theorem 3.2.8 states that any function with these three properties must be a scalar multiple of the area function. That is, if  $f : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$  satisfies linearity in the first argument, reverses orientation, and has  $f(\vec{e}_1, \vec{e}_2) = 1$ , then there exists a scalar  $c \in \mathbb{R}$  such that  $f = c \cdot A$ , where  $A$  is the area function.

If you have another function that multilinear (Definition 3.2.1) and alternating (Definition 3.2.2), how can you tell if it is the area function or not?

If you have another function that is multilinear and alternating, you can check if it satisfies the property of being equal to the area function  $A(\vec{x}, \vec{y})$  on the standard basis vectors  $\vec{e}_1$  and  $\vec{e}_2$ . In other words, you can evaluate the function on the pairs of vectors  $(\vec{e}_1, \vec{e}_2)$  and  $(\vec{e}_2, \vec{e}_1)$  and see if the values match the known values of  $A(\vec{e}_1, \vec{e}_2) = 1$  and  $A(\vec{e}_2, \vec{e}_1) = -1$ . If the function gives the same values as the area function, then it must be the same as the area function, since the area function is the unique multilinear, alternating function that satisfies these properties.

## §2 3.2 Determinant of an nxn Matrix

What does it mean for a function to be multilinear? How is this different from the definition of linear that we have been using up until now?

A function  $f : V_1 \times V_2 \times \cdots \times V_k \rightarrow W$  is said to be multilinear if it is linear in each of its arguments  $v_i \in V_i$  when all other arguments are held constant. More precisely, for any  $i \in 1, 2, \dots, k$ , if we hold all the vectors  $v_j$  for  $j \neq i$  fixed and let  $v_i$  vary, then the function  $f$  is linear with respect to  $v_i$ .

This is different from the definition of linear that we have been using, which deals with only one vector space. A linear transformation  $T : V \rightarrow W$  between two vector spaces  $V$  and  $W$  satisfies the properties  $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$  and  $T(c\vec{v}) = cT(\vec{v})$  for all  $\vec{v}, \vec{w} \in V$  and  $c \in \mathbb{R}$ . In contrast, a multilinear function deals with multiple vector spaces and satisfies a different set of properties.

What does alternating mean?

A multilinear function  $f$  on a vector space  $V$  is called alternating if  $f(v_1, \dots, v_n) = 0$  whenever  $v_i = v_j$  for some  $i \neq j$ . In other words, if two of its arguments are equal, then the function evaluates to 0. This property is also called skew-symmetry or anti-symmetry.

Use Definition 3.2.4 and Proposition 3.2.5 to write down step-by-step instructions for calculating the determinant of a 4x4 matrix. Include an example with a matrix that has at least 3 zeros in the row or column along which you are expanding.

1. Identify a row or column of the matrix to expand along. Let's say we choose to expand along the first row.
2. For each entry in the row or column, create a new matrix by deleting the row and column containing that entry. In our example, we would create three 3x3 matrices by deleting the first row and the column containing each of the entries in the first row of  $A$ .
3. Calculate the determinant of each of the three 3x3 matrices using the same method (by expanding along a row or column). Let's say we choose to expand along the first row for each of the three 3x3 matrices as well.
4. Multiply each determinant by the corresponding entry in the original row or column we are expanding along, and alternate the signs of these products according to the alternating property. Specifically, the first entry in the row or column is positive, the second entry is negative, the third entry is positive, and so on.
5. Add up the products from step 4 to get the determinant of the original matrix  $A$ .

$$A = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

$$\det(A)_{1,1} = 0 - 0 + 0 = 0$$

$$\det(A)_{1,2} = -2 \cdot \det \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix} = 0$$

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

Finally, we multiply each determinant by the corresponding entry in the first row of  $A$ , alternating the signs as we go:

$$1 \cdot 0 - 0 \cdot 0 + 0 \cdot 6 + 2 \cdot 0 = 0$$

Therefore, the determinant of  $A$  is 0.

When we combine Proposition 3.1.6 and Theorems 3.2.10, 3.2.13 and 3.2.14, what do we know about invertibility and determinants?

1. A square matrix  $A$  is invertible if and only if its reduced row echelon form is the identity matrix  $I_n$ , and in this case, the determinant of  $A$  is nonzero.
2. If the determinant of a square matrix  $A$  is nonzero, then  $A$  is invertible.
3. The determinant of a matrix  $A$  is equal to the product of the determinants of its elementary matrices, and the determinant of an elementary matrix is either 1 or  $-1$ , depending on the type of elementary row operation used to obtain it.

## §3 3.3 Further Properties of the Determinant

List four properties of determinants of matrices.

1. The determinant of the transpose of a matrix is equal to the determinant of the original matrix, i.e.,  $\det(A^T) = \det(A)$  for any matrix  $A$ .
2. If a matrix  $A$  has a row or column consisting entirely of zeros, then its determinant is zero, i.e.,  $\det(A) = 0$  if  $A$  has a row or column of zeros.
3. If two rows or columns of a matrix  $A$  are interchanged, then the determinant of the resulting matrix  $B$  is equal to the negation of the determinant of  $A$ , i.e.,  $\det(B) = -\det(A)$ .
4. If a matrix  $A$  is singular, i.e.,  $\det(A) = 0$ , then its inverse does not exist. On the other hand, if a matrix  $A$  is invertible, i.e.,  $\det(A) \neq 0$ , then its inverse exists and can be found using the formula  $A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$ , where  $\text{adj}(A)$  is the adjugate matrix of  $A$ .

Let  $T$  be a linear transformation. Which theorem allows us to talk about “the determinant of a linear transformation”  $\det(T)$  instead of  $\det([T]_{\alpha}^{\alpha})$  for different bases  $\alpha$ ? Explain.

Let  $V$  be a finite-dimensional vector space with bases  $\alpha$  and  $\beta$ , and let  $T : V \rightarrow V$  be a linear transformation. Then the determinants of the matrices  $[T]_{\alpha}^{\alpha}$  and  $[T]_{\beta}^{\beta}$  are equal, i.e.,  $\det([T]_{\alpha}^{\alpha}) = \det([T]_{\beta}^{\beta})$ .

This theorem is sometimes called the “invariance of the determinant” or the “determinant formula,” and it holds for any linear transformation on a finite-dimensional vector space. It tells us that the determinant of a linear transformation is well-defined and does not depend on the choice of basis.

How can the determinant function  $\det(T)$  help us decide if a linear transformation  $T$  is an isomorphism? Explain why your answer works.

- If  $\det(T) \neq 0$ , then  $T$  is invertible, and hence an isomorphism. This is because if  $\det(T) \neq 0$ , then  $T$  has an inverse  $T^{-1}$ , and the composition  $T \circ T^{-1} = T^{-1} \circ T = I$ , where  $I$  is the identity transformation. This shows that  $T$  is bijective, and hence an isomorphism.
- If  $\det(T) = 0$ , then  $T$  is singular, and hence not an isomorphism. This is because if  $\det(T) = 0$ , then the kernel of  $T$  is non-trivial, i.e., there exist nonzero vectors in the domain of  $T$  that are mapped to the zero vector in the range of  $T$ . This means that  $T$  is not injective, and hence not an isomorphism.

To see why this works, note that the determinant of a matrix representation of a linear transformation measures the scaling factor of the transformation. If the determinant is nonzero, then the transformation does not collapse any nonzero vectors to zero, and hence is one-to-one. Moreover, the scaling factor is not zero, so the transformation does not "squeeze" any nonzero vectors to zero, and hence is onto. This implies that the transformation is a bijection, and hence an isomorphism.

On the other hand, if the determinant is zero, then the transformation does collapse some nonzero vectors to zero, and hence is not one-to-one. Moreover, the scaling factor is zero, so the transformation "squeezes" some nonzero vectors to zero, and hence is not onto. This implies that the transformation is not a bijection, and hence not an isomorphism.

Give one example for each of part (i) and part (ii) of Proposition 3.3.12 that demonstrates those properties are true.

(i) Let  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  be defined by  $T(x, y, z) = (x + y, x + z)$ , and let  $U$  be the subspace of  $\mathbb{R}^3$  spanned by  $(1, 0, 0)$  and  $(0, 1, 0)$ . Then  $\dim(U) = 2$ , and  $\ker(T)$  is the subspace of  $\mathbb{R}^3$  spanned by  $(0, -1, 1)$ . We have  $T(U) = \text{span}((1, 1), (0, 1))$ , which has dimension 2. Therefore, by part (i) of Proposition 3.3.12, we have  $\dim(U) + \dim(\ker(T)) = 2 + 1 = 3 = 2 + 2 = \dim(T(U)) + \dim(\ker(T))$ .

(ii) Let  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  be defined by  $T(x, y, z) = (x + y, x + z)$ , and let  $U$  be the subspace of  $\mathbb{R}^3$  spanned by  $(1, 0, 0)$  and  $(0, 1, 0)$ , and let  $W'$  be the subspace of  $\mathbb{R}^2$  spanned by  $(1, 0)$  and  $(0, 1)$ . Then  $T(U) = \text{span}((1, 1), (0, 1))$ , which is a subspace of  $W'$ . We have  $\dim(U) = 2$ , and  $\ker(T)$  is the subspace of  $\mathbb{R}^3$  spanned by  $(0, -1, 1)$ . We have  $\dim(\ker(T)) = 1$  and  $\dim(W') = 2$ , so by part (ii) of Proposition 3.3.12, we have  $\dim(U) \leq \dim(\ker(T)) + \dim(W')$ , which is equivalent to  $2 \leq 1 + 2$ , which is true.

What calculation are we doing if we can we use Cramer's Rule to do it?

Cramer's Rule is a method for solving systems of linear equations by using determinants. Specifically, if we have a system of  $n$  linear equations in  $n$  variables, then Cramer's Rule provides a formula for the solution of the system in terms of determinants.

In general, if we can use Cramer's Rule to solve a system of linear equations, then we are calculating the values of the variables in the system. More precisely, Cramer's Rule gives us the unique solution to the system of equations, provided that the determinant of the coefficient matrix is nonzero.