1111: Linear Algebra I

Dr. Vladimir Dotsenko (Vlad)

Lecture 3

Vector product and volumes

Theorem. For three 3D vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, the volume of the parallelepiped defined by these three vectors is equal to $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$.

Proof. The volume is equal to the product of the height of the parallelepiped by the area of the base. We may take for the base the parallelogram defined by the vectors $\mathbf{v}$ and $\mathbf{w}$. The theorem from the previous slide tells us how to compute the area of the base, so it remains to compute the height. The height is equal to the length of the projection of $\mathbf{u}$ on the direction perpendicular to the base.

By an observation we made earlier, the direction perpendicular to the base is given by the vector product $\mathbf{n} = \mathbf{v} \times \mathbf{w}$. The length of the projection of $\mathbf{u}$ on the direction of $\mathbf{n}$ is equal to $|\mathbf{u}||\cos\theta|$, where θ is equal to the angle between $\mathbf{u}$ and $\mathbf{n}$. Multiplying it by the area of the base, that is $|\mathbf{v} \times \mathbf{w}|$, we conclude that the volume is equal to $|\mathbf{u}||\mathbf{v} \times \mathbf{w}||\cos\theta| = |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$. □

Vector product and volumes

Let us remark that the formula for the volume that we just proved does agree very well with the property

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = -\mathbf{v} \cdot (\mathbf{u} \times \mathbf{w})$$

we proved earlier: it tells that these two quantities do have the same absolute value, the volume of the corresponding parallelepiped.

This is another example of a "sanity check" one can perform on a formula. It is very useful to incorporate doing things like that in your repertoire of mathematical skills, it often allows to catch silly mistakes / misprints in solutions.

An unexpected application

Suppose that a cube $a \times a \times a$ is positioned in 3D in such a way that all its vertices have integer coordinates (but edges do not have to be parallel to the grid lines). It turns out that then a is an integer. (This is so not true in 2D).

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be the vectors connecting one of the vertices of the cube with its neighbours. All these vectors have integer coordinates. Therefore, $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ is an integer, so a^3, the volume of the cube, is an integer M. Also, $\mathbf{u} \cdot \mathbf{u}$ is an integer, so a^2, the square of the length of the side of our cube, is an integer N. But this implies that $a = \frac{a^3}{a^2} = \frac{M}{N}$, a fraction with both numerator and the denominator being integers! A fraction whose square is an integer must be an integer (otherwise the denominator will never cancel).

Outline of another application

Another important application of vector products is in the definition of *quaternions*, an omnipresent algebraic system discovered by William Rowan Hamilton while walking along the Royal Canal on October 16, 1843. This system consists of formal expressions of the form $c + \mathbf{v}$, where c is a scalar, and $\mathbf{v}$ is a 3D vector. (Talk about adding apples to oranges. . .)

The product $q_1 q_2$ of two quaternions $q_1 = c_1 + \mathbf{v}_1$ and $q_2 = c_2 + \mathbf{v}_2$ is defined as follows:

$$(c_1 + \mathbf{v}_1)(c_2 + \mathbf{v}_2) = (c_1 c_2 - \mathbf{v}_1 \cdot \mathbf{v}_2) + (c_1 \mathbf{v}_2 + c_2 \mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2).$$

It turns out that the equation $(q_1 q_2) q_3 = q_1 (q_2 q_3)$ holds for any three quaternions. This, and the fact that for each quaternion q one can find a multiplicative inverse $1/q$, makes them a truly exceptional algebraic system. There are notable appearances of quaternions in mathematical physics, geometry, computer science etc.

Vectors and lines

Using vectors, it is very easy to describe straight lines algebraically. If we are given a point P belonging to the line ℓ, and $\mathbf{v}$ is a nonzero vector parallel to ℓ, then for each point X on ℓ, the vector $\overrightarrow{PX}$ is parallel to $\mathbf{v}$. In other words, we have

$$\overrightarrow{PX} = t\mathbf{v}$$

for some number t.

Let us fix a point O, and consider, for each point A, the vector $\overrightarrow{OA}$, the *position vector of A relative to O*. Then the previous equation can be written as

$$\overrightarrow{OX} = \overrightarrow{OP} + t\mathbf{v}.$$

If the point O is *the origin* of the coordinate system, that is all its coordinates are equal to zero, this equation expresses coordinates of X as functions of parameter t, and is called the parametric equation of the line ℓ.

Scalar product and planes

Using vectors, we can also describe 2D planes in 3D. If we are given a point P belonging to the line α, and $\mathbf{n}$ is a nonzero vector perpendicular to α, then for each point X in α, the vector $\overrightarrow{PX}$ is perpendicular to $\mathbf{n}$. In other words, we have

$$\overrightarrow{PX} \cdot \mathbf{n} = 0 \ .$$

Using the relative position vectors, we can rewrite that equation as $(\overrightarrow{OX} - \overrightarrow{OP}) \cdot \mathbf{n} = 0$, or

$$\overrightarrow{OX} \cdot \mathbf{n} = \overrightarrow{OP} \cdot \mathbf{n} \ .$$

If the point O is the origin of the coordinate system, this equation becomes an equation of the form $Ax + By + Cz = D$, where x, y, z are coordinates of the varying point X. This is called the standard equation of the plane α.

Systems of linear equations

Geometrically, we are quite used to the fact that if we take two planes in 3D which are not parallel, their intersection is a line. With our new algebraic approach, this means that if we take a system of two equations

$$\begin{cases} A_1x + B_1y + C_1z = D_1, \\ A_2x + B_2y + C_2z = D_2, \end{cases}$$

for which the triples (A_1, B_1, C_1) and (A_2, B_2, C_2) are not proportional, then the solution set of this system can be described parametrically

$$\begin{cases} x = x_0 + at, \\ y = y_0 + bt, \\ z = z_0 + ct. \end{cases}$$

Our next goal is to develop algebraic formalism for systems of linear equations with arbitrary numbers of equations and unknowns that would allow to extend our 3D geometric intuition to higher dimensional spaces.

Systems of linear equations

A linear equation with unknowns $x_1, \ldots, x_n$ is an equation of the form

$$A_1x_1 + A_2x_2 + \cdots + A_nx_n = B,$$

where A_1, ..., A_n, and B are known numbers.
We shall develop a method for solving systems of m simultaneous linear equations with n unknowns

$$\begin{cases} A_{1,1}x_1 + A_{1,2}x_2 + \cdots + A_{1,n}x_n = B_1, \\ A_{2,1}x_1 + A_{2,2}x_2 + \cdots + A_{2,n}x_n = B_2, \\ \qquad \qquad \qquad \qquad \qquad \qquad \cdots \\ A_{m,1}x_1 + A_{m,2}x_2 + \cdots + A_{m,n}x_n = B_m, \end{cases}$$

where $A_{i,j} (1 \le i \le m, 1 \le j \le n)$, and $B_i (1 \le i \le m)$ are known numbers. To save space, we shall often write A_{ij} instead of $A_{i,j}$, implicitly assuming the comma between i and j (and of course taking care to never multiply i by j in this context!)

Gauss–Jordan elimination

The most common technique for solving simultaneous systems of linear equations is *Gauss–Jordan elimination*. Anyone who ever tried to solve a system of linear equations probably did something of that sort, carefully eliminating one variable after another. We shall formulate this recipe in the form of an algorithm, that is a sequence of instructions that a person (or a computer) can perform mechanically, ending up with a solution to the given system.

For convenience, we shall not carry around the symbols representing the unknowns, and will encode the given system of linear equations by $m \times (n+1)$-matrix of coefficients

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} & B_1 \\ A_{21} & A_{22} & \cdots & A_{2n} & B_2 \\ \vdots & \dots & \ddots & \dots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} & B_m \end{pmatrix}$$

From equations to matrices

For example, if we consider the system of equations

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 + x_5 = 1, \\ -3x_1 - 6x_2 - 2x_3 - x_5 = -3, \\ 2x_1 + 4x_2 + 2x_3 + x_4 + 3x_5 = -3, \end{cases}$$

then the corresponding matrix is

$$A = \begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ -3 & -6 & -2 & 0 & -1 & -3 \\ 2 & 4 & 2 & 1 & 3 & -3 \end{pmatrix}$$

(note the zero entry that indicates that x_4 is not present in the second equation).

Elementary row operations

We define *elementary row operations* on matrices to be the following moves that transform a matrix into another matrix with the same number of rows and columns:

- *Swapping rows*: literally swap the row i and the row j for some $i \neq j$, keep all other rows (except for these two) intact.
- *Re-scaling rows*: multiply all entries in the row i by a nonzero number c, keep all other rows (except for the row i) intact.
- *Combining rows*: for some $i \neq j$, add to the row i the row j multiplied by some number c, keep all other rows (except for the row i) intact.

Let us remark that elementary row operations are clearly reversible: if the matrix B is obtained from the matrix A by elementary row operations, then the matrix A can be recovered back. Indeed, each individial row operation is manifestly reversible.

From the equations viewpoint, elementary row operations are simplest transformations that do not change the set of solutions, so we may hope to use them to simplify the system enough to be easily solved.