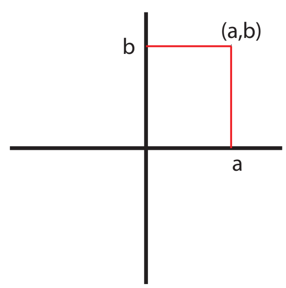
Research Methods in Mathematics
Lecture 13: Vectors and matrices

T. PERUTZ

Vectors in the plane

The plane $\mathbb{R}^2$ is the set of all pairs (a,b) of real numbers. We can picture (a,b) as the point on a coordinate plane where $x = a$ and $y = b$.



A *vector* $\mathbf{v}$ is the same thing as a point in the plane, i.e., a pair of real numbers (a,b). But we use different typography, writing our vectors as columns:

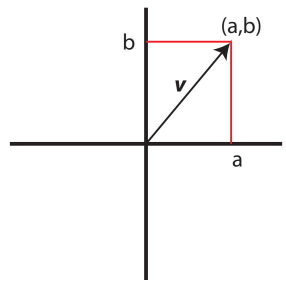
$$\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$$

We picture $\mathbf{v}$ as an arrow beginning at the origin whose tip is at the point (a,b). Notice that this arrow contains exactly the same information as the point (a,b).

This arrow has a length $\|\mathbf{v}\|$, given by

$$\|\mathbf{v}\| = \sqrt{a^2 + b^2}.$$

The length is zero if and only if $\mathbf{v}$ is the zero vector $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.



If $\mathbf{v} \neq \mathbf{0}$, the arrow also has a direction (more on that later).

We declare a 'scalar multiplication' rule for multiplying a vector $\mathbf{v}$ by a real number (scalar) s to get a vector $s\mathbf{v}$:

$$s\mathbf{v} = s\begin{bmatrix} a \\ b \end{bmatrix} := \begin{bmatrix} sa \\ sb \end{bmatrix}.$$

If $s > 0$, then this leaves the direction of $\mathbf{v}$ unchanged, while if $s < 0$ it reverses the direction. The effect on the length is given by the formula

$$\|s\mathbf{v}\| = |s|\|\mathbf{v}\|.$$

Some simple algebraic properties:

$$(st)\mathbf{v} = s(t\mathbf{v}), \quad 1\mathbf{v} = \mathbf{v}, \quad 0\mathbf{v} = \mathbf{0}.$$

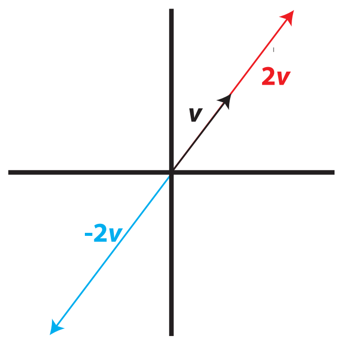
We also declare an 'addition rule' for adding two vectors $\mathbf{u}$ and $\mathbf{v}$ to get another vector $\mathbf{u} + \mathbf{v}$:

$$\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} a' \\ b' \end{bmatrix} := \begin{bmatrix} a + a' \\ b + b' \end{bmatrix}.$$

We can easily check the following algebraic properties:

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}, \quad (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}), \quad \mathbf{u} + \mathbf{0} = \mathbf{u},$$

$$s(\mathbf{u} + \mathbf{v}) = s\mathbf{u} + s\mathbf{v}, \quad (s + t)\mathbf{u} = s\mathbf{u} + t\mathbf{u}.$$



Notice also that

$$(-1)\mathbf{v} + \mathbf{v} = \mathbf{0}.$$

We usually write $-\mathbf{v}$ for $(-1)\mathbf{v}$.

The arrow picture gives a geometric interpretation of vector addition. To find $\mathbf{u} + \mathbf{v}$, represent both vectors as arrows. Move the arrow $\mathbf{v}$, leaving its length and direction unchanged, so that it begins at the tip of $\mathbf{u}$. Now draw an arrow (vector) $\mathbf{w}$ from the origin to the tip of this translated version of $\mathbf{v}$. Then $\mathbf{w} = \mathbf{u} + \mathbf{v}$.

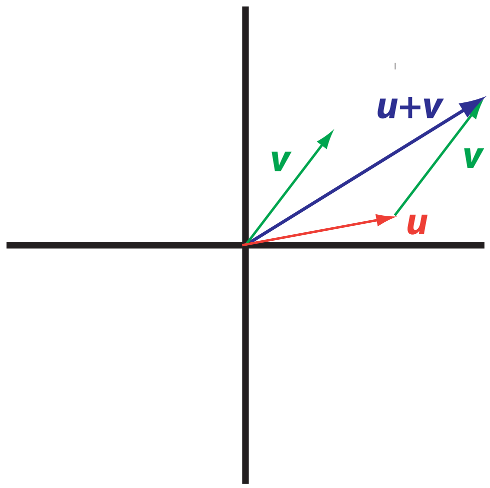
The triangle that's involved in the geometric interpretation of vector addition can be read in a second way. If we put $\mathbf{x} = \mathbf{u}+\mathbf{v}$, we have $\mathbf{u} = \mathbf{x}-\mathbf{v}$. Geometrically, to obtain the difference $\mathbf{x} - \mathbf{v}$ we begin at the origin, go out along $\mathbf{x}$, and then go backwards along a copy of $\mathbf{v}$ that begins where $\mathbf{x}$ ends. We end up at the tip of the arrow $\mathbf{u}$.

Two-by-two matrices

A 2×2 matrix is an array of real numbers

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The theme of this part of this course will be that matrices are not simply tables of data; they also have geometric and algebraic significance. The geometry illuminates the algebra, and vice versa.



The most important thing we can do with a 2×2 matrix A is that for any vector $\mathbf{v}$ we can 'multiply' it by A to obtain another vector denoted by $A\mathbf{v}$. The definition of 'multiplication' says that

$$A\mathbf{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} av_1 + bv_2 \\ cv_1 + dv_2 \end{bmatrix}.$$

For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \mathbf{v}.$$

So the matrix

$$I := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

leaves every vector unchanged. This matrix I is called the *identity matrix*.

For another example,

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -v_2 \\ v_1 \end{bmatrix}.$$

Geometrically, this means that the point (v_1, v_2) is rotated by 90° counterclockwise about the origin.

If we have two matrices

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix},$$

we can ask what happens if we form the product $B\mathbf{v}$, then apply A to that to obtain $A(B\mathbf{v})$. Let's compute:

$$\begin{aligned} A(B\mathbf{v}) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \left(\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a'v_1 + b'v_2 \\ c'v_1 + d'v_2 \end{bmatrix} \\ &= \begin{bmatrix} (aa' + bc')v_1 + (ab' + bd')v_2 \\ (ca' + dc')v_1 + (cb' + dd')v_2 \end{bmatrix} \\ &= \begin{bmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}. \end{aligned}$$

We learn from this calculation that $A(B\mathbf{v})$ is the equal to $C\mathbf{v}$, for a certain matrix C:

$$C = \begin{bmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{bmatrix}.$$

We choose to call C the *matrix product* of A and B, and we denote it by AB, so that the result of our computation is the natural-looking equation

$$A(B\mathbf{v}) = (AB)\mathbf{v}.$$

Here are two important things to note about matrix products. The first is that they are associative. It's possible to prove this by a big computation, but this is tedious and utterly unenlightening. We can also prove it without doing any computation, as follows. For matrices X, Y and Z, we have

$$((XY)Z)\mathbf{v} = (XY)(Z\mathbf{v}) = X(Y(Z\mathbf{v})) = (X(YZ))\mathbf{v}.$$

Two matrices A and B such that $A\mathbf{v} = B\mathbf{v}$ for every vector $\mathbf{v}$ are equal: $A = B$. (Can you prove this?) Applying this to $A = (XY)Z$ and $B = X(YZ)$, we find that

$$(XY)Z = X(YZ).$$

Secondly, they are not always commutative: for certain A and B, we have $AB \neq BA$. For example, take

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$