Vectors for Beginners

Leo Dorst*

September 6, 2010

1 Three ways of looking at linear algebra

We will always try to look at what we do in linear algebra at three levels:

- *geometric*: drawing a picture. This is often related to the application.
- *computational*: giving a formula. This is related to the implementation (in Matlab or other code).
- *symbolic*: specifying the essence of the computation, using the language elements of linear algebra.[1]

To use linear algebra effectively in computer science, you have to try and encode the geometric pictures of your application, into the symbolic level; the computational level then follows automatically. Thinking at the computational level is tiresome, and will make linear algebra appear like a bag of tricks to memorize, rather than as a language to master fluently.

So the order in the practice of computer science is typically: geometry → symbolic → computation. Often the order of definition in the Bretscher book will be: geometry → computation → symbolic; or even computation → symbolic→ geometry.

2 Vector properties

What are vectors?

- *geometric*: For now, vectors are like arrows starting from a common point (the *origin*) in an n-dimensional space. In our examples we use 2-D and 3-D. A vector has a direction, and a length. That length is called its *norm*.

  Vectors are used to represent many quantities in the real world. Concrete examples are displacements, movements, velocities and forces.

- *symbolic*: We will denote a vector as a variable with an arrow overhead: the vector $\vec{x}$. The norm (length) is written as $\|\vec{x}\|$. We can only specify the direction relative to standard vectors. In n dimensions, you need n standard vectors, called a *basis*. We will denote those standard vectors as $\vec{e}_1, \cdots \vec{e}_n$. We often choose them perpendicular and of unit length; then they point in the direction of the coordinate axes that in high school you used to call the x-axis, y-axis, et cetera. Since we run out of the alphabet soon (already after 3 dimensions, and that is not enough for computer science!), and because programs typically run over the indices of the dimensions, we will use $\vec{e}_i$ instead, with i running from 1 to n.

- *computational*: You can think of an n-D vector as n numbers. These numbers give the components relative to the standard vectors. Adding these components (see below) gives the vector. It is

*AI student Eva Greiner helped improve this document in 2009 through her many helpful comments on the earlier version.

[1] Vectors and hyperplanes are the nouns, products and matrices (we'll meet those later) are the verbs.

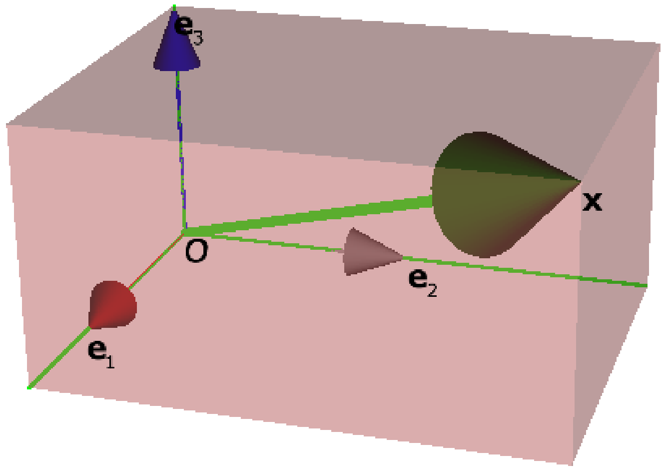


Figure 1: *The vector* $\vec{x} = 1.5\,\vec{e}_1 + 2\,\vec{e}_2 + 1\,\vec{e}_3$. *From LAcoord(). (These indications in figure captions refer to GAViewer scripts used to generate them. GAViewer and the scripts are available via the BlackBoard pages of this class. In these and other figures, the vectors are denoted by bold font rather than by an overhead arrow.)*

traditional to write the n numbers in a column, with square brackets, and label them with the basis vector they refer to. So

$$\vec{x} \mapsto \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{1}$$

is the computational representation of the 3-dimensional vector $\vec{x}$ in the basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$.

You see that the geometry and the computational representation have to be specific on the dimensionality, whereas the symbolic representation $\vec{x}$ need not. That is an advantage of working symbolically, for we are going to need spaces of many different dimensionalities in computer science.

3 Basic operations on vectors: scaling and adding

3.1 Scaling

You can change the length of a vector, making it λ times longer. Here λ is a real number, allowed to be positive, zero or negative. In linear algebra, numbers are often called *scalars*, because they can 'scale' the vectors.

- *geometric*: Multiplying the vector $\vec{x}$ by λ makes the vector λ times as long. The resulting vector still points in the same direction if $\lambda > 0$, in the opposite direction if $\lambda < 0$, and has lost its direction when $\lambda = 0$ (it then gives the null vector: no direction, no length, so hardly a vector at all).
- *symbolic*: We denote the product of a scalar (number) λ and a vector $\vec{x}$ as $\lambda\,\vec{x}$.
- *computational*: The computational representation of $\lambda\vec{x}$ is obtained by multiplying each of its components by λ:

$$\lambda\,\vec{x} \mapsto \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \end{bmatrix}. \tag{2}$$

  That reduces the implementation of vector scaling to the multiplication of real numbers.

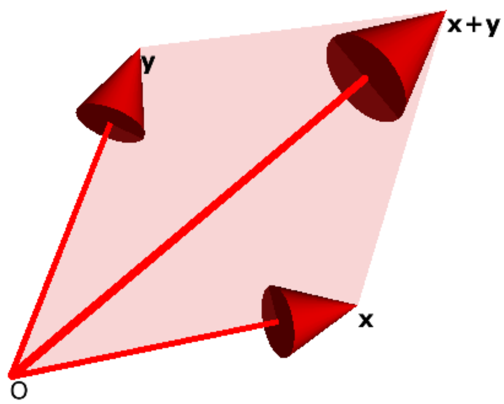


Figure 2: *Addition of the vectors $\vec{x}$ and $\vec{y}$ to form the vector $\vec{x}+\vec{y}$. (In these figures, we label the vectors with bold font, rather than overhead arrows.) From LAadd().*

Using the computational approach, you should now be able to prove:

$$\textit{distributivity:} \quad (\lambda + \mu)\,\vec{x} \quad = \quad \lambda\,\vec{x} + \mu\,\vec{x}, \tag{3}$$

since this property reduces completely to the analogous property for the individual scalars of the components of the vectors.

3.2 Addition

You can also add two vectors $\vec{x}$ and $\vec{y}$ to make a new vector which is called their *sum*.

- *geometric*: Vector addition is done by completing the parallelogram of which $\vec{x}$ and $\vec{y}$ are the sides, and using the diagonal from the origin as the resulting sum vector (see Figure 2).
- *symbolic*: We denote the sum as $\vec{x} + \vec{y}$.
- *computational*: We compute the sum as:

$$\vec{x} + \vec{y} \mapsto \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix}. \tag{4}$$

  That reduces the vector addition to n additions of real numbers for the components.

Since the computational definition defines the sum of vectors, it can be used to prove the following symbolic properties:

$$\begin{array}{lrcl} \textit{commutativity:} & \vec{x}+\vec{y} & = & \vec{y}+\vec{x} \\ \textit{associativity:} & (\vec{x}+\vec{y})+\vec{z} & = & \vec{x}+(\vec{y}+\vec{z}) \end{array} \tag{5}$$

Prove this for yourself. The geometry of the commutativity can be inferred from Figure 2, for the parallelogram can be made as $\vec{x}+\vec{y}$ or as $\vec{y}+\vec{x}$ with the same 'diagonal' as a result. The associativity property is illustrated in Figure 3 (analyze it!).

The scaling of a vector behaves nicely relative to the addition:

$$\textit{distributivity:} \quad \lambda\,(\vec{x}+\vec{y}) \quad = \quad \lambda\,\vec{x} + \lambda\,\vec{y} \tag{6}$$

Prove this for yourself.

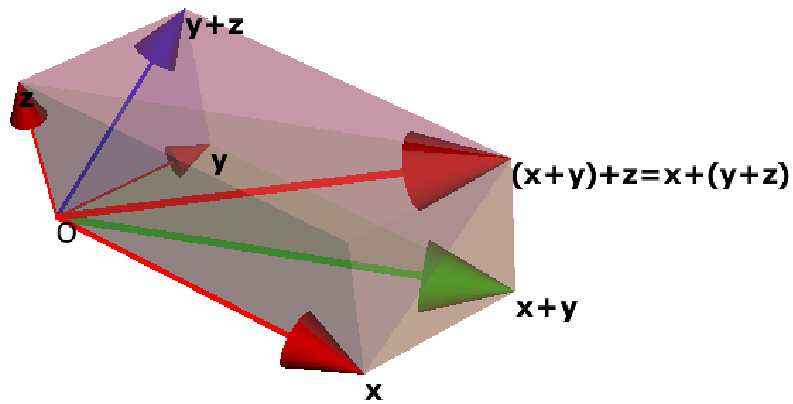


Figure 3: *The associativity of the addition of 3 vectors $\vec{x}$, $\vec{y}$ and $\vec{z}$. From LAasso().*

3.3 The meaning of coordinates

Now that we have scaling and addition, we can specify the connection between the coordinates and the symbolic notation, through the basis vectors. It is that a vector with coordinates x_1, x_2 and x_3 can be composed using a sum of scaled basis vectors $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ as:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\,\vec{e}_1 + x_2\,\vec{e}_2 + x_3\,\vec{e}_3 \tag{7}$$

Read the other way, we can say that a 3-D vector can be *decomposed* into 3 components along the coordinate axis directions.

We can write the vector $\vec{x}$ in Figure 1 in several different ways:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} = 1.5 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 1.5\,\vec{e}_1 + x_2\,\vec{e}_2 + x_3\,\vec{e}_3.$$

The trick in these conversions is to realize that the basis coordinate vectors can be represented in coordinates as:

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

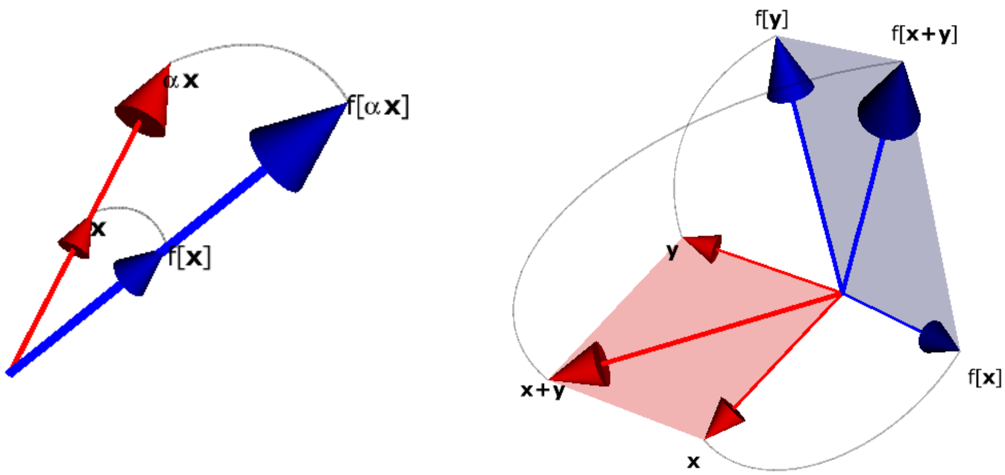
(It should be clear how that extends to more dimensions.) The conversions then just follow from the basis vector properties of multiplication and addition.

With these conversions between the two representations, we can use either to specify our computations (but as we hinted before, the symbolic approach is what you should use in n-D).

3.4 Representing points

Points of the world can now be represented by choosing an origin (arbitrary, but fixed), and using the vector $\vec{x}$ from the origin to the point X to represent the point mathematically. (We will meet better ways to represent points by vectors in robotics, computer vision and in computer graphics, but this is good enough for now.)

If you need a bit of intuition to see how we can then represent lines of points, take a peek at Section 5.1. You can understand the first part of that now and may find it motivating at this point.



(a) (b)

Figure 4: *The linear transformation* $\mathsf{f}[\cdot]$ *applied to vectors satisfies by definition the linearity properties eq.(8). This is the geometry of those properties:* (a) *scaling by* α *can be done before or after the transformation* f*;* (b) *addition can be done before or after.*

3.5 Linearity

In linear algebra, we are going to be interested in operations on vectors. For example, rotating a vector to have it point to another direction. For the moment, we denote the rotation of a vector $\vec{x}$ around an axis through the origin as $R(\vec{x})$. We are not going to define it computationally yet, but you may check geometrically that this rotation operation has the properties:

$$\begin{array}{lrcl} \textit{scaling:} & R(\lambda\,\vec{x}) & = & \lambda\,R(\vec{x}) \\ \textit{addition:} & R(\vec{x}+\vec{y}) & = & R(\vec{x})+R(\vec{y}) \end{array} \tag{8}$$

This just means that you get the same result if you first scale and then rotate, or first rotate and then scale; and the same for addition. The pictures associated with these identities are as in Figure 4.

These are called *the linearity properties*, and an operation with these properties is called a *linear operation* (or linear mapping, or linear transformation). Linear algebra is all about these kinds of operations, and gives handy symbolic theory and powerful computational techniques for them.

As an example of the linearity properties, let us take the clockwise rotation of 90 degrees around the $\vec{e}_3$-axis. It transforms a vector in the following way, as you can see from a quick sketch:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \mapsto \quad R(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}) = \begin{bmatrix} -x_2 \\ x_1 \\ x_3 \end{bmatrix}.$$

To test the linearity properties we take two vectors, for instance:

$$\vec{x} = \begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{y} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \tag{9}$$

Now if we scale $\vec{x}$ by a factor 2, the scaling property holds for the rotation:

$$R(2\,\vec{x}) = R(2\begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix}) = R(\begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}) = \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} = 2\begin{bmatrix} -2 \\ 1.5 \\ 1 \end{bmatrix} = 2\,R(\vec{x}).$$

Make sure you understand each step here, by demonstrating the scaling property for the vector $\vec{y}$ yourself. For the sum, we have:

$$R(\vec{x}+\vec{y}) = R(\begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}) = R(\begin{bmatrix} 0.5 \\ 2 \\ 2 \end{bmatrix}) = \begin{bmatrix} -2 \\ 0.5 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1.5 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = R(\vec{x}) + R(\vec{y}).$$

So the addition property also holds for these two vectors $\vec{x}$ and $\vec{y}$. To actually show the linearity properties for this $R(\cdot)$, we would need to show that they hold for *all* vectors. We will get into that later, when we use the Bretscher book.

Linear operations occur a lot in practice, and that makes linear algebra very useful. In fact, it is so useful that if an operation is not linear, we try to *linearize* it, i.e., make it linear. Sometimes we succeed exactly, by a clever trick of choosing our vectors; sometimes we can only linearize approximately. But linearization enormously extends the usefulness of linear algebra. Because this is such a powerful approach to all sorts of problems (not only obviously geometrical ones!), we teach linear algebra in computer science.

When you have a hammer, everything looks like a nail.

4 The Dot Product

There is a product for vectors, called the *dot product* (or *inner product*), which produces a scalar number from two vectors. We define how to compute it first, and then try to interpret the meaning.

- *computational*: The dot product of two 3-D vectors $\vec{x}$ and $\vec{y}$ is defined as:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = x_1\, y_1 + x_2\, y_2 + x_3\, y_3. \tag{10}$$

It is a number. In linear algebra, a number is called a *scalar*, since it has only a scale, not a direction (in contrast to vectors).

Let us compute the dot product of the two vectors of eq.(9). From eq.(10) we get:

$$\vec{x}\cdot\vec{y} = \begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = (1.5)\times(-1) + 2\times 0 + 1\times 1 = -0.5. \tag{11}$$

Note that the dot product of two vectors is indeed a number.

For the dot product of vectors in n-D, just add n terms in definition eq.(10).

- *symbolic*: From the computational definition, you can prove for yourself the following properties at the symbolic level:

$$\begin{array}{lrcl} \textit{symmetry:} & \vec{x}\cdot\vec{y} & = & \vec{y}\cdot\vec{x} \\ \textit{scaling:} & (\lambda\,\vec{x})\cdot\vec{y} & = & \lambda\,(\vec{x}\cdot\vec{y}) \\ \textit{distributivity:} & (\vec{x}+\vec{y})\cdot\vec{z} & = & (\vec{x}\cdot\vec{z}) + (\vec{y}\cdot\vec{z}) \end{array} \tag{12}$$

The dot product of vectors therefore seems to behave very much like multiplication of numbers.

You can use the properties of eq.(12) to derive more complicated equations, by 'pattern matching'. We can show that process explicitly with brackets. For example,

$$\begin{aligned} (\vec{a}-\vec{b})\cdot(\vec{c}+2\,\vec{d}) &= \left(\vec{a}+(-\vec{b})\right)\cdot\left(\vec{c}+(2\,\vec{d})\right) \\ &\overset{\text{distr.}}{=} \Big(\vec{a}\cdot\left(\vec{c}+(2\,\vec{d})\right)\Big) + \Big((-\vec{b})\cdot\left(\vec{c}+(2\,\vec{d})\right)\Big) \\ &\overset{\text{symm.}}{=} \Big(\left(\vec{c}+(2\,\vec{d})\right)\cdot\vec{a}\Big) + \Big(\left(\vec{c}+(2\,\vec{d})\right)\cdot(-\vec{b})\Big) \\ &\overset{\text{dist.}}{=} \Big(\vec{c}\cdot\vec{a} + (2\,\vec{d})\cdot\vec{a}\Big) + \Big(\vec{c}\cdot(-\vec{b}) + (2\,\vec{d})\cdot(-\vec{b})\Big) \\ &\overset{?}{=} \vec{a}\cdot\vec{c} \;+\; 2\,\vec{a}\cdot\vec{d} \;-\; \vec{b}\cdot\vec{c} \;-\; 2\,\vec{b}\cdot\vec{d}. \end{aligned}$$

We combined some steps in the final simplification, spell those out for yourself. You see that the result is just as you would expect if we had been combining numbers, not vectors; then we would have $(a - b)(c + 2d) = ac + 2ad - bc - 2bd$. So in practice, it is not as tedious as this example might suggest: you can write out the products as easily as for numbers.

However, the dot product is *not associative*, i.e., it does *not* satisfy $(\vec{x} \cdot \vec{y}) \cdot \vec{z} = \vec{x} \cdot (\vec{y} \cdot \vec{z})$ (whereas scalars do). We cannot even express associativity properly, for $(\vec{x} \cdot \vec{y}) \cdot \vec{z}$ is not defined: $(\vec{x} \cdot \vec{y})$ is a scalar, and we do not know how to take the dot product of a scalar and a vector $\vec{z}$. Because of such difficulties, *you should be very accurate in your notation of the products*. Do not use $\cdot$ for scalar multiplication. Do not use $\times$ either (we will give it a different meaning later). Just use a space for regular multiplication of scalars, and reserve $\cdot$ for a true dot product, of vectors only.

- *geometric*: It turns out that the geometric interpretation of the dot product is an interesting combination of angles and lengths. Let the norm (length) of the vector $\vec{x}$ be $\|\vec{x}\|$, similar for $\vec{y}$, and let the angle from $\vec{x}$ to $\vec{y}$ be ϕ. Then it can be shown that:[2]

$$\vec{x} \cdot \vec{y} = \|\vec{x}\| \, \|\vec{y}\| \, \cos(\phi). \tag{13}$$

So two things that we normally think of as different come together here: distance (norm) and angle (at least its cosine).

As a 2-D example, let us take the vectors $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. A sketch shows that these vectors clearly make an angle of 45 degrees, of which the cosine is $\frac{1}{2}\sqrt{2}$, and that their lengths are $\|\vec{x}\| = 1$ and $\|\vec{y}\| = \sqrt{2}$. And indeed:

$$\vec{x} \cdot \vec{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \times 0 + 1 \times 1 = 1 = 1 \times \sqrt{2} \times \tfrac{1}{2}\sqrt{2} = \|\vec{x}\| \, \|\vec{y}\| \, \cos(\phi).$$

When working with vectors, you should use the dot '$\cdot$' when you mean the dot product. To prevent confusion, it is customary to leave out a multiplication symbol when doing a multiplication of scalars. So $\vec{a} \cdot \vec{b}$ is clearly a dot product between vectors $\vec{a}$ and $\vec{b}$, and $\alpha\beta$ is the classical scalar multiplication of two scalars (numbers) α and β. (We will later define $\vec{a} \times \vec{b}$ for vectors, so the '$\times$' symbol should not be used for scalar multiplication either.)

4.1 Lengths and Angles

Let us play some more with the interpretations. If we take the dot product of a vector with itself, we get computationally:

$$\vec{x} \cdot \vec{x} \mapsto x_1^2 + x_2^2 + x_3^2 \tag{14}$$

and from (13), since the angle between $\vec{x}$ and itself is zero (and its cosine therefore equal to 1):

$$\vec{x} \cdot \vec{x} = \|\vec{x}\|^2. \tag{15}$$

Combined, this gives us a formula to compute the norm (length) of a vector $\vec{x}$:

$$\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x_1^2 + x_2^2 + x_3^2}. \tag{16}$$

This is of course Pythagoras' theorem, applied to the coordinates of $\vec{x}$, now in 3-D with the basis vectors $\vec{e}_1, \vec{e}_2, \vec{e}_3$. Since $\vec{x} \cdot \vec{x}$ can be computed in n-D, the symbolic formula $\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}}$ is valid in n-D, and the computational formula can be extended simply to that case.

The norm (length) of the vector $\vec{x}$ in eq.(9) is (using $\times$ for scalar multiplication here):

$$\|\vec{x}\| = \left\| \begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} \right\| = \sqrt{\begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix}} = \sqrt{1.5 \times 1.5 + 2 \times 2 + 1 \times 1} = \sqrt{7.25} \approx 2.69. \tag{17}$$

Compute yourself that the norm of the vector $\vec{y}$ in eq.(9) is: $\|\vec{y}\| = \sqrt{2} \approx 1.41$.

[2] But we will not show this here, we just give the result. We will encounter it later in Bretscher. You may attempt a computational proof now, steered by some geometrical insights...

Another special case of (13) is when $\vec{x}$ and $\vec{y}$ are perpendicular. Then $\phi = \pm\pi/2$ (in high school, you called this ± 90 degrees but from now on you should work in radians), so $\cos(\phi) = 0$. Conversely, if the dot product of two non-zero vectors is zero, they must be perpendicular. So we get:

$$\vec{x} \text{ perpendicular to } \vec{y} \;\Leftrightarrow\; \vec{x} \cdot \vec{y} = 0.$$

Please realize that this is very unlike what we know of multiplication of numbers: if the product of numbers satisfies $x\,y = 0$, then $x = 0$ or $y = 0$ (or both). Vectors can have a zero dot product without themselves being zero!

We can also use (13) to make a symbolic or computational formula for the angle ϕ between two vectors:

$$\cos(\phi) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\|\,\|\vec{y}\|} = \frac{x_1 y_1 + x_2 y_2 + x_3 y_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}\,\sqrt{y_1^2 + y_2^2 + y_3^2}}. \tag{18}$$

This can also be extended to n-dimensional space, in the obvious way, by just adding more terms.

Example: We have computed the dot product and the norms of $\vec{x}$ and $\vec{y}$ in eq.(11) and eq.(17), so we can easily compute the cosine of their angle:

$$\cos(\phi) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\|\,\|\vec{y}\|} = \frac{-0.5}{\sqrt{7.25}\,\sqrt{2}} = -0.1313.$$

Using the acos or $\cos^{-1}$ function, the angle is then about 1.605 radians.

4.2 Projections and Components

A useful special case is if $\vec{y}$ is a *unit vector* $\vec{e}$, which is a vector with unit norm $\|\vec{e}\| = 1$. You can obtain the unit vector in the $\vec{y}$ direction from a non-unit vector $\vec{y}$ through dividing it by its norm:

$$\vec{e} = \frac{\vec{y}}{\|\vec{y}\|} = \frac{\vec{y}}{\sqrt{\vec{y} \cdot \vec{y}}} \qquad \left(\text{and in 2-D this would give } \vec{e} = \begin{bmatrix} y_1/\sqrt{y_1^2 + y_2^2} \\ y_2/\sqrt{y_1^2 + y_2^2} \end{bmatrix}\right).$$

It is easy to prove that $\vec{e}$ is a unit vector:

$$\|\vec{e}\|^2 = \vec{e} \cdot \vec{e} = \frac{\vec{y}}{\|\vec{y}\|} \cdot \frac{\vec{y}}{\|\vec{y}\|} = \frac{\vec{y} \cdot \vec{y}}{(\sqrt{\vec{y} \cdot \vec{y}})^2} = \frac{\vec{y} \cdot \vec{y}}{\vec{y} \cdot \vec{y}} = 1. \tag{19}$$

Make sure you can relate all steps to properties of dot products and vectors! So the length of $\vec{e}$ is 1, as we wanted. We also need to check that $\vec{e}$ is in the $\vec{y}$ direction, for $\vec{e} \cdot \vec{y} = \sqrt{\vec{y} \cdot \vec{y}} = \|\vec{y}\| = \|\vec{e}\|\,\|\vec{y}\|$, so their cosine is 1 and their angle therefore 0.

Computing the dot product of $\vec{x}$ with $\vec{e}$ rather than $\vec{y}$ gives:

$$\vec{x} \cdot \vec{e} = \|\vec{x}\|\,\cos(\phi). \tag{20}$$

This is precisely the length of the perpendicular projection of $\vec{x}$ on $\vec{e}$, see Figure 5 (remember the geometric meaning of the cosine, see also Figure 8). Therefore, if we multiple the unit vector $\vec{e}$ by this length, we get the component $(\vec{x} \cdot \vec{e})\,\vec{e}$ of $\vec{x}$ along the $\vec{e}$ direction.

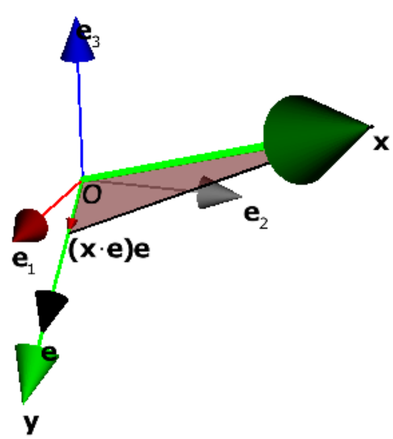
Example: To compute the projection of $\vec{x}$ onto $\vec{y}$ (of eq.(9)), we first make a unit vector $\vec{e}$ in the $\vec{y}$-direction:

$$\vec{e} = \frac{\vec{y}}{\|\vec{y}\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{bmatrix}.$$

Then we get for the projection of $\vec{x}$ onto the $\vec{y}$-direction:

$$\begin{aligned} (\vec{x} \cdot \vec{e})\,\vec{e} &= \left(\begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{bmatrix}\right) \begin{bmatrix} -\frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{bmatrix} \\ &= (-\tfrac{1}{4}\sqrt{2}) \begin{bmatrix} -\frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ 0 \\ -\frac{1}{4} \end{bmatrix}. \end{aligned}$$

We will later learn how to do this more conveniently.



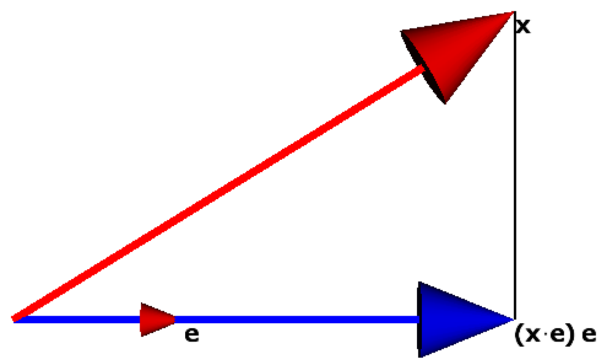


Figure 5: *Projection of $\vec{x}$ onto a unit vector $\vec{e}$ using the inner product, resulting in $(\vec{x} \cdot \vec{e})\,\vec{e}$. From LAip().*

4.3 Proof of Pythagoras

Using the definition of the norm eq.(16) in terms of the dot product, and the properties of the dot product given in eq.(12), you can now prove Pythagoras' theorem for vectors, in Exercise 8 below.

At the symbolic level, Pythagoras' theorem reads:

$$\vec{x} \text{ is perpendicular to } \vec{y} \quad \Leftrightarrow \quad \|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2. \tag{21}$$

(Here $\Leftrightarrow$ means 'if and only if', sometimes written as 'iff'). This is a powerful result, simply expressed. Note that it is true in arbitrary dimensions!

4.4 Exercises:

1. What is the length of the vector $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$?
2. What is the length of the vector $2\vec{e}_1 + 3\vec{e}_2$?
3. What is the cosine of the angle between the vectors $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$?
4. What is the angle between the vectors $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$?
5. Make a unit vector in the same direction as $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$.
6. Make a unit vector in the same direction as $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.
7. In 2D, you compute the length of a vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ by Pythagoras as $\sqrt{x_1^2 + x_2^2}$. If you add a third dimension, your vector gets an extra component $x_3\vec{e}_3$. Show that by again applying 2D Pythagoras in a well-chosen plane, you find that the length of $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ equals $\sqrt{x_1^2 + x_2^2 + x_3^2}$ (which is eq.(16)). The same trick generalizes to n-D.

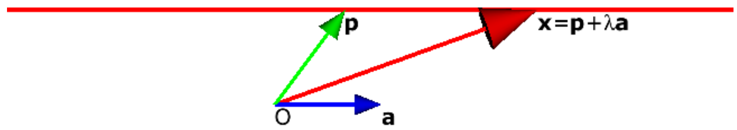


Figure 6: *Generating the vectors $\vec{x}$ of points on a line with position vector $\vec{p}$ and direction vector $\vec{a}$.* *From LAparline().*

8. Prove (21) at the symbolic level, so encode both sides in terms of things at that level (hint: use the dot product through eq.(16) and eq.(12) by means of 'pattern matching'). Avoid using the computational approach, do not spell everything out in coordinates, so that your proof holds in any number of dimensions!

9. Draw a picture of (21), for the geometrical approach. It is not quite the Pythagoras triangle, for we have said that vectors should start at the origin!

5 Lines and Planes in Linear Algebra

With what we have so far, we can represent lines and planes (and more).

5.1 A line in the plane

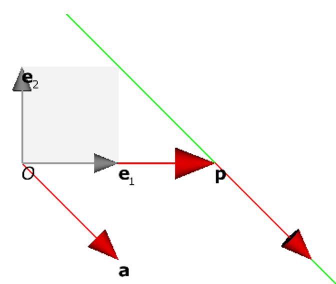
A line in the plane can be characterized using vectors.

- *geometric*: Geometrically, we can characterize points by the vectors pointing to them. A line travels in a certain direction, and in doing so passes through many points. Any of those points can be used to characterize the location of the line.

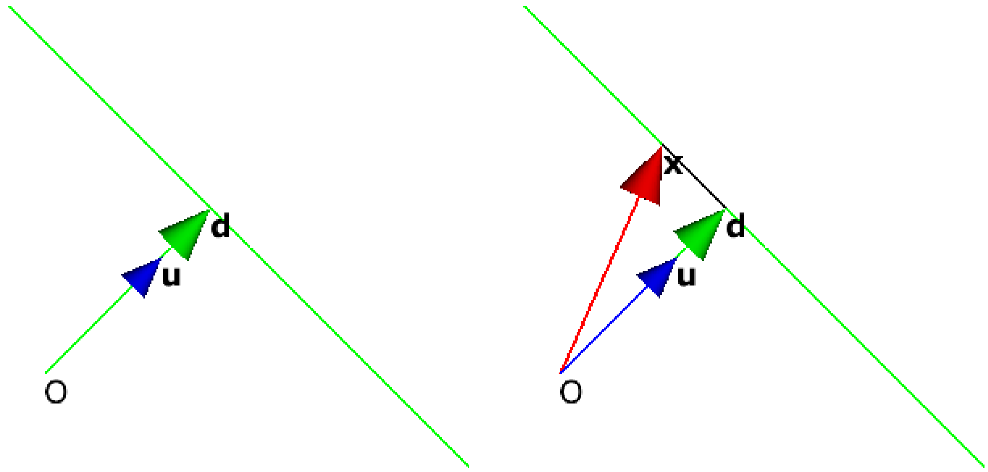
- *symbolic*: Any point vector $\vec{x}$ on a line can be made from two vectors: a *position vector* $\vec{p}$ and a *direction vector* $\vec{a}$, as:
$$\vec{x} = \vec{p} + \lambda\,\vec{a}, \tag{22}$$
where λ is some number. If $\lambda = 0$, you get $\vec{x} = \vec{p}$, and as λ gets bigger you get the points $\vec{p}+\vec{a}, \vec{p}+2\vec{a}$, et cetera. For negative λ, you go from $\vec{p}$ in the other direction. See Figure 6.

- *computational*: You can make the points on the line simply by plotting their coordinates, which are:
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} p_1 + \lambda a_1 \\ p_2 + \lambda a_2 \end{bmatrix}$$

This figure shows the line
$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{23}$$
It contains points like $(3, -1)$, $(1, 1)$, $(0, 2)$ – find out how to choose λ to get those points!



Usually, graphics packages contain functions that plot line segments between two points. You may want to try your hand at determining between what bounds λ should vary to generate the points from $\vec{p}$ to $\vec{q}$ (the symbolic derivation saves paper!). The answer is most conveniently written as $\vec{x} = (1-\mu)\,\vec{p} + \mu\,\vec{q}$, with μ varying from 0 to 1.



(a) (b)

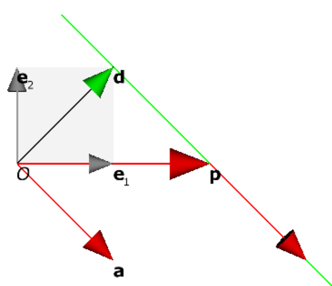
Figure 7: *Characterizing a line in 2D by a support vector $\vec{d}$ or a normal vector $\vec{u}$ and a distance δ such that $\vec{d} = \delta\,\vec{u}$. From LAline().*

(Stop here if you are reading this as a sideline to Section 3.4.)

We can also characterize the complete 2-D line by only 1 vector, and the method we use will be especially useful when we want to describe hyperplanes in higher-dimensional spaces.[3] For lines, this seems strange at first, but we will see that the computational description is recognizable.

- *geometric*: Look at Figure 7a. We have an origin O, and a green line which we want to characterize relative to this origin. A good idea is to use the vector $\vec{d}$, pointing from the origin perpendicular to the line. All points of the line are obtained by "going to the end of $\vec{d}$ and looking left and right".

  As you may be able to tell from the figure, the vector $\vec{d}$ for this line (same as the previous example) is $\vec{d} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. (But we should of course develop the proper symbolic and computational techniques.)



- *symbolic*: What does a vector $\vec{x}$ need to satisfy to point exactly to a point on the line? To encode that with previous constructions, let us make a unit vector $\vec{u}$ in the direction of $\vec{d}$, so that $\vec{d} = \delta\,\vec{u}$. Then δ is the distance of the line to the origin, in the $\vec{u}$-direction.

  Now we use (20): for any $\vec{x}$ on the line, the projection of $\vec{x}$ onto $\vec{u}$ should be have the length δ. You can see this clearly by comparing Figure 5 and Figure 7(b).

  When we express that geometric insight symbolically, we simply get:

$$\vec{x} \cdot \vec{u} = \delta, \quad \text{with } \vec{u} \cdot \vec{u} = 1 \tag{24}$$

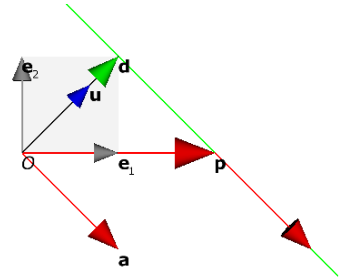
  as the characterization of the line.

[3] A hyperplane is an element of dimension $(n-1)$ in n-dimensional space, so a line (1D) in 2D is an example, and so is a plane (2D) in 3D space.

For the line in the figure, you see that $\vec{u} = \begin{bmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix}$ and $\delta = \sqrt{2}$, so that

$$\vec{x} \cdot \begin{bmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix} = \sqrt{2}. \tag{25}$$

is the characterization.



Now we can get back to the original vector $\vec{d}$ by multiplying both sides by δ. The left hand side then gives $\vec{x} \cdot \vec{d}$ (why?), the right hand side δ^2, which is the same as $\vec{d} \cdot \vec{d}$ (why?).

So $\vec{x}$ points to a point on the line characterized by $\vec{d}$ if and only if:

$$\vec{x} \cdot \vec{d} = \vec{d} \cdot \vec{d}. \tag{26}$$

Actually, (24) is a bit more general, for it still works when $\delta = 0$, whereas (26) then has problems. How would you characterize the problem geometrically (take a line through the origin; what property of the line can (26) not encode?).

In either (24) or (26), the vector $\vec{u}$ (or $\vec{d}$) are called a *normal vector* for the line. This is jargon, 'normal' here means 'perpendicular'. Since $\vec{u}$ is a unit vector, it is called a *unit normal (vector)* for the line, and since $\vec{d}$ 'carries' the line it is sometimes called the *support vector* of the line.

The relationship with the previous description (22) of the line is that $\vec{d}$ is like a position vector $\vec{p}$ (since it points to a point on the line), but moreover perpendicular to $\vec{a}$. So $\vec{x} = \vec{p} + \lambda\,\vec{a}$ becomes $\vec{x} = \vec{d} + \lambda\,\vec{a}$, and then $\vec{x} \cdot \vec{d} = \vec{d} \cdot \vec{d} + \lambda\,\vec{d} \cdot \vec{a} = \vec{d} \cdot \vec{d} + 0 = \vec{d} \cdot \vec{d}$.

> For the line above, the characterization analogous to eq.(26) leads to the computational expression
>
> $$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2. \tag{27}$$
>
> We can substitute eq.(23) to verify that this is indeed the same line we had originally:
>
> $$\begin{bmatrix} 2+\lambda \\ -\lambda \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = (2+\lambda) - (\lambda) = 2.$$

- *computational*: Now let us look at his computationally, taking (24) in 2-D. Then the unit vector $\vec{u}$ can be specified relative to some fixed coordinate basis $\{\vec{e}_1, \vec{e}_2\}$, where it has components u_1 and u_2. That means that the line equation $\vec{x} \cdot \vec{u} = \delta$ is expressed in coordinates as:

$$x_1\,u_1 + x_2\,u_2 = \delta, \quad \text{with } u_1^2 + u_2^2 = 1. \tag{28}$$

Such an equation is known as the 'normal equation of a line' (Dutch: normaalvergelijking), since it is based on the (unit) normal vector $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$.

> Working out eq.(25) we get $\frac{1}{2}\sqrt{2}\ x_1 + \frac{1}{2}\sqrt{2}\ x_2 = \sqrt{2}$ as the computational version of the line equation.

This begins to look like a line equation you may have seen in high school if you realize that x_1 is our notation for the x-coordinate and x_2 for the y-coordinate, and that u_1 and u_2 are given constants.

5.2 Lines in High School

You may not recognize eq.(28) and think of it as completely new. But in high school, you called x_1 (which is the component of $\vec{x}$ in the $\vec{e}_1$-direction) the x-coordinate of the point to which the vector $\vec{x}$ points, and x_2 the y-coordinate. So we get:

$$x\,u_1 + y\,u_2 = \delta, \quad \text{with } u_1^2 + u_2^2 = 1. \tag{29}$$

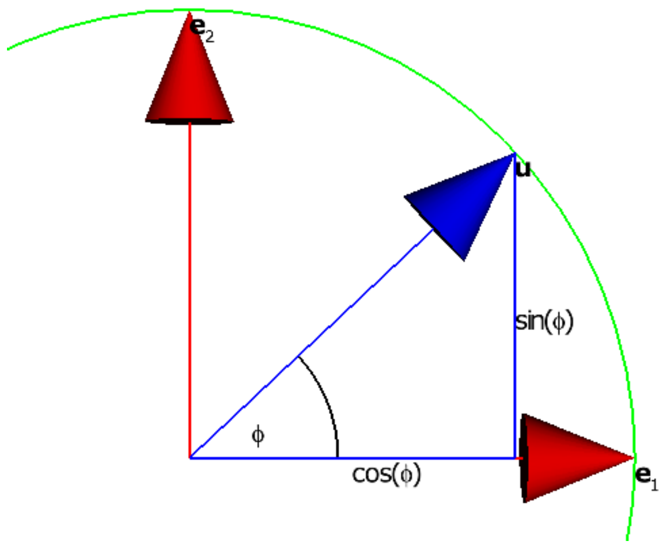


Figure 8: *The geometry of sine and cosine.* *From LAtriangle().*

Also in high school, you were used to having things of the form $y = f(x)$, since you were dealing with functions. We are dealing with geometry, but let us write our equation in the same form:

$$y = -\frac{u_1}{u_2}\,x + \frac{\delta}{u_2}, \quad \text{with } u_1^2 + u_2^2 = 1. \tag{30}$$

This is clearly something of the form

$$y = a\,x + b, \tag{31}$$

if we make the correct identifications by defining a and b in terms of u_1, u_2 and δ. So indeed, the original (28) derived from (24) defines a line in its more classical form.

The line in the example would be written as $y = -x + 2$, since $a = -\frac{u_1}{u_2} = -\frac{1}{1} = -1$, and $b = \frac{\delta}{u_2} = \frac{\sqrt{2}}{\frac{1}{2}\sqrt{2}} = 2$.

If the two representations are equivalent, why would we prefer (28) to (31)? For geometrical lines, which may occur in any orientation, the high school form (31) has problems. It is not defined when the line is 'vertical' (in the y-direction). Then a would have to be infinite, or we would have to switch over to a line of the form $x = c$. For hand computations that is OK, but in a computer we would prefer to have a representation that always works. So (24) is much better, for it has no problems with lines in any direction or location. It simply happens to be vertical when $u_2 = 0$, so that the line equation then is $x\,u_1 = \delta$. No different from how a horizontal line is represented through $u_1 = 0$, giving $y\,u_2 = \delta$.

5.3 Detailed Correspondence (For the Curious)

You should skip this section at first reading, it is rather technical!

To study the representation of lines with different orientations in general, let us have $\vec{u}$ make an angle ϕ with $\vec{e}_1$. Then its components are:

$$\vec{u} \mapsto \begin{bmatrix} \cos(\phi) \\ \sin(\phi) \end{bmatrix}, \tag{32}$$

as you can easily verify by drawing the triangle formed by the vector $\vec{u}$ and its components (see Figure 8). As a consequence (24) becomes, computationally:

$$x_1 \cos(\phi) + x_2 \sin(\phi) = \delta. \tag{33}$$

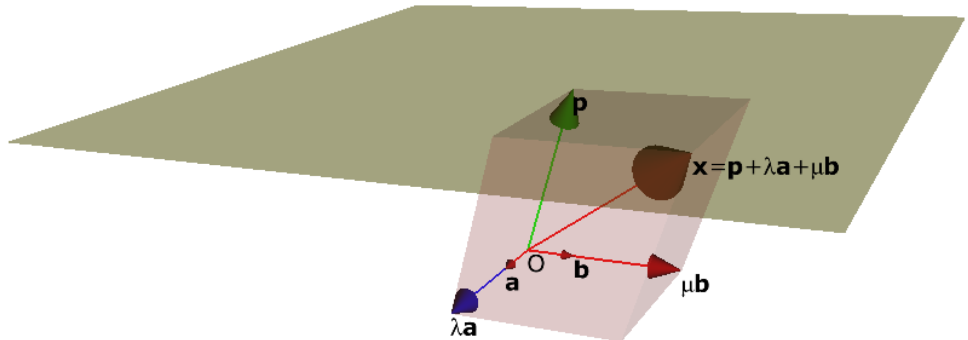


Figure 9: *A plane in space determined by three vectors. From LAparplane().*

(Note that $u_1^2 + u_2^2 = 1$ is automatically satisfied.) This may be rewritten in the form of (31):

$$y = -\frac{\cos(\phi)}{\sin(\phi)}\,x + \frac{\delta}{\sin(\phi)}, \tag{34}$$

though we can only do this when $\sin\phi \neq 0$. So indeed, we should avoid a line with horizontal normal vector (i.e., a vertical line), in this classical representation, whereas in (33), it was fine.[4] That is why we prefer (33) (or (24) or (28)).

5.4 Exercises

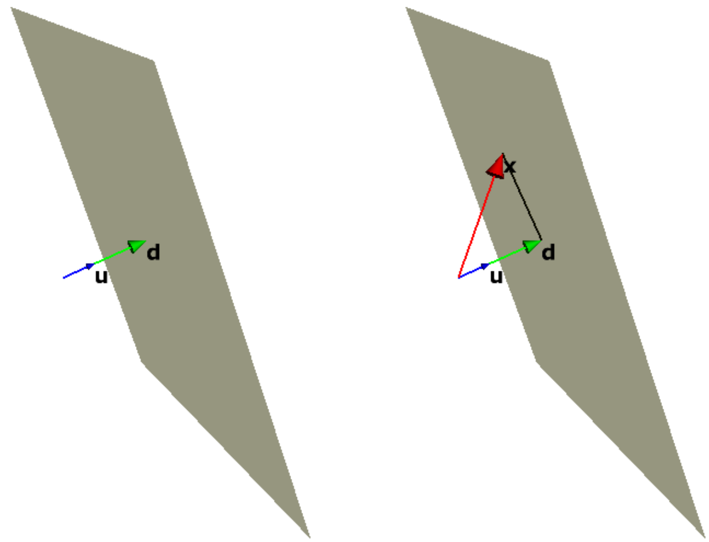
9. Express the high school line $y = x + 1$ using vectors (like eq.(24)).
10. Express the high school line $y = -x - 1$ using vectors (like eq.(24)), but using $\vec{u} = (1/\sqrt{2}, 1/\sqrt{2})$.
11. Draw the line with $\vec{u} = (1/\sqrt{2}, 1/\sqrt{2})$ and $\delta = 1$.
12. Draw the line with $\vec{d} = (1, -1)$.
13. What is the distance to the origin (δ) of the line $y = x + 1$?
14. What is the distance to the origin (δ) of the line $y = a\,x + b$, uitgedrukt in a en b?
15. Write the line $x_1 - 2x_2 = 1$ in the form of (28). What is its distance to the origin (δ) ?
16. What is the distance to the origin (δ) of the line $a_1x_1 + a_2x_2 = c$?

5.5 A plane in space

For planes in space, we also have two methods, both very similar to what we do for lines in the plane.

- *geometric*: A plane can be characterized as stretching in not one vector direction, but in many vector directions. Still, all those directions can be made from two direction vectors $\vec{a}$ and $\vec{b}$. They are not unique: many choices of $\vec{a}$ and $\vec{b}$ lead to the same plane. A plane also has a location, which is determined once we know one point $\vec{p}$ on it.

[4]If you compute a from (31) and (34), you get $a = -\frac{\cos(\phi)}{\sin(\phi)}$. Perhaps you had expected to recognize a tangent, like $a = \tan(\phi) = \frac{\sin(\phi)}{\cos(\phi)}$. But remember that ϕ is the angle of the normal vector, not of the line. The line angle is $\phi \pm \pi/2$, and you find $\tan(\phi \pm \pi/2) = \sin(\phi \pm \pi/2)/\cos(\phi \pm \pi/2) = \pm\cos(\phi)/\mp\sin(\phi) = -1/\tan(\phi) = a$. So all is well.



(a) (b)

Figure 10: *The normal representation of a plane in 3-D. From LAplane().*

- *symbolic*: Once we have chosen the direction vectors $\vec{a}$ and $\vec{b}$, and a point $\vec{p}$ on the plane, the position vector $\vec{x}$ of any point on the plane can be written as:

$$\vec{x} = \vec{p} + \lambda\,\vec{a} + \mu\,\vec{b}, \tag{35}$$

  see Figure 9.

- *computational*: The symbolic description translates immediately into a recipe for coordinates:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} p_1 + \lambda a_1 + \mu b_1 \\ p_2 + \lambda a_2 + \mu b_2 \\ p_3 + \lambda a_3 + \mu b_3 \end{bmatrix}$$

  In computer graphics, you often want to plot only a part of a plane, for instance the triangle with corner points $\vec{p}$, $\vec{q}$, $\vec{r}$. You might want to try to derive how to find the vectors of the points of that triangle. The answer is: $\vec{x} = (1 - \lambda - \mu)\,\vec{p} + \lambda\,\vec{q} + \mu\,\vec{r}$, with both λ and μ varying between 0 and 1.

As we did for lines, we can also characterize a plane by only one vector.

- *symbolic*: In the symbolic equations we got, (24) and (26), there is nothing stating that we are working with vectors in the plane. So we can apply these equations in 3-D space as well. Therefore a plane with unit normal vector $\vec{u}$ at a distance δ from the origin is given by all points with vectors $\vec{x}$ satisfying:

$$\vec{x} \cdot \vec{u} = \delta, \quad \text{with } \vec{u} \cdot \vec{u} = 1, \tag{36}$$

  or the conversion to the counterpart of eq.(26):

$$\vec{x} \cdot \vec{d} = \vec{d} \cdot \vec{d}, \tag{37}$$

  with $\vec{d}$ the support vector of the plane: $\vec{d} = \delta\,\vec{u}$.

- *geometric*: Of course eq.(36) is still interpretable as stating that the perpendicular projection of $\vec{x}$ along $\vec{u}$ has length δ. That suggests the picture of Figure 10(b). It therefore seems to characterize a plane in space, with support vector $\vec{d}$ as in Figure 10(a).

- *computational*: It is more difficult to define directions and angles in space, and we will not do that here. But it is simple to verify that the symbolic equation resulting from the dot product has the form:
$$a_1\, x_1 + a_2\, x_2 + a_3\, x_3 = c, \tag{38}$$
for some particular values of a_1, a_2, a_3 and c. This represents a plane with a normal vector (i.e., vector perpendicular to the plane)
$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}. \tag{39}$$
The norm of this vector is $\sqrt{a_1^2 + a_2^2 + a_3^2}$. Dividing both sides by this norm to go to the unit normal vector gives:
$$\frac{a_1\, x_1 + a_2\, x_2 + a_3\, x_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}} = \frac{c}{\sqrt{a_1^2 + a_2^2 + a_3^2}}. \tag{40}$$
This is an equation of the form $\vec{x} \cdot \vec{u} = \delta$ if we define the unit vector $\vec{u} = \vec{a}/\|\vec{a}\|$. It follows that the distance from the origin of the plane eq.(38) is by eq.(36)
$$\delta = \frac{c}{\sqrt{a_1^2 + a_2^2 + a_3^2}}. \tag{41}$$

Of course, this approach does not stop with 3-D. The equation (24) defines a *hyperplane* an n-dimensional space, i.e. a flat element of dimensionality $(n-1)$. In 2-D, the hyperplane is 1-D, and we are used to calling it a line; in 3-D, the hyperplane is 2-D and we call that a plane.

5.6 Exercises:

17. Give the equation of a plane with normal vector $\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ at distance -1 from the origin.
18. Where does the plane with equation $2x_1 + 3x_2 - 5x_3 = 6$ cut the $\vec{e}_1$-axis?
19. The plane with equation $2x_1 + 3x_2 - 5x_3 = 6$ cuts the plane with equation $x_3 = 0$ in a line. What is the equation of that line?

6 The 3D Cross Product

You may wonder how the two plane representations eq.(35) and eq.(37) (or eq.(36)) relate, one in terms of $\vec{a}$, $\vec{b}$ and $\vec{p}$ and the other with perpendicular support vector $\vec{d}$ (or unit normal vector $\vec{u}$ and distance δ). In particular, how can we make a normal vector $\vec{d}$ (or $\vec{u}$) that is perpendicular to the direction vectors $\vec{a}$ and $\vec{b}$?

For this, people have introduced a special product between 3D vectors. It is called the *cross product* because it is written as $\vec{a} \times \vec{b}$; its definition is most easily given in coordinates:
$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}. \tag{42}$$
Note that the cross product on two vectors is a vector (whereas the dot product of two vectors is a scalar).

You can verify that eq.(42) works: show that the result is perpendicular to $\vec{a}$ and $\vec{b}$ by demonstrating that $\vec{a} \cdot (\vec{a} \times \vec{b}) = 0$ and $\vec{b} \cdot (\vec{a} \times \vec{b}) = 0$. These equations show again that you have to very careful with your multiplication symbols: '$\cdot$' is the dot product, '$\times$' is the totally different cross product.

We state without proof that the norm of the resulting vector is proportional to the norms of the original vectors:
$$\|\vec{a} \times \vec{b}\| = \|\vec{a}\|\, \|\vec{b}\|\, |\sin(\phi)|,$$

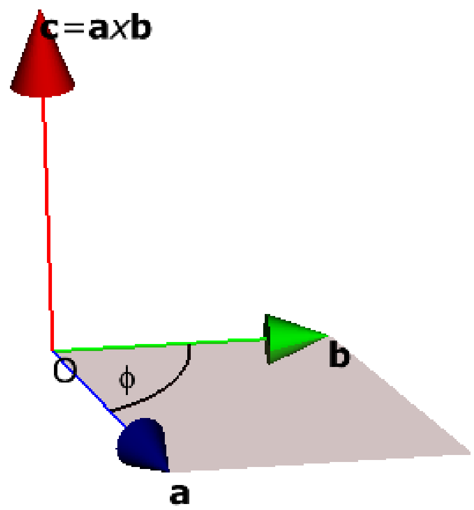


Figure 11: *The cross product $\vec{a} \times \vec{b}$ of two vectors.* From LAcross().

where ϕ is the angle between the vectors. This means that for parallel vectors $\vec{a}$ and $\vec{b}$, the cross product is zero. Geometrically that should be interpreted as: there is then no unique perpendicular direction to $\vec{a}$ and $\vec{b}$.

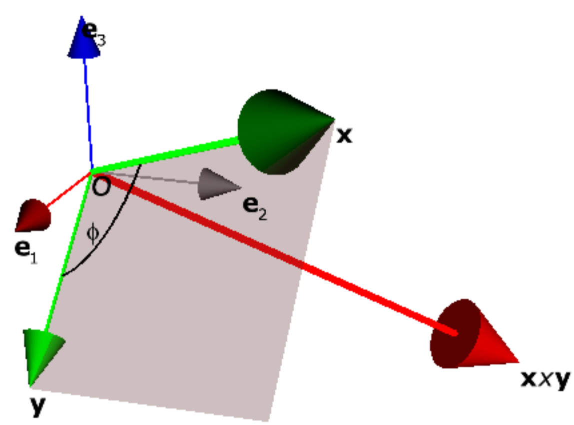
The direction of $\vec{a} \times \vec{b}$ is determined by the 'right-hand rule': if you hold $\vec{a} \times \vec{b}$ in your right hand with the arrow pointing up, then your fingers point from $\vec{a}$ to $\vec{b}$. Realize that this, combined with the norm, means that $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$; that result of course also follows from the computational definition above. If you get the order wrong, you get the opposite answer, for (42) immediately shows that $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$.

An example of the cross product computation is the cross product between the vectors

$$\vec{x} = \begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{y} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Following the above recipe, we find (using $\times$ to denote the scalar multiplication of coordinates for the moment):

$$\begin{aligned} \vec{x} \times \vec{y} &= \begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \times 1 & - & 1 \times 0 \\ 1 \times (-1) & - & 1.5 \times 1 \\ 1.5 \times 0 & - & 2 \times (-1) \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ -2.5 \\ 2 \end{bmatrix}. \end{aligned}$$



Note that the cross product of two vectors is indeed a vector (whereas the dot product of two vectors is a scalar number).

It is good to check whether the vector $\vec{x} \times \vec{y}$ is indeed perpendicular to both $\vec{x}$ and $\vec{y}$, by verifying that the corresponding dot products are zero. Do this yourself.

In contrast to the dot product, which was defined in n-D, the cross product is limited to 3-D. The reason is that only in 3D is the 'direction' perpendicular to two vectors characterizable by a single vector. We will learn what to do for perpendicularity in higher dimensions in the Bretscher book.

Back to the plane representation. The support vector $\vec{d}$ is perpendicular to $\vec{a}$ and $\vec{b}$, so it should be proportional to $\vec{a} \times \vec{b}$, and its length should be such that $\vec{p}$ is on the plane. Let $\vec{u} = (\vec{a} \times \vec{b})/\|\vec{a} \times \vec{b}\|$, the unit vector in the cross product direction. Then $\vec{d} = (\vec{p} \cdot \vec{u})\,\vec{u}$, the projection of $\vec{p}$ on that vector. So that is how the one vector $\vec{d}$ can be computed from the three others $\vec{p}$, $\vec{a}$ and $\vec{b}$, to get the plane equation in the form eq.(37). The representation of the plane by $\vec{d}$ is of course more efficient and unambiguous than by $\vec{p}, \vec{a}, \vec{b}$. For instance, it is much easier to test if two planes are the same (Question: How?).