

Mathematics for Graphics and Vision

Steven Mills

March 3, 2016

Contents

1	Introduction	5
2	Scalars	6
2.1	Visualising Scalars	6
2.2	Operations on Scalars	6
2.3	A Note on Multiplication Symbols	7
3	Trigonometric Functions	8
3.1	Degrees and Radians	9
3.2	Sine, Cosine, and Tangent	9
3.3	Plots of Trigonometric Functions	10
3.4	Inverse Functions	10
3.5	Some Common Angles	11
4	Vectors	13
4.1	Visualising Vectors	13
4.2	Some Special Vectors	13
4.3	Vector Addition and Subtraction	14
4.4	Scalar Multiplication	15
4.5	The Dot Product	16
4.6	Vector Lengths and Normalisation	16
4.7	Vector Cross Product	17
4.8	Geometric Interpretation of Products of Vector	17
5	Matrices	19
5.1	Matrix Addition and Subtraction	19
5.2	Scalar Multiplication	21
5.3	Matrix Multiplication	21

5.4	Matrix Transpose	22
5.5	Matrices as Transforms of Vectors	23
5.6	The Identity Matrix	24
5.7	Inverse Matrices	24

Notation Summary

General Notation

$a, b, \dots$	Lower case letters in italics are scalars (simple numbers)
$\mathbf{a}, \mathbf{b}, \dots$	Lower case letters in bold are vectors (1-D arrays of numbers)
$A, B, \dots$	Upper case letters are matrices (2-D arrays of numbers)
$\sin^2(\theta)$	Means $(\sin(\theta))^2$, not $\sin(\sin(\theta))$. Similarly for cos and tan.
$\sin^{-1}(x)$	Arcsine, the angle whose sine is x . Similarly for cos and tan.

Matrices and Vectors

v_i	The i th element of a vector, $\mathbf{v}$
m_{ij}	The element of a matrix, M , in the i th row and j th column
$\mathbf{v}^T, M^T$	Transpose of a vector or matrix
$\mathbf{0}$	The zero-vector, $\mathbf{0} = [0 \ 0 \ \dots \ 0]^T$.
$\mathbf{u} \cdot \mathbf{v}$	Dot product (a scalar value) of two vectors
$\mathbf{u} \times \mathbf{v}$	Cross product (a 3-vector) of two 3-vectors
$\ \mathbf{v}\ $	The length of a vector, $\ \mathbf{v}\ = \sqrt{\mathbf{v} \cdot \mathbf{v}}$
$\hat{\mathbf{v}}$	A unit vector in the same direction as $\mathbf{v}$, $\hat{\mathbf{v}} = \frac{\mathbf{v}}{\ \mathbf{v}\ }$.
$\mathbf{u} \perp \mathbf{v}$	The vectors $\mathbf{u}$ and $\mathbf{v}$ are perpendicular to one another.
$\mathbf{u} \parallel \mathbf{v}$	The vectors $\mathbf{u}$ and $\mathbf{v}$ are parallel to one another.
I	The identity matrix, with diagonal entries of 1, and off-diagonal entries of 0.
M^{-1}	The inverse of a matrix, M . $MM^{-1} = I = M^{-1}M$.

Quick Reference

A few results and identities are gathered here for easy reference:

- Trigonometric functions of some common angles:

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
Degrees	0°	30°	45°	60°	90°
Opposite	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
Adjacent	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0
Hypotenuse	1	1	1	1	1
$\sin \theta$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
$\cos \theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0
$\tan \theta$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	—

- $\sin^2(\theta) + \cos^2(\theta) = 1$.
- The angle between two vectors is $\cos^{-1} \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \right)$.
- Two non-zero vectors are perpendicular if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.
- Two unit vectors are parallel if and only if $\hat{\mathbf{u}} \cdot \hat{\mathbf{v}} = 1$.
- Two non-zero 3-vectors are parallel if and only if $\mathbf{u} \times \mathbf{v} = 0$.
- $\mathbf{u} \times \mathbf{v}$ is perpendicular to $\mathbf{u}$ and $\mathbf{v}$ unless they are zero or parallel.
- Matrix multiplication is *not commutative*, so $\mathbf{AB} \neq \mathbf{BA}$ in general.
- Inverses and transposes of products reverse the order of multiplication, so $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$ and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$.

1 Introduction

This document provides mathematics which is be useful for computer graphics and vision students. It is mainly aimed at students taking COSC342 (Computer Graphics) at Otago, but may be of interest to those taking related papers such as COSC360 (Computer Game Design) or COSC450 (Computer Graphics and Vision) as well.

We will be dealing with three main types of mathematical object – scalars, vectors, and matrices. Scalars are just ordinary numbers like 1, $-2\frac{3}{7}$, or π . Vectors are one-dimensional arrays of numbers, and matrices are two-dimensional arrays of numbers. It is important to keep the type of values in mind, because it affects what you can do with them and what the result is. For example, while it makes sense to divide two numbers (as long as the second is not zero), division is not defined for vectors or matrices.

Vectors will be used to represent geometric objects like points and lines. Matrices will represent transformations of these objects, like scaling, translation, and rotation. Using vector and matrix notation gives us a set of operations (an algebra) to work with. Representing these geometric ideas algebraically is very useful, since it allow us to easily write computer programs that work with them.

2 Scalars

We'll start by revising some properties of scalars (ordinary numbers). This may be familiar to you, but it is important to make sure we have a common language and have defined a few important terms.

2.1 Visualising Scalars

Scalar values can be visualised on the number line, as shown in Figure 1.

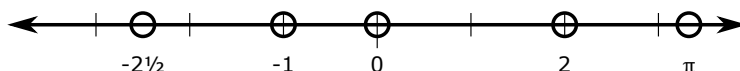


Figure 1: The values $-2\frac{1}{2}$, -2 , 0 , 2 , and π shown on the number line

Scalars can be thought of as either a point on the line, or as a distance along the line from zero. If thought of as distances, negative numbers are a step to the left, and positive numbers to the right. Addition can be thought of as combining steps, and multiplication as repeating a step. So $2 - 3$ is a step of 2 units to the right, followed by a step of 3 units to the left. Instead 2×3 can be thought of as taking two steps to the right three times (although 6 is not shown in Figure 1).

2.2 Operations on Scalars

The usual operations that are applied to numbers are addition, subtraction, multiplication, and division. For this discussion, it helps to think of subtraction as adding negative numbers, and as division as multiplying by reciprocals, so when we see $1 - 3$, we think $1 + (-3)$, and by $1 \div 2$ we mean $1 \times \frac{1}{2}$. This makes the order of operations easier, because the meaning of $1 - 2 + 3$, is not immediately clear – is it $1 - (2 + 3) = -4$ or $(1 - 2) + 3 = 2$? If we think of this as $1 + (-2) + 3$ then the order of operations doesn't matter. Likewise $1 \div 3 \times 2$ is not clear, but $1 \times \frac{1}{3} \times 2$ is.

This leads us to the first properties of addition and multiplication – they are *associative* and *commutative*, which means that when you add a list of numbers or multiply a list of numbers the order in which do you things doesn't matter.

More formally the commutative rule for addition says that

$$a + b = b + a \quad \forall a, b. \quad (1)$$

A few words about notation are probably useful here. Firstly, we use lower case italic letters to denote scalars. Secondly, the symbol $\forall$, which is usually read ‘for all’, means that this is true for any values of a and b . The commutative rule for multiplication tells us that

$$a \times b = b \times a \quad \forall a, b. \quad (2)$$

The associative rule tells us that we can apply multiple instances of addition or multiplication in any order. For addition this gives us

$$a + (b + c) = (a + b) + c \quad \forall a, b, c, \quad (3)$$

and for multiplication

$$a \times (b \times c) = (a \times b) \times c \quad \forall a, b, c. \quad (4)$$

Things get a little more complex when we mix multiplication and addition, but the key rule is to apply multiplication first, unless brackets tell you otherwise. So $1 + 2 \times 3$ means $1 + (2 \times 3)$, and $1 \times 2 + 3$ means $(1 \times 2) + 3$. Combining operators also gives us a third property – *distributive*. We say that multiplication *distributes* over addition to mean that

$$a \times (b + c) = (a \times b) + (a \times c) \quad \forall a, b, c. \quad (5)$$

Note that addition does *not* distribute over multiplication, since

$$1 + (2 \times 3) \neq (1 + 2) \times (1 + 3). \quad (6)$$

2.3 A Note on Multiplication Symbols

There are many ways to write multiplication of two values. The most usual is form is probably $a \times b$, which we’ve used so far, but there are two other common forms used when dealing with symbols (rather than numbers). The first is a dot, $a \cdot b$ and the second is just writing two symbols side-by-side, ab . Note that when dealing with actual numbers, $\times$ is preferred, since $1 \cdot 2$ can look like 1.2, and 12 is a different number all together. This will become important when we come to vectors and matrices where multiplication is a little more complicated.

We’ll typically just place symbols side by side to indicate multiplication, but use $\times$ when needed for clarity. Note that this is also why mathematics typically uses single-letter variables, and why precise notation is important. For example, *sin* should be read as $s \times i \times n$, where *sin* is the trigonometric function ‘sine’. (Shortening ‘sine’ to *sin* doesn’t save much space, but mathematicians do this to make the abbreviation the same length as the shortening of ‘cosine’ to *cos* and ‘tangent’ to *tan*.)

3 Trigonometric Functions

We will be dealing with angles quite a lot, and this leads to trigonometry. We will need to be familiar with measuring angles in degrees and radians, and with the basic trigonometric functions, sine, cosine, and tangent.

Trigonometry is often introduced in terms of right-angled triangles, but it is also helpful to keep in mind the relationship with circles. These relationships are illustrated in [Figure 2](#).

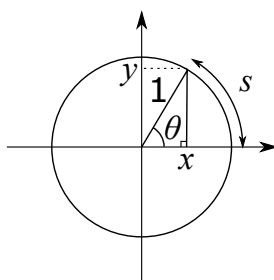


Figure 2: Angles can be thought about as rotation around a circle, or as part of a right-angled triangle.

Some points to note from the figure:

- The angle of interest is θ . We will often use Greek letters for angles, but this is just convention, and not always followed.
- We have a unit circle (a circle with radius 1), and measure the angle anti-clockwise from the positive x -axis (the positive x -axis is pointing to the right in [Figure 2](#)).
- The angle then relates to some point (x, y) on the circle.
- The length of the arc around the circle to (x, y) is s .
- This point defines a right-angled triangle with sides x , y , and 1.
- The length of the longest side (the *hypotenuse*) of the triangle is 1.
- The length of the side of the triangle *opposite* the angle is y .
- The length of the shorter side *adjacent* the angle is x .

For angles greater than 90° this still holds, but the x and/or y values may be negative.

3.1 Degrees and Radians

Angles are commonly measured in degrees, but mathematically it is easier to work in radians. Radians measure the distance around a circumference of circle of unit radius ($r = 1$) associated with a given angle, labelled s in [Figure 2](#). Since the total circumference of a unit circle is 2π , that is equivalent to 360° , and the conversion from an angle, s , measured in radians to and from the equivalent value, d , in degrees are

$$s = \frac{\pi}{180^\circ} d, \quad d = \frac{180^\circ}{\pi} s. \quad (7)$$

Note that radians are a unitless quantity – you do not have ‘degrees radian’, and while we may talk about an angle being ‘ $\frac{\pi}{2}$ radians’ for clarity, it is equally correct to just say that the angle is $\frac{\pi}{2}$.

3.2 Sine, Cosine, and Tangent

The trigonometric functions we will use are sine, cosine, and tangent. These are denoted $\sin$, $\cos$, and $\tan$. If we think of things in terms of a right-angled triangle, then

$$\sin \theta = \frac{O}{H}, \quad \cos \theta = \frac{A}{H}, \quad \tan \theta = \frac{O}{A}, \quad (8)$$

where H is the length of the longest side of the triangle (the hypotenuse), O is the length of the side opposite the angle, and A is the length of the side adjacent to the angle. I learned this as *SOH CAH TOA*, but why a string on nonsense syllables helps, I don’t know.

If we return to the view of a unit circle, and [Figure 2](#), then we can substitute $H = 1$, $O = y$, and $A = x$, giving

$$\sin \theta = y, \quad \cos \theta = x, \quad \tan \theta = \frac{y}{x}. \quad (9)$$

For a more general circle, with a radius of r , this can be rearranged to show that

$$(x, y) = (r \cos \theta, r \sin \theta), \quad (10)$$

for any point, (x, y) on a circle. The combination of these definitions with Pythagoras’ theorem (i.e., $a^2 + b^2 = c^2$, for a right angled triangle with sides of length a , b , and hypotenuse c) leads directly to the identity

$$\sin^2 \theta + \cos^2 \theta = 1, \quad (11)$$

where the conventional notation is that $\sin^2 \theta = (\sin \theta)^2$, etc.

3.3 Plots of Trigonometric Functions

Plots of the sine and cosine functions are shown in **Figure 3**. Note that both functions are periodic, so $\sin(\theta + 2\pi) = \sin \theta$. Also, $\cos(\theta) = \sin(\theta + \frac{\pi}{2})$.

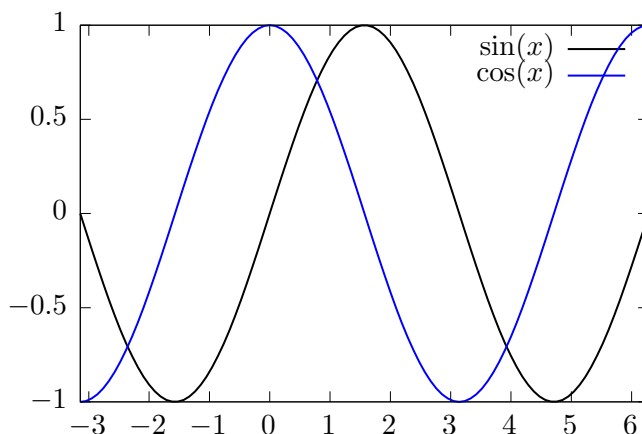


Figure 3: Plot of sine and cosine functions.

Figure 4 shows the tangent function. Like sine and cosine it is periodic, with a period of 2π . Note also that $\tan(\theta)$ is undefined at $\frac{\pi}{2}$, and therefore at $\frac{(2k+1)\pi}{2}$ for any integer, k . This comes about directly from the definition in **Equation 8**, since the length of the ‘side’ adjacent to the angle is 0. It is important to remember that $\tan(\theta)$ is *undefined* at these points, not ‘infinite’ – the function diverges so it gets larger (more positive) without bound from one side and smaller (more negative) from the other.

3.4 Inverse Functions

When working with lines and vectors, we often end up computing the sine or cosine of an angle. In some cases we need to find the angle from these values, and so the inverse functions are needed. These are written as $\sin^{-1}(x)$, etc. which means ‘the angle which has a sine of x ’.

This notation is a little unfortunate, since it conflicts a bit with the idea of $\sin^2(\theta)$ being $(\sin(\theta))^2$. Following that logic, $\sin^{-1}(x)$ could be interpreted as $(\sin(\theta))^{-1} = \frac{1}{\sin(\theta)}$. This is not the usual convention, however.

The inverse functions are often referred to as the arcsine, arccosine, and arctangent. This is why the corresponding functions in languages like C are

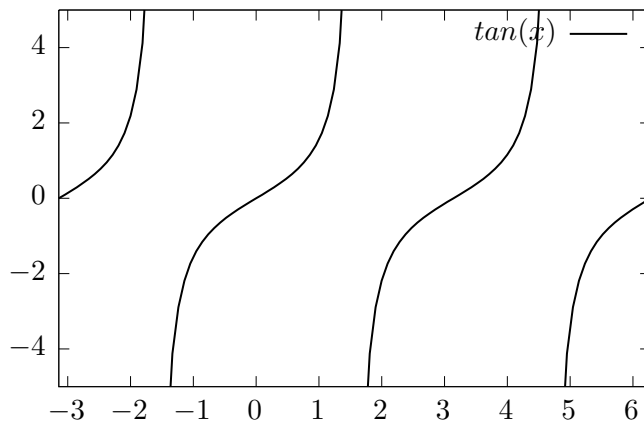


Figure 4: Plot of the tangent functions.

called `asin`, `acos`, and `atan`.

Note that the inverse is not unique, since the trigonometric functions are periodic. Furthermore, there are generally two angles in each cycle with the same value for each function. For example, $\cos(-\theta) = \cos(\theta)$ for any θ . By convention, $\cos^{-1}(x)$ is treated as the value in the range $[0, \pi]$, while $\sin^{-1}(x)$ and $\tan^{-1}(x)$ return a value in the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

3.5 Some Common Angles

While in computation we will generally use any value for angles, computing trigonometric functions for an arbitrary angle by hand is not generally recommended. However, for a set of common angles, the trigonometric functions can be easily expressed. These arise from two special right-angled triangles – one with two 45° angles, and one with a 30° and 60° angles.

The 45° – 45° – 90° triangle has two equal sides, and the longest side has a length of 1. Pythagoras' theorem tells us that if the length of the other two sides is a , then $a^2 + a^2 = 1$, so $a = \frac{1}{\sqrt{2}}$.

The 30° – 60° – 90° triangle is half of an equilateral triangle with sides of length 1. This means that the shortest side (adjacent to the 60° angle) has length $\frac{1}{2}$. Therefore the remaining side length, b is found from

$$\left(\frac{1}{2}\right)^2 + b^2 = 1, \quad (12)$$

which tells us that $b = \frac{\sqrt{3}}{2}$.

These results, and the trigonometric functions that follow directly from them are summarised in **Table 1**.

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
Degrees	0°	30°	45°	60°	90°
Opposite	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
Adjacent	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0
Hypotenuse	1	1	1	1	1
$\sin \theta$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
$\cos \theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0
$\tan \theta$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	–

Table 1: Trigonometric function values for some common angles. Note that $\tan \frac{\pi}{2}$ is undefined.

4 Vectors

A vector is a one-dimensional array of numbers, and we write them in bold face, such as $\mathbf{v}$ to distinguish them from scalars. Other notations you might see for vectors include an arrow or bar over or above a value, such as $\vec{v}$, $\bar{v}$, or $\underline{v}$. These alternative notations are often used when writing by hand.

A vector, being an array, has a size, which is the number of elements it contains. A vector with n elements is called an n -vector, so we talk about 2-vectors, 3-vectors, and so on. Note that the *size* of a vector refers to the number of elements in a vector, not how big or small the elements are. By convention, the i th element of a vector $\mathbf{v}$ is denoted v_i , using the same letter or symbol as the vector, but as a scalar.

Vectors may be written out in a number of different ways, and we will follow the common convention of using *column vectors*, where the elements are arranged vertically,

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}. \quad (13)$$

This convention will be useful when we come to combine matrices and vectors, but takes up a lot of room on the page. When writing vectors in text, it is convenient to convert them to rows. The *transpose* operator converts a column vector to a row, or vice-versa, and so we can write $\mathbf{v}^T = [v_1 \ v_2 \ \dots v_n]$, or equivalently $\mathbf{v} = [v_1 \ v_2 \ \dots v_n]^T$. The transpose is a matrix operation, and so will be discussed later, but other notations you might see are a prime, $\mathbf{v}'$, or an asterisk, $\mathbf{v}^*$.

4.1 Visualising Vectors

Just as scalars can be thought of as points or steps along a line, vectors can be thought of as points or steps in the plane (for 2-vectors), 3D space (for 3-vectors) and so on. Some 2-vectors are shown in [Figure 5](#), and can be thought of as either steps along the x - and y -axes, or as points in 2D space. Likewise, a 3-vector can be thought of as steps along the x -, y -, and z -axes in 3D space.

4.2 Some Special Vectors

A few special cases of vectors arise, and lead to some additional notation. The first is a vector where every element is zero. This is called a *zero vector*,

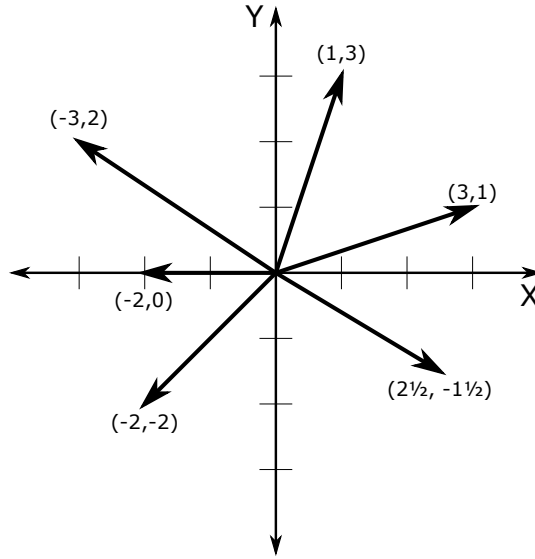


Figure 5: Some 2-vectors visualised in the plane. The vector $\mathbf{v} = \begin{bmatrix} x & y \end{bmatrix}^T$ is seen as either the point (x, y) , or the step from the origin to (x, y) .

and often written $\mathbf{0}$. There is a whole family of zero vectors of different length so we have the zero 2-vector, the zero 3-vector, etc. When the size of the vector is clear from context, we'll just talk about *the* zero vector.

You might also see $\mathbf{1}$ for a vector of all ones, but this is less common. Note that $\mathbf{1}$ is *not* called the 'one-vector'. A 1-vector is a vector with one element, and is equivalent to a scalar.

4.3 Vector Addition and Subtraction

Vector addition and subtraction are defined simply by adding the corresponding elements of a vector, so

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1+3 \\ 2+4 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}. \quad (14)$$

This only makes sense if there are the same number of elements in each vector. If the two vectors have different sizes, then their sum is undefined.

More generally, the sum of two n -vectors, $\mathbf{u}$ and $\mathbf{v}$, is

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}. \quad (15)$$

The difference of two vectors is defined in a similar way,

$$\mathbf{u} - \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} - \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \\ \vdots \\ u_n - v_n \end{bmatrix}. \quad (16)$$

As with scalars, it is sometimes convenient to think of subtraction as adding a negative, so $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$, where $-\mathbf{v} = [-v_1 \ -v_2 \ \dots \ -v_n]^T$.

The definitions of addition and subtraction can be represented more compactly if we think about the individual elements. If $\mathbf{w} = \mathbf{u} + \mathbf{v}$, we can see that $w_i = u_i + v_i$. Likewise if $\mathbf{d} = \mathbf{u} - \mathbf{v}$, then $d_i = u_i - v_i$.

Vector addition has the same properties as scalar addition. Vector addition is commutative and associative, so

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \quad \forall \mathbf{u}, \mathbf{v}, \quad (17)$$

and

$$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w}. \quad (18)$$

4.4 Scalar Multiplication

Multiplication of vectors is more complex than for scalars, and there are several different types of product that can be formed. The first, and simplest, is *scalar multiplication*, which is the result of multiplying a vector by a scalar. Like vector addition, this is done element-wise, so multiplying a vector by 2 simply doubles each element. More formally, given a scalar, s , and an n -vector, $\mathbf{v}$, their scalar product, $\mathbf{u} = s\mathbf{v}$ is also an n -vector, with elements $u_i = s \times v_i$.

Scalar multipliers are conventionally written on the left of the vector, but this is not essential, and $\mathbf{v}s = s\mathbf{v}$. Multiplication by several scalars is associative and commutative, so $st\mathbf{v} = ts\mathbf{v}$ and $s(t\mathbf{v}) = (st)\mathbf{v}$. Finally, multiplication by 0 yields a zero vector of the same size as the original vector, $0\mathbf{v} = \mathbf{0}$.

4.5 The Dot Product

The most common form of multiplication with vectors is the dot product, written $\mathbf{u} \cdot \mathbf{v}$. This is also called the scalar product (not to be confused with scalar multiplication) or sometimes the inner product.

Given two vectors of the same size, their dot product is a scalar value. The corresponding elements of the two vectors are multiplied together, and these are added up to give the result.

More formally,

$$\mathbf{u} \cdot \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n, \quad (19)$$

which can be written more compactly as

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^n u_i v_i. \quad (20)$$

The dot product is commutative, so $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$. It is *not* associative, because the dot product of three vectors is undefined. Taking the dot product of two vectors gives a scalar result, and the dot product of a scalar and a vector is not defined.

The dot product is distributive over vector addition, so

$$\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w}. \quad (21)$$

Scalar multiplication also factors out, so $(r\mathbf{u}) \cdot (s\mathbf{v}) = rs(\mathbf{u} \cdot \mathbf{v})$.

4.6 Vector Lengths and Normalisation

The length of a vector is written $\|\mathbf{v}\|$, or $|\mathbf{v}|$, and can be computed as

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} = \sqrt{\mathbf{v} \cdot \mathbf{v}}. \quad (22)$$

The length of the vector is also called its *magnitude* or *norm*, or more precisely its Euclidean norm or 2-norm.

A *unit vector* is a vector with length 1. Any vector can be converted into a unit vector by dividing by its length, and we use a ‘hat’ over a vector to denote this,

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}. \quad (23)$$

This process is called *normalisation*.

4.7 Vector Cross Product

The final form of multiplication we need for vectors is the cross product. The cross product is *only defined for 3-vectors*. Given two 3-vectors, $\mathbf{u}$ and $\mathbf{v}$, their cross product is also a 3-vector given by

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \times \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix} \quad (24)$$

It can be easily shown from this definition that $\mathbf{v} \times \mathbf{v} = \mathbf{0}$ for any 3-vector, $\mathbf{v}$.

The cross product is *not* associative, so $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} \neq \mathbf{v} \times (\mathbf{u} \times \mathbf{w})$ in general. The cross product is *anticommutative*. This means that changing the order of the vectors does change the result, but in a predictable way, since $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$.

The cross product distributes over addition, so $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$. Scalar multiplication also interacts in the expected manner so $(r\mathbf{u}) \times (s\mathbf{v}) = rs(\mathbf{u} \times \mathbf{v})$.

4.8 Geometric Interpretation of Products of Vector

The various products of two vectors have useful geometric interpretations. These are related to the angle between the vectors, and we introduce some notation for special cases. If two vectors are at right angles to each other, they are perpendicular and we write $\mathbf{u} \perp \mathbf{v}$. If two vectors point in the same direction, they are parallel, and we write $\mathbf{u} \parallel \mathbf{v}$. The zero-vector, $\mathbf{0}$, is not parallel or perpendicular to any vector, and is a special case in all of the discussions that follow.

Two (non-zero) vectors are parallel if, and only if, one is a scalar product of the other. So if $\mathbf{u} \parallel \mathbf{v}$, then $\mathbf{u} = s\mathbf{v}$ for some non-zero s , and vice versa.

The dot product tells us about the angle between vectors, since

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta, \quad (25)$$

where θ is the angle between the two vectors. If we normalise the vectors, this simplifies to

$$\hat{\mathbf{u}} \cdot \hat{\mathbf{v}} = \cos \theta. \quad (26)$$

Since $\cos \theta = 0$ at $\pm \frac{\pi}{2}$, this means that if the dot product of two non-zero vectors is 0 if, and only if, they are perpendicular to one another (or one of them is $\mathbf{0}$).

In general, the cross product of two vectors is perpendicular to both of them, so $\mathbf{u} \perp (\mathbf{u} \times \mathbf{v})$, and $\mathbf{v} \perp (\mathbf{u} \times \mathbf{v})$. There are two exceptions to this rule. Firstly, if either vector is $\mathbf{0}$, then their cross product is $\mathbf{0}$. This follows directly from the definition of the cross product (Equation 24). The second is if the two vectors are parallel, in this case $\mathbf{u} = s\mathbf{v}$, and it follows fairly directly that $\mathbf{u} \times \mathbf{v} = \mathbf{0}$.

These facts help us to solve some common problems that come up in computer graphics, so to summarise, given two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$:

- The angle between the vectors is $\cos^{-1} \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \right)$
- They are perpendicular if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.
- They are parallel if and only if $\mathbf{u} \times \mathbf{v} = \mathbf{0}$.
- A vector perpendicular to both vectors is $\mathbf{u} \times \mathbf{v}$ (unless $\mathbf{u} \parallel \mathbf{v}$).

5 Matrices

Many of the properties of vectors are special cases of matrix operations, so we'll cover matrices first. A matrix is a two-dimensional array of numbers, and we write them in uppercase upright letters, such as A . The size of matrices is often important, and is expressed as [number of rows] $\times$ [number of columns], so a 3×4 matrix has 3 rows and 4 columns.

Since a matrix is an array of numbers, it is often convenient to talk about individual elements of a matrix. We'll use $a_{i,j}$ to refer to the element of a matrix A in the i th row and j th column. Following mathematical convention, we'll start indices from 1, although in code you'll often start from 0. When expressed in full, matrices are written inside square brackets, so a $m \times n$ matrix can be written as

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}. \quad (27)$$

Matrices have many of the same operators as scalars, although some function a little differently. One new operator is the *transpose*, which involves swapping the rows and columns of a matrix. Given the $m \times n$ matrix A in [Equation 27](#), its transpose is the $n \times m$ matrix

$$A^T = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}. \quad (28)$$

5.1 Matrix Addition and Subtraction

Matrix addition and subtraction are defined element-wise. That is, you add (or subtract) two matrices by adding (or subtracting) the values at each

position in the matrix. Written out more formally we have

$$A + B = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n} \\ b_{2,1} & b_{2,2} & \dots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \dots & b_{m,n} \end{bmatrix} \quad (29)$$

$$= \begin{bmatrix} (a_{1,1} + b_{1,1}) & (a_{1,2} + b_{1,2}) & \dots & (a_{1,n} + b_{1,n}) \\ (a_{2,1} + b_{2,1}) & (a_{2,2} + b_{2,2}) & \dots & (a_{2,n} + b_{2,n}) \\ \vdots & \vdots & \ddots & \vdots \\ (a_{m,1} + b_{m,1}) & (a_{m,2} + b_{m,2}) & \dots & (a_{m,n} + b_{m,n}) \end{bmatrix} \quad (30)$$

$$A - B = \begin{bmatrix} (a_{1,1} - b_{1,1}) & (a_{1,2} - b_{1,2}) & \dots & (a_{1,n} - b_{1,n}) \\ (a_{2,1} - b_{2,1}) & (a_{2,2} - b_{2,2}) & \dots & (a_{2,n} - b_{2,n}) \\ \vdots & \vdots & \ddots & \vdots \\ (a_{m,1} - b_{m,1}) & (a_{m,2} - b_{m,2}) & \dots & (a_{m,n} - b_{m,n}) \end{bmatrix} \quad (31)$$

This only makes sense if all the values line up, so you can only add or subtract matrices of the same size. Addition and subtraction of matrices of different sizes is not defined.

We can summarise addition and subtraction by considering how each element in the result is computed. If $C = A + B$, then the i, j th element of C is given by

$$c_{i,j} = a_{i,j} + b_{i,j}, \quad (32)$$

likewise, if $D = A - B$, then

$$d_{i,j} = a_{i,j} - b_{i,j}. \quad (33)$$

Addition of matrices is associative and commutative, just like addition of scalars. Given any matrices A , B and C all of the same size, then

$$A + B = B + A \quad (34)$$

and

$$A + (B + C) = (A + B) + C. \quad (35)$$

5.2 Scalar Multiplication

Multiplying a matrix by a scalar, like addition, is defined on a per-element basis. The product of a scalar, s , and an $m \times n$ matrix, A is

$$sA = \begin{bmatrix} sa_{1,1} & sa_{1,2} & \dots & sa_{1,n} \\ sa_{2,1} & sa_{2,2} & \dots & sa_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ sa_{m,1} & sa_{m,2} & \dots & sa_{m,n} \end{bmatrix}, \quad (36)$$

or in element-wise form, if $B = sA$,

$$b_{i,j} = sa_{i,j}. \quad (37)$$

Note that it is customary to write scalars on the left, but this is not necessary since

$$sA = As. \quad (38)$$

Scalar multiplication distributes over matrix addition, so

$$s(A + B) = sA + sB \quad \forall s, A, B. \quad (39)$$

Likewise, it distributes over scalar addition, giving

$$(s + t)A = sA + tA \quad \forall s, t, A. \quad (40)$$

5.3 Matrix Multiplication

Finally we come to multiplying two matrices. As with addition, this is not defined for all pairs of matrices, but the rules are a little more complicated. The product of two matrices is defined if, and only if, the number of columns in the left-hand matrix, and the number of rows in the right-hand matrix are the same. Given matrices A of size $m \times p$ and B of size $p \times n$, their product is a $m \times n$ matrix,

$$\begin{matrix} A & \times & B & = & C \\ \text{of size} & & \text{of size} & & \text{of size} \\ m \times p & & p \times n & & m \times n \end{matrix} \quad (41)$$

As with scalar multiplication, we'll usually just write AB for the product of two matrices, and leave out the multiplication sign.

The i, j th element of $C = AB$ is found from the i th row of A and the j th column of B . The corresponding elements of the row and column are multiplied together then summed, giving

$$c_{i,j} = \sum_{k=1}^p a_{i,k} b_{k,j}. \quad (42)$$

Note that this is only defined if the rows of A and columns of B have the same number of elements. This is why the number of columns in A must be the same as the number of rows in B.

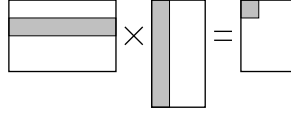


Figure 6: Matrices are multiplied by combining the rows of the left-hand side with the columns of the right-hand side

Matrix multiplication is associative, but *not commutative*. Assuming that A, B, and C have appropriate sizes,

$$A(BA) = (AB)C, \quad (43)$$

but in general

$$AB \neq BA. \quad (44)$$

Note that unless A and B are square, BA won't even be defined. Even if they are square, the products will usually be different in each case

5.4 Matrix Transpose

The transpose of a matrix was introduced briefly when talking about vectors. Given an $m \times n$ matrix, M, its transpose, M^T is an $n \times m$ matrix where the i th row of M^T is the i th column of M. Alternatively we can think about individual elements and (with some abuse of notation), $m_{ij}^T = m_{ji}$. The transpose ‘undoes’ itself, so $(M^T)^T = M$.

Transposes of sums and scalar products are fairly straightforward, since $(A + B)^T = A^T + B^T$, and $(sA)^T = s(A^T)$. The transpose of a product of two matrices is a little surprising, since the order of multiplication reverses, giving

$$(AB)^T = B^T A^T. \quad (45)$$

With this definition of the transpose and matrix multiplication, we can revisit the dot product. Viewing vectors as column matrices, it is easy to show that $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$. The vectors are treated as $n \times 1$ matrices, so $\mathbf{u}^T$ is a $1 \times n$ matrix. The matrix product $\mathbf{u}^T \mathbf{v}$ is then a 1×1 matrix, which is equivalent to a scalar, and expanding out its definition shows that it is the same as the dot product.

5.5 Matrices as Transforms of Vectors

We can think of a vector as a matrix with only one column. This means that we can multiply matrices by vectors, and the result is itself a vector, so multiplying a $m \times n$ matrix by a n -vector gives a m vector as a result. Usually we'll use square matrices, which take n -vectors and transform them to new n -vectors, but this is not always the case.

If we think about the vectors as points, then this lets us think about matrices in terms of how they transform (or move) those points. This is the key to the use of matrices in computer graphics – they give us a mathematical way to represent transformations, and therefore a way to implement these in our programs.

More detail about this will be covered in lectures, but as a simple example, consider the matrix

$$M = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}. \quad (46)$$

If we apply this to some 2-vector, $\mathbf{v} = \begin{bmatrix} x & y \end{bmatrix}^T$, we get a transformed vector $\mathbf{v}' = M\mathbf{v} = \begin{bmatrix} 2x & \frac{1}{2}y \end{bmatrix}^T$. So we can think about M as having the effect of stretching points by a factor of 2 horizontally, and compressing it by half vertically. We can apply M to lines, curves, or shapes by applying it to each point on the line curve, or shape. **Figure 7** shows the effect of applying M to a square positioned at the origin.

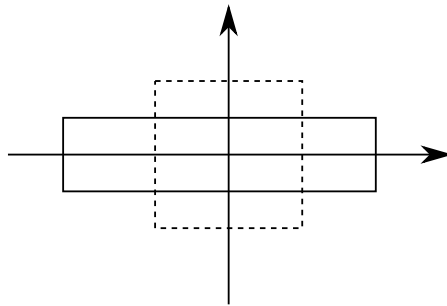


Figure 7: Applying the matrix, M , from **Equation 46** to a square (dotted lines), stretches it horizontally and squashes it vertically (dashed lines).

5.6 The Identity Matrix

An important matrix is the identity matrix, written I . The diagonal entries of the identity matrix (those in the i th row and i th column) are 1, and the off-diagonal entries are zero. As with zero-vectors, it is more accurate to talk about *an* identity matrix because there is one for each size of matrix. Where it is not clear, a subscript can be used to indicate the size of an identity matrix,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \dots \quad (47)$$

The identity matrix is a transform with no effect, so $I\mathbf{v} = \mathbf{v}$, for any $\mathbf{v}$. Likewise, I has no effect when applied to other matrices, $IM = M = MI$ for any matrix M .

5.7 Inverse Matrices

The inverse of a square matrix, M , is a matrix, M^{-1} , such that

$$MM^{-1} = M^{-1}M = I. \quad (48)$$

Not all matrices have an inverse, but for those that do the inverse is unique. If a matrix has an inverse it is called invertible, nonsingular, or nondegenerate. Matrices without an inverse are called singular or degenerate.

If we have a product of two $n \times n$ matrices, say $C = AB$ then the inverse of the product reverses the order of multiplication, giving

$$C^{-1} = (AB)^{-1} = B^{-1}A^{-1}. \quad (49)$$

This generalises to any number of matrices, so

$$(A_1A_2 \dots A_{k-1}A_k)^{-1} = A_k^{-1}A_{k-1}^{-1} \dots A_2^{-1}A_1^{-1}. \quad (50)$$

For scalar multiplication we have $(sA)^{-1} = \frac{1}{s}A^{-1}$, for non-zero s . There is no general formula for the inverse of the sum of two matrices.

The transpose and inverse can be applied in any order, since $(A^{-1})^T = (A^T)^{-1}$. This may be written as A^{-T} as a short-hand.

In terms of transforms, the inverse matrix ‘undoes’ the original transform. For example, the inverse of the matrix from [Equation 46](#) is

$$M^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix}, \quad (51)$$

which can be verified by computing MM^{-1} and $M^{-1}M$. While M stretches points horizontally and compresses vertically, M^{-1} compresses horizontally and stretches vertically.

As an example of a matrix that is not invertible, consider the matrix

$$S = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}. \quad (52)$$

Suppose this did have an inverse, say $S^{-1} = T$, where

$$T = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}. \quad (53)$$

Now we have

$$ST = I \quad (54)$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (55)$$

$$\begin{bmatrix} t_{11} + 2t_{21} & t_{12} + 2t_{22} \\ 2t_{11} + 4t_{21} & 2t_{12} + 4t_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (56)$$

The top left element tells us that $t_{11} + 2t_{21} = 1$, and the bottom left tells us that $2t_{11} + 4t_{21} = 0$. Dividing the second equation by 2 gives us $t_{11} + 2t_{21} = 0$, but that has the same left-hand side as the first equation. Since $0 \neq 1$, there can be no such matrix, and S has no inverse.

There are many ways of expressing the general property that a $n \times n$ matrix, M has to have to be invertible. Some of the more useful ones to us are:

- The columns (or rows) of M must be linearly independent – it must not be possible to write one as a weighted sum of the others.
- The columns of M must *span* or form a *basis* for n -dimensional space. That is any n -vector can be written as a weighted sum of the columns of M .
- The equation $M\mathbf{x} = \mathbf{0}$ has only the trivial solution, $\mathbf{x} = \mathbf{0}$.

For the matrix S from [Equation 52](#) these properties do not hold:

- The second column is twice the first.
- No weighted sum of the columns of S can give $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$.
- $\mathbf{x} = \begin{bmatrix} 2 & -1 \end{bmatrix}^T$ is a non-trivial solution to $S\mathbf{x} = \mathbf{0}$.