Tutorial Letter 103/3/2016
Computer Graphics

COS3712

Semesters 1 and 2

School of Computing

This tutorial letter contains important information about your module.

1   Introduction

This tutorial letter contains additional notes that should help you get into COS3712. The prescribed book assumes that you have a working understanding of certain mathematical concepts, especially linear algebra. This tutorial letter is intended as a linear algebra primer for computer graphics programming.

2 Notes on Linear Algebra

The following notes are intended to cover most of the mathematical prerequisites, i.e. the necessary concepts of linear algebra, needed for COS3712. Please note that they do not cover all the mathematics required for Computer Graphics in general, or COS3712 in particular. The prescribed book covers the additional topics you need to understand and to be able to apply.

2.1   3-dimensional space

Most of the computer graphics in COS3712 is done in 3-dimensional space. We assume there is a set of three axes, *x*, *y* and *z*, at right angles to each other, according to which any position can be specified. Angel (2009) defines an *affine space* as a vector space that also includes points. By *3-dimensional space* we mean an affine space.

2.2   Points, vectors and scalars

A *point* in 3-dimensional space, generally represented graphically by a single dot, is a position or location specified by three values placed in parentheses, representing the *x*, *y* and *z* coordinates of the point with respect to the set of axes. For example, (1, 2, 3) represents the point with 1 as its *x*-coordinate, 2 as its *y*-coordinate and 3 as its *z*-coordinate. The point (0, 0, 0) represents the point where the *x*, *y* and *z* axes intersect each with the value 0, and is referred to as the *origin*. We generally use lowercase letters of the alphabet, eg. *p*, *q* and *r* to refer to points.

A *vector* in 3-dimensional space, generally represented graphically by an arrow, is a direction with a certain magnitude. A vector is normally also specified by three values placed in parentheses, but should not be confused with a point which specifies a position. The vector (1, 2, 3) merely represents the direction from the origin (0, 0, 0) to the point (1, 2, 3). Its magnitude is the length of the line segment from (0, 0, 0) to (1, 2, 3). Note that a vector does not have a position. The vector (1, 2, 3) is not positioned at the origin, neither is it positioned at (1, 2, 3). Its position can be anywhere. It is just handy to specify a vector as if it started at the origin and ended at (1, 2, 3). We generally use lowercase letters of the alphabet, eg. *u*, *v* and *w*, to refer to vectors.

Also, do not confuse a vector with a *line segment*. A line segment in 3-dimensional space has a begin and endpoint (whereas a vector does not) and a line segment doesn't point in one of two possible directions (whereas a vector does).

A *scalar* is a value or a magnitude. It has neither position nor direction. Examples of scalars are 1, &37, 3.141579... and 0. We generally use lower case letters of the Greek alphabet, eg. α, β or γ, to represent arbitrary scalars, but sometimes also lowercase letters of the alphabet.

2.3   Expressing points and vectors in homogenous coordinates

We can distinguish between points and vectors by expressing them in four dimensions, i.e. by adding a fourth coordinate. For a point, the fourth coordinate is 1 and for a vector it is 0. In other words, the point (1, 2, 3) is represented in homogenous coordinates by (1, 2, 3, 1) and the vector (1, 2, 3) is represented by (1, 2, 3, 0).

The above explanation is really an oversimplification of the theory behind what points and vectors are and how they are related to each other, and in particular why we use 1 as the fourth coordinate for points and 0 for vectors. However, this simple explanation is all you need to know to express points and vectors in homogenous coordinates and use them in computer graphics.

2.4   Adding and subtracting vectors and points

We can add two vectors together, or subtract two vectors from one another. We can add a vector to a point (or subtract a vector from a point, although this is the same as adding a vector pointing in the opposite direction). We cannot add two points together, but we can subtract one point from another.

The sum of two vectors is another vector. We simply add the corresponding coordinates. For example if *u* is the vector (1, 2, 3) and *v* is the vector (4, 5, 6), then *u* + *v* is the vector (5, 7, 9). This can be seen graphically by drawing the two vectors end-to-end. The sum of the two vectors is a vector drawn from the beginning of the first vector to the end of the second.

The difference of two vectors is another vector. We simply subtract the corresponding coordinates.

For example if u is the vector (1, 2, 3) and v is the vector (4, 5, 6), then u - v is the vector (-3, -3, -3) and v - u is the vector (3, 3, 3).

The sum of a point and a vector (or a vector and a point) is a point. We simply add the corresponding coordinates. For example if p is the point (1, 2, 3) and v is the vector (4, 5, 6), then $p + v$ is the point (5, 7, 9). You can think of this as moving the point in the direction and distance specified by the vector.

The difference between two points is a vector. We simply subtract the corresponding coordinates. For example if p is the point (1, 2, 3) and q is the point (4, 5, 6), then p - q is the vector (-3, -3, -3) and q - p is the vector (3, 3, 3). You can think of subtraction of two points as the vector specifying the direction and distance between the two points.

2.5 Multiplying a vector by a scalar

The product of a vector and a scalar is a vector. We simply multiply each of the coordinates by the scalar. For example, if v is the vector (0, 1, 2) and α is the scalar 3, then αv is the vector (0, 3, 6). You can think of this as increasing the magnitude of a vector without changing its direction.

Similarly, we can divide a vector by a scalar.

2.6 Magnitude of a vector

The magnitude of a vector is a scalar. It is the square root of the sum of the squares of its coordinates. In other words, if v is the vector (v_x, v_y, v_z), then the magnitude of v, written as $|v|$, is

$$\sqrt{v_x^2 + v_y^2 + v_z^2}.$$

For example, if v is the vector (3, 0, 4), then $|v| = \sqrt{3^2 + 0^2 + 4^2} = \sqrt{9 + 0 + 16} = \sqrt{25} = 5$

The magnitude of a vector represents its size, without specifying its direction. It can be thought of the length of the arrow used to represent it.

2.7 Normalising a vector

A normalised vector is a vector with a magnitude of 1. To normalise a vector, we divide it by its magnitude. In other words, if v is a vector, then v normalised is $v / |v|$.

For example, say v is the vector (3, 0, 4) then v normalised is (3, 0, 4) / 5 = (0.6, 0, 0.8). We can check that this vector indeed has a magnitude of 1:

$$|(0.6, 0, 0.8)| = \sqrt{0.6^2 + 0^2 + 0.8^2} = \sqrt{0.36 + 0 + 0.64} = \sqrt{1} = 1$$

2.8 Dot product of two vectors

The dot product (also called the inner product) of two vectors is a scalar. It is the sum of the products of the corresponding coordinates of the two vectors. In other words, if vectors u and v are defined as (v_x, v_y, v_z) and (u_x, u_y, u_z) respectively, then their dot product, written as $u \cdot v$, is $v_x.u_x + v_y.u_y + v_z.u_z$.

For example, say vectors u and v are (1, 2, 3) and (-1, 0, 1) respectively.

Then $u \cdot v$ is $1.-1 + 2.0 + 3.1 = -1 + 0 + 3 = 2$.

The dot product of two vectors is useful for calculating the angle θ between two vectors by means of the following formula:

$$\cos \theta = (u \cdot v) / (|u|/|v|)$$

2.9 Cross product of two vectors

The cross product of two vectors is a vector. Each coordinate is calculated as the difference between the products of the other two coordinates of the two vectors. In other words, if u and v are vectors (u_x, u_y, u_z) and (v_x, v_y, v_z) respectively, then their cross product, written as $u \times v$, is ($u_y.v_z - u_z.v_y$, $u_z.v_x - u_x.v_z$, $u_x.v_y - u_y.v_x$).

For example, say vectors u and v are (1, 2, 3) and (-1, 0, 1) respectively.
Then $u \times v$ is $(2.1 - 3.0, 3.-1 - 1.1, 1.0 - 2.-1) = (2, -4, 2)$.

The cross product is useful for calculating a vector normal (i.e. at right angles) to two vectors.

You can see that $u \times v = n$ is at right angles to u and v by calculating the dot product of n and u and n and v.

$n \cdot u = 2.1 - 4.2 + 3.2 = 0$
If $\cos \theta = 0$, then $\theta = 90$.

Similarly, you can check that $n \cdot v = 0$.

2.10 Lines

Apart from vectors and points, we often have to deal with lines in 3-dimensional space. There are various ways of specifying them.

In 2-dimensions, the equation of a line is
        $y = mx + c$
It is sometimes handy to rearrange this to the form
        $ax + by + c = 0$

To specify a line in 3-dimensions by means of an equation is somewhat trickier. This can done by a set of so-called Cartesian equations:

$$\frac{x - q_1}{d_1} = \frac{y - q_2}{d_2} = \frac{z - q_3}{d_3}$$

where q_1, q_2, q_3, d_1, d_2 and d_3 are scalars, and x, y and z are the coordinates of any point on the line.

An alternative way of specifying lines in 3-dimensions is by means of *parametric equations*:

Say we have a point q on the line and a vector d pointing in the direction of the line. Then the following parametric equation specifies the coordinates of any point p on the line:
        $p = q + \gamma d$
where γ is a scalar value.

To see how this equation relates to the Cartesian equations of the line, we write it in expanded form, namely

$$(x, y, z) = (q_1, q_2, q_3) + \gamma(d_1, d_2, d_3)$$

where x, y and z are the coordinates of point p, and q_1, q_2 and q_3 are the coordinates of point q and d_1, d_2 and d_3 are the coordinates of vector d.

If we separate out the coordinates of these points and vectors, we get three equations:

$$x = q_1 + \gamma d_1$$
$$y = q_2 + \gamma d_2$$
$$z = q_3 + \gamma d_3$$

Then if we express each of these equations in terms of γ, we get the Cartesian equations above.

For example, say we have the point $q = (1, 2, 3)$ and the vector $d = (1, 4, -1)$, then the line passing through point q and in the direction of d is specified by the Cartesian equations

$$\frac{x-1}{1} = \frac{y-2}{4} = \frac{z-3}{-1}$$

Note that if one of the components of the direction vector is zero, then the Cartesian equations turn out to be the equation of a line in 2-dimensions. For example, say the direction vector $d = (1, 4, 0)$, then the Cartesian equations are

$$\frac{x-1}{1} = \frac{y-2}{4}$$

i.e.

$$y = 4x - 2$$

We can also specify a line by means of a parametric equation if we have two points q and r on the line:

$$p = q + \gamma(r - q)$$

Here, $r - q$ determines the direction vector. (Remember that the difference between two points is a vector.) For values of γ between 0 and 1 we have all the points on the line segment between points q and r.

Sometimes the parametric equation of a line is rearranged in the form

$$p = \gamma r + (1-\gamma)q$$

2.11 Planes

A plane (in 3-dimensions) can be specified by means of an equation of the form

$$ax + by + cz + d = 0$$

where a, b, c and d are scalars, and x, y and z are the coordinates of any point in the plane. (As it turns out, a, b and c are the components of a vector normal to the plane.)

For example, the equation $x + 2y + 3z - 4 = 0$ is the equation of a plane with normal vector $(1, 2, 3)$, i.e. at right-angles to it.

A plane can be determined in various ways: (i) by a point on the plane and a normal vector, (ii) a point and two (linearly independent) vectors in the plane, (iii) three non-collinear points (i.e. not in a straight line), or (iv) a line and a point (not on the line). From each of these combinations, the equation of the plane can be determined:

(i) Given a point q on the plane and a vector n normal to the plane (with components n_1, n_2 and n_3) then (as stated above) the equation

$n_1 x + n_2 y + n_3 z + d = 0$

will define the plane. To get a value for scalar d, we substitute the coordinates of point q:

$n_1 q_x + n_2 q_y + n_3 q_z + d = 0$

so

$d = -(n_1 q_x + n_2 q_y + n_3 q_z)$

(ii) Given a point q on the plane and two vectors u and v in the plane, the cross product $u \times v$ will give a normal to the plane. Then using the method above, we can determine the equation of the plane.

(iii) Given three non-collinear points q, r and s in the plane, we can use two pairs of these points to determine two vectors in the plane by subtracting the points from one another. For instance, $q - r$ and $q - s$ will give two vectors in the plane. Then using the methods described above we can determine a normal to the plane (in this instance $(q - r) \times (q - s)$) and hence determine the equation of the plane.

For example, say we have three points $q = (1, 2, 0)$, $r = (0, 1, 2)$ and $s = (2, 0, 1)$ and we wish to determine the equation of the plane through them. One vector in the plane will be $q - r = (1, 1, -2)$ and another will be $q - s = (-1, 2, -1)$. The cross product of these two vectors is $n = (1.-1 - -2.2, -2.-1 - 1.-1, 1.2 - 1.-1) = (3, 3, 3)$ which will be a normal to the plane. Then the equation of the plane passing through the point q (or r or s) is $3x + 3y + 3z - 9 = 0$ or $x + y + z - 3 = 0$.

(iv) Say we have a line in the plane and point r in the plane. The line must be specified in some way, say a set of Cartesian equations as above, or a parametric equation. Then two vectors in the plane will be the direction vector $d = (d_1, d_2, d_3)$ and the vector $q - r$ (where q is the point (q_1, q_2, q_3) on the line and r is the other point on the plane). Now we have two vectors and a point on the plane (either q or r), and we can follow the above methods to determine the equation of the plane.

Planes can also be specified by means of parametric equations. These can be formed easily for cases (ii) and (iii) above:

(ii) Given a point q and two (linearly independent) vectors u and v in the plane, the equation

$p = q + \alpha u + \beta v$

will specify all points p on the plane.

(iii) Given three non-collinear points q, r and s, $q - r$ and $q - s$ will give two linearly independent vectors in the plane. The equation

$p = q + \alpha (q - r) + \beta (q - s)$

will then specify all points p on the plane. For values of α and β between 0 and 1, this equation gives all the points in the triangle formed by q, r and s.

2.12 Two-dimensional matrices, including the identity matrix

A two-dimensional *matrix* is a table of values with a fixed number of rows and columns. The values in the table are generally placed in a single pair of extended square brackets.

For example, $M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ has 4 rows and 4 columns, and is called a 4 by 4 matrix.

$N = \begin{bmatrix} 5 & -7 & 3 \\ 0.6 & 0 & -1 \end{bmatrix}$ has 2 rows and 3 columns, and is called a 2 by 3 matrix.

A matrix with only one row is called a row matrix, and a matrix with only one column is called a column matrix (see transpose of a matrix below for examples).

A matrix with the same number of rows and columns is called a square matrix. Matrix M above is an example of a square matrix.

A square matrix filled with zeroes except the main diagonal which is filled with ones is called an identity matrix.

For example, $I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ is a 4 by 4 identity matrix.

Matrices do no represent any specific geometric objects in 3-dimensional space. They are sometimes used to represent objects, for example a row or column matrix can be used to represent a point or vector. 4 by 4 matrices are used to represent transformations, for example a rotation or a translation that can be applied to an object (see matrix multiplication below). We generally use uppercase letters of the alphabet, eg. M and N, to refer to matrices.

2.13 Transpose of a matrix

The transpose of a matrix is another matrix with all its row and column values swapped. To be more precise, if A is an m by n matrix, then the transpose of A, written as A^T, is the n by m matrix with $\alpha^T_{ij} = \alpha_{ji}$.

For example, the transpose of the row matrix [1 2 3 4] is the column matrix $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$

The transpose of the 4 by 4 matrix $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ is the 4 by 4 matrix $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 \\ 4 & 5 & 7 & 1 \end{bmatrix}$

2.14 Matrix addition and subtraction

Matrix addition is only defined for two matrices that have the same number of rows and columns, and their sum is another matrix, also with the same number of rows and columns. Each value of the resulting matrix is simply the sum of the corresponding values of the two matrices.

For example, if $M = \begin{bmatrix} 5 & -7 & 3 \\ 6 & 0 & -1 \end{bmatrix}$ and $N = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ then the sum of M and N is

$$M + N = \begin{bmatrix} 5+1 & -7+2 & 3+3 \\ 6+4 & 0+5 & -1+6 \end{bmatrix} = \begin{bmatrix} 6 & -5 & 6 \\ 10 & 5 & 5 \end{bmatrix}$$

Similarly, one matrix can be subtracted from another (provided they once again have the same number of rows and columns). Each value of the resulting matrix is simply the difference between the corresponding values of the two matrices.

2.15 Multiplying a matrix by a scalar

The result of multiplying a matrix by a scalar is a matrix with each value multiplied by the scalar.

For example, say $\alpha = 2$ and $M = \begin{bmatrix} 5 & -7 & 3 \\ 6 & 0 & -1 \end{bmatrix}$ then $\alpha M = \begin{bmatrix} 10 & -14 & 6 \\ 12 & 0 & -2 \end{bmatrix}$

2.16 Dot product of two matrices

The product of one matrix with another is a matrix. In particular, the product of an n by l matrix A with an l by m matrix B, written $A\,B$, is the n by m matrix $D = [\delta_{ij}]$ where

$$\delta_{ij} = \sum_{k=1}^{l} \alpha_{ik}\beta_{kj}$$

For example, say we have two 4 by 4 matrices A and B:

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{bmatrix}$$

$$B = \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \\ \beta_{31} & \beta_{32} & \beta_{33} & \beta_{34} \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} \end{bmatrix}$$

Then their product is

$$A \cdot B = \begin{bmatrix} \cdots & & \cdots & & \cdots & \cdots \\ \cdots & & \cdots & & \cdots & \cdots \\ \cdots & \alpha_{31}\beta_{12} + \alpha_{32}\beta_{22} + \alpha_{33}\beta_{32} + \alpha_{34}\beta_{42} & & \cdots & \cdots \\ \cdots & & \cdots & & \cdots & \cdots \end{bmatrix}$$

(To keep things simple, we have only shown the value in row 3, column 2 of the product matrix. It is determined from the product of row 3 of matrix A and column 2 of matrix B as indicated. Similarly, all the other values of $A\ B$ are the products of the respective rows of A and columns of B.)

As a more concrete example, say we want to multiply a 4 by 4 matrix R with a 4 by 1 column matrix V where

$$R = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$V = [5\ \ 6\ \ 7\ \ 8]^T$$

Then the product of R and V is the 4 by 1 column matrix

$$R \cdot V = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix} = \begin{bmatrix} 1*5 + 2*6 + 3*7 + 4*8 \\ 0*5 + 1*6 + 2*7 + 3*8 \\ 0*5 + 0*6 + 1*7 + 2*8 \\ 0*5 + 0*6 + 0*7 + 1*8 \end{bmatrix} = \begin{bmatrix} 70 \\ 44 \\ 23 \\ 8 \end{bmatrix}$$

In Computer Graphics, the product of two matrices is used for so-called transformations. For example, if we want to translate, scale or rotate an object, we multiply each of its points (specified as column matrices) by the applicable transformation matrix. See Section 4.10 of Angel (2009) for a discussion of the different transformation matrices that are used.

2.17 Determinant of a matrix

The determinant of a matrix is a scalar and can only be calculated for square matrices. In particular, the determinant Δ of an n by n matrix A is defined as $|A| = \sum_{i=1}^{n} \alpha_{ij} C_{ij}$ (for some j) where

$C_{ij} \equiv (-1)^{i+j} M_{ij}$ and where M_{ij} is the minor of matrix A (i.e. the matrix formed by eliminating row i and column j from A).

For example, say we have the following 3 by 3 matrix:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Then

$$|A| = 1. \begin{Vmatrix} 5 & 6 \\ 8 & 9 \end{Vmatrix} - 2. \begin{Vmatrix} 4 & 6 \\ 7 & 9 \end{Vmatrix} + 3. \begin{Vmatrix} 4 & 5 \\ 7 & 8 \end{Vmatrix}$$

$$= 1.(5.9 - 8.6) - 2.(4.9 - 6.7) + 3.(4.8 - 5.7)$$

$$= -3 + 12 - 9$$

$$= 0$$

The determinant of a matrix is used for various purposes, for example to determine whether a collection of vectors are linearly independent. For example, the fact that the determinant of above matrix is 0 means that the vectors (1, 2, 3), (4, 5, 6) and (7, 8, 9) are linearly dependent (i.e. not linearly independent).

3   References

Angel, E. (2012) *Interactive Computer Graphics: A Top-Down Approach Using Shader Based OpenGL*. 6th edition. Addison-Wesley.

Astle, D. & Hawkins, K. (2002) *OpenGL Games Programming* First edition. Course Technology PTR.

Deitel, H.M & Deitel, P.J. (2000) *C How to Program*. Third edition. Prentice Hall