Introductory Vector Calculus
Version 1.04

Norman Wittels
121 Clark Road
Brookline MA 02146
wittels@world.std.com

Introductory Vector Calculus

Introduction
These notes summarize some of the vector calculus used in computer graphics and machine vision. They are not intended to supplant mathematics courses or texts nor are they intended to be complete or rigorous. The main purposes of these notes are to show a few examples at a proficiency level useful for designing computer graphics and machine vision algorithms and to provide references which may be helpful in achieving that level.

Vectors
We use *coordinate systems* to denote positions of points in space. All computer graphics/vision is done using *affine coordinate systems*. In such a system, each point is determined by a unique set of *coordinates*. The size of the set – the number of coordinates required – is called the *dimensionality* of the coordinate system. A 2D system requires 2 coordinates, a 3D system requires 3 coordinates, etc. Most computer graphics/vision applications use 2D, 3D, or 4D systems. We humans perceive that the world we live in can be described by 3D systems.

An affine coordinate system also has a unique point, called the *origin*, in which all coordinate values are zero. A point in the space is referenced by denoting the set of coordinates in the system being used. For example, the origin in a 3D space is

$$O = (0,0,0)$$

For the rest of these notes, the number of dimensions, if important, will be stated or can be deduced from the context. Different systems can be used to denote points in the space and the origins need not coincide. A *vector* is a direction in space defined as direction of the line connecting one point with another. If the beginning point is the origin, it is useful to associate a point P with the vector $\mathbf{P}$ that points from the origin to it.

$$\mathbf{P} = \overline{OP}$$

Vectors add. One consequence is that a group of vectors calculated from a sequence of points sum to the vector from the first to the last points:

$$\mathbf{P}_M = \overline{OP_1} + \overline{P_1P_2} + \cdots + \overline{P_{M-1}P_M} = \overline{OP_M}$$

It is convenient to represent a vector as a linear combination of a *basis set* of vectors

$$P = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_N\mathbf{v}_N$$

The most frequently used basis sets are made of *orthonormal* vectors. That means that the basis vectors are *orthogonal* – – and *normalized* – the length of each vector is unity (one). In the familiar cartesian coordinate system, these are just the unit vectors along each axes:

$$\hat{\mathbf{x}} = \overline{O(1,0,0)}, \quad \hat{\mathbf{y}} = \overline{O(0,1,0)}, \quad \hat{\mathbf{z}} = \overline{O(0,0,1)}$$

Note the use of the caret above a vector to indicate that it is normalized. Then the point P and its associated vector $\mathbf{P}$ can both be expressed in terms of the point's coordinates

Version 1.04

Introductory Vector Calculus

$$P = (x, y, z), \quad \mathbf{P} = (x\hat{\mathbf{x}}, y\hat{\mathbf{y}}, z\hat{\mathbf{z}})$$

When using the cartesian coordinate system, some authors use a mixed notation in which the linear sum of the unit vectors is implied by the coefficients (the coordinates)

$$\mathbf{P} = (x, y, z)$$

Be sure you understand the notation whenever you are reading a document in which both points and vectors occur. These notes will provide practice.

We will explore how to solve geometry problems using vectors. You may wonder why we don't simply use analytical geometry. One reason is that most of the equations of analytical geometry require careful attention to special cases. For example, the slope-intercept form of a line in 2D

$$y = y_0 + mx$$

does not work for vertical lines in which x is constant and the slope infinite. Vector forms of lines, however, work for all orientations without special cases. Thus it is easier to write algorithms that perform geometrical calculations using vectors instead of analytical geometry.

Although vector equations are general (independent of coordinate system), we will use cartesian coordinates for the rest of these notes because they are most generally used in algorithms that implement vector calculations. From the way vectors were associated with points above, we can see that linearity applies

$$\mathbf{P} = (x, y, z, w) \Rightarrow k\mathbf{P} = (kx, ky, kz, kw) \tag{1a}$$
$$\mathbf{P}_1 = (x_1, y_1), \quad \mathbf{P}_2 = (x_2, y_2) \Rightarrow a\mathbf{P}_1 + b\mathbf{P}_2 = (ax_1 + bx_2, ay_1 + by_2) \tag{1b}$$

Note, as a consequence of Equation (1a), a vector still points in the same direction when it is multiplied by a positive constant; it points in the opposite direction when the multiplicative constant is negative.

Vector Scalar Product
There are two products of vectors that are useful in solving geometric problems. One of them, is the *scalar product*, combines two vectors, of the same arbitrary dimension, to produce a scalar.

$$\mathbf{P}_1 = (x_1, y_1, z_1), \quad \mathbf{P}_2 = (x_2, y_2, z_2) \Rightarrow \mathbf{P}_1 \bullet \mathbf{P}_2 = x_1 x_2 + y_1 y_2 + z_1 z_2 \tag{2}$$

Because of the notation used, the scalar product is also known as the *dot product*. The number of terms in Equation (2) is adjusted to match the number of dimensions in the space. This equation also shows how to calculate the *magnitude* of a vector, which corresponds to the cartesian distance between any two points whose separation is described by the vector

$$\mathbf{P} = (x, y, z, w) \Rightarrow |\mathbf{P}| = \sqrt{\mathbf{P} \bullet \mathbf{P}} = \sqrt{x^2 + y^2 + z^2 + w^2} \tag{3a}$$

Introductory Vector Calculus

Note that the vector part of Equation (3a) is general – it applies to vectors with any number of dimensions

$$|\mathbf{P}| = \sqrt{\mathbf{P} \bullet \mathbf{P}} \tag{3b}$$

We can use this last result to normalize a vector.

$$\hat{\mathbf{P}} = \frac{\mathbf{P}}{|\mathbf{P}|} = \frac{\mathbf{P}}{\sqrt{\mathbf{P} \bullet \mathbf{P}}} \tag{3c}$$

From Equation (3a), it can be shown that the dot product is commutative and distributive with respect to addition

$$\mathbf{P}_1 \bullet \mathbf{P}_2 = \mathbf{P}_2 \bullet \mathbf{P}_1, \tag{4a}$$

$$\mathbf{P}_1 \bullet (\mathbf{P}_2 + \mathbf{P}_3) = \mathbf{P}_1 \bullet \mathbf{P}_2 + \mathbf{P}_1 \bullet \mathbf{P}_3 \tag{4b}$$

In the right side of Equation (4b), the associativity is implied: the dot products must be done first since there is no triple dot product – the dot product is a binary operation and a scalar cannot be dotted with a vector.

To help develop a geometric interpretation of the dot product, consider the 2D vector diagrams shown in Figure 1.
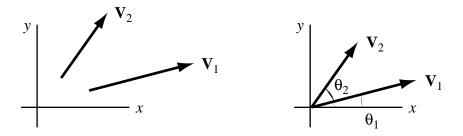

Figure 1. Equivalent 2D vector diagrams.

Although this Figure shows the vectors $\mathbf{V}_1$ and $\mathbf{V}_2$ lying in the first quadrant and generally pointing out from the origin, our analysis will work for vectors of arbitrary location and orientation. In the calculation we will use the angle $\theta = \theta_2 - \theta_1$ between the directions of $\mathbf{V}_1$ and $\mathbf{V}_2$. Since the vectors are pure directions, we can place them tail to tail, move the tails to the origin, and normalize them without changing the value of θ. The coordinates of the points corresponding to the two vectors are:

$$\mathbf{V}_1 = |\mathbf{V}_1|(\cos\theta_1, \sin\theta_1), \quad \mathbf{V}_2 = |\mathbf{V}_2|(\cos\theta_2, \sin\theta_2)$$

Thus

$$\mathbf{V}_1 \bullet \mathbf{V}_2 = \mathbf{V}_{1,x}\mathbf{V}_{2,x} + \mathbf{V}_{1,y}\mathbf{V}_{2,y} = \left|\mathbf{V}_1\right|\left|\mathbf{V}_2\right|\left(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2\right) = \left|\mathbf{V}_1\right|\left|\mathbf{V}_2\right|\cos\theta \qquad (5)$$

If the vectors are normalized, the dot product is just the perpendicular projection of each onto the other. The expression Equation (5) is even with respect to the angle θ, as it must be for the dot product to be commutative. Thus the dot product is useful for calculating the magnitude of the angle between vectors, but not the direction.

Example 1
Calculate the angle that the vector $\mathbf{V}=(-1.0,-2.0)$ makes with the x axis.

If we apply Equation (5) directly, we obtain.

$$\theta = \cos^{-1}\frac{\mathbf{V}\bullet\hat{\mathbf{x}}}{\left|\mathbf{V}\right|\left|\hat{\mathbf{x}}\right|} = \cos^{-1}\frac{-1}{\sqrt{5}} = 116.6°$$

This answer implies that $\mathbf{V}$ points to the upper left (x negative, y positive) whereas it actually points to the lower left (both x and y negative). The problem is that cos is an even function so there is no way to deduce that the correct answer is $-116.6°$. A better way to solve this type of problem is to calculate the inverse tangent using projections on both the x and y axes, Figure 2.
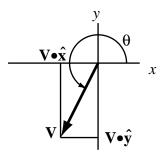

Figure 2. Projections of a vector on to two coordinate axes.

From this Figure, it can be seen that the angle is

$$\theta = \tan^{-1}\frac{\mathbf{V}\bullet\hat{\mathbf{y}}}{\mathbf{V}\bullet\hat{\mathbf{x}}} = \tan^{-1}\frac{-2}{-1} = 243.4°$$

This answer is equivalent to $-116.6°$. Note that any algorithm to calculate the angle should use a four-quadrant inverse tangent function, such as atan2() in the standard C library package[1].

---

1 PJ Plauger, *The Standard C Library* (1992, Prentice Hall, Englewood Cliffs NJ).

Equation (5) shows that the dot product of orthogonal vectors is zero. This makes it natural to consider using the dot product whenever orthogonality needs to be proven or satisfied.

Example 2
Find a vector orthogonal to the vector $\mathbf{V}=(-1.0,-2.0)$.

We expect that the answer will not be unique because there are two opposite directions orthogonal to the direction of $\mathbf{V}$ and because there is an infinity of vectors pointing in these two directions but differing in magnitude.

Let the unknown vector be $\mathbf{U}=(x,y)$. Then the values x and y have to satisfy the relationship

$$\mathbf{U} \bullet \mathbf{V} = -x - 2y = 0 \tag{6a}$$

There are, indeed, infinitely many vectors that satisfy Equation (6a): $(1,-0.5)$, $(-1,0.5)$, $(10,-5)$, $(-2\pi,\pi)$, etc. Note that the first two have the same magnitude but point in opposite directions. We can restrict the range of acceptable answers by appending other conditions. For example, if we require that $\mathbf{U}$ be a unit vector,

$$\mathbf{U} \bullet \mathbf{U} = x^2 + y^2 = 1 \tag{6b}$$

then the solution to Equations (6a) and (6b) becomes

$$y^2 = \frac{1}{5} \quad \Rightarrow \quad \mathbf{U} = \left(\frac{2}{\sqrt{5}}, \frac{-1}{\sqrt{5}}\right), \quad \mathbf{U} = \left(\frac{-2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right) \tag{6c}$$

The remaining directional ambiguity cannot be resolved without imposing further constraints.

If you are used to doing degrees-of-freedom analysis on systems of equations, note that Equation (6a) only imposes a single constraint because it is only one equation unlike other vector expressions such as Equation (1b), which really represents two equations (in a 2D space). Again, it is important to keep track of how many equations are embedded in each vector expression.

Lines
A line has infinite extent and infinitesimal width. There are two ways that lines are specified in vector algebra, Figure 3.
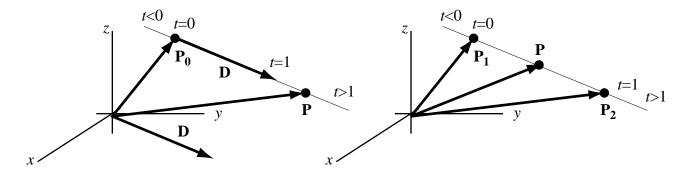
Figure 3. Two ways to define a line.

In the first definition, a point $\mathbf{P}$ is specified by a point $\mathbf{P_0}$ on the line and direction vector $\mathbf{D}$ whose magnitude is non-zero. It may be easier to visualize what is happening if the tail of the direction vector is moved to the point $\mathbf{P_0}$ as shown. Then the equation of the point $\mathbf{P}$ is

$$\mathbf{P} = \mathbf{P_0} + t\mathbf{D} \tag{7a}$$

There are infinitely many points $\mathbf{P}$ which lie on the line and each has a unique value of t. As shown in the Figure, $t=0$ describes the point $\mathbf{P_0}$, $t<0$ describes points in the direction opposite $\mathbf{D}$ from point $\mathbf{P_0}$, and $t>0$ describes points in the direction of $\mathbf{D}$ from point $\mathbf{P_0}$.

The second definition of a line is the same as the first except that two points are used to define the line and the direction vector is calculated as the difference between them

$$\mathbf{P} = \mathbf{P_1} + t(\mathbf{P_2} - \mathbf{P_1}) = (1-t)\mathbf{P_1} + t\mathbf{P_2} \tag{7b}$$

Example 3
Find the intersection of the line passing through $(1,1,1)$ in the $(-1,-1,1)$ direction with the xy plane.

From Equation (7a), the equation of the line is

$$\mathbf{P} = (1,1,1) + t(-1,-1,1) = (1-t, 1-t, 1+t) \tag{8}$$

Notice that we have three equations in one unknown, t. To find its value we use the fact that the intersection point must lie in the xy plane, where its z value is zero. Setting the z portion of Equation (8) to zero yields the solution.

$$1 + t = 0 \quad t = -1, \quad P = (2,2,0)$$

The final step was to turn the vector $\mathbf{P}$ back into the point P. This is another example of a notation that requires care to prevent confusion.

In the last example we used the fact that we already knew one of the coordinates of the

intersection point to decide which of the three equations in Equation (8) to solve. One usually does not have this kind of knowledge and it is necessary to solve all of the vector equations to find a solution.

Example 4

The vector $\mathbf{V}$ connects the point $(1,1)$ with the point P. The length of $\mathbf{V}$ is $\sqrt{20}$ and it is perpendicular to the vector $\mathbf{D} = (-1,-2)$. What are the coordinates of P ?

We first need to find $\mathbf{V}$ and then we can add it to $(1,1)$ to find $\mathbf{P}$. We know two equations for $\mathbf{V}$

$$\mathbf{V} \bullet \mathbf{D} = 0 \quad \Rightarrow \quad -x_\mathbf{V} - 2y_\mathbf{V} = 0 \tag{9a}$$

$$|\mathbf{V}|^2 = x_\mathbf{V}^2 + y_\mathbf{V}^2 = 20 \tag{9b}$$

Solving these two equations simultaneously provides two values for $\mathbf{V}$ and two values for $\mathbf{P}$.

$$\mathbf{V} = (-4,2), \quad \mathbf{V} = (4,-2) \quad \Rightarrow \quad \mathbf{P} = (-3,3), \quad \mathbf{P} = (5,-1)$$

Again, we have two solutions because there are two vectors orthogonal to $\mathbf{D}$. Additional constraints will be necessary if we wish to eliminate one of the solutions.

The techniques of Example 4 would not have worked had there been more than two dimensions in the space. That is because in three or more dimensions, there are infinitely many vectors perpendicular to $\mathbf{D}$; Equations (9a) and (9b) then would have represented two equations in three or more unknowns and there would have been no general solution. Again, it is important to keep track of the number of equations when solving vector problems.

Vector Product

The second product of vectors is the *vector product* or *cross product*.

$$\mathbf{A} \times \mathbf{B} = \left(y_a z_b - y_b z_a\right)\hat{\mathbf{x}} + \left(z_a x_b - z_b x_a\right)\hat{\mathbf{y}} + \left(x_a y_b - x_b y_a\right)\hat{\mathbf{z}} \tag{10a}$$

Note the cyclic nature of the variables: *xyz*, *yzx*, *zxy*. For convenience, Equation (10a) is sometimes written as the determinant of a 3×3 matrix.

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ x_a & y_a & z_a \\ x_b & y_b & z_b \end{vmatrix} \tag{10b}$$

The cross product differs from the other vector operators we have seen in several respects. First, it only applies to 3D vectors. As shown in Equation (10a); the result has exactly three coordinates. While some generalizations of the cross product to other dimensions have been presented in the research literature, for purposes of computer graphics/vision it is advised that

you restrict the cross product to 3D problems only. Second, although the result *looks* like a vector, it has different properties. For example, if **A** and **B** are both forces measured in N, their cross product has units N^2, a quantity with no physical significance.

In a similar manner used in the dot product, it can be shown that

$$\mathbf{A} \times \mathbf{B} = \hat{\mathbf{N}}\left(|\mathbf{A}||\mathbf{B}|\sin\theta\right) \tag{10c}$$

where θ is the angle between **A** and **B** and $\hat{\mathbf{N}}$ is a unit vector perpendicular to both **A** and **B**. The definitions in Equations (10) only apply if neither **A** nor **B** has magnitude zero and they are neither parallel nor anti-parallel (non-intersecting and pointing in opposite directions). If any of those cases obtain, the magnitude of the cross product is zero and its direction is undefined. The fact that parallel vectors have a zero cross product can be used to produce another definition of the line that is equivalent to Equation (7a) but not always as useful since it only applies in 3D systems and it contains three equations instead of one.

$$\left(\mathbf{P} - \mathbf{P_0}\right) \times \mathbf{D} = 0 \tag{11}$$

As with the dot product operator, the cross product can be used to determine the angle between vectors, but caution is advised – the odd nature of the sin causes a quadrant ambiguity similar to that discussed in Example 1. The cross product is used for testing whether vectors are parallel – if so, their cross product will have magnitude zero. And, since the cross product of two vectors is perpendicular to both of them, it can be used to find vectors perpendicular to pairs of vectors. Thus, for example, if **A** and **B** are vectors in a plane, Equation (10c) can be used to find the direction of the normal to the plane. Before we explore this further, it will be helpful to establish a few more properties of the cross product.

Unlike the dot product, the cross product is not commutative

$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A} = \mathbf{B} \times -\mathbf{A} = -(\mathbf{B} \times \mathbf{A}) \tag{12a}$$

This result is consistent with Equation (10c) which showed that the cross product is odd with regard to the angle between the vectors, and thus their order. Also, the cross product is associative with addition

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$$

but not with the cross product. The expression

$$\mathbf{A} \times \mathbf{B} \times \mathbf{C}$$

has no meaning since the two possible interpretations have different expansions which point in different directions (assuming **A**, **B**, and **C** are arbitrary directions)

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \bullet \mathbf{C})\mathbf{B} - (\mathbf{A} \bullet \mathbf{B})\mathbf{C}, \quad (\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = (\mathbf{A} \bullet \mathbf{C})\mathbf{B} - (\mathbf{B} \bullet \mathbf{C})\mathbf{A} \tag{12b}$$

Introductory Vector Calculus

There is an intuitive explanation for this result. **B**×**C** is perpendicular to both **B** and **C** so **A**×(**B**×**C**) must be perpendicular both to it and to **A**. This means **A**×(**B**×**C**) is coplanar with **B** and **C**. Similarly, (**A**×**B**)×**C** is coplanar with **A** and **B**. Thus they cannot, in general, be equal.

The mixed product involving three vectors, one cross product, and one dot product is unambiguous, is formed of nine terms, and can be written as a determinant:

$$\mathbf{A} \bullet \mathbf{B} \times \mathbf{C} = \mathbf{A} \bullet (\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} x_a & y_a & z_a \\ x_b & y_b & z_b \\ x_c & y_c & z_c \end{vmatrix} \tag{12c}$$

From Equation (12a) and identities involving the interchange of rows in the determinant, these identities follow

$$\mathbf{A} \bullet \mathbf{B} \times \mathbf{C} = \mathbf{B} \bullet \mathbf{C} \times \mathbf{A} = \mathbf{C} \bullet \mathbf{A} \times \mathbf{B} = -\mathbf{A} \bullet \mathbf{C} \times \mathbf{B} = -\mathbf{B} \bullet \mathbf{A} \times \mathbf{C} = -\mathbf{C} \bullet \mathbf{B} \times \mathbf{A} \tag{12d}$$

We can apply Equation (12c) with the observation that determinants with equal rows are zero, to prove the following statements

$$\mathbf{A} \bullet \mathbf{A} \times \mathbf{B} = \mathbf{A} \bullet \mathbf{B} \times \mathbf{A} = \mathbf{B} \bullet \mathbf{A} \times \mathbf{B} = \mathbf{B} \bullet \mathbf{B} \times \mathbf{A} = 0 \tag{12e}$$

Equations (12e) are the basis for the assertion that the cross product of **A** and **B** is orthogonal to both, as stated earlier. If **A** is a linear combination of **B** and **C** then it is co-planar with both. Thus, if we know that any mixed product of three vectors is zero, such as

$$\mathbf{A} \bullet \mathbf{B} \times \mathbf{C} = 0,$$

that is sufficient to prove that the three vectors are coplanar.

As a final note about the cross product, let's consider its geometrical interpretation. Figure 4 and Equation (10c) shows the magnitude of the cross product **A**×**B** equals the area of the parallelogram and twice the area of the triangle formed by the origin and the two vectors.
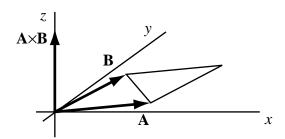

Figure 4. Geometrical interpretation of **A**×**B**

Similarly the magnitude of the mixed product of three vectors can be seen to be the volume of the parallelepiped defined by the three vectors.

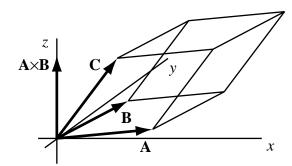Version 1.04

Introductory Vector Calculus

Figure 5. Geometrical interpretation of **C•A×B**

Planes

The two-point definition of a line, Figure 3 and Equation (7b) can be generalized to produce a three-point definition of a plane. Assume that a plane is defined by three non-conlinear points A,B, and C. Then the vector to point **P** in the plane can be described as a linear combination of the vectors **A**, **B**, and **C**, Figure 6.

$$\mathbf{P} = \mathbf{A} + t(\mathbf{B} - \mathbf{A}) + u(\mathbf{C} - \mathbf{A}) \quad \Rightarrow \quad \mathbf{P} = (1 - t - u)\mathbf{A} + t\mathbf{B} + u\mathbf{C} \tag{13a}$$

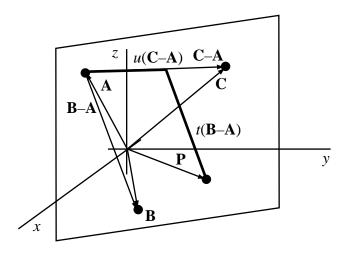The parameters t and u are arbitrary real numbers.

Figure 6. Three point definition of a plane.

Example 5

Find the point in the plane that passes through the points (1,0,0,1), (0,1,1,0), and (0,–1,1,1) where that plane intersects the line which passes through the points (0,0,0,0) and (1,1,1,1).

Using Equation (13a), the equation of a point in the plane is

$$\mathbf{P} = (1 - t - u)(1,0,0,1) + t(0,1,1,0) + u(0,-1,1,1)$$

and using Equation (7b) equation of a point on the line is

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

Equating these two definitions to force the point to lie both in the plane and on the line,

$$\mathbf{P} = (1 - t - u, t - u, t + u, 1 - t) = (v, v, v, v)$$

we obtain four equations to solve for the three unknowns

$$\left.\begin{array}{rrrr} -t & -u & -v & = -1 \\ t & -u & -v & = 0 \\ t & +u & -v & = 0 \\ -t & & -v & = -1 \end{array}\right\} \implies t = \frac{1}{2}, \quad u = 0, \quad v = \frac{1}{2} \implies \mathbf{P} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$$

Note that the selection of the five points in this problem was fortuitous; most sets of overdetermined equations have no solutions. That is equivalent to saying that most 4D lines and planes do not intersect.

_____

A second approach for deriving the equation of the plane is based on the normal to the plane **N** and a point in the plane **P₀** , Figure 7.
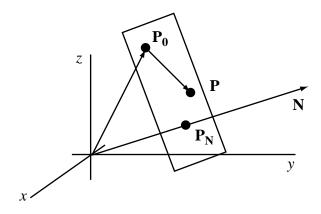

Figure 7. Point-normal definition of a plane.

Assume, for now, that **N** and **P₀** are known. Then, for any point in the plane **P**, the vector from **P₀** to **P** is orthogonal to **N**

$$(\mathbf{P} - \mathbf{P_o}) \bullet \mathbf{N} = 0 \tag{13b}$$

An alternate form can be obtained by dividing Equation (13b) by the magnitude of **N** to normalize the normal.

$$\mathbf{P} \bullet \hat{\mathbf{N}} = \mathbf{P_O} \bullet \hat{\mathbf{N}} = D \tag{13c}$$

Note that the right side of Equation (13c) is just the length of $\mathbf{P_0}$ projected on to the normal. The symbol $\mathbf{P_N}$ has been placed in Figure 7 to show where the point $\mathbf{P_0}$ projects. D is significant because it is the normal distance (the minimal distance) from the origin to the plane. This Figure also shows that any point in the plane projected on to the normal produces the vector $\mathbf{P_N}$

$$\mathbf{P_N} = \left(\mathbf{P} \bullet \hat{\mathbf{N}} \right) \hat{\mathbf{N}}$$

Given three non-collinear points on the plane, A, B, and C, the normal **N** can be found by using the cross product of two vectors in the plane (**B**–**A**) and (**C**–**A**).

$$\mathbf{N} = (\mathbf{B} - \mathbf{A}) \times (\mathbf{C} - \mathbf{A}) = (\mathbf{A} \times \mathbf{B}) + (\mathbf{B} \times \mathbf{C}) + (\mathbf{C} \times \mathbf{A}) \tag{14}$$

Note that the sign of the normal is indeterminate; of the six possible orders for selecting points A,B, and C when finding the two vectors in the plane, three have each sign. Note that the second form in Equation (14) is obtained from the first by means of vector identities. For the known point in the plane in Equations (13b) or (13c) $\mathbf{P_0}$ any of the three vectors **A**, **B**, or **C** can be used. Since the cross product is only defined in three dimension,s Eq. (14) only applies in 3D.

Example 6
Find the point at which the line passing through the points (1,–1,0) and (1,1,1) intersects the plane passing through the points (1,0,0), (0,1,0) and (0,0,1). This is the generalization of Example 3.

The equation of the line is

$$\mathbf{P} = (1 - t)(1,-1,0) + t(1,1,1) = (1,-1+2t,t) \tag{15a}$$

Use the first point as the known point and calculate the normal using Equation (14)

$$\mathbf{N} = \left[(0,1,0) - (1,0,0) \right] \times \left[(0,0,1) - (1,0,0) \right] = (1,1,1)$$

to find the plane

$$\left[\mathbf{P} - (1,0,0) \right] \bullet (1,1,1) = 0 \tag{15b}$$

Combine Equations (15a) and (15b) and solve for t

$$\left[(1,-1+2t,t) - (1,0,0) \right] \bullet (1,1,1) = 3t - 1 = 0 \quad \Rightarrow \quad t = \frac{1}{3} \quad \Rightarrow \quad \mathbf{P} = \left(1, \frac{-1}{3}, \frac{1}{3} \right)$$

It is important to remember that, becaue of the cross product, Equation (14) only applies in 3D. In 2D, all vectors must be coplanar so there is no meaning to a vector normal to the plane. With any number of dimensions above three, some other method will have to be used to find the normal. Thus, Example 5 could not be solved using the method we used for Example 6.

Example 7
Find the distance from the point $\mathbf{A}=(-1,1,1)$ to the plane passing through the points $(0,0,0)$, $(1,1,1)$, and $(1,1,0)$.

Begin by finding the equation of the plane.

$$\mathbf{N} = \left[(1,1,1)-(0,0,0)\right] \times \left[(1,1,0)-(0,0,0)\right] = (-1,1,0) \quad \Rightarrow \quad \hat{\mathbf{N}} = \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$$

$$\left[\mathbf{P}-(0,0,0)\right] \bullet \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) = 0$$

We only know how to find the distance from the origin to the plane. If we subtract $(-1,1,1)$ from every point in the problem, we move $\mathbf{A}$ to the origin but the geometry otherwise remains the same. The plane will move by the same amount; it's normal will remain the same but the known point will change. The new equation of the plane is

$$\left[\mathbf{P}-(-1,1,1)\right] \bullet \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) = 0$$

Rearrange this equation to put it into the form of Equation (13c)

$$\mathbf{P} \bullet \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) = (-1,1,1) \bullet \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) = D \quad \Rightarrow \quad D = \sqrt{2}$$

If we had moved all four points to place $\mathbf{A}$ at the origin before calculating the equation of the plane, we would have obtained the same result.

The vector methods we have outlined can be used to solve any problem of analytical geometry, but vector analysis is a tool, not an automatic method. Thought and understanding are also required.

Example 8
What is the equation of the line of intersection of the two planes described by the following equations?

$$\left[\mathbf{P}-(1,1,1)\right] \bullet (1,1,1) = 0, \quad \left[\mathbf{P}-(1,0,-1)\right] \bullet (1,1,0) = 0 \tag{16a}$$

Introductory Vector Calculus

A line can be defined by two points or one point and a vector. Since we don't know any points on the line, let's start by calculating the vector describing its direction. Any vector perpendicular to a plane's normal must lie in the plane. Thus the cross product of the two normals must lie in both planes, that is on the line of intersection.

$$\mathbf{V} = (1,1,1) \times (1,1,0) = (-1,1,0) \tag{16b}$$

The fact that this vector has non-zero magnitude assures us that the planes are not parallel so an intersection exists. We still need to find a point in both planes. Assume the point's coordinates are (x,y,z) and use Equation (16a) to find equations for the coordinates.

$$(x,y,z) \bullet (1,1,1) = (1,1,1) \bullet (1,1,1) \quad \Rightarrow \quad x+y+z=3$$
$$(x,y,z) \bullet (1,1,0) = (1,0,-1) \bullet (1,1,0) \quad \Rightarrow \quad x+y=1$$

Thus we have two equations in three unknowns, not enough information to find a unique solution. That is not surprising since we know there are infinitely many points on the line. We only need to find a single point, any point, to be able to write the equation for the line, so let's guess one of the coordinates and see if we can find a solution. The obvious choice is to set one of the coordinates to zero. Let's look at all three cases:

Try x=0:

$$y+z=3, \quad y=1 \quad \Rightarrow \quad \mathbf{P}_1 = (0,1,2) \tag{16c}$$

Try y=0:

$$x+z=3, \quad x=1 \quad \Rightarrow \quad \mathbf{P}_2 = (1,0,2) \tag{16d}$$

Try z=0:

$$x+y=3, \quad x+y=1 \quad \Rightarrow \quad \text{no solution!}$$

This last result is telling us that the line is parallel to the z axis so it never crosses it. Thus, in general, we expect to find one, two, or three possible intersection points, depending on the direction of the line. We can use this method to find one point on the line, but not necessarily two. That is why we chose to use the point-vector definition of the line and not the two-point definition. Either of the points in Equations (16c) or (16d) will suffice – different but equally valid solutions will result. One equation of the intersection line is

$$\mathbf{P} = (0,1,2) + t(-1,1,0) = (-t, 1+t, 2) \tag{16e}$$

To verify that this point lies in both planes, substitute Equation (16e) into Equation (16a). Both parts are satisfied for all values of t, as required.

_____

References
In these notes, we have presented a summary of the vector calculus used in computer graphics and vision. If you want to have more complete explanations, some of these references will be useful. Warning: this list is not exhaustive and the annotations are subjective.

B Hoffmann, *About Vectors* (1966, Dover, New York). This is an excellent, inexpensive introductory text book about vectors and vector operations in 3D.

GE Hay, *Vector and Tensor Analysis* (1953, Dover, New York). This is an inexpensive introductory text book about vectors and tensors with applications to problems in areas other than computer graphics and vision. Chapter 2 is an excellent discussion of how to solve geometry problems using vectors.

K Rektorys, *Survey of Applicable Mathematics* (1969, MIT Press, Cambridge MA). This is a massive handbook with summaries of large areas of mathematics. The writing is compressed and prior knowledge is assumed. The section on vector algebra is complete and the section on analytical geometry is excellent. This book is useful for understanding how to solve geometry problems as a precursor for writing vector solutions to geometry problems.

HP Manning, *Geometry of Four Dimensions* (1956, Dover, New York). This is a complete, rigorous book of axioms, theorems, and problems of 4D geometry. This is not a reference book – it must be read cover -to-cover to be useful.

FS Woods, *Higher Geometry* (1961, Dover, New York). This is a complete, rigorous book about 2D, 3D, 4D and higher geometry. It provides ways to think inductively about higher geometry (how to develop intuitions about (N+1)D geometry from your understanding of ND geometry).