ARE Math Review Lecture 4: Euclidean Spaces

Stephen M. Stohs

August 16, 2004

1 Vectors in Euclidean Space

I have already mentioned vectors in my discussion of matrix algebra, where I suggested that a matrix restricted to either a single row or a single column could be construed as a vector. The material at hand focuses on the special properties of vectors in Euclidean space.

The material in sections 10.1 and 10.2 of Simon and Blume looks like background reading, so I merely offer a few clarifications, then begin the discussion with the material in section 10.3.

1. A vector in n-dimensional Euclidean space can be written several different ways. For this chapter, we will write them as ordered n-tuples $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, but keep in mind that we could just as well write them as "row matrices"

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}$$

or "column matrices"

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Not much changes in these cases.

For example, given the $m \times n$ matrix A, we can interpret each column of A as a vector in R^m, and each row of A as a vector in R^n. The notation and context are different but the defining properties for vectors in a Euclidean space still apply.

2. Some treatments make a sharp distinction between vectors and points. We will abstract from this dichotomy by noting that Euclidean n-tuples subject to the algebraic properties discussed below lend themselves to a variety of interpretations. With this proviso, we can think of vectors as representing points, displacements, forces, portfolios, quantities, prices, etc., depending upon the context, and exploit the symmetry that obtains from describing all of them as vectors in a Euclidean space.

3. The main conceptual difference between thinking of a Euclidean n-tuple as a vector or a point is that we generally think of points as fixed with respect to the origin, whereas a vector has magnitude and direction, and may be moved around (displaced) by translation to a new origin.

4. There is no stipulation in the theory of vectors which limits them to be Euclidean n-tuples; rather, the properties pertain to any set of scalars and set of vectors which, taken together, obey them. Euclidean n-tuples, which are the focus of this discussion, are but one of myriad different mathematical objects which behave according to the properties which define vectors as such.

2 The Algebra of Vectors

Vector addition and scalar multiplication are defined as follows: Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ be two vectors in R^n and t a scalar. Then addition is defined by

$$\mathbf{x}+\mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y + n) \tag{1}$$

and scalar multiplication, by

$$t\mathbf{x} = (tx_1, tx_2, \ldots, tx_n). \tag{2}$$

Vectors have similar properties to matrices under the operations of scalar multiplication and addition. In short, vectors obey all of the same laws of addition and scalar multiplication which matrices do.

A set of vectors subject to the laws of addition and scalar multiplication, and which is closed under both addition and scalar multiplication, is called a

vector space. Page 751 in Simon and Blume provides the full list of defining properties of a vector space, including all the laws governing addition and scalar multiplication. It is a simple exercise to verify that these laws hold for Euclidean n-tuples, subject to addition and scalar multiplication.

3 The Geometry of Vectors

Corresponding to the notions of addition and scalar multiplication of vectors are simple diagrams which provide insight. To add vectors $\mathbf{x} + \mathbf{y}$, first draw $\mathbf{x}$ from the origin, then draw $\mathbf{y}$ translated so that its tail is positioned at the head of $\mathbf{x}$. The sum $\mathbf{x} + \mathbf{y}$ connects the origin to the head of the (translated) version of $\mathbf{y}$.

Scalar multiplication $t\mathbf{x}$ has the effect of either lengthening $\mathbf{x}$ ($|t| > 1$), shortening it ($|t| < 1$) or leaving its length unchanged ($|t| = 1$). Further, the direction of $t\mathbf{x}$ changes if $t < 0$, and is otherwise unchanged.

4 Properties of the Inner Product

I introduced inner product in the examples from the lecture on matrix operations. This section will take an extensive look at the properties of the (Euclidean) inner product. The "Euclidean" qualifier suggests that we are considering a special case here: The notions of length and inner product (or dot product) defined on vectors in a finite-dimensional Euclidean space. However, the properties actually apply to a wider class of mathematical objects, provided there is a meaningful way to define the inner product.

4.1 Dot Product and Euclidean Norm

As already mentioned in class, the inner product applied to a Euclidean space is referred to as the **dot product** or **Euclidean inner product**, defined for vectors $\mathbf{x}$ and $\mathbf{y}$ in R^n by

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i. \tag{3}$$

A related concept to the dot product is that of the Euclidean norm, defined by

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}. \tag{4}$$

It is clear from comparing definitions that

$$\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2 . \tag{5}$$

It is easy to show that scalar multiplication of a vector has the effect of scaling up the norm proportionally:

$$\|r\mathbf{x}\| = |r| \, \|\mathbf{x}\| \tag{6}$$

.

The algebraic properties of the inner product are easy to verify, and are listed on p. 214 of Simon and Blume. Let $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$ be any vectors and t be any scalar.

They include commutative and distributive laws for dot product multiplication, plus the following additional properties:

$$\mathbf{x} \cdot (t\mathbf{y}) = t(\mathbf{x} \cdot \mathbf{y}) = (t\mathbf{x}) \cdot \mathbf{y}, \tag{7}$$

which shows how scalar multiplication may be factored out of the dot product, and the requirement that

$$\mathbf{x} \cdot \mathbf{x} \geq 0, \tag{8}$$

with equality if and only if $\mathbf{x} = 0$.

Note that the commutative law

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x} \tag{9}$$

and the distributive law

$$\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z} \tag{10}$$

together imply the rule for expanding

$$(\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} + 2(\mathbf{x} \cdot \mathbf{y}) + \mathbf{y} \cdot \mathbf{y}. \tag{11}$$

4.2 Length and Distance

Next we consider the question of how to find the distance between two points. Consider two representative n-tuples in R^n; call them $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. We can either interpret these as points in n-dimensional Euclidean space, or as vectors with tails at the origin and heads at these points. We will review both interpretations below, and see that the same notions of length and distance apply in either case.

4.2.1 Euclidean Distance

Recall from your first exposure to analytic geometry the distance formula, which applies the Pythagorean theorem to a right triangle in the Cartesian (x-y) plane[1]. We wish to find the distance between arbitrary points $\mathbf{u} = (x_1, y_1)$ and $\mathbf{v} = (x_2, y_2)$. To do so, we construct the right triangle whose acute angles have vertices at $\mathbf{u}$ and $\mathbf{v}$, and whose right angle has its vertex at $\mathbf{w} = (x_2, y_1)$. We then note that the segment connecting $\mathbf{u}$ and $\mathbf{v}$ is the hypotenuse, and hence its length, which is the distance we seek, may be found by applying the Pythagorean theorem. Noting that the horizontal leg has length $|x_1 - x_2|$ and the vertical leg has length $|y_1 - y_2|$, we can write

$$d(\mathbf{u}, \mathbf{v}) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \tag{12}$$

where $d(\cdot, \cdot)$ denotes the distance function.

Next extend the example to three dimensions. The standard left-handed three dimensional coordinate system[2] introduced in a first course in vector calculus three mutually-perpendicular coordinate axes with a common origin: a positive x-axis pointing to the right, a positive y-axis coming out of the page towards the reader, and a positive z-axis pointing vertically upwards. To measure the distance between points $\mathbf{u} = (x_1, y_1, z_1)$ and $\mathbf{v} = (x_2, y_2, z_2)$ we first measure the distance between the points $\mathbf{u}$ and $\mathbf{w} = (x_2, y_2, z_1)$. Noting that the two points both lie in the plane $z = z_1$, we see there is no movement in the z direction, and hence the two-dimensional distance formula applies to show $d(\mathbf{u}, \mathbf{w}) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$. Next note that the segment connecting $\mathbf{w}$ to $\mathbf{v}$ is perpendicular (orthogonal) to the plane $z = z_1$, and hence there is only movement in the z direction along this segment; its length is clearly $|z_1 - z_2|$. Finally, noting that the segments connecting $\mathbf{u}$ to $\mathbf{w}$ and $\mathbf{w}$ to $\mathbf{v}$ form the legs of a right triangle with hypotenuse connecting $\mathbf{u}$ to $\mathbf{v}$, we can apply the Pythagorean theorem to obtain the three-dimensional distance formula:

$$d(\mathbf{u}, \mathbf{v}) = \sqrt{d(\mathbf{u}, \mathbf{w})^2 + (z_1 - z_2)^2}$$

[1] In the language of this course, we are talking about two-dimensional Euclidean space, R^2.

[2] Note that we could change the setup to utilize a right-handed coordinate system by a change of orientation; for example, interchanging the labels on the x-axis and the y-axis would suffice. Objects in this world are a mirror image of those in the corresponding left-handed coordinate system.

$$= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}. \quad (13)$$

We see from the above examples that adding another dimension requires adding another squared difference under the square root sign in the distance formula to account for movement in the direction of the newly added dimension. This is the basis for extending the distance formula to points in n-dimensional Euclidean space. We define the distance between $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ to be

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}. \quad (14)$$

Referring back to the definition of the Euclidean norm, we see that Euclidean distance can be defined in terms of the norm:

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}. \quad (15)$$

Setting $\mathbf{y} = \mathbf{0}$, we see that $\|\mathbf{x}\|$ measures the distance between the point described by $\mathbf{x}$ and the origin.

4.2.2 Vector Geometry

Now consider the geometry of $\mathbf{x}$ and $\mathbf{y}$ interpreted as vectors. We can draw them as arrows emanating from a common origin, each generally having different lengths and pointing in separate directions. The heads of these two vectors (represented by arrow heads in the drawing) may be connected by another arrow from $\mathbf{x}$ to $\mathbf{y}$. This new arrow depicts $\mathbf{y} - \mathbf{x}$, the vector which must be added to $\mathbf{x}$ to obtain $\mathbf{y}$. Conversely, the arrow in the same position but with the opposite orientation depicts $\mathbf{x} - \mathbf{y}$.

The length of $\mathbf{x} - \mathbf{y}$ and $\mathbf{y} - \mathbf{x}$ is the same, and is found using the Euclidean norm

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}. \quad (16)$$

Note that this formula is identical to the one that measures the distance between $\mathbf{x}$ and $\mathbf{y}$ regarded as points; in other words, the norm of a vector equals the length of the line segment which describes its path. Similar, the Euclidean norm $\|\mathbf{x}\|$ of vector $\mathbf{x}$ is the length of $\mathbf{x}$ measured from the origin to its head.

4.3 The Cosine of the Angle Between Two Vectors

The cosine of the angle θ between two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in n-dimensional Euclidean space is given by

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}. \tag{17}$$

The argument to verify this relationship relies on the assumption that the laws of trigonometry extend to vectors in n-dimensional Euclidean space. Under this assumption, it will always be possible to find a scalar multiple[3] of $\mathbf{v}$ such that the vector $\mathbf{u} - t\mathbf{v}$ is orthogonal (perpendicular) to $\mathbf{v}$.

The argument follows from imposing two conditions:

1. The definition of the cosine says that it equals the ratio of the lengths of the adjacent leg of a right triangle to the length of the hypotenuse:

$$\cos\theta = \frac{\|t\mathbf{v}\|}{\|\mathbf{u}\|}. \tag{18}$$

2. The Pythagorean theorem applies to yield

$$\|\mathbf{u}\|^2 = \|t\mathbf{v}\|^2 + \|\mathbf{u} - t\mathbf{v}\|^2. \tag{19}$$

Using $\|t\mathbf{v}\| = t\,\|\mathbf{v}\|$ and expanding the right term in the Pythagorean relationship leads to an expression for t. Substituting this into the expression for $\cos\theta$ leads to the result.

4.4 The Cauchy-Bunyakovsky-Schwarz (CBS) Inequality

Although many treatments refer to this result as the Schwarz inequality or the Cauchy-Schwarz inequality, Bunyakovsky had precedence in its discovery by twenty-five years; thus we will acknowledge his role[4].

[3]For the argument given in the book, which is incorporated here, it appears that we need to assume $0 \le \theta \le \pi/2$. The result seems to hold without this restriction, but it would take some extra work to consider the case $t < 0$.

[4]Perhaps there is a bias against four-syllable Russian names in the mathematics community?

The inequality says that for any vectors $\mathbf{x}$ and $\mathbf{y}$,

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\| \, . \tag{20}$$

This inequality follows immediately from the rule which defines the cosine of the angle between two vectors, by noting that $-1 \leq \cos\theta \leq 1$.

Exercises:

1. The sample correlation coefficient of elementary statistics is defined by

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{21}$$

   where $\{x_1, x_2, \ldots, x_n\}$ and $\{y_1, y_2, \ldots, y_n\}$ are statistical samples of observations on the random variables X and Y, and the sample means are defined by

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{22}$$

   and

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}. \tag{23}$$

   Recast this calculation in terms of inner product and the Euclidean norm, by first writing $\tilde{x}_i = x_i - \bar{x}$ and $\tilde{y}_i = y_i - \bar{y}$ and defining $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ to be the vectors with such entries[5].

2. Use the CBS inequality to show that the sample correlation coefficient[6] must lie on the range $-1 \leq r \leq 1$.

3. Based on this result, explain the following statements:

   (a) The sample correlation coefficient may be interpreted as the cosine of the angle between the vectors of deviations from the mean for the two samples.

   (b) The sample correlation will equal $r = 1$ if and only if $\tilde{\mathbf{x}}$ is a (nonzero) scalar multiple of $\tilde{\mathbf{y}}$. (What are the possible angles between $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ in this case?)

[5]In statistical terminology, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ are the vectors of **deviations** of the items in each sample about their respective sample means.

[6]Actually, the same argument applies to the theoretical correlation coefficient, once the relevant inner product concept is defined.

(c) The sample correlation will equal $r = 0$ if and only if the vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ are orthogonal (meaning that the angle between them is $\frac{\pi}{2}$).

4.5 The Triangle Inequality

The triangle inequality states that for any two vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n,

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|. \tag{24}$$

The simple intuition of the triangle inequality is that the length of one side of a triangle, $\|\mathbf{u} + \mathbf{v}\|$, is always less than the sum of the lengths of the other two sides, $\|\mathbf{u}\| + \|\mathbf{v}\|$. It plays an important role in the limit arguments presented in Chapter 12.

The easiest way to prove this is to

1. square both sides,
2. replace the norm on the left-hand side with a dot product,
3. expand the dot product, and
4. simplify.

Finally, one obtains a statement which may be written

$$\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \, \|\mathbf{v}\|} = \cos\theta \leq 1, \tag{25}$$

which is always true. Since all steps in the proof may be reversed, the result is verified.

5 Nonparametric Representation of Lines, Planes, and Hyperplanes

Recall from elementary analytical geometry that the standard form of a line in x-y space is given by

$$ax + by = c. \tag{26}$$

The line crosses the y-axis at the height $\frac{c}{b}$ (found by setting $x = 0$ and solving for y), while the slope is $m = -\frac{a}{b}$.

The three-dimensional analogue of the two-dimensional formula for a line describes a plane:

$$ax + by + cz = d. \tag{27}$$

In vector notation, a plane may be written $\mathbf{n} \cdot (\mathbf{x} - \mathbf{x_0}) = 0$, which says that the plane consists of all vectors with tails at $\mathbf{x_0}$ which are orthogonal to the normal vector, $\mathbf{n}$. To translate the standard linear equation form given above into vector notation[7], we could write $\mathbf{n} = (a, b, c)$, then take $\mathbf{x_0} = (0, 0, d/c)$.

The extension to n-dimensional Euclidean space is straightforward: The $(n-1)$-dimensional hyperplane has linear equation

$$c_1x_1 + c_2x_2 + \ldots + c_nx_n = d, \tag{28}$$

and the vector representation may be written in the same form as previously, $\mathbf{n} \cdot (\mathbf{x} - \mathbf{x_0}) = 0$, but now with $\mathbf{n} = (c_1, c_2, \ldots, c_n)$ and $\mathbf{x_0} = (0, 0, \ldots, 0, d/c_n)$.

The linear (nonparametric) equations for a line, plane, or hyperplane all may be thought of as constituting a system of one equation in n unknowns (where n depends on the dimensions of the space under consideration). This suggests that we might also use Gaussian elimination to analyze the solutions to these systems. We will take a further look at this idea when we cover the unit on linear independence.

6 Parametric Representations

Think of starting at the origin in R^3, then moving away from the origin by some displacement $\mathbf{x_0}$. We can regard the point at the head of of $\mathbf{x_0}$ as a new origin, which in this context will serve as a reference point for constructing either a line or a plane. We can extend this idea to R^n, for $n > 3$, by noting that $\mathbf{x_0}$ could be the reference point for construction of a line, a plane, or a hyperplane[8] in this higher dimension Euclidean space. This discussion sets the stage for describing the parametric equation of either a line or a plane in R^3.

---

[7] We can also apply this setup to the two-dimensional case, taking $\mathbf{n} = (a, b)$ and $\mathbf{x_0} = (0, c/b)$.

[8] We can regard a line or a plane as special cases of the class of hyperplanes which can be described in a Euclidean space of dimension greater than three.

6.1 Lines

A line through the point $\mathbf{x_0}$ is described by the **parametric representation**

$$\mathbf{x}(t) = \mathbf{x_0} + t\mathbf{v}, \tag{29}$$

where $\mathbf{x_0}$ is a displacement from the origin, t is an unrestricted scalar parameter, and $\mathbf{v}$ describes the direction of the line through $\mathbf{x_0}$. There is no restriction of this description of a line to two or three dimensional Euclidean spaces where they can be literally depicted, but it is insightful to consider these two cases.

If we take $\mathbf{v} = \mathbf{x_1} - \mathbf{x_0}$, we get the following representation of the line through $\mathbf{x_0}$ and $\mathbf{x_1}$:

$$\mathbf{x}(t) = (1 - t)\mathbf{x_0} + t\mathbf{x_1}. \tag{30}$$

Setting $t = 0$ shows that $\mathbf{x_0}$ is on the line, while setting $t = 1$ shows that $\mathbf{x_1}$ is on the line. Restricting $0 \leq t \leq 1$ results in restricting the parametric representation to the line segment joining $\mathbf{x_0}$ to $\mathbf{x_1}$.

6.2 Planes

The parametric equation of a plane through the point $\mathbf{x_0}$ is given by

$$\mathbf{x}(t, s) = \mathbf{x_0} + t\mathbf{v} + s\mathbf{w}, \tag{31}$$

where $\mathbf{x_0}$ is a displacement from the origin, t and s are unrestricted scalar parameters, and the vectors $\mathbf{v}$ and $\mathbf{w}$ do not point in either the same or opposite directions[9].

To describe the plane through the three points $\mathbf{x_0}, \mathbf{x_1}$ and $\mathbf{x_2}$, we can take $\mathbf{v} = \mathbf{x_1} - \mathbf{x_0}$ and $\mathbf{w} = \mathbf{x_2} - \mathbf{x_0}$. The parametric equation becomes

$$\mathbf{x}(t, s) = (1 - t - s)\mathbf{x_0} + t\mathbf{x_1} + s\mathbf{x_2}. \tag{32}$$

Considering the restrictions $s = t = 0$, $t = 1$ with $s = 0$, and $t = 0$ with $s = 1$ shows that the three specified points all lie in the plane.

[9] I.e., we require that it is not possible to write $\mathbf{v} = c\mathbf{w}$ for any scalar c.

6.3 Hyperplanes

The parametric representation of a line or a plane extends naturally to the parametric representation of a hyperplane in a higher dimension Euclidean space. For illustration, consider the situation in four dimensions. The vectors look like $\mathbf{v} = (v_1, v_2, v_3, v_4)$, $\mathbf{w} = (w_1, w_2, w_3, w_4)$, etc. The parametric representation for a hyperplane in R^4 is

$$\mathbf{x}(t, s, r) = \mathbf{x_0} + t\mathbf{v} + s\mathbf{w} + r\mathbf{u}, \tag{33}$$

where $\mathbf{x_0}$ is once again a displacement from the origin, t, s and r are unrestricted scalar parameters, and the vectors $\mathbf{v}, \mathbf{w}$ and $\mathbf{u}$ are **linearly independent** vectors[10].

We can explore this definition by considering a few special cases.

1. Restricting both s and r to zero would limit us to a line in four dimensional Euclidean space.

2. Relaxing the restriction to only $r = 0$ would result in the parametric equation of a plane in R^4.

3. We can think of moving from away from the restriction $r = 0$ to cases where $r \neq 0$ as projecting, in the direction of the vector $\mathbf{u}$, off points in the plane defined by the zero restriction.

4. If we restrict $\mathbf{x_0} = \mathbf{0}$, then we obtain the parametric equation of a hyperplane through the origin in R^4.

[10]This means that it is only possible to have $b\mathbf{v}+c\mathbf{w}+d\mathbf{u} = \mathbf{0}$ for the case $b = c = d = 0$.