R.P. McAfee

September 1986

Economics 241

HANDOUT

3.1 VECTOR SPACES

A vector space is a set X, along with two operations on members of X.

Definition 3.1: X is a real vector space under $\oplus$ and $\odot$ if, $(\forall x \in X)(\forall y \in X)(\forall z \in X)(\forall \alpha \in R)(\forall \beta \in R)$ the following properties hold:

$$x \oplus y \in X \tag{3.1}$$

$$x \oplus y = y \oplus x \tag{3.2}$$

$$(x \oplus y) \oplus z = x \oplus (y \oplus z) \tag{3.3}$$

$$(\exists\ \bar{0} \in X)\ x \oplus \bar{0} = \bar{0} \tag{3.4}$$

$$(\forall\ x \in X)(\exists\ w \in X)\ x \oplus w = \bar{0} \tag{3.5}$$

$$\alpha \odot x \in X \tag{3.6}$$

$$\alpha \odot (x \oplus y) = (\alpha \odot x) \oplus (\alpha \odot y) \tag{3.7}$$

$$(\alpha\beta) \odot x = \alpha \odot (\beta \odot x) \tag{3.8}$$

$$(\alpha + \beta) \odot x = (\alpha \odot x) \oplus (\beta \odot x) \tag{3.9}$$

$$1 \odot x = x \tag{3.10}$$

$\oplus$ is referred to as addition of vectors, while $\odot$ is called scalar multiplication. Generally we shall write $\lambda \odot x$ as λx in what follows. The symbols $\oplus$ and $\odot$ are chosen to remind the reader that $\oplus$ and $\odot$ are abstract versions of addition and multiplication. Effectively, properties (3.1)-(3.10) are the only properties we'll need for addition and mulitiplication, at a very abstract (and hence very generally applicable) level.

EXAMPLE 3.1 $X = \{\bar{0}\}$

Define $\oplus$, $\odot$ by

$$\bar{0} \oplus \bar{0} = \bar{0}$$

$$\lambda \odot \bar{0} = \bar{0}$$

It is easily established that $\{\bar{0}\}$ under +, • is a vector space, easily the most trivial.

EXAMPLE 3.2 $X = R^n$, with

$$(x_1,\dots,x_n) + (y_1,\dots,y_n) = (x_1 + y_1,\dots,x_n + y_n)$$

$$\lambda(x_1,\dots,x_n) = (\lambda x_1,\dots,\lambda x_n)$$

$$\bar{0} = (0,\dots,0)$$

These are the usual definitions of + and • for vectors.

EXAMPLE 3.3 Let C[0,1] be the set of continuous functions $f: [0,1]\to R$.

For $f,g \in c\ [0,1]$,

$$(f \oplus g)(x) = f(x) + g(x)$$

$$(\lambda \odot f)(x) = \lambda f(x)$$

Finally, the zero function satisfies $\bar{0}(x) = 0$.

As these examples illustrate, X being a vector space requires defining the operations addition and scalar multiplication. In example 3.1, these took on unusual meanings. Basically, since $\bar{0}$ is the only element, we may satisfy all of the properties by defining every operation to yield $\bar{0}$.

Properties (3.1) and (3.6) require that addition of vectors (as elements of X will be called vectors) and multiplication of vectors by scalars results in vectors. (3.2) is commutativity: order is irrelevant. (3.3) is associativity: when addition is to be performed twice, we may perform either addition first. (3.4) requires the existence of a zero element. (3.5) provides subtraction. To see this, consider

$$z - x = y$$

to mean

$$z = y \oplus x.$$

Since there is a w satisfying

$$x \oplus w = \bar{0}$$

we have $y = z \oplus w$, since

$$(z \oplus w) \oplus x \overset{(3.3)}{=} z \oplus (w \oplus x) \overset{(3.5)}{=} z \oplus \bar{0} \overset{(3.4)}{=} z$$

Thus, (3.5), by forcing w to exist, allows us to subtract x by adding w. In addition

THEOREM 3.1: $0 \odot x = \bar{0}$

Proof: From (3.9) $0 \odot x = (0+0) \odot x \overset{(3.9)}{=} 0 \odot x \oplus 0 \odot x$

Now let y satisfy $0 \odot x \oplus y = \bar{0}$, which exists by (3.5). Then

$$\bar{0} = 0 \odot x \oplus y = (0 \odot x \oplus 0 \odot x) \oplus y \overset{(3.3)}{=} 0 \odot x \oplus (0 \odot x \oplus y) = 0 \odot x \oplus \bar{0} = 0 \odot x$$

as desired.

From this result, we see immediately that $(-1) \odot x$ is the w of (3.5), since

$$\bar{0} = 0 \odot x = (1+(-1)) \odot x = 1 \odot x \oplus (-1) \odot x \qquad (3.11)$$

Because of (3.11), we shall refer to $(-1) \odot x$ as $-x$, noting

$$x \oplus (-x) = \bar{0} \qquad (3.12)$$

Example 3.2 gives a general version of vectors in the plane, the case n=2. A vector $x = (x_1, x_2)$ is viewed as an arrow from the origin to the point x (see figure 3.1). Vector addition is accomplished by starting the vector y at the termination point of x, so that one has added the two arrows. Scalar multiplication merely extends the vector (see λx in figure 3.1).

The value of such an abstract definition of addition and scalar multiplication is illustrated by example 3.3. The space of continuous functions is not very similar to R^n, and yet both are vector spaces. Consequently, any properties we can prove about vector spaces in general apply equally to both examples. The operations of addition and scalar multiplication in C[0,1] are illustrated in figure 3.2.

3.2 Inner Product Spaces

Suppose $x = (x_1,\ldots,x_n)$ is the consumption bundle discussed in Section 1.?, and $p = (p_1,\ldots,p_n)$ are the prices of these goods. Then

$$y = \sum_{i=1}^{n} p_i x_i = p \cdot x$$

represents the expenditure on these goods. The notation $p \cdot x$ (which the reader should not confuse with scalar multiplication of the previous section) for $\sum_{i=1}^{n} p_i x_i$ is an example of a quantity called an inner product or dot product.

Definition 3.2: Let X be a vector space. A function $f\colon X \times X \to R$ is an inner product if $\forall\, x \in X \ \ \forall\, y \in X \ \ \forall\, \lambda \in R$

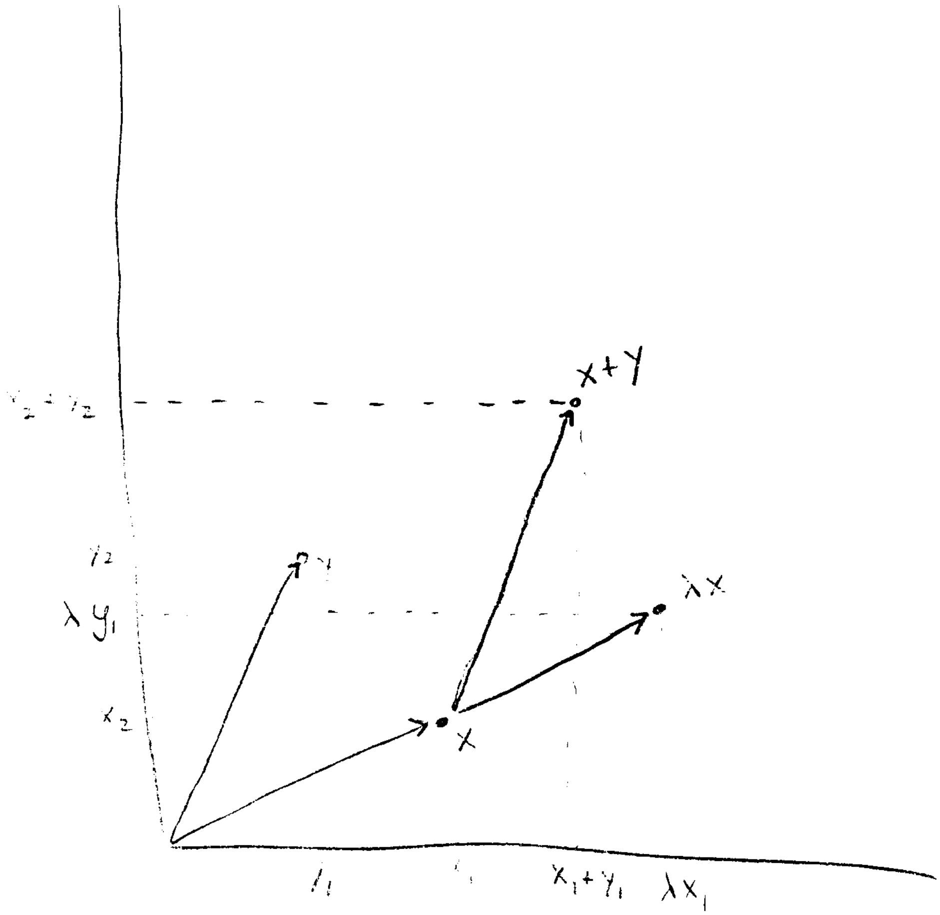

Figure 3.1 : The plane : Vector Addition and Scalar Multiplication.

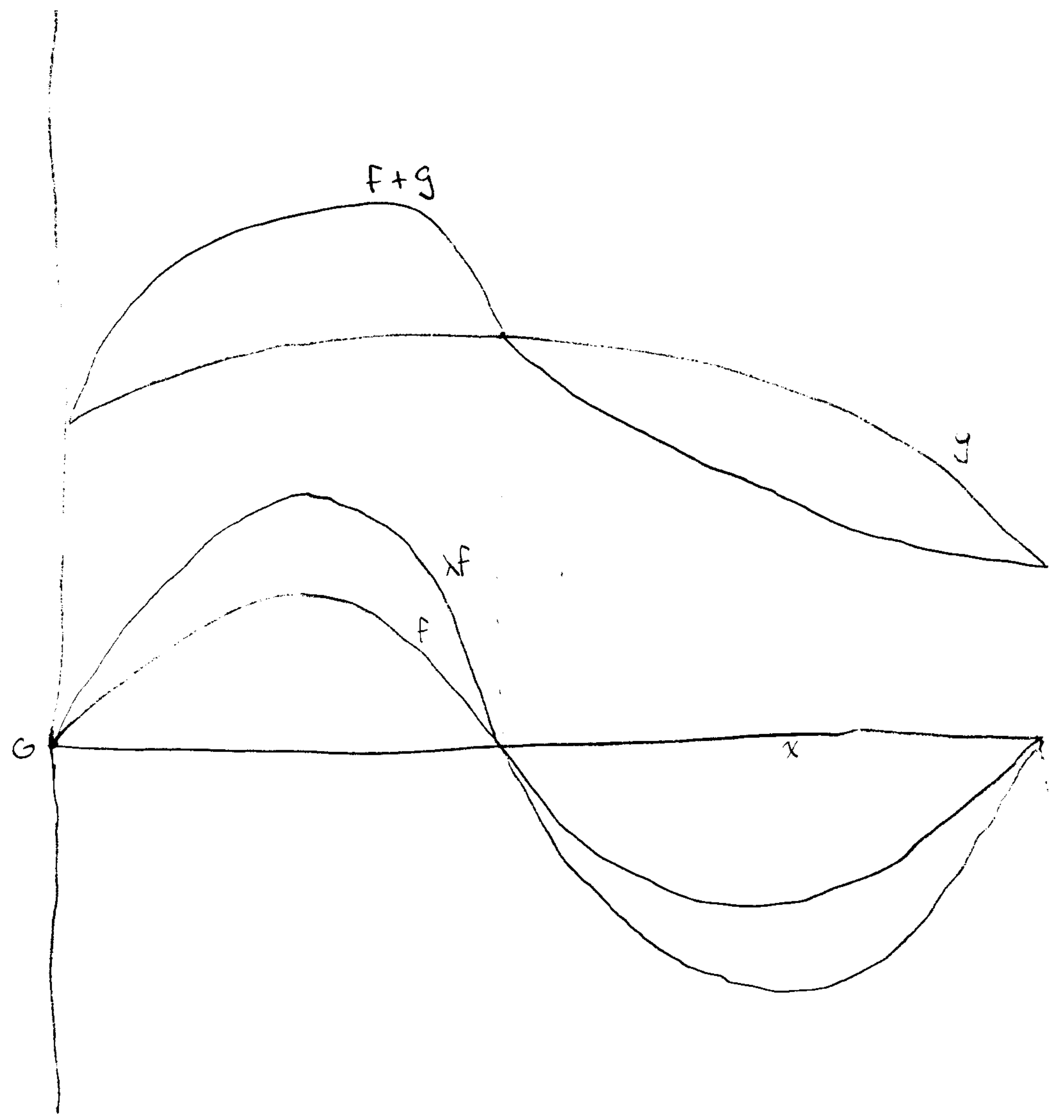

Figure 3.2 : Scalar multiplication (λf)
and addition ($f+g$) in $C[0,1]$.

$$f(x,x) \geq 0 \tag{3.12}$$

$$f(x,x) = 0 \quad \text{IFF } x = \bar{0} \tag{3.13}$$

$$f(\lambda x,y) = \lambda f(x,y) \tag{3.14}$$

$$f(x,y) = f(y,x) \tag{3.15}$$

$$f(x+y,\ z) = f(x,z) + f(y,z) \tag{3.16}$$

Generally, we shall denote inner products by $f(x,y) = \langle x,y\rangle$. This chapter will only concern inner product spaces, and we shall assume all vectors are members of inner product spaces.

EXAMPLE 3.4: For $X = R^n$ and vector addition and scalar multiplication as in example 3.2, and for $\alpha_1, \alpha_2, \ldots, \alpha_n$ real, positive numbers $\langle x,y\rangle = \sum_{i=1}^{n} \alpha_i x_i y_i$ is an inner product.

In particular, for $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 1$, the dot product $x \cdot y = \sum_{i=1}^{n} x_i y_i$ is an inner product.

EXAMPLE 3.5: For the example 3.3 vector space, and for any continuous function α: $[0,1] \to (0,\infty)$, $\langle f,g\rangle = \int_0^1 f(x)g(x)\alpha(x)dx$ is an inner product.

Inner products are a kind of vector multiplication (which is why they are called dot products on occasion), but the outcome of this kind of multiplication is a real number.

THEOREM 3.2 (Cauchy-Schwarz inequality):

$$\langle x,y\rangle \leq \sqrt{\langle x,x\rangle\langle y,y\rangle} \qquad (3.17)$$

Proof: Note (3.17) is true if $x = \bar{0}$ or $y = \bar{0}$. So suppose $x \neq \bar{0}$ and $y \neq \bar{0}$.

$$0 \overset{(3.12)}{\leq} \langle \alpha x - \beta y,\ \alpha x - \beta y\rangle \overset{(3.16)}{=} \langle \alpha x,\ \alpha x - \beta y\rangle + \langle -\beta y,\ \alpha x + \beta y\rangle \overset{(3.14)}{=}$$

$$\alpha\langle x,\alpha x - \beta u\rangle - \beta\langle y,\ \alpha x - \beta y\rangle \overset{(3.15)}{=} \alpha\langle \alpha x - \beta y,\ x\rangle - \beta\langle \alpha x - \beta y,\ y\rangle \overset{(3.16)}{=}$$

$$\alpha(\langle \alpha x,x\rangle + \langle -\beta y,\ x\rangle) - \beta(\langle \alpha x,\ y) + \langle -\beta y,\ y\rangle) \overset{(3.14)}{=}$$

$$\alpha^2\langle x,x\rangle - \alpha\beta\langle y,x\rangle - \beta\alpha\langle x,y\rangle + \beta^2\langle y,y\rangle \overset{(3.15)}{=}$$

$$\alpha^2\langle x,x\rangle - 2\alpha\beta\langle x,y\rangle + \beta^2\langle y,y\rangle.$$

Now let $\alpha = \sqrt{\langle y,y\rangle}$, $\beta = \sqrt{\langle x,x\rangle}$. Then

$$0 \leq \langle y,y\rangle\langle x,x\rangle - 2\sqrt{\langle x,x\rangle}\ \sqrt{\langle y,y\rangle}\ \langle x,y\rangle + \langle x,x\rangle\langle y,y\rangle$$

or, dividing by $2\sqrt{\langle x,x\rangle}\ \sqrt{\langle y,y\rangle}$,

$$0 \leq \sqrt{\langle x,x\rangle\langle y,y\rangle} - \langle x,y\rangle$$

Q.E.D.

COROLLARY 3.3: For scalars α, β:

$$\langle \alpha x + \beta y,\ \alpha x + \beta y\rangle = \alpha^2\langle x,x\rangle + 2\alpha\beta\langle x,y\rangle + \beta^2\langle y,y\rangle \qquad (3.18)$$

$$|\langle x,y\rangle| \leq \langle x,x\rangle\langle y,y\rangle \qquad (3.19)$$

The equation (3.18) is the first step in the proof of Theorem 3.2, while (3.19) follows from 3.18 with $-y$ substituted for y.

For R^n under the euclidean dot product:

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

The Cauchy Schwarz inequality reduces to

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) \qquad (3.20)$$

Define the norm of x by

$$\|x\| = \sqrt{\langle x,x\rangle} \qquad (3.21)$$

Norms may be thought of as length, since, for the plane under the euclidean dot product:

$$\|(x_1,x_2)\| = \sqrt{x_1^2 + x_2^2} \qquad (3.22)$$

which is the usual notion of the length of the line segment connecting (0,0) to (x_1,x_2) (see figure 3.3) by the pythagorean theorem.

Any norm, or notion of length, must satisfy three properties given in the following theorem.

THEOREM 3.4: $\|x\| = \sqrt{\langle x,x\rangle}$ satisfies

i). $x \neq \bar{0} \Rightarrow \|x\| > 0$ (3.23)

ii). $\forall \lambda \in R \ \|\lambda x\| = |\lambda|\|x\|$ (3.24)

iii). $\|x + y\| \leq \|x\| + \|y\|$ (3.25)

Proof: i). follows immediately from (3.12) and (3.13)

$$\|\lambda x\| \overset{(3.21)}{=} \sqrt{\langle \lambda x,\lambda x\rangle} \overset{(3.18)}{=} \sqrt{\lambda^2\langle x,x\rangle} = \sqrt{\lambda^2}\ \sqrt{\langle x,x\rangle} \overset{(3.21)}{=} |\lambda|\ \|x\|$$

Note the Cauchy Schwarz inequality can be expressed as:

$$\langle x,y\rangle \leq \|x\|\ \|y\| \qquad (3.26)$$

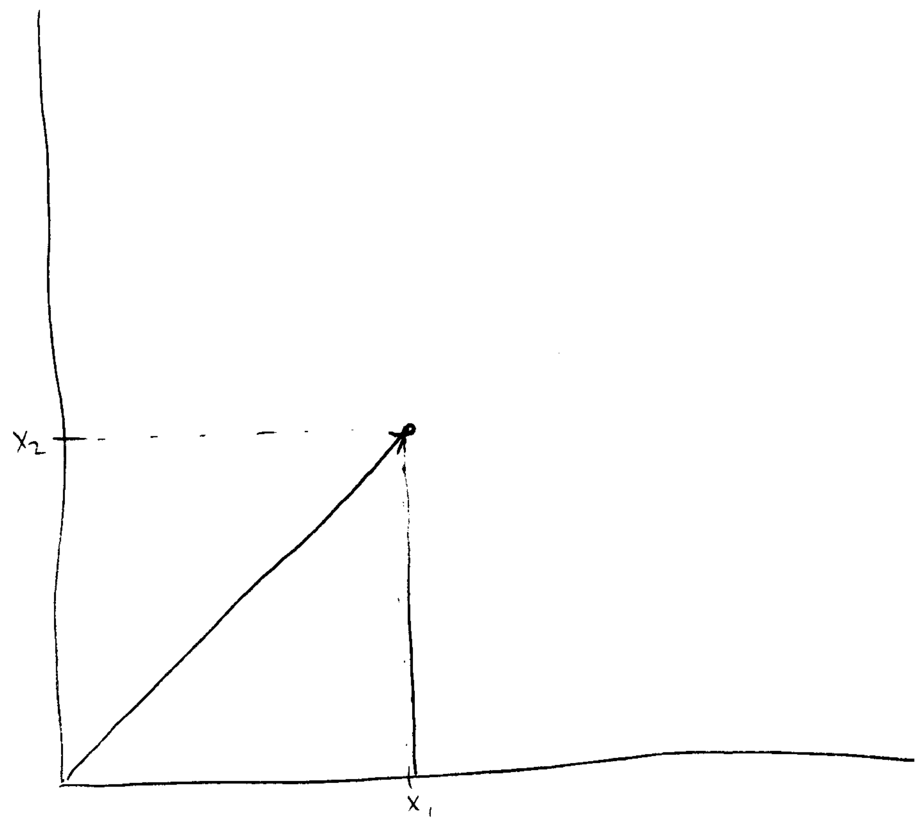

Figure 3.3 The length of the vector x is $\sqrt{x_1^2 + x_2^2}$.

forcing

$$\langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle \leq \|x\|^2 + 2\|x\|\ \|y\| + \|y\|^2$$

or

$$\langle x + y,\ x + y\rangle \leq (\|x\| + \|y\|)^2$$

or

$$\|x + y\| \leq \|x\| + \|y\|$$

Q.E.D.

Equation (3.23) states that the length of a vector x (or, equivalently, the distance from x to $\bar{0}$) is positive unless x = $\bar{0}$. (3.24) forces that, as we scale a vector up by a factor λ, the length of the vector increases accordingly. Finally, (3.25) says that the distance from $\bar{0}$ to x + y is not longer than the distance to x plus the distance to y. For this reason, (3.25) is sometimes called the triagle inequality, since it is shorter to go directly from zero to x + y rather than go by way of x (figure 3.1 illustrates this). Going to x + y by way of x requires travelling two sides of the triangle formed by $\bar{0}$, x and x + y.

When is it no extra distance to go by way of x? The next theorem answers this:

THEOREM 3.5: $\|x+y\| = \|x\| + \|y\|$ if and only if $\langle x.y\rangle = \sqrt{\langle x,x\rangle\langle y,y\rangle}$ if and only if $x = \bar{0}$ or $y = \lambda x$ for scalar $\lambda \geq 0$.

Proof: Examining the proof of Theorem 3.4, equation (3.26), we see that

$\langle x,y\rangle = \|x\|\ \|y\|$ iff

$\langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle = \|x\|^2 + 2\ \|x\|\ \|y\| + \|y\|^2$ iff

$\|x+y\| = \|x\| + \|y\|$.

This establishes the first equivalence.

Now, examining the proof of Theorem 3.2, we see that if $x=\bar{0}$ or $y=\bar{0}$ (the latter meaning $y = 0x$), the Cauchy-Schwarz inequality holds with equality. If $x \neq \bar{0}$ and $y \neq \bar{0}$, the only way the Cauchy Schwarz holds with equality is

$$0 = \langle \alpha x - \beta y,\ \alpha x - \beta y\rangle$$

which, by (3.13), requires

$$\bar{0} = \alpha x - \beta y$$

or $y = \frac{\alpha}{\beta}x$. Letting $\lambda = \frac{\alpha}{\beta} = \sqrt{\frac{\langle y,y\rangle}{\langle x,x\rangle}} \geq 0$, we are done.

Q.E.D.

Theorem 3.5 gives the nice result that the only time the distance from $\bar{0}$ to x+y equals the distance from $\bar{0}$ to x plus the length of y is when x is along the vector from $\bar{0}$ to x + y, i.e. $x + y = (1 + \lambda)x$.

LEMMA 3.6. If y and z satisfy $(\forall x)\langle x,y\rangle = 0 \Rightarrow \langle x,z\rangle = 0$

Then $\exists\lambda\in R\ \ z = \lambda y$

Proof: Case i): $y=\bar{0}$ Then, by (3.14), $(\forall x)\ \langle x,z\rangle = 0$. Thus, in particular, $\langle z,z\rangle = 0$, and by (3.13), $z = \bar{0} = 0y$.

Case ii). $y \neq \bar{0}$. Consider

$$x = z - \frac{\langle y,z\rangle}{\langle y,y\rangle} y$$

$$\langle x,y\rangle = \langle z,y\rangle - \frac{\langle y,z\rangle}{\langle y,y\rangle} \langle y,y\rangle = 0.$$

Thus

$$\langle x,z\rangle = \langle z,z\rangle - \frac{\langle y,z\rangle}{\langle y,y\rangle} \langle y,z\rangle = 0$$

or

$$\langle z,z\rangle\langle y,y\rangle - \langle y,z\rangle^2 = 0.$$

By theorem 3.5, $\exists\lambda\in R$

$z = \lambda y.$ Q.E.D.

In words, we have that, if everything perpendicular to y is also perpendicular to z, then y and z are parallel.

LEMMA 3.7: If y,z satisfy $(\forall x)\ \langle x,y\rangle \leq 0 \Rightarrow \langle x,z\rangle \leq 0$, then $\exists\lambda \geq 0$, $z = \lambda y$.

Proof: First, we show the hypothesis of Lemma 3.6 holds.

Suppose $\langle x,y\rangle = 0$. Then $\langle -x,y\rangle = -\langle x,y\rangle = 0$.

$\langle x,y\rangle = 0 \Rightarrow \langle x,y\rangle \leq 0 \Rightarrow \langle x,z\rangle \leq 0$

$\langle -x,y\rangle = 0 \Rightarrow \langle -x,y\rangle \leq 0 \Rightarrow -\langle x,z\rangle = \langle -x,z\rangle \leq 0$, or $\langle x,z\rangle \geq 0$. But this implies $\langle x,z\rangle = 0$, and Lemma 3.6 applies, yielding $z = \lambda y$.

We now show $\lambda \geq 0$.

If $y = \bar{0}$, then $(\forall x)\ \langle x,z\rangle = 0 \Rightarrow \langle z,z\rangle = 0 \Rightarrow z = \bar{0}$ and $\lambda = 0$ works.

If $y \neq \bar{0}$, $\langle -y,y\rangle = -\langle y,y\rangle < 0 \Rightarrow -\lambda\langle y,y\rangle = -\langle y,\lambda y\rangle = \langle -y,z\rangle \leq 0$

Thus $-\lambda \leq 0$, since $\langle y,y\rangle > 0$, or $\lambda \geq 0$ as desired. Q.E.D.

Theorem 3.8

$$\big|\|x\| - \|y\|\big| \leq \|x + y\| \tag{3.27}$$

$$\big|\|x\| - \|y\|\big| \leq \|x - y\| \tag{3.28}$$

Proof: Let $b = -y$ and $a = x + y$. Then

$\|x\| = \|a + b\| \leq \|a\| + \|b\| = \|x + y\| + \|-y\| = \|x + y\| + \|y\|.$

Thus

$\|x\| - \|y\| \leq \|x + y\|.$

Analogously,

$\|y\| - \|x\| \leq \|x + y\|$

or

$\|x\| - \|y\| = -(\|y\| - \|x\|) \geq -\|x + y\|.$

Thus

$-\|x + y\| \leq \|x\| - \|y\| \leq \|x + y\|$

or

$|\|x\| - \|y\|| \leq \|x + y\|$

(3.28) is (3.27), substituting -y for y. Q.E.D.

The final property of norms to be established is the parallelogram law:

THEOREM 3.9:

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad (3.29)$$

Proof: $\|x + y\|^2 + \|x - y\|^2 \overset{(3.21)}{=}$

$\langle x + y, x + y\rangle + \langle x - y, x - y\rangle \overset{(3.18)}{=}$

$\langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle + \langle x,x\rangle - 2\langle x,y\rangle + \langle y,y\rangle =$

$2(\langle x,x\rangle + \langle y,y\rangle) \overset{(3.21)}{=} 2(\|x\|^2 + \|y\|^2)$ Q.E.D.

The parallelogram law will be used in Chapter 5, and is illustrated in figure 3.4.

In this section, we have derived a number of properties of norms. The motivation for this is that the norm is the notion of length. The euclidean norm

$$\|(x_1, \ldots, x_n)\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

is exactly the usual notion of length arising (at least for n=2 or n=3) from

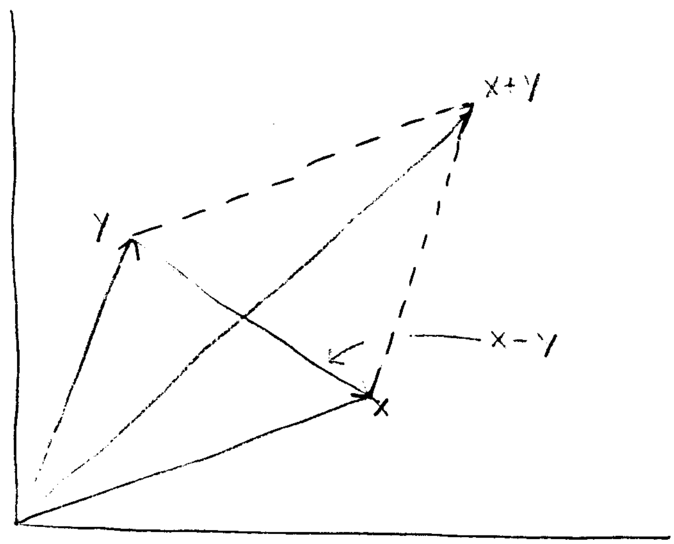

figure 3.4 : x, y, x+y and x - y

the pythagorean theorem. Because $\|x\|$ can be interpreted as the length of the vector from $\bar{0}$ to the point x, and hence the distance from $\bar{0}$ to x, generally we may consider $\|x-y\|$ as the distance from y to x. Thus, we have succeeded in defining the notion of distance between two points in any vector space. This is a significant gain in generality over distance in R^n, because, as we will see in later chapters, there are other vector spaces of interest to the economist.

One vector space of essential interest was described in example 3.5. Consider a firm which must choose it's level of investment I(t) for each point in time t. This firm, then, is choosing a function I: $R \to R$. Thus, if we are to make headway in solving any investment problem, we must have some knowledge of the properties of spaces of functions. Thus, it is of use to know continuous functions form a vector space under the norm

$$\|f\| = (\int (f(x))^2 dx)^{1/2}.$$

This will allow many of our optimization results of chapter 5 to extend beyond maximization in R^n.

At the same time, the reader is cautioned that it is very difficult to obtain a geometric intuition for these results, beyond their application in the plane. Thus, it is often useful to translate each theorem into the two dimensional special case. In the next section, we take up the geometry of vectors, inner products and norms.

3.3 The Geometry of Inner Product Spaces

Consider any inner product space X. For x, y $\in$ X, we consider the distance between x and y to be $\|x-y\|$. Thus, from (3.25)

$$\|x-z\| \leq \|x-y\| + \|y-z\| \tag{3.30}$$

That is, the distance from x to z is no more than the distance from x to y plus the distance from y to z. This is certainly required, since one way to get from x to z is to first go to y and then go from y to z.

The use of the symbol $\| \;\|$ is to remind the reader that norms work like absolute value $| \;|$. Thus, while we may be dealing with very complicated or abstract vector spaces, we can remember the properties of norms by their relationship to absolute value. In particular (3.23), (3.24) and (3.25) are all satisfied by absolute value, and this is the R^n case, for n=1.

Further intuition is gained by considering inner products to concern angles. Consider figure 3.5. The vectors x, y and x-y are arranged to form a triangle. From the well known pythagorean theorm, if the angle α is 90°, then the length of x squared plus the length of y squared equals the length of x-y squared, that is

$$\|x\|^2 + \|y\|^2 = \|x-y\|^2.$$

By definition, this means

$$\langle x,x\rangle + \langle y,y\rangle = \langle x-y,\ x-y\rangle,$$

or, by (3.18)

$$\langle x,x\rangle + \langle y,y\rangle = \langle x,x\rangle - 2\langle x,y\rangle + \langle y,y\rangle$$

or

$$0 = -2\langle x,y\rangle.$$

Thus, when $\langle x,y\rangle = 0$, the vectors x and y form a right angle. In this case, they are said to be orthogonal.

In a similar vein, if the angle $\alpha < 90^\circ$, then

$$\|x-y\|^2 < \|x\|^2 + \|y\|^2.$$

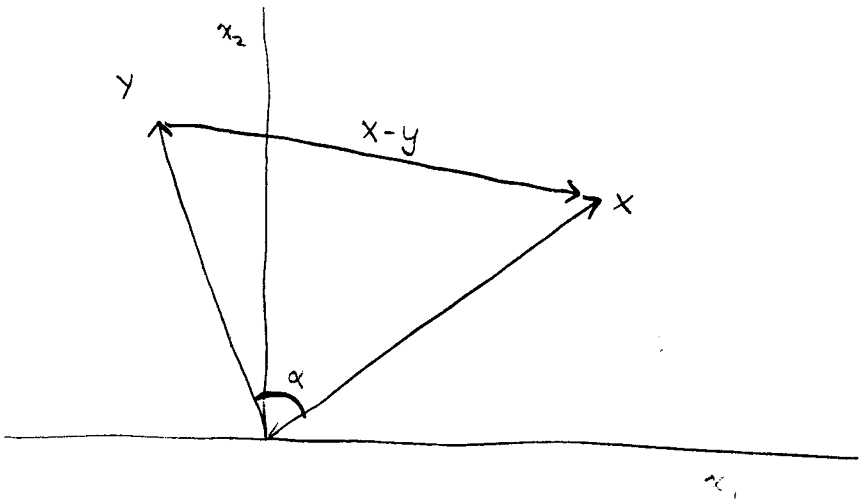

Figure 3.5

But this forces, as above:

$$0 > -2 \langle x,y \rangle$$

or

$$\langle x,y \rangle > 0.$$

Finally, since the case of $\alpha > 90^\circ$ is similar, we have proved

THEOREM 3.10: The angle formed by x and y equals 90° if $\langle x,y \rangle = 0$, exceeds 90° if $\langle x,y \rangle < 0$ and is less than 90° if $\langle x,y \rangle > 0$.

Inner products concern angles, in the sense that the sign of $\langle x,y \rangle$ (0, positive or negative) informs us of the type of angle (right, acute or obtuse). However, the units of $\langle x,y \rangle$ are denominated in areas, that is, $\langle x,y \rangle$ is telling us something precise about the area of the parallelogram formed by $\bar{0}$, x, x+y, y. This is shown in the following theorem.

THEOREM 3.11: The parallelogram formed by the points 0, x, x+y and y has an area equal to

$$P = \sqrt{\langle x,x \rangle \langle y,y \rangle - \langle x,y \rangle^2} \tag{3.31}$$

Proof: Consider a triangle with sides of length a,b,c. In general, the area of this triangle is:

$$A = 1/4 \sqrt{2(a^2b^2 + a^2c^2 + b^2c^2) - (a^4 + b^4 + c^4)} \tag{3.32}$$

This is easily established by dropping a perpendicular from a to the line segment connecting b and c, and using the pythagorean theorem to find the length of this perpendicular (whose length, times ½c, is the area of the triangle).

Now consider the triagle formed by $\bar{0}$, x and y, with sides of length $\|x\|$, $\|y\|$, and $\|x-y\|$. Substituting into (3.32), we obtain half of (3.31). However, the triangle is half of the parallelogram, by symmetry.

Q.E.D.

Theorem 3.9 demonstrates that the cauchy-schwarz inequality, suitably rewritten as

$$\sqrt{\langle x,x\rangle\langle y,y\rangle - \langle x,y\rangle^2} \geq 0 \qquad (3.33)$$

refers to an area, and is merely the expression that, suitably measured areas can not be negative. In addition, we see that $\langle x,y\rangle$ is effectively measuring the deviation of the $\bar{0}$, x, y, x+y parallelogram from a rectangle. As the parallelogram's angle at the origin becomes smaller, holding the lengths of the sides ($\|x\|$ and $\|y\|$) constant, its area decreases and $\langle x,y\rangle$ increases. This is illustrated in figure 3.6.

Thus, our analysis of vector spaces has taken two very ordinary notions, of parallelness and perpendicularity and found a general expression for these in terms of an inner product. To summarize, x and y are parallel if there is a scalar λ so that $y = \lambda x$, and this occurs when $\langle x,y\rangle = \|x\|\|y\|$. In addition, x and y are perpendicular if $\langle x,y\rangle = 0$. In this case, the pythagorean theorem holds, and $\|x\|^2 + \|y\|^2 = \|x-y\|^2$.

The alert reader may note that, in this section, the analysis went "backward," as we started with intuitive notions (right angles, perpendiculars, parallelness) and figured out what the inner product says about these notions. A more formal, but less intelligible, treatment would start with the inner product, since it was axiomatically developed independently of our understanding of the plane, and then show that the concepts so developed correspond to our understanding. Such a treatment begins with:

Definition 3.3: x and y are orthogonal if $\langle x,y\rangle = 0$.

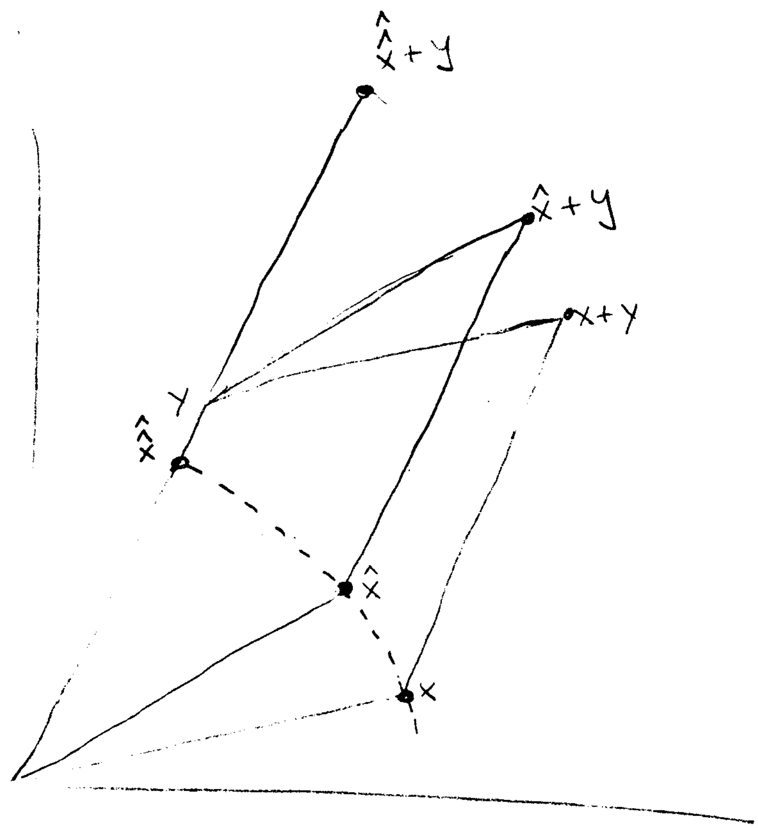

Figure 3.6: As x goes to $\hat{x}$, holding $||x|| = ||\hat{x}||$, the area of the parallelogram diminishes. It reaches zero when x is parallel to y, at $\hat{\hat{x}}$.

THEOREM 3.12 (Pythagorean Theorem): If x and y are orthogonal, $\|x\|^2 + \|y\|^2 = \|x-y\|^2$.

DEFINITION 3.4: x and y are parallel if $x = \bar{0}$ or $(\exists\lambda\in R)\ y = \lambda x$.

THEOREM 3.13: x and y are parallel if and only if $\langle x,y\rangle = \|x\|\ \|y\|$.

3.4 BASES

The study of bases concerns finding an economical means of expressing vectors in an inner product space.

DEFINITION 3.5 Let X be an inner product space. $B \subseteq X$ is a basis for X if every $x \in X$, there is a unique set of scalars α_v so that

$$x = \sum_{v\in B} \alpha_v v$$

DEFINITION 3.6: A set of vectors B is linearly independent if the only solution (for scalars α_v) to:

$$\sum_{v\in B} \alpha_v v = \bar{0}$$

is $(\forall v\in B)(\alpha_v = 0)$.

otherwise, B is linearly dependent.

EXAMPLE 3.6: Consider vectors in R^2, with the inner product of example 3.4. Then two vectors (x_1, x_2) and (y_1, y_2) are linearly independent if

$\alpha(x_1, x_2) + \beta(y_1, y_2) = (0,0)$

implies $\alpha = \beta = 0$.

This, in turn, is equivalent to

$$\alpha x_1 + \beta y_1 = 0$$

$$\alpha x_2 + \beta y_2 = 0.$$

Multiply the first by x_2 and the second by x_1 to obtain

$$\alpha x_1 x_2 + \beta y_1 x_2 = 0$$

$$\alpha x_1 x_2 + \beta x_1 y_2 = 0$$

and thus

$$\beta(y_1 x_2 - x_1 y_2) = 0.$$

Similarly, obtain

$$\alpha x_1 y_2 + \beta y_1 y_2 = 0$$

$$\alpha y_1 x_2 + \beta y_1 y_2 = 0$$

forcing

$$\alpha(x_1 y_2 - y_1 x_2) = 0.$$

Thus, either $\alpha = \beta = 0$ or $x_1 y_2 - y_1 x_2 = 0$. However, if $x_1 y_2 - y_1 x_2 = 0$, the vectors (x_1, x_2) and (y_1, y_2) are parallel.

Thus, in the plane, x and y are linearly independent if and only if they are not parallel.

THEOREM 3.14: If B is a basis for X, B is linearly independent.

Proof: By contrapositive. Suppose B is linearly dependent. Then there are scalars α_v, not all zero, satisfying

$$\sum_{v \in B} \alpha_v \cdot v = \bar{0}.$$

Since not all are zero, there is a $v_o \in B$ with $\alpha_{v_o} \neq 0$. Let:

$$\beta_v = \begin{cases} 0 & v \neq v_o \\ 1 & v = v_o \end{cases}$$

$$\gamma_v = \begin{cases} -\alpha_v & v \neq v_o \\ 0 & v = v_o \end{cases}$$

Then

$$v_o = \sum_{v\in B} \beta_v v = \sum_{v\in B} \gamma_v v,$$ which provides two expressions for v_o,

contradicting the uniqueness of expression in a basis. Q.E.D.

We shall return to the analysis of bases and linear independence in the next chapter, in terms of functions. However, we find one theorem relating inner products and bases of use.

Definition 3.7: B is an orthonormal basis for X if B is a basis of X and

$$\forall u,v\in B \ \langle u,v\rangle = \begin{matrix} 1 & u = v \\ 0 & u \neq v \end{matrix}$$

THEOREM 3.15 (Gram-Schmidt): Suppose X has a countable basis. Then X has an orthonormal basis.

Proof: Let $B = \{b_1, b_2, b_3, \dots\}$ be a basis for X.

Define

$$v_1 = b_1$$

and

$$v_n = b_n - \sum_{i=1}^{n-1} \frac{\langle b_n, v_i\rangle}{\langle v_i, v_i\rangle} v_i \qquad n=2,3,\dots$$

Finally, define $y_n = \dfrac{v_n}{\|v_n\|}$.

Note $\langle y_n, y_n\rangle = \left\langle \dfrac{v_n}{\|v_n\|}, \dfrac{v_n}{\|v_n\|}\right\rangle = \dfrac{\langle v_n, v_n\rangle}{\|v_n\|^2} = 1.$

Further, for $n \neq m$

$$\langle y_n, y_m \rangle = \frac{\langle v_n, v_m \rangle}{\|v_n\| \, \|v_m\|}$$

so we need only establish that $n \neq m \Rightarrow \langle v_n, v_m \rangle = 0$, and then show $\{y_1, y_2, \ldots\}$ is a basis.

Suppose, for an induction, that $\langle v_i, v_n \rangle = 0$ for all $i < n$. This is true, trivially, for $n = 1$. Then, for $i < n+1$

$$\langle v_{n+1}, v_i \rangle = \langle b_{n+1} - \sum_{j=1}^{n} \frac{\langle b_{n+1}, v_j \rangle}{\langle v_j, v_j \rangle} v_j, v_i \rangle$$

$$= \langle b_{n+1}, v_i \rangle - \frac{\langle b_{n+1}, v_i \rangle}{\langle v_i, v_i \rangle} \langle v_i, v_i \rangle - \sum_{j \neq i}^{n} \frac{\langle b_{n+1}, v_j \rangle}{\langle v_j, v_j \rangle} \langle v_j, v_i \rangle$$

$= 0.$

$\{y_1, y_2, \ldots\}$ is linearly independent. Consider

$$\sum \alpha_j y_j = \bar{0}.$$

Then

$$\alpha_i = \alpha_i \langle y_i, y_i \rangle = \sum \alpha_j \langle y_j, y_i \rangle = \langle \sum \alpha_j y_j, y_i \rangle =$$

$$\langle \bar{0}, y_i \rangle = 0.$$

Thus all α_i's are zero, and the set is linearly independent. Since $\{b_1, b_2, \ldots\}$ is a basis:

$$x = \sum_j \alpha_j b_j$$

$$= \sum_j \alpha_j (v_j + \sum_{i=1}^{n-1} \frac{<b_j, v_i>}{<v_i, v_i>} v_i)$$

$$= \sum_j (\alpha_j + \sum_{i>j} \alpha_i \frac{<b_i, v_j>}{<v_j, v_j>}) v_j$$

$$= \sum_j [(\alpha_j + \sum_{i>j} \alpha_i \frac{<b_i, v_j>}{<v_j, v_j>})(\|v_j\|)] y_j$$

$$= \sum_j \beta_j y_j .$$

Uniqueness of the expression of x follows from linear independence.

Q.E.D.

Thus, inner product spaces with a countable basis have the nice property that the vectors can be expressed as tuples. If $\{y_1, y_2, \ldots\}$ is an orthonormal basis for X, than any $x \in X$ can be written as

$$x = \sum \beta_j y_j$$

and thus, we may express x as $x = (\beta_1, \beta_2, \ldots)$. In addition, the inner product takes the nice form of

$$<(\alpha_1, \alpha_2, \ldots), (\beta_1, \beta_2, \ldots)> = \sum_i \alpha_i \beta_i .$$

Thus, there is another way of expressing the vectors in X so that the vectors, expressed in the new way, act like tuples of real numbers.

EXAMPLE 3.7: Constructing an orthonormal basis from (1,1,1), (1,-1,3) and (5,1,-3) for R^3.

$$v_1 = (1,1,1)$$

$$v_2 = (1,-1,3) - \frac{\langle 1,-1,3),\ (1,1,1)\rangle}{\langle (1,1,1),\ (1,1,1)\rangle}(1,1,1)$$

$$= (0,-2,2)$$

$$v_3 = (5,1,-3) - \frac{\langle (5,1,-3),\ (1,1,1)\rangle}{\langle (1,1,1),\ (1,1,1)\rangle}(1,1,1)$$

$$- \frac{\langle (5,1,-3),\ (0,-2,2)\rangle}{\langle (0,-2,2),\ (0,-2,2)\rangle}(0,-2,2)$$

$$= (4,-2,2).$$

Finally, our orthonormal basis is

$$y_1 = (\sqrt{3}/3,\ \sqrt{3}/3,\ \sqrt{3}/3)$$

$$y_2 = (0,\ -\sqrt{2}/2,\ \sqrt{2}/2)$$

$$y_3 = (\sqrt{6}/3,\ -\sqrt{6}/6,\ -\sqrt{6}/6)$$

The Gram Schmidt procedure, which constructed an orthonormal basis in the proof of Theorem 3.15, essentially tailors the basis for a given inner product, to make the inner product, expressed in terms of the new basis, look like the euclidean inner product. In addition, the norm, expressed in terms of the new basis, is the euclidean norm:

$$\|x\| = \sqrt{\sum x_i^2}$$

where $x = (x_1, x_2, \ldots)$ is the description of the point in terms of the orthonormal basis $\{y_1, y_2, \ldots\}$, that is, the point in question is $\sum x_i\ y_i$. Note that x_i is a scalar, the i^{th} component of x, while y_i is a vector, the i^{th} member of the basis.

For the remainder of this manuscript, we shall take $\langle x,y\rangle = \sum x_i y_i$, and write it using the conventional dot product notation:

$$x \cdot y = \sum x_i y_i.$$

For the case when the number of elements of a basis B for an inner product space X is finite, note that our expressions in the orthornormal basis are vectors in R^n. That is, in the orthonormal representation, $x = (x_1,\ldots,x_n) \in R^n$, and the inner product is the usual euclidean dot product. This essentially proves that all finite dimensional (i.e. having a finite basis) inner product spaces are copies of R^n, or, more strictly speaking, the properties of finite dimensional inner product spaces are the same as the properties of R^n. Formally, we say that finite dimensional inner product spaces are isomorphic to R^n. Because of the revealed importance of R^n under the euclidean inner product, we shall develop some properties of R^n in detail, in the next section and next chapter.

3.5 Sequences in R^n

In this section, we shall replicate the Theorems of section 2.2 for sequences of real vectors. To prevent confusion, we shall use superscripts to denote the order in the sequence, while subscripts shall denote the components of a vector. That is

$$x^i = (x_1^i, x_2^i,\ldots,x_n^i)$$

will represent the i^{th} member of a sequence of vectors. A sequence will be denoted $\{x^i\}_{i=1}^{\infty}$ or merely x^i.

Definition 3.8 x^i converges to x^o if $(\forall\ \varepsilon > 0)\ (\exists N\varepsilon \in \mathbb{N})\ i \geq N_\varepsilon \Rightarrow \|x^i - x^o\| < \varepsilon$.

Note that this definition is the same as Definition 2.6, except that the distance of x^i and x^o is now a norm instead of absolute value. As before, if x^i converges to x^o, we write

$$\lim_{i\to\infty} x^i = x^o$$

or $x^i \to x^o$.

If x^i converges to something, we say x^i converges.

Remarkably, a sequence converges if and only if every component converges.

THEOREM 3.16: $x^i \to x^o$ if and only if $\forall j\ x_j^i \to x_j^o$.

Proof: ($\Rightarrow$) Let $\varepsilon > 0$. Since $x^i \to x^o$, $\exists N$, $i \geq N \Rightarrow \|x^i - x^o\| < \varepsilon$.

But

$$|x_j^i - x_j^o| = \sqrt{(x_j^i - x_j^o)^2} \leq \sqrt{\textstyle\sum_{i=1}^n (x_j^i - x_j^o)^2} = \|x^i - x^o\| < \varepsilon \qquad (3.34)$$

Thus, if $i \geq N$, $|x_j^i - x_j^o| < \varepsilon$, and $x_j^i \to x_j^o$.

Since $x_j^i \to x_j^o$, $\exists N_j$, $i \geq N_j \Rightarrow |x_j^i - x_j^o| < \varepsilon/\sqrt{n}$.

But then, if $i \geq N = \max\{N_1, N_2, \ldots, N_n\}$,

$$\|x^i - x_o\| = \sqrt{\sum_{j=1}^n (x_j^i - x_j^o)^2} < \sqrt{\sum_{j=1}^n (\varepsilon/\sqrt{n})^2} = \varepsilon$$

Q.E.D.

DEFINITION 3.9: x^i is cauchy if $\forall\ \varepsilon > 0\ \exists N \in \mathbb{N}$, $i, k \geq N \Rightarrow \|x^i - x^k\| < \varepsilon$.

Again, this is Definition 2.7 in its vector clothes.

THEOREM 3.17: x^i converges if and only if x^i is cauchy.

Proof: x^i converges iff $\forall_j\ x_j^i$ converges iff $\forall_j\ x_j^i$ is cauchy. The first equivalence by Theorem 3.14, the second by Theorem 2.6. It remains to be shown that x^i is cauchy iff $\forall_j\ x_j^i$ is cauchy.

Suppose x^i is cauchy. Let $\varepsilon > 0$.

Then $\exists N$, $i,k \geq N \Rightarrow \|x^i - x^k\| < \varepsilon$.

Thus, by (3.34):

$$|x_j^i - x_j^k| \leq \|x^i - x^k\| < \varepsilon$$

and hence x_j^i is cauchy.

Suppose $\forall_j\ x_j^i$ is cauchy. Then $\exists N_j\ i,k \geq N_j \Rightarrow |x_j^i - x_j^k| < \varepsilon/\sqrt{n}$. Thus, if $i,k \geq N = \max\{N_1,\ldots,N_n\}$,

$$\|x^i - x^k\| = \sqrt{\sum_{j=1}^{n}(x_j^i - x_j^k)^2} < \sqrt{\sum_{j=1}^{n}(\varepsilon/\sqrt{n})^2} = \varepsilon$$

Q.E.D.

Theorems 3.14 and 3.15 denomstrate that much of what we learned about sequences of reals applies to sequences of real vectors. In particular, the sequence of vectors converges if and only if the components converge, and, if so, the j^{th} component of $x^o = \lim_{i\to\infty} x^i$ is $\lim_{i\to\infty} x_j^i$.

This puts us on comfortable, well explored ground, since we analyzed these sequences in chapter 2 in great detail. In figure 3.7, a sequence in R^2 is plotted, converging to x^o. As this illustrates, the set of x satisfying

$$\|x - x^o\| < \varepsilon$$

forms an open disk around x^o. In R^3, $\|x - x^o\| < \varepsilon$ creates a ball (or sphere) of x's, called an ε-ball or ε-neighborhood.

THEOREM 3.18: Suppose $x^i \to x^o$, $y^i \to y^o$ and $\lambda^i \to \lambda^o$, for scalars λ^i.

(i). $(x^i + y^i) \to (x^o + y^o)$

(ii) $\lambda^i(x^i) \to \lambda^o x^o$.

The proof is left as an exercise.

THEOREM 3.19: Suppose $x^i \to x_o$, $y^i \to x^o$ and $(\forall i \in \mathbb{N})\ (\forall j \leq n)\ x_j^i \leq z_j^i \leq y_j^i$.

Then $z^i \to z^o$.

Proof: Follows immediately from Theorems 2.10 and 3.16.

Q.E.D.

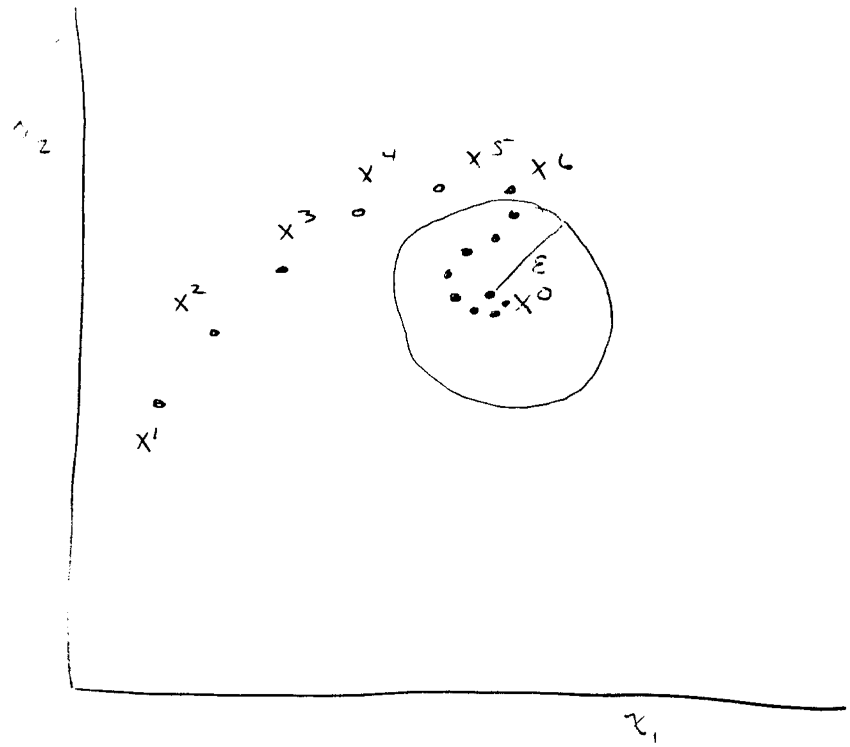

Figure 3.7 : For $x^i \to x^0$, all but finitely many points must lie within ε of x^0.

It is clear that the hard work concerning sequences was done in Chapter 2, and theorems 3.16-19 are obtained virtually for free. However, some mention of what goes wrong when the space is infinite dimensional is warranted. Consider vectors

$x = (x_1, x_2, \ldots), \ x_i \in R$

of infinite length, and define a norm

$\|x\|_\infty = \sup\{|x_i|\}.$

Now let $X = \{x = (x_1, x_2, \ldots) / \|x\|_\infty < \infty\}$. This is clearly a norm (i.e. it satisfies the conditions stated in Theorem 3.4) when

$(x_1, x_2, \ldots) + (y_1, y_2, \ldots) = (x_1 + y_1, x_2 + y_2, \ldots)$

$\lambda(x_1, x_2, \ldots) = (\lambda x_1, \lambda x_2, \ldots).$

Now consider the sequence

$x^i = (0, 0, \ldots, 0, 1, 1, 1, \ldots)$

so that the first i components of x^i are zero, and the remaining components are 1. Clearly $x^i_j \to 0$, since $\forall\ \varepsilon > 0$, if $i \geq j$ $|x^i_j - 0| = |0-0| = 0$.

However,

$\|x^i - \bar{0}\| = 1$ for all i.

That is, all of the components converge, but the sequence does not. Such problems in infinite dimensional spaces will not be discussed further in this manuscript.

EXERCISES

1. Show $\|x\|_\infty = \max\{|x_i| \mid i=1,\ldots,n\}$ is a norm for $x \in R^n$, that is, show

   i. $\|x\|_\infty = 0$ IFF $x = \bar{0}$

   ii. $\|\lambda x\|_\infty = |\lambda|\,\|x\|_\infty$

   iii. $\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty$.

   For n=2, find the "circle", that is, the points $\|(x_1,x_2)\|_\infty = \varepsilon$ for constant ε.

2. Show that R^n with the operations of example 3.2 is a vector space.

3. Show that C[0,1] in example 3.3 is a vector space.

4. Complete example 3.4, by showing $\sum_{i=1}^{n} \alpha_i x_i y_i$ is an inner product.

   What norm arises from this inner product? Draw its "circle" ($\|x\| = 1$) for n=2.

5. Complete example 3.5, by showing $\langle f,g\rangle$ is an inner product.

6. What does the cauchy-schwarz inequality say for the example 3.4 inner product?

7. Prove (3.32).

8. From definitions 3.1, 3.2 and 3.3 only, prove Theorem 3.12.

9. Prove Theorem 3.13

10. Show any three vectors in R^2 are linearly dependent.

11. Show any n linearly independent vectors in R^n form a basis for R^n.

12. Show that, if an inner product space X has a basis B with $n < \infty$ elements, then any set $A \subseteq X$ with more than n elements is linearly dependent.

13. Show that, if an inner product space X has a basis B with $n < \infty$ elements, then any set $A \subseteq X$ with strictly less elements than n elements is not a basis.

14. Show $\{x \mid \|x-x_o\| < \varepsilon\}$ is convex for every x_o and $\varepsilon > 0$. That is show $\|x-x_o\| < \varepsilon$ and $\|y-y_o\| < \varepsilon \Rightarrow \|\lambda x + (1-\lambda)y-x_o\| < \varepsilon$.

15. Consider the inner product in R^2

$<(x_1,x_2), (y_1,y_2)> = x_1y_1 + 3x_2y_2$ for $\alpha, \beta > 0$.

Consider the basis $\{(1,3), (1,0)\}$. Use the Gram Schmidt procedure to produce an orthonormal basis.

16. Consider the inner product in R^n

$<x,y>_\infty = \max_i x_i y_i$

the maximum of the product of the components of x and y.

i). Show this is an inner product.

ii). Draw the circules for $\|x\|_\infty = (<x,x>_\infty)^{1/2}$.

iii). What vectors are orthogonal in this inner product? Illustrate with a diagram.