I. Introduction

A. Euclidean space, which is where most economic analysis takes place, has lots of ‘features.’

1. Many of these features have been given axiomatic descriptions, which have in turn became the definitions of general structures of great importance, because they encompass many important applications.

2. Here we will emphasize the definitions and terminology, since this is an effective way to convey the basic facts concerning these structures.

B. An underlying of great importance in 20th century mathematics is the idea of developing the theory in a “coordinate free” way. In order to do explicit computations, one needs to impose some sort of coordinate system, but for many applications (especially physics) there are many equally good coordinate systems, and it is much simpler, when possible, to express things in a way that does not depend on which coordinate system is chosen.

II. Euclidean Space

A. Fix an integer $n \geq 1$. By definition $\mathbb{R}^n$ is the set of ordered $n$-tuples $x = (x_1, \ldots, x_n)$ of real numbers. The elements of $\mathbb{R}^n$ are called vectors, and the number $x_i$ is called the $i^{th}$ component of $x$.

B. The two fundamental operations on vectors are:
1. *addition*: for \( x, y \in \mathbb{R}^n \) define

\[
x + y := (x_1 + y_1, \ldots, x_n + y_n).
\]

2. *scalar multiplication*: for \( x \in \mathbb{R}^n \) and \( \alpha \in \mathbb{R} \), define

\[
\alpha x := (\alpha x_1, \ldots, \alpha x_n).
\]

C. We now talk about what kind of sets of vectors we can get using these operations.

1. For \( v_1, \ldots, v_k \in \mathbb{R}^n \), a *linear combination* is a sum of the form

\[
\alpha_1 v_1 + \ldots + \alpha_k v_k
\]

where \( \alpha_1, \ldots, \alpha_k \in \mathbb{R} \).

2. A set of vectors \( B \) is *linearly independent* if, whenever

\[
\alpha_1 v_1 + \ldots + \alpha_k v_k = 0
\]

where \( v_1, \ldots, v_k \in B \) and \( \alpha_1, \ldots, \alpha_k \in \mathbb{R} \), it is the case that

\[
\alpha_1 = \ldots = \alpha_k = 0.
\]

D. From an algorithmic point of view, the key idea of linear algebra is an algorithm testing for linear independence based on the following two ideas.

1. If the first component of \( v_1 \) is nonzero, and the first components of \( v_2, \ldots, v_k \) are all nonzero, then \( v_1, v_2, \ldots, v_k \) are linearly independent if and only if \( v_2, \ldots, v_k \) are linearly independent. (Of course it doesn’t have to be the first component.)

2. If \( \beta_2, \ldots, \beta_k \in \mathbb{R} \), then \( v_1, v_2, \ldots, v_k \) are linearly independent if and only if

\[
v_1, v_2 + \beta_2 v_1, \ldots, v_k + \beta_k v_1
\]

are linearly independent.
3. Algorithm: a) if \( v_1 = (0, \ldots, 0) \) return ‘Not linearly independent;’ b) otherwise find \( \beta_2, \ldots, \beta_k \) such that there is a component of \( v_1 \) that is nonzero and that component is zero for each of \( v_2 + \beta_2 v_1, \ldots, v_k + \beta_k v_1 \).

Test whether \( v_2 + \beta_2 v_1, \ldots, v_k + \beta_k v_1 \) are linearly independent.

III. The Inner Product

A. The inner product of two vectors \( x \) and \( y \) is \( x \cdot y := x_1 y_1 + \ldots + x_n y_n \).

B. It is also popular to write \( \langle x, y \rangle \) to denote the inner product, and we will do so now, while discussing the basic properties of the operation, but revert to \( x \cdot y \) later, at least most of the time.

C. The critical properties of the inner product are:

**Theorem:** For all vectors \( x, y, z \in \mathbb{R}^n \) and all scalars \( \alpha \in \mathbb{R} \):

(a) \( \langle x, y \rangle = \langle y, x \rangle \). (Symmetry)

(b) \( \langle x, \alpha y + z \rangle = \alpha \langle x, y \rangle + \langle x, z \rangle \). (Bilinearity)

(c) \( \langle x, x \rangle \geq 0 \), with strict inequality unless \( x = 0 \). (Positive definiteness)

**Proof:** (a) follows from the fact that multiplication is commutative, (b) is a matter of applying the distributive law and the commutativity of addition, and (c) follows from the fact that, for each \( i \), \( x_i^2 \geq 0 \), with strict inequality unless \( x_i = 0 \). ■

D. We say that \( x, y \in \mathbb{R}^n \) are perpendicular or orthogonal is \( \langle x, y \rangle \). Just why this definition matches up with more geometric intuitions concerning perpendicularity will emerge as we go along.

E. The following, very important, theorem is the Cauchy-Schwartz Inequality.

**Theorem:** For all vectors \( x, y \in \mathbb{R}^n \),

\[
\langle x, y \rangle^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle.
\]

**Proof:** For any scalar \( \alpha \), the positive definiteness, bilinearity, and symmetry of the inner
product imply that
\[ 0 \leq \langle \alpha x + y, \alpha x + y \rangle \]
\[ = \alpha^2 \langle x, x \rangle + \alpha \langle x, y \rangle + \alpha \langle y, x \rangle + \langle y, y \rangle \]
\[ = \alpha^2 \langle x, x \rangle + 2\alpha \langle x, y \rangle + \langle y, y \rangle. \]
If \( x = 0 \) then
\[ \langle x, y \rangle^2 = 0 = \langle x, x \rangle \cdot \langle y, y \rangle, \]
and otherwise we may set \( \alpha = -\langle x, y \rangle / \langle x, x \rangle \), obtaining
\[ 0 \leq \frac{\langle x, y \rangle^2}{\langle x, x \rangle} - 2 \frac{\langle x, y \rangle^2}{\langle x, x \rangle} + \langle y, y \rangle, \]
which simplifies to
\[ 0 \leq -\langle x, y \rangle^2 + \langle x, x \rangle \langle y, y \rangle, \]
as desired. ❄

E. The inner product expresses the concept of angle through the formula
\[ \cos \angle(x, y) = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}. \]

IV. The Norm
A. The norm of \( x \in \mathbb{R}^n \) is
\[ \|x\| := \langle x, x \rangle^{1/2} = (x_1^2 + \ldots + x_n^2)^{1/2}. \]
1. The geometric interpretation is that \( \|x\| \) is the distance from \( x \) to the origin.

**Theorem:** The norm has the following properties:
(a) \( \|x\| \geq 0 \), with strict inequality unless \( x = 0 \);
(b) \( \|\alpha x\| = |\alpha| \cdot \|x\| \) for all \( \alpha \in \mathbb{R} \);
(c) \( \|x + y\| \leq \|x\| + \|y\| \) for all \( x, y \in \mathbb{R}^n \).
Proof: Here (a) and (b) are simple, and (c) follows from the Cauchy-Schwartz inequality:
\[
\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\
\leq \langle x, x \rangle + 2\sqrt{\langle x, x \rangle \cdot \langle y, y \rangle} + \langle y, y \rangle \\
= \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \]

B. The *distance* between two vectors $x$ and $y$ in $\mathbb{R}^n$ is $d(x, y) = \|x - y\|$. Here are the key properties:

**Theorem:** For all $x, y, z \in \mathbb{R}^n$

(a) $d(x, y) \geq 0$, with strict inequality if $x \neq y$;

(b) $d(x, y) = d(y, x)$; (Symmetry)

(c) $d(x, z) \leq d(x, y) + d(y, z)$. (Triangle Inequality)

**Proof:** These are direct consequences of the corresponding properties of the norm.  

V. We now present axiomatic concepts that abstract the key properties identified above.

Except to whatever extent we study these in greater detail later (mostly we will stick to $\mathbb{R}^n$) you are not responsible for these concepts. This is background information whose purpose is to give you some sense of what is important.

**Definition:** A *vector space* is a set $V$ with an operation $+: V \times V \to V$ called *vector addition*, and an operation taking pairs $(\alpha, x) \in \mathbb{R} \times V$ to elements $\alpha x \in V$, called *scalar multiplication*, with the following properties:

(a) $(x + y) + z = x + (y + z)$ for all $x, y, z \in V$;

(b) there is a vector $0 \in V$ (called the *origin*) that is an identity element for addition: $x + 0 = x$ for all $x \in V$, and

(c) for any $x \in V$ there is $-x \in V$ such that $x + (-x) = 0$;

(d) $x + y = y + x$ for all $x, y \in V$;

(e) $1x = x$ for all $x \in V$;

(f) for all $x, y \in V$ and all $\alpha, \beta \in \mathbb{R}$:
(i) \((\alpha \beta)x = \alpha(\beta x)\);
(ii) \((\alpha + \beta)x = \alpha x + \beta x\);
(iii) \(\alpha(x + y) = \alpha x + \alpha y\).

Remark: Note that (a)-(e) amount to the assertion that \(V\) with the binary operation \(+\) is a commutative group.

A. In the next lecture we will see that every \(n\)-dimensional is “just like” \(\mathbb{R}^n\). (We will have to say precisely what this means.) From this point of view the virtue of approaching things abstractly is that we do not presuppose any commitment to a particular coordinate system. Reasoning that is expressed in terms of an abstract vector space gets our minds closer to “what is really going on.”

Definition: A inner product space is a vector space \(V\) with a binary operation
\[
\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}
\]
called the inner product such that for all vectors \(x, y, z \in V\) and all scalars \(\alpha \in \mathbb{R}\):
(a) \(\langle x, y \rangle = \langle y, x \rangle\). (Symmetry)
(b) \(\langle x, \alpha y + z \rangle = \alpha \langle x, y \rangle + \langle x, z \rangle\).
(c) \(\langle x, x \rangle \geq 0\), with strict inequality unless \(x = 0\).

B. Again, it will turn out that every \(n\)-dimensional inner product space is “just like” \(\mathbb{R}^n\) with the inner product defined earlier.

Definition: A normed space is a vector space \(V\) with a function \(\langle \cdot, \cdot \rangle : V \to [0, \infty)\) called the norm such that for all \(x, y \in V\)
(a) \(\|x\| \geq 0\), with strict inequality unless \(x = 0\);
(b) \(\|\alpha x\| = |\alpha| \cdot \|x\|\) for all \(\alpha \in \mathbb{R}\);
(c) \(\|x + y\| \leq \|x\| + \|y\|\).

C. Exercise: Prove that
\[
\|x\|_1 := |x_1| + \ldots + |x_n| \quad \text{and} \quad \|x\|_\infty := \max\{|x_1|, \ldots, |x_n|\}
\]
are norms on $\mathbb{R}^n$. Thus there are other norms that the one defined above, which is sometimes denoted by $\| \cdot \|_2$. For any number $p$ between 1 and $\infty$ there is a norm $\| \cdot \|_p$. Can you guess what its formula is? (For more information google “Minkowski’s inequality.”)

**Definition:** A *metric space* is a pair $(X, d)$ in which $X$ is a set and $d : X \times X \rightarrow \mathbb{R}$ is a function called the *metric* such that for all $x, y, z \in X$:

(a) $d(x, y) \geq 0$, with strict inequality if $x \neq y$;

(b) $d(x, y) = d(y, x)$;

(c) $d(x, z) \leq d(x, y) + d(y, z)$.

D. There are lots and lots of examples of metric spaces. They are important in mathematics because they are a framework in which the concepts of limit and continuity can be expressed in a general way, and these concepts are nicely behaved.

VI. Dimension

A. We now wish to define what it means for a vector space to be finite dimensional, and to define the dimension of a finite dimensional vector space. We begin with definitions we have seen before, now for general vector spaces.

1. For $v_1, \ldots, v_k \in V$, a *linear combination* is a sum of the form $\alpha_1 v_1 + \ldots + \alpha_k v_k$ where $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$.

2. A set of vectors $B$ is *linearly independent* if, whenever $\alpha_1 v_1 + \ldots + \alpha_k v_k = 0$ where $v_1, \ldots, v_k \in B$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, it is the case that $\alpha_1 = \ldots = \alpha_k = 0$.

3. The *span* of a set $S \subset V$ is the set of all linear combinations $\alpha_1 v_1 + \ldots + \alpha_k v_k$ where $v_1, \ldots, v_k \in S$.

4. A *basis* for $V$ is a subset $B \subset V$ that spans $V$ and is minimal for this property, meaning that there are no proper subsets of $B$ that span $V$.  

a. A basis is linearly independent, since if there was a linear dependence, one of the basis elements could be expressed as a linear combination of the others, and this would show that the subset obtained by removing this element also spanned $V$, contradicting minimality.

b. Conversely, a linearly independent set that spans $V$ is a basis, since for any element that we might remove, that element is not spanned by the remaining elements, so no proper subset spans $V$.

c. A very important point is that if $v_1, \ldots, v_m$ is a finite basis of $V$, then every point in $v \in V$ is representable as a linear combination

$$v = \alpha_1 v_1 + \ldots + \alpha_m v_m$$

in exactly one way. If there were multiple representations, equating them would yield a linear dependence of $v_1, \ldots, v_m$.

**Definition:** A vector space $V$ is *finite dimensional* if it has a finite basis. If $V$ is finite dimensional, its *dimension* is the number of elements of any basis.

The next result shows that this definition makes sense.

**Theorem:** If $V$ has a finite basis, then all bases have the same number of elements.

**Proof:** Let $v_1, \ldots, v_m$ and $w_1, \ldots, w_n$ be finite bases of $V$. Since we can swap these two bases, we can assume without loss of generality that $m \geq n$. We claim that, possibly after some reassignment of indices,

$$w_1, \ldots, w_k, v_{k+1}, \ldots, v_m$$

is a basis for $V$ for each $k = 0, 1, \ldots, n$. In particular, $w_1, \ldots, w_n, v_{n+1}, \ldots, v_m$ is a basis, but if $m > n$ this is impossible because $w_1, \ldots, w_n$ spans $v_{n+1}$. Therefore it suffices to establish the claim.
We argue by induction, noting that for $k = 0$ the claim is simply the assumption that $v_1, \ldots, v_m$ is a basis. Suppose that, for some $k < n$, we have already shown that $w_1, \ldots, w_k, v_{k+1}, \ldots, v_m$ is a basis. Then

$$w_{k+1} = \alpha_1 w_1 + \ldots + \alpha_k w_k + \beta_{k+1} v_{k+1} + \ldots + \beta_n v_n$$

for some numbers $\alpha_1, \ldots, \alpha_k, \beta_{k+1}, \ldots, \beta_n$. It cannot be the case that $\beta_{k+1} = \ldots = \beta_n = 0$, since then we would have a violation of the linear independence of $w_1, \ldots, w_k, w_{k+1}$. Therefore (possibly after rearrangement of indices) we may assume that $\beta_{k+1} \neq 0$.

Suppose that

$$0 = \gamma_1 w_1 + \cdots + \gamma_k w_k + \gamma_{k+1} w_{k+1} + \gamma_{k+2} v_{k+2} + \cdots + \gamma_m v_m.$$ 

Substituting the equation above gives

$$0 = \left( \gamma_1 + \alpha_1 \gamma_{k+1} \right) w_1 + \cdots + \left( \gamma_k + \alpha_k \gamma_{k+1} \right) w_k + \gamma_{k+1} \beta_{k+1} v_{k+1}$$

$$+ \left( \gamma_{k+2} + \beta_{k+2} \gamma_{k+1} \right) v_{k+2} + \cdots + \left( \gamma_m + \alpha_m \gamma_{k+1} \right) v_m.$$ 

Since $b_{k+1} \neq 0$, linear independence of $w_1, \ldots, w_k, v_{k+1}, \ldots, v_m$ implies that $\gamma_{k+1} = 0$, after which linear independence also implies that $0 = \gamma_1 = \cdots \gamma_k = \gamma_{k+2} = \cdots = \gamma_m$. Thus $w_1, \ldots, w_k, w_{k+1}, v_{k+2}, \ldots, v_m$ are linearly independent.

Since $\beta_{k+1} \neq 0$, the equation above gives

$$v_{k+1} = -\frac{\alpha_1}{\beta_{k+1}} w_1 - \cdots - \frac{\alpha_k}{\beta_{k+1}} w_k + \frac{1}{\beta_{k+1}} w_{k+1} - \frac{\beta_{k+2}}{\beta_{k+1}} v_{k+2} - \cdots - \frac{\beta_m}{\beta_{k+1}} v_m.$$ 

Substituting this into an equation expressing an element of $V$ as a linear combination of $w_1, \ldots, w_k, v_{k+1}, \ldots, v_m$ gives an equation expressing that element as a linear combination of $w_1, \ldots, w_k, w_{k+1}, v_{k+2}, \ldots, v_m$. Therefore $w_1, \ldots, w_k, w_{k+1}, v_{k+2}, \ldots, v_m$ span $V$. ■