Module 1: Linear spaces and linear maps

"The introduction of numbers as coordinates is an act of violence" - Hermann Weyl

1 Introduction

Economics is a multivariate affair – trade-offs exist only when there are two or more outcomes to entertain. Linear (or vector) spaces comprise the simplest environment capable of handling multivariate, i.e. multi-dimensional problems. In fact, linear spaces are the *only* environments in which equations generically admit closed-form solutions. Indeed, almost the entire "analysis" playbook involves reduction to some type of local linear approximation so that the powerful tools of linear algebra may be applied.

While it is common to begin a discussion of vector spaces by referring to the canonical, finite-dimensional case, i.e. $\mathbb{R}^n$ or $\mathbb{C}^n$, I intend to avoid this. Many of the important concepts related to vector spaces have nothing to do with dimension, finite or otherwise, and many of the vector spaces encountered in economics are necessarily infinite-dimensional. Furthermore, many of the initial concepts are unrelated to coordinate systems (i.e. choice of basis). For this reason, I prefer to begin the conceptual discussion without reference to dimension, bases, or coordinates (hence the quote at the top of the page), and instead introduce the simple, abstract setting of a linear space.

2 Vector spaces and direct sums

A (real) *vector space* is a set V, closed under a commutative binary operation that is referred to as *addition* and indicated with the symbol "$+$", and closed under *scalar multiplication* by elements of $\mathbb{R}$, indicated by juxtaposition. These operations are related by the distributive property:

$$\alpha \in \mathbb{R}, v, w \in V \implies \alpha(v+w) = \alpha v + \alpha w.$$

Finally, there is a special element of V, called the *additive identity* and usually denoted by 0, such that $v \in V$ implies $v + 0 = v$. A *complex vector space* is analogously defined. Our vector spaces, sadly, will be real unless otherwise noted.

Exercise 1 *Is $\mathbb{R}$ a vector space?*

Exercise 2 *Define $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ (i.e. the Cartesian productive of $\mathbb{R}$ with itself) and $\mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{n-1}$ for $n > 2$. Show $\mathbb{R}^n$ is a vector space under coordinate-wise operations.*

Exercise 3 *Let Ω be a set and define $\mathscr{F}(\Omega)$ to be the set of all functions from Ω to $\mathbb{R}$. Show that $\mathscr{F}(\Omega)$ is a vector space.*

Exercise 4 *Let Ω be any set with exactly two elements. Describe the vector space $\mathscr{F}(\Omega)$ conceptually.*

If V and W are vector spaces then denote by $V \oplus W$ the Cartesian product of V and W, together with "coordinate-wise" operations. Thus $x, y \in V \oplus W$ implies $x = (v_x, w_x)$ and $y = (v_y, w_y)$, for some $v_x, v_y \in V$ and $w_x, w_y \in W$; and, $x + y = (v_x + v_y, w_x + w_y)$, and analogously for scalar multiplication. It is straightforward to show that $V \oplus W$, which is called the *direct sum* of V and W, is a vector space. Observe that $\mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$.

Exercise 5 *Does $\mathbb{R}^2 \oplus \mathbb{R} = \mathbb{R} \oplus \mathbb{R}^2$? I wonder what "=" means in this context...*

2.1 Subspaces and spans

A subspace of a vector space V is a subset $W \subset V$ that is also a vector space. If V is a vector space and $A \subset V$ then the *span* of A, denoted $\operatorname{span}(A)$, is the set of all *finite linear combinations* of elements of A. More formally

$$\operatorname{span}(A) = \left\{ \sum_{k=1}^{m} \alpha_k a_k \text{ such that } m \in \mathbb{N}, \alpha_k \in \mathbb{R}, a_k \in A \right\}.$$

Intuitively, $\operatorname{span}(A)$ is the subspace of V generated by (i.e. containing all linear combinations of the information in) A: see exercise 7 below.

Exercise 6 *Show that $W \subset V$ is a subspace of V if and only if it is closed under addition and scalar multiplication, and it contains the additive identity.*

Exercise 7 *If $A \subset V$, show that $\operatorname{span}(A)$ is the smallest subspace of V containing A. Note: you'll need to define what "smallest" means.*

2.2 Linear independence and bases

A subset A of non-zero elements of a vector space V is *linearly independent* if no element of A can be written as a finite linear combination of other elements of A. More formally, A is linearly independent if whenever $\{a_1, \ldots, a_n\} \subset A$ and $\sum_{k=1}^{n} \alpha_k a_k = 0$ it follows that $\alpha_1, \ldots, \alpha_n = 0$. A subset A is a *basis* for V provided it is linearly independent and $\operatorname{span}(A) = V$. The following exercise asks you to show that the basis representation of a vector is unique.

Exercise 8 *Recall that a permutation of a set is a reordering. More formally, and in the context relevant here, let $N_n = \{1,\dots,n\} \subset \mathbb{N}$. A permutation of N_n is a function $\tau : N_n \to N_n$ such that $\tau(i) = \tau(j) \Leftrightarrow i = j$. Now show that if $\mathscr{B}$ is a basis for V and if $\sum_{i=1}^{n} \alpha_i a_i = \sum_{j=1}^{m} \beta_j b_j$ for some $a_i, b_j \in \mathscr{B}$ then $n = m$ and there is a permutation τ so that $a_i = b_{\tau(j)}$ and $\alpha_i = \beta_{\tau(j)}$.*

It is common for an area of mathematics to have one central result – a fundamental theorem – from which most other results follow. For example, the fundamental theorem of algebra was discussed in the previous chapter, and you are familiar with the fundamental theorem of calculus, which relates relates global averages (integrals) to local behavior (derivatives). Often, the fundamental theorem is deep, in that its proof requires more than manipulations and the observations of simple connections; but, with the fundamental theorem in hand, most other important results in the area follow from manipulations and connections *and* an appeal to the fundamental theorem. Finally, I should note that the fundamental theorem of an area may not be unique: it could be, say, *"result A"* from which all results, including, say, *"result B"* follow relatively easily, or it could be *"result B"* from which all results, including *"result A"* follow relatively easily. The point is this: you have to do some hard work somewhere – it may not matter so much where – and after that it's mostly bookkeeping.

I think of the following result as the fundamental theorem of *finite-dimensional* vector spaces. Of course we haven't defined *dimension* yet, but we will get to that.

Theorem 1 *If $A = \{a_1,\dots,a_n\} \subset V$ is a basis for V and $B = \{b_1,\dots,b_m\} \subset V$ is linearly independent then $m \le n$.*

Proof (sketch). The proof is tedious and I won't go through the details, but I will sketch the argument, which is by contradiction. Suppose $m > n$. Let $\hat{B} = \{b_1,\dots,b_n\}$, and observe that $\text{span}(\hat{B}) \subset V = \text{span}(A)$. Next, show that $A \subset \text{span}(\hat{B})$. This is the tedious part. The idea is to write each b_i as a linear combination of elements of A, i.e.

$$b_i = \sum_{j=1}^{n} \beta_{ij} a_j, \text{ for } i = 1,\dots n, \tag{1}$$

and then use the linearly independence of A to solve for the a_j. This shows that we may write

$$a_j = \sum_{i=1}^{n} \alpha_{ji} b_i, \text{ for } j = 1,\dots n, \tag{2}$$

whence $A \subset \text{span}(\hat{B})$.[1] But this means $V = \text{span}(A) \subset \text{span}(\hat{B})$, and thus $V = \text{span}(\hat{B})$. It follows that $b_{n+1} \in \text{span}(\hat{B})$ which contradicts the linear independence of B, and completes the sketch of the proof. ■

[1] The notation used here is intended to suggest the use of matrices – we will come back to these equations in the next chapter.

An immediate consequence of Theorem 1 is that if a vector space V has a basis with n elements, then *every* basis has n elements. In this case we say that the *dimension* of V is n, and write $\dim(V) = n$. If V does not have a finite basis then we write $\dim(V) = \infty$ – many of the important vector spaces in economics are infinite-dimensional.

Here, and in the sequel, when acting on sets, we use $|\cdot|$ to indicate cardinality. Thus if A is a finite set then $|A|$ is the number of elements. If A is not a finite set we will write $|A| = \infty$, though the notion of infinite cardinals can be greatly refined.[2]

Exercise 9 *Assume V is a vector space with dimension n. Establish the following consequences of Theorem 1.*

1. *If $B \subset V$ is linearly independent then B is a basis for* $\operatorname{span}(B)$ *and* $\dim(\operatorname{span}(B)) = |B|$.
2. *If $B \subset V$ is linearly independent and $|B| = m < n$ then there exists $C \subset V$ so that $B \cup C$ is a basis for V.*
3. *If W is a subspace of V then* $\dim(W) \leq \dim(V)$. *Be careful here. In particular, how do you know W has a finite basis?*

Item 2 of this exercise is very important: is says that if you start with a linearly independent set, you can extend that set to be a basis for the whole space. More generally, Theorem 1 and its corollaries indicate that, at least in the finite-dimensional case, it doesn't really matter what basis you choose. On the other hand, for applied purposes, the choice of basis can be crucial, as we will see.

Exercise 10 *Let Ω be any set with exactly n elements. Compute the dimension of the vector space $\mathscr{F}(\Omega)$. I wonder what infinite-dimensional vector spaces look like?*

3 Linear functions

The formal language mathematicians use to discuss areas of mathematics at an abstract level – sort of the math of math – is *category theory*. A category is a collection of *objects* and *arrows* (also called *morphisms*). Intuitively you can think of the objects as the items of interest and the arrows as the relations between these items. Commonly, the objects are sets with additional structure and the morphisms are functions that preserve that structure. For the case at hand, the objects are vector spaces and the morphisms are the functions between vector spaces that preserve their vector-space (i.e. linear) structure: linear maps.[3]

[2]For example, while it is certainly true that both $\mathbb{Q}$ and $\mathbb{R}$ are not finite sets, there is a very important sense in which $\mathbb{R}$ is much larger than $\mathbb{Q}$.

[3]The terms "map" and "function" are used interchangeably.

If V and W are vector spaces then a function $f : V \to W$ is *linear* provided that

$$f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2).$$

The map f is an *isomorphism* if it is invertible, that is, if there exists a linear map $g : W \to V$ so that $g(f(v)) = v$ and $f(g(w)) = w$ for all $v \in V$ and $w \in W$. In this case, we say that V and W are isomorphic, written $V \simeq W$, which means that, *as vectors spaces,* there is nothing interesting to distinguish them. When it exists, the inverse of f is often denoted f^{-1}, though this notation can be confusing.

Exercise 11 *Some practice exercises:*

1. *Show that if $f : V \to W$ is a linear map then $f(0) = 0$.*
2. *Suppose $V = \mathbb{R}$ and $W = \mathbb{R}$. Show that $v \to 1 + v$ is not a linear map.*
3. *Let $V = \mathbb{R}$ and $W = \mathbb{R}$. Show that $f : V \to W$ is linear if and only if there is an $\alpha \in \mathbb{R}$ such that $f(v) = \alpha v$. (Hint: sufficiency is straightforward. For necessity, let your candidate α be $f(1)$.)*
4. *Suppose $V = \mathbb{R} \oplus \mathbb{R}$ and $W = \mathbb{R}$. Show $(v_1, v_2) \to v_1 + 2v_2$ is a linear map.*

If $f : V \to W$ is a linear map then the *kernel* (also called *nullspace*) and *range* of f are defined as follows:

$$\begin{aligned} \ker(f) &= \{v \in V : f(v) = 0_W\} \subset V \\ f(V) &= \{w \in W : \exists v \in V \text{ with } f(v) = w\} \subset W. \end{aligned}$$

Here I wrote 0_W to emphasize that I am referring to the additive identity in W. I will suppress this reference in the sequel cause it looks stupid. So the kernel of f is the collection of vectors that f sends to zero, and the range of f is the image of V in W under the action of f.

Exercise 12 *Let $f : V \to W$ be a linear map. Show that* $\ker(f)$ *is a subspace of V and $f(V)$ is a subspace of W.*

If $f : V \to W$ is a linear map then the *rank* of f is the dimension of the range, and the *nullity* of f is the dimension of the kernel.

Theorem 2 (rank-nullity theorem) *If V is finite dimensional and $f : V \to W$ is a linear map then the rank of f plus the nullity of f equals the dimension of V. Formally*

$$\dim(\ker(f)) + \dim(f(V)) = \dim(V).$$

Proof (sketch). Pick a basis B for $\ker(f)$ (how?). Find linearly independent $C \subset V$ so that $B \cup C$ is a basis for V. Show that $f(C)$ is a basis for $f(V)$. Note that this proof is fairly easy exactly because we already have Theorem 1, our fundamental theorem. ■

Exercise 13 *Some more practice exercises:*

1. *Suppose $V = \mathbb{R} \oplus \mathbb{R}$ and $W = \mathbb{R}$. If $f : V \to W$ is linear, what is the minimum value for the nullity of f?*
2. *Suppose $V = \mathbb{R} \oplus \mathbb{R}$ and $W = \mathbb{R}$. Define $f : V \to W$ by $(v_1, v_2) \to v_1 + 2v_2$. Find a basis for the kernel of f.*

The importance of this theorem is that it summarizes in a very simple way the structural information that is lost by a map f. Specifically, if $v_1, v_2 \in \ker(f)$ then $f(v_1) = f(v_2)$, i.e. f does not distinguish v_1 and v_2. In this sense, the information distinguishing v_1 and v_2 has been lost. Also, if $v_1 \in \ker(f)$ and $v_2 \notin \ker(f)$ then $f(v_2) = f(v_1 + v_2)$, i.e. f does not distinguish v_2 and $v_1 + v_2$. More generally, f partitions V into a collection of equivalence classes which present as affine translations of the kernel. More specifically, the associated equivalence relation is

$$v_1 \sim v_2 \Leftrightarrow f(v_1) = f(v_2), \text{ i.e. } v_1 - v_2 \in \ker(f).$$

In light of exercise 14 below, each equivalence class has the form $v + \ker(f)$ for some $v \in V$, i.e. an affine translation of the kernel. This partition can be viewed as a *coarsening* of V into components that can be distinguished by f.

Exercise 14 *Suppose $v_1 \sim v_2$. Show that, as sets, $v_1 + ker(f) = v_2 + ker(f)$.*

The powerful relationship between linear maps and bases is captured by the following observation: the behavior of linear map is entirely characterized by its behavior on a basis. More specifically, in the finite dimensional case, if $A = \{a_1, \ldots, a_n\}$ is a basis for V and $\{w_1, \ldots, w_n\}$ are n not-necessarily linearly independent, or even distinct vectors in W then we may define a linear map f from V to W by sending a_i to w_i, and extending linearly:

$$v = \sum_{k=1}^{n} \alpha_k a_k \to \sum_{k=1}^{n} \alpha_k w_k = f(v).$$

Note that this is possible by exercise 3. And conversely, if we know how a given linearly function f acts on the basis, i.e. we know $f(a_k)$, then we know how it acts everywhere. The following theorem exploits this relationship.

Theorem 3 *If $\dim(V) = \dim(W) = n$ then V is isomorphic to W.*

Proof (sketch). Let $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_n\}$ be bases for V and W respectively. Define $f : V \to W$ by $a_i \to b_i$ and extending, as in previous discussion. Define $g : W \to V$ analogously, and show they are inverses. ■

While this theorem is simple to prove, in light of all the results we already have, it is also stunning in its implication: there is, up to isomorphism, only one vector space of dimension n.

Exercise 15 *Show that the dimension of* $\mathbb{R}^n$ *is n.*

4 Dynamics and decompositions

A common application of linear maps will be to the analysis of dynamic systems, and a brief introduction is warranted here. Thus let V have dimension n and let $f : V \to V$ be linear. We think of f as imparting dynamics on V as follows: given any initial condition v_0, we may recursively define $v_{t+1} = f(v_t)$, where t is interpreted as time. In this way, given an initial condition, the map f traces a path, or an *orbit* in the vector space V.

If W is a subspace of V and if $f(W) \subset W$ then W is called an *invariant* subspace. This is a *very* important concept: to understand the behavior of a given linear map, we often decompose the space V into invariant subspaces, and study the map on these subspaces. We discuss this in detail in the next chapter. The main theorem is

Theorem 4 (Schur's decomposition) *Let V be finite dimensional and* $f : V \to V$ *be linear. Then there is a collection of invariant subspaces* $\{V_k\}_{k=1}^n$ *such that*

$$\dim(V_k) = k \text{ and } V_k \subset V_{k+1}. \tag{3}$$

The power of the theorem comes from the conditions (3): we can study the function f recursively, and each step in the recursion adds only one dimension i.e. one new basis element on which f must be examined.

While the decomposition above is entirely general, a considerably simple case very often presents itself. To discuss this case, we must first introduce eigenvalues. Geometrically, a scalar $\lambda \in \mathbb{R}$ is an *eigenvalue* of the linear map f provided there exists $v \in V$ such that $f(v) = \lambda v$, where v is called an associated *eigenvector*. Intuitively, if v is an eigenvector associated to λ then f scales v by λ, but doesn't change its "direction." The sad part is that eigenvalues can be complex even when the vector spaces are real. In the next chapter we will establish the existence of eigenvalues, and, in fact, characterize them as roots of a certain polynomial.

Exercise 16 *Show that the map $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ given by counterclockwise rotation through angle θ is linear. What do the eigenvectors look like? (Ask yourself: which vectors have directions that are unchanged by T_θ?)*

Exercise 17 *Show that the collection of all eigenvalues associated to λ form a subspace of V. This space is called an* eigenspace

It turns out that in most (though not all!) applications, the collection of eigenvectors will span V. In this case, a very useful decomposition presents itself. Denote by $V(\lambda)$ the eigenspace associated to λ, and assume the map f has m distinct eigenvalues, λ_i. Then we can decompose V into the m invariant subspaces $V(\lambda_i)$; further, unlike in the Schur decomposition, we can "change coordinates" so that these subspaces are unrelated (orthogonal – we will discuss this later) to each other. More formally, if the eigenvectors span V then they can be used to construct an isomorphism $\phi : V \to \oplus_{i=1}^m V(\lambda_i)$. This isomorphism can, in turn, be used to study f; more specifically, we can use the following diagram, and study $\hat{f}$ restricted to each subspace $V(\lambda_i)$. While I realize this all sounds quite abstract, in the next chapter, we make it entirely concrete.

$$
\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, "\phi"'] & V \\
\bigoplus_{i=1}^m V(\lambda_i) \arrow[r, "\hat{f}"'] & \bigoplus_{i=1}^m V(\lambda_i) \arrow[u, "\phi^{-1}"']
\end{tikzcd}
$$

Here, $\hat{f} \equiv \phi \circ f \circ \phi^{-1}$, i.e. $\hat{f}(\dot{}) = \phi\left(f\left(\phi^{-1}(\dot{})\right)\right)$