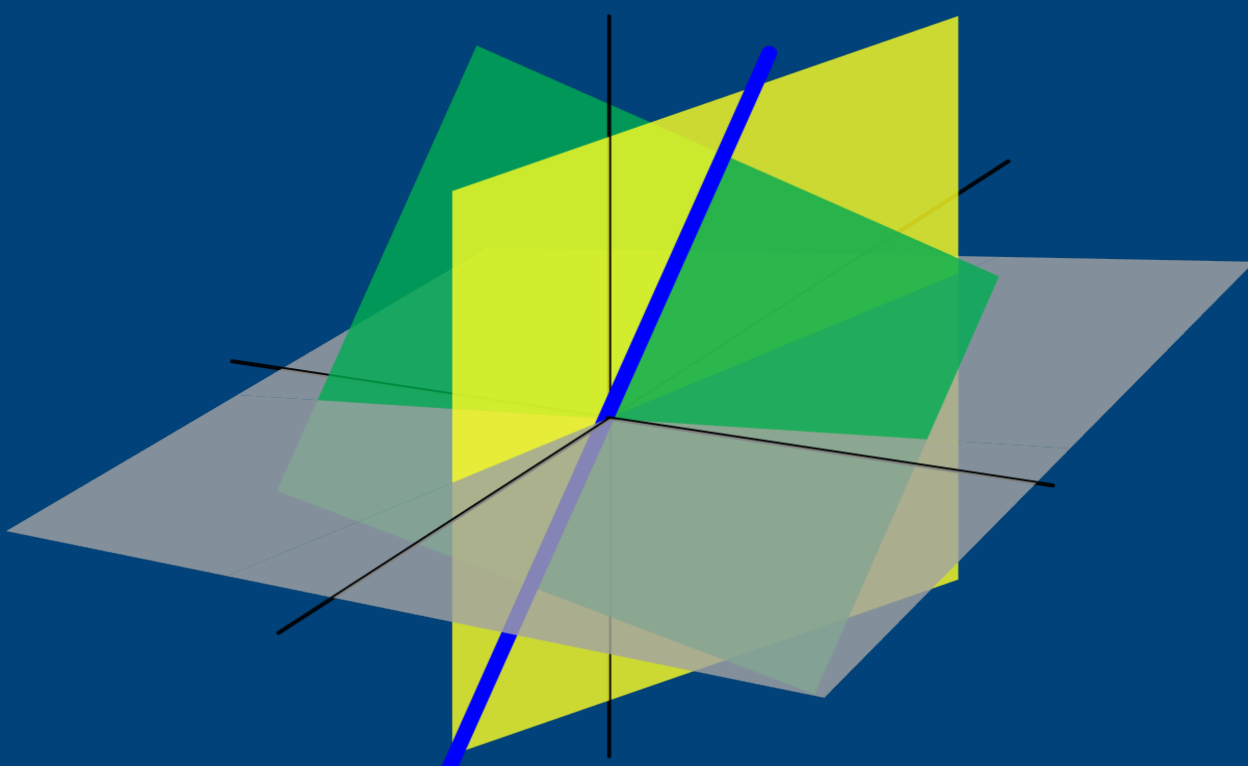
Linear Algebra Done Right

Lecture Notes

"It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment."

CARL FRIEDRICH GAUSS



Lectures by:
Sheldon Axler
Notes by:
Rohit Agarwal
rohaga@berkeley.edu

Contents

1 Vector Spaces **2**
1.1 $\mathbb{R}^n$ and $\mathbb{C}^n$ 2
1.2 Definition of a Vector Space 3
1.3 Subspaces 5

2 Finite-Dimensional Vector Spaces **7**
2.1 Span and Linear Independence 7

1 Vector Spaces

1.1 $\mathbb{R}^n$ and $\mathbb{C}^n$

Definition 1.1 $\mathbb{R}$ denotes the field of real numbers.

Some nonconstant polynomials with real coefficients have no real zeroes. Example: the equation:

$$x^2 + 1 = 0$$

has no real solutions. Thus, we invent a solution called i, such that $i^2 = -1$.

Definition 1.2 Complex Numbers

- A complex number is an ordered pair (a, b), where $a, b \in \mathbb{R}$, but we will write this as $a + bi$.
- The set of all complex numbers is denoted by $\mathbb{C}$:

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$$

- Addition and multiplication on $\mathbb{C}$ are defined as follows

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

Note 1.1 If $a \in \mathbb{R}$, we identify $a + 0i$ with the real number a. Thus we think of $\mathbb{R}$ as a subset of $\mathbb{C}$. We also usually write $0 + bi$ as just bi, and we usually write $0 + 1i$ as just i. From the definition of multiplication above, we have that $i^2 = -1$.

Note 1.2 Properties of Complex Arithmetic: $\forall \alpha, \beta \in \mathbb{C}$

- Commutativity

$$\alpha + \beta = \beta + \alpha \text{ and } \alpha\beta = \beta\alpha$$

- Associativity

$$(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda) \text{ and } (\alpha\beta)\lambda = \alpha(\beta\lambda)$$

- Identities

$$\lambda + 0 = \lambda \text{ and } \lambda 1 = \lambda$$

- Additive Inverse

$$\text{For every } \alpha \in \mathbb{C} \text{ there exists a unique } \beta \in \mathbb{C} \text{ such that } \alpha + \beta = 0$$

- Multiplicative Inverse

$$\text{For every } \alpha \in \mathbb{C} \setminus \{0\}, \text{ there exists a unique } \beta \in \mathbb{C} \text{ such that } \alpha\beta = 1$$

- Distributivity

$$\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$$

Definition 1.3 $\mathbb{F}$ denotes either $\mathbb{R}$ or $\mathbb{C}$

Elements of $\mathbb{F}$ are sometimes called scalars. We call it $\mathbb{F}$ because those are both fields.

Now we discuss the idea of a "list." To understand the idea, here some examples of simple sets we have already seen in other mathematics:

- The set $\mathbb{R}^2$, which you can think of as a plane, is the set of all ordered pairs of real numbers:

$$\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$$

- The set $\mathbb{R}^3$, which you can think of as ordinary space, is the set of all ordered triples of real numbers:

$$\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$$

Definition 1.4 A *list* of *length* n is an ordered collection of n numbers separated by commas and surrounded by parenthesis.

$$\text{i.e. } (x_1, \ldots, x_n)$$

Two lists are equal if an only if they have the same length and the same elements.

Here are some examples of lists from sets we are familiar with:

1. $(7, 3)$ is a list of length 2. Thus $(7, 3) \in \mathbb{R}^2$.
2. $(5, 9, -2)$ is a list of length 3. Thus $(5, 9, -2) \in \mathbb{R}^3$

Definition 1.5 $\mathbb{F}^n$ is the set of all lists of length n of elements of $\mathbb{F}$:

$$\mathbb{F}^n = \{(x_1, \ldots, x_n) : x_j \in \mathbb{F} \text{ for } j = 1, \ldots, n\}$$

Elements of $\mathbb{F}^n$ are often called *points* or *vectors*.

It does not matter if these sets have geometric sense. We can manipulate them algebraically. This is where the name linear algebra comes from.

Definition 1.6 Addition in $\mathbb{F}^n$ is defined by adding the corresponding coordinates:

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

Definition 1.7 The product of a number $\lambda \in \mathbb{F}$ and a vector $\mathbb{F}^n$ is defined by multiplying each coordinate of the vector by λ:

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

Single letters can denote elements of $\mathbb{F}^n$ efficiently. You can say $x + y = z$ instead of saying e.g.

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (z_1, \ldots, z_n)$$

Definition 1.8 Let 0 denote the list of length n whose coordinates are all 0:

$$0 = (0, \ldots, 0)$$

It should always be clear from context which 0 you're talking about. For example: we have the following:

Theorem If $x \in \mathbb{F}^n$, then $0x = 0$.

The 0 on the LHS is a scalar in $\mathbb{F}$. The 0 on the RHS is a vector in $\mathbb{F}^n$.

1.2 Definition of a Vector Space

The motivation for the definition of a vectors space comes from the properties of addition and scalar multiplication in $\mathbb{F}^n$:

- Addition is commutative, associative, and has an identity.
- Every element has an additive inverse.
- Scalar multiplication is associative.
- Scalar multiplication by 1 acts as expected.
- Addition and scalar multiplication are connected by distributive properties.

First, let us define what addition/scalar multiplication is.

Definition 1.9
- An *addition* on a set V is a function that assigns an element $u + w \in V$ to each pair of elements $u, w \in V$
- A *scalar multiplication* on a set V is a function that assigns an element $\lambda u \in V$ to each $\lambda \in \mathbb{F}$ and each $u \in V$

Example 1.1 Suppose V is the set of real valued functions on the interval $[0, 1]$. For $f, g \in V$ and $\lambda \in \mathbb{R}$, define $f + g$ and λf by:

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

and

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

Thus $f + g \in V$ and $\lambda f \in V$.

Now, we can define a vector space V. These are based off the properties of $\mathbb{F}^n$:

Definition 1.10 A *vector space* is a set V along with an addition on V and a scalar multiplication on V such that the following properties hold:

- $u + w = w + u$ for all $u, w \in V$
- $(u + v) + w = u + (v + w)$ and $(ab)u = a(bu)$ for all $u, v, w \in V$ and all $a, b \in \mathbb{F}$
- There exists $0 \in V$ such that $u + 0 = u$ for all $u \in V$
- For every $u \in V$, there exists $w \in V$ such that $u + w = 0$
- $1u = u$ for all $u \in V$
- $a(u + w) = au + aw$ and $(a + b)u = au + bu$ for all $a, b \in \mathbb{F}$ and all $u, w \in V$

Example 1.2 Vector Spaces:

- $\mathbb{F}^n$ with the usual operations of addition and scalar multiplications is a vector space.
- $\mathbb{F}^\infty$ is defined to be the set of all sequences of elements of $\mathbb{F}$:

  $$\{(x_1, x_2, \ldots) : x_j \in \mathbb{F} \text{ for } j = 1, 2, \ldots\}$$

  Addition and scalar multiplication are also defined coordinate-wise. This is also a vector space.
- More generally, if S is a set, let F^S denote the set of functions from S to $\mathbb{F}$. For $f, g \in \mathbb{F}^S$, the sum $f + g \in \mathbb{F}^S$ is the function defined by:

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

  for all $x \in S$. For $\lambda \in \mathbb{F}$ and $f \in \mathbb{F}^S$, the product $\lambda f \in \mathbb{F}^S$ is the function defined by

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

  for all $x \in S$. With these definitions, $\mathbb{F}^S$ becomes a vector space.

Our first theorem then follows:

Theorem 1.1 If V is a vector space, $\forall u \in V$, $0u = 0$.

Proof For arbitrary $u \in V$, we have:

$$\begin{aligned} 0u &= (0+0)u \\ &= 0u + 0u \end{aligned}$$

Adding the additive inverse of $0u$, denoted $-0u$, to both sides of the equation above gives:

$$\begin{aligned} 0u + (-0u) &= 0u + 0u + (-0u) \\ 0 &= 0u \end{aligned}$$

as desired. ■

Advantages of the abstract approach to vector spaces:

- Can apply what was done in multiple new situations.
- Stripping away inessential properties leads to greater understanding.

If V is a vector space, it would be incorrect to prove that $0u = 0$ for $u \in V$ by writing: Let $u = (x_1, \ldots, x_n)$, thus...

Note 1.3 An element of V is not necessarily of the form $(x_1, \ldots, x_n)$.

1.3 Subspaces

Let's add a new convention. From now on, V denotes a vector space over $\mathbb{F}$ for brevity.

Definition 1.11 A subset U of V is called a subspace of V if U is also a vector space (using the same addition and scalar multiplication as on V).

Example 1.3 $\{(x_1, x_2, 0) : x_1, x_2, \in \mathbb{F}\}$ is a subspace of $\mathbb{F}^3$

Definition 1.12 A subset U of V is a subspace of V if and only if U satisfies the following three conditions:

- $0 \in U$
- $u, w \in U \implies u + w \in U$
- $\lambda \in \mathbb{F}, u \in U \implies \lambda u \in U$

Note that we do not need to check any of the other properties of a vector space because we know that they will hold. Most of these properties are related to the addition and multiplication properties, which we know hold since we're using the same ones.

Example 1.4 Examples of subspaces:

- If $b \in \mathbb{F}$, then
$$\{(x_1, x_2, x_3, x_4) \in \mathbb{F}^4 : x_3 = 5x_4 + b\}$$
is a subspace of $\mathbb{F}^4$ if and only if $b = 0$, in order to have the additive identity in the set.
- The set of continuous real-valued functions on the interval [0,1] is a subspace of $\mathbb{R}^{[0,1]}$. (The zero function is the identity in this case)

- The set of differentiable real-valued functions on $\mathbb{R}$ is a subspace of $\mathbb{R}^{\mathbb{R}}$ (a sum of two differentiable functions is differentiable)
- The set of all sequences of complex numbers with limit 0 is a subspace of $\mathbb{C}^{\infty}$
- The subspaces of $\mathbb{R}^2$ are precisely $\{0\}, \mathbb{R}^2$ and all lines in $\mathbb{R}^2$ through the origin.

Definition 1.13 Suppose $U_1, \ldots, U_m$ are subsets of V. The sum of $U_1, \ldots, U_m$, denoted $U_1 + \ldots + U_m$, is the set of all possible sums of elements of $U_1, \ldots, U_m$.

$$U_1 + \ldots + U_m = \{u_1 + \cdots + u_m : u_1 \in U_1, \ldots, u_m \in U_m\}$$

Theorem 1.2 Suppose $U_1, \ldots, U_m$ are subspaces of V. Then $U_1 + \ldots U_m$ is the smallest subspace of V containing $U_1, \ldots, U_m$.

Definition 1.14 Suppose $U_1, \ldots, U_m$ are subspaces of V.
The sum $U_1 + \ldots + U_m$ is called a direct sum if each element of $U_1 + \ldots + U_m$ can be written in only one way as a sum $u_1 + \ldots + u_m$ where each u_j is in U_j
If the sum is indeed a direct sum, we use $\oplus$ between the symbols to denote that it is a direct sum.

Example 1.5 Suppose

$$U = \{(x, y, 0) \in \mathbb{F}^3 : x, y \in \mathbb{F}\}, W = \{(0, 0, z) \in \mathbb{F}^3 : z \in \mathbb{F}\}$$

Thus, $\mathbb{F}^3 = U \oplus W$.

These two theorems make it easy to see whether something is a direct sum.

Theorem 1.3 Suppose $U_1, \ldots, U_m$ are subspaces of V. Then $U_1 + \ldots + U_m$ is a direct sum if and only if the only way to write 0 as a sum $u_1 + \ldots + u_m$, where each u_j is in U_j, is by taking each u_j equal to 0.

Theorem 1.4 Suppose U and W are subspaces of V. Then $U + W$ is a direct sum if and only if $U \cap W = \{0\}$.

2 Finite-Dimensional Vector Spaces

2.1 Span and Linear Independence

Definition 2.1 A list of length n is an ordered collection of n elements (which might be numbers, other lists, or more abstract entities) separated by commas (and perhaps surrounded by parentheses).

Definition 2.2 A linear combination of a list $v_1, \ldots, v_m$ of vectors in V is a vector of the form

$$a_1 v_1 + \ldots + a_m v_m$$

where $a_1, \ldots, a_m \in \mathbb{F}$

- $(13, -1, 7)$ is a linear combination of $(2, 1, -1), (1, -2, 4)$ because:

$$5(2, 1, -1) + 3(1, -2, 4) = (13, -1, 7)$$

- $(13, -1, 6)$ is a linear combination of $(2, 1, -1), (1, -2, 4)$ because: there do not exist numbers $a_1, a_2 \in \mathbb{F}$ such that:

$$a_1(2, 1, -1) + a_2(1, -2, 4) = (13, -1, 7)$$

Definition 2.3 The set of all linear combinations of a list of vectors $v_1, \ldots, v_m$ in V is called the span of $v_1, \ldots, v_m$, denoted span$(v_1, \ldots, v_m)$. In other words

$$span(v_1, \ldots, v_m) = \{a_1 v_1, + \ldots + a_m v_m : a_1, \ldots, a_m \in \mathbb{F}\}$$

The previous example shows that in $\mathbb{F}^3$:

- $(13, -1, 7) \in \text{span}((2, 1, -1), (1, -2, 4))$
- $(13, -1, 6) \notin \text{span}((2, 1, -1), (1, -2, 4))$

Definition 2.4 The span of a list of vectors in V is the smallest subspace of V containing all the vectors in the list.

Definition 2.5 A vector space is called **finite-dimensional** if the span of some list of vectors in it is the entire vector space.

Example 2.1 $\mathbb{F}^3$ is finite-dimensional because:

$$\mathbb{F}^3 = \text{span}((1, 0, 0), (0, 1, 0), (0, 0, 1))$$

Note that we have defined all lists to be finite.

Definition 2.6 A vector space is called **infinite-dimensional** if it is not finite-dimensional.

Example 2.2 $\mathbb{F}^\infty$ is infinite-dimensional.

Linear algebra is the study of linear maps on finite-dimensional vector spaces.

Definition 2.7 A list $v_1, \ldots, v_m$ of vectors in V is called **linearly independent** if the only choice of $a_1, \ldots, a_m \in \mathbb{F}$ that makes $a_1v_1 + \cdots + a_mv_m$ equal 0 is $a_1 = \cdots = a_m = 0$.

Example 2.3 Examples of linearly independent lists:

- A list v of one vector $v \in V$ is linearly independent iff $v \neq 0$
- A list of two vectors in V is linearly independent iff neither vector is scalar multiple of the other.
- The list $1, x, \ldots, x^m$ is linearly independent in $\mathbb{R}^{\mathbb{R}}$ for each nonnegative integer m. The reason is that the subspace spanned by these vectors represent all the polynomials of degree up to m. The only way that a polynomial can be identically 0 (the identity) is if all the coefficients are 0.

Definition 2.8 A list of vectors in V is called **linearly dependent** if it is not linearly independent.
Alternatively, a list $v_1, v_2, \ldots, v_m$ of vectors in V is linearly dependent if there exist $a_1, a_2, \ldots, a_m \in \mathbb{F}$, not all 0, such that $a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0$.

Examples of linearly dependent lists:

- $(2, 3, 1), (1, -1, 2), (7, 3, 8)$ is linearly dependent in $\mathbb{F}^3$ because
$$2(2, 3, 1) + 3(1, -1, 2) + (-1)(7, 3, 8) = (0, 0, 0)$$
- Every list of vectors in V containing the 0 vector is linearly dependent.
- If some vector in a list of vectors in V is a linear combination of the other vectors, then the list in linearly dependent.

Theorem 2.1 Suppose $v_1, v_2, \ldots, v_m$ is a linearly dependent list in V. Then there exists $j \in \{1, 2, \ldots, m\}$ such that the following hold:

- $v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$
- If the jth term is removed from $v_1, v_2, \ldots, v_m$, the span of the remaining list equals $\operatorname{span}(v_1, v_2, \ldots, v_m)$.

This captures the idea of redundancy. Let's look at an example:

$(2, 3, 1), (1, -1, 2), (7, 3, 8)$ is linearly dependent in $\mathbb{F}^3$.

So, $(7, 3, 8) \in \operatorname{span}(2, 3, 1), (1, -1, 2)$ and we also see that without it, the span remains the same.

Theorem 2.2 In a finite-dimensional vector space, the length of every linearly independent list of vectors is less than or equal to the length of every spanning list of vectors.

Proof Suppose $u_1, u_2, \ldots, u_m$ is linearly independent in V. Suppose also that $w_1, w_2, \ldots, w_n$ spans V. We need to prove that $m \leq n$. We do so through the multi-step process described below.
Step 1
Let B be the list $w_1, w_2, \ldots, w_n$, which spans V. Thus adjoining any vector in V to this list produces a linearly dependent list (because the newly adjoined vector can be written as a linear combination of the other vectors). In particular

$$u_1, w_1, w_2, \ldots, w_n$$

is linearly dependent. Thus by the Linear Dependence Lemma, we can remove one of the w's so that the new list B (of length n) consisting of u_1 and the reamining w's spans V.
Step j
The list B (of length n) from step $j - 1$ spans V. Thus adjoining any vector to this list produces a linearly dependent list. In particular, the list of length $n+1$ is obtained by adjoining u_j to B, placing it just after $u_1, \ldots, u_{j-1}$, is linearly dependent. By the Linear Dependence Lemma, one of the vectors in this list is in the span of the previous ones, and because $u_1, u_2, \ldots, u_j$ is linearly independent, this vector is one of the w's, note one of the u's. We can remove that w from B so that the new list B (of length n) consisting of $u_1, \ldots, u_j$ and the remaining w's spans V.
After step m, we have added all the u's and the process stops. At each step as we add a u to B, the Linear Dependence Lemma implies that there is some w to remove. Thus there are at least as many w's as u's. ■

Let's apply the theorem to claim the following:

Example 2.4
- The list $(1, 2, 3), (4, 5, 8), (9, 6, 7), (-3, 2, 8)$ is not linearly independent because the list $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ spans $\mathbb{R}^3$. The theorem is applied here because something bigger than a spanning list cannot possibly be a linearly independent list.
- The list $(1, 2, 3, -5), (4, 5, 8, 3), (9, 6, 7, -1)$ does not span $\mathbb{R}^4$ because the list $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)$ is linearly independent in $\mathbb{R}^4$. The theorem applied here because the contrapositive is that no spanning list can be smaller than a linearly independent one, which shows that the smaller list spanning is impossible.

Theorem 2.3 Every subspace of a finite-dimensional vector space is finite-dimensional.