Linear Algebra Notes

Lance Remigio

December 6, 2023

Contents

1 Vector Spaces **5**

1.1 Vector Spaces . . . 5

1.1.1 Basics . . . 5

1.1.2 Non-examples . . . 8

1.1.3 Basic Extensions from Definition . . . 8

1.2 Subspaces . . . 9

1.3 Linear Combinations and Systems of Linear Equations . . . 12

1.4 Linear Dependence and Linear Independence . . . 16

1.4.1 Motivation . . . 16

1.4.2 Examples of Linearly Dependent Sets . . . 17

1.4.3 Examples of Linear Independent Sets . . . 17

1.5 Bases And Dimension . . . 19

Chapter 1

Vector Spaces

1.1 Vector Spaces

1.1.1 Basics

Definition 1.1.1 (Vector Spaces). A **vector space** (or **linear space**) over a field F consists of a set on which two operations (called **addition** and **scalar multiplication**, respectively) are defined so that for each pair of elements x, y, in V there is a unique element ax in V, such that the following conditions hold:

(VS 1) For all $x, y \in V$, $x + y = y + x$ (commutativity of addition).

(VS 2) For all $x, y, z \in V$, $(x + y) + z = x + (y + z)$ (associativity of addition).

(VS 3) There exists an element in V denoted by O such that $x + O = x$ for each $x \in V$

(VS 4) For each element $x \in V$, there exists an element $y \in V$ such that $x + y = O$.

(VS 5) For each element $x \in V$, we have $1x = x$.

(VS 6) For each $a, b \in F$ and each element $x \in V$, then $(ab)x = a(bx)$.

(VS 7) For each element $a \in F$ and each pair $x, y \in V$, we have $a(x + y) = ax + ay$.

(VS 8) For each pair $a, b \in F$ and each $x \in V$, we have $(a + b)x = ax + bx$.

The elements $x + y$ and ax are called the **sum** of x and y and the **product** of a and x, respectively.

- The elements of a field F are called **scalars** and the elements of a vector space V are called **vectors** (these should not be confused!).
- Every vector space will always be defined over a given field, mostly defined over the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$ unless otherwise noted.
- Every vector space should specify the operations of addition and scalar multiplication.

Definition 1.1.2 (n-tuples). An object of the form $(a_1, a_2, \ldots, a_n)$, where the entries $a_1, a_2, \ldots, a_n$ are elements of a field F, is called an **n-tuple** with entries from F. The elements $a_1, a_2, \ldots, a_n$ are called **entries** or **components** of the n-tuple.

Definition 1.1.3. We say that two n-tuples, $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$, are **equal** if $a_i = b_i$ for $i = 1, 2, \ldots, n$.

Example 1. The set of all n-tuples with entries from a field F denoted by F_n is a vector space. To see why, suppose $u, v \in F_n$ where $u = (a_1, a_2, \ldots a_n)$ and $v = (b_1, b_2, \ldots, b_n)$. If we take term-by-term addition of the entries in both u and v, then we end up with

$$u + v = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$

and likewise,

$$cu = (ca_1, ca_2, \ldots, ca_n).$$

These same set of operations define $\mathbb{R}^3$ as a vector space over $\mathbb{R}$ and likewise, $\mathbb{C}^2$ is a vector space over $\mathbb{C}$.

- Note that vectors in F^n can be written as **column vectors**

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

  rather than **row vectors** $(a_1, a_2, \ldots, a_n)$.

- 1-tuples are are just scalars or an just an element from F.

Definition 1.1.4. An $m \times n$ **matrix** with entries from a field F is a rectangular array of the form

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

where each entry a_{ij} with $(1 \leq i \leq m, 1 \leq j \leq n)$ is an element of F. We call the entries a_{ij} with $i = j$ the **diagonal entries** of the matrix. The entries $a_{i1}, a_{i2}, \ldots a_{in}$ compose the i**th row**of the matrix, and the entries $a_{1j}, a_{2j}, \ldots a_{mj}$ compose the jth column of the matrix.

- The rows make a vector space which we denote F^n.
- Likewise, the columns make a vector space we denote F^m

Definition 1.1.5 (Zero Matrix)**.** The $m \times n$ matrix in which each entry equals zero is called the **zero matrix** and is denoted by O.

Definition 1.1.6 (Square Matrix)**.** A matrix is **square** if the number of rows and columns of a matrix are equal.

- Just like our tuple example, the set of all $m \times n$ matrices with entries from a field F form a vector space. Denote this vector space as $M_{m\times n}(F)$ endowed with two operations; that is, **matrix addition** and **scalar multiplication**. Suppose for $A, B \in M_{m\times n}(F)$ and $c \in F$, we have

$$(A + B)_{ij} = A_{ij} + B_{ij}$$

  and

$$(cA)_{ij} = cA_{ij}$$

  for $1 \leq i \leq m$ and $1 \leq j \leq n$. In other words, the two operations can be performed entry-wise.

- The operations from our tuple case extends very naturally to $M_{m\times n}(F)$. In other words, if we add two matrices A_{ij} and B_{ij}, then we would expect to that $A_{ij} + B_{ij} \in M_{m\times n}(F)$ as well and likewise for the scalar multiplication case.

Definition 1.1.7 (Set of All Functions). Let S be any nonempty set and F be any field, and let $\mathcal{F}(S, F)$ denote the set of all functions from S to F.

Definition 1.1.8. Two functions $f, g \in \mathcal{F}(S, F)$ are called **equal** if $f(s) = g(s)$ for each $s \in S$.

The set of functions above forms a vector space with the operations of addition and scalar multiplication defined for every $f, g \in \mathcal{F}(S, F)$ and $c \in F$ with

$$(f + g)(s) = f(s) + g(s) \text{ and } (cf)(s) = cf(s)$$

Definition 1.1.9 (Set of All Polynomials). A **polynomial** with coefficients from a field F is an expression of the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

where n is a nonnegative integer and each $a_k \in F$ is called the **coefficient** of x_k.

Definition 1.1.10 (Zero Polynomial). We call $f(x) = 0$ the **zero polynomial** if $a_n = a_{n-1} = \cdots = a_0 = 0$.

Definition 1.1.11 (Degree). The **degree** of a given polynomial f is defined to be the largest exponent of x that appears in the representation

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

Definition 1.1.12 (Equality of Polynomials). We call two polynomials f, g, where

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

and

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0,$$

are **equal** if $m = n$ and $a_i = b_i$ for all $i = 0, 1, \ldots, n$.

Suppose we have $c \in F$ and say we evaluated the polynomial $f \in F$ at c. Then we would have the following

$$f(c) = a_n c^n + a_{n-1} c^{n-1} + \cdots + a_1 c + a_0$$

where $f(c) \in F$.

Definition 1.1.13 (Basic Operations of Polynomials). Define polynomial **addition** $f + g$ as the following:

$$f(x) + g(x) = (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0).$$

Let $c \in F$. Let scalar **multiplication** be defined by

$$cf(x) = ca_n x^n + ca_{n-1} x^{n-1} + \cdots + ca_1 x + ca_0.$$

The operations above form a vector space for $P(F)$ (the set of all polynomials).

Definition 1.1.14 (Sequences). A **sequence** in F is a function $\sigma : \mathbb{Z}^+ \to F$. A given sequence σ such that $\sigma(n) = a_n$ for $n = 1, 2, \ldots$ is denoted (a_n).

Let V be the set of all sequences $\sigma(n) \in F$. For every $(a_n), (b_n) \in V$ with $t \in F$, we have the following operations

$$(a_n) + (b_n) = (a_n + b_n) \text{ and } t(a_n) = (ta_n).$$

1.1.2 Non-examples

Example 2. Let $S = \{(a_1, a_2) : a_1, a_2 \in R\}$ where R is a field. For every $(a_1, a_2), (b_1, b_2) \in S$ and $c \in R$, define

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 - b_2) \text{ and } c(a_1, a_2) = (ca_1, ca_2).$$

Note that S is **NOT** a vector space since (VS 1), (VS 2), and (VS 8) fail.

Example 3. Let S be the same set as in the last example. For $(a_1, a_2), (b_1, b_2) \in S$ and $c \in R$, define

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, 0)$$

and

$$c(a_1, a_2) = (ca_1, 0).$$

Note that S is **NOT** a vector space with these operations since (VS 3), (VS 4), and (VS 5) fail.

1.1.3 Basic Extensions from Definition

Theorem 1.1.1 (Cancellation Law for Vector Addtion)**.** If $x, y, z \in V$ such that $x + z = y + z$, then $x = y$.

Proof. There exists a vector $v \in V$ such that $z + v = O$ (VS 4). Thus, we have

$$\begin{aligned} x &= x + O \\ &= x + (z + v) \\ &= (x + z) + v \\ &= (y + z) + v \\ &= y + (z + v) \\ &= y + O \\ &= y. \end{aligned}$$

Hence, we have $x = y$. ■

Corollary. The vector described O described in (VS 3) is unique.

Proof. Let $x \in V$. Suppose there exists two elements $O, O' \in V$ such that

$$x + O = x \tag{1}$$

and

$$x + O' = x. \tag{2}$$

Our objective is to show that $O = O'$. Equating (1) and (2), we find that

$$x + O = x + O' \tag{3}$$

Using the Cancellation Law for Vector Addition, (3) implies that $O = O'$. Hence, O is a unique vector in V. ■

Corollary. The vector y described in (VS 4) is unique.

Proof. Let $x \in V$. Suppose there exists $y, y' \in V$ such that

$$x + y = O \text{ and } x + y' = O.$$

Equating the two equations above, we get that

$$x + y = x + y'.$$

Applying the Cancellation Law for Vector Addition once again, we get that $y = y'$. Hence, y is a unique vector in V. ■

Theorem 1.1.2. In any vector space V, the following statements are true:

(a) $0x = O$ for any $x \in V$.

(b) $(-a)x = -(ax) = a(-x)$ for each $a \in F$ and each $x \in V$.

(c) $a0 = 0$ for any $a \in F$.

Proof. (a) Using (VS 8), (VS 3), and (VS 1), it follows that

$$\begin{aligned} 0x + 0x &= (0 + 0)x \\ &= 0x \\ &= 0x + O \\ &= O + 0x. \end{aligned}$$

Hence, $0x = O$ by Theorem 1.1.

(b) The vector $-(ax)$ is the unique element of V such that $ax + [-(ax)] = O$. If we know that $ax + (-a)x = O$, we can use Corollary 2 to Theorem 1.1 to state that $(-a)x = -(ax)$. But by (VS 8), we have

$$ax + (-ax) = [a + (-a)]x = 0x = O$$

by (a). Hence, we have that $(-a)x = -(ax)$. In particular, we have $(-1)x = -x$. Thus, (VS 6) implies that

$$a(-x) = a[(-1)x] = [a(-1)]x = (-a)x.$$

(c) The proof of (c) is similar to the proof of (a).

■

1.2 Subspaces

Definition 1.2.1 (Subspaces). A subset W of a vector space V over a field F is called a **subspace** of V if W is a vector space over F with the operations of addition and scalar multiplication.

The most simple examples of subspaces of V is V itself and the set containing just the zero vector. The latter is denoted as the **zero subspace** of V.

We don't have to check all the vector space properties to prove that a subset W of V is a subspace of V. This is because vectors in any subset of V already satisfy all the properties of a vector space. Hence, we need only check that W is closed under

1. **Addition**: $x + y \in W$ whenever $x \in W$ and $y \in W$.
2. **Scalar Multiplication**: $cx \in W$ whenever $c \in F$ and $x \in W$.
3. W contains a zero vector.
4. Each vector in W has an additive inverse in W.

Theorem 1.2.1 (Subspaces). Let V be a vector space and W a subset of V. Then W is a subspace of V if and only if the following three conditions hold for the operations defined in V.

(a) $O \in W$.

(b) $x + y \in W$ whenever $x \in W$ and $y \in W$.

(c) $cx \in W$ whenever $c \in F$ and $x \in W$.

Proof. Suppose W is a subspace of V. Since W is also a vector space with the operations of addition and scalar multiplication defined in V. Hence, we know that W is closed under addition and scalar multiplication. Note that W contains a zero vector $O' \in W$ such that for any $x \in W$, we have $x + 0' = x$. Since x is also in V (since $W \subseteq V$), we know that $x + 0 = x$. Using the cancellation , we can see that $x + 0' = x + 0$ implies $0' = 0$. Hence, properties (a), (b), and (c) are satisfied.

Conversely, suppose conditions (a), (b), and (c) hold. We need to show that W is a subspace. This tells us that W is closed under addition and scalar multiplication in addition to containing the zero vector. Let $x \in W$. Since W is closed under scalar multiplication, we know that $(-1)x \in W$. By part (b) of theorem 2, we know that $1(-x) = -x \in W$. Hence, W contains an additive inverse and we are done. ■

The theorem above provides a simpler way to determine whether a given subset of a vector space is a subspace. Instead of using the definition to show that a give subset is a subspace, it is more common to use the result above.

Definition 1.2.2 (Tranpose). The **transpose** of A^t of an $m \times n$ matrix A is the $n \times m$ matrix obtained from A by interchanging the rows with the columns; that is, $(A^t)_{ij} = A_{ji}$.

Example 4.

$$\begin{pmatrix} 1 & -2 & 3 \\ 0 & 5 & -1 \end{pmatrix}^t = \begin{pmatrix} 1 & 0 \\ -2 & 5 \\ 3 & -1 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^t = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}.$$

Definition 1.2.3 (Symmetric Matrices). A **symmetric matrix** is a matrix A such that $A^t = A$.

- The easiest example of a symmetric matrix would be a square matrix where $i = j$.
- The set W of all symmetric matrices in $M_{n \times n}(F)$ is a subspace of $M_{n \times n}(F)$ since the conditions of Theorem 1.3 hold.

We can show that W is indeed a subspace.

1. The zero matrix is equal to its transpose and hence belongs to W.
2. Suppose $A \in W$ and $B \in W$. Hence, $A^t = A$ and $B^t = B$. Hence, we have
$$(A + B)^t = A^t + B^t = A + B$$
which implies that $A + B \in W$.
3. We have $A \in W$ implies $A^t = A$. Now, let $a \in F$. Then we have that $(aA)^t = aA^t = aA$. Hence, $aA \in W$.

Example 5. Let n be non-negative integer, and let $P_n(F)$ consist of all polynomials in $P(F)$ having degree less than or equal to n. We get that $P_n(F)$ is a subspace because:

1. Zero polynomial has degree -1, it is in $P_n(F)$.
2. The sum of two polynomials with degrees less than or equal to n is another a polynomial less than or equal to n.
3. The product of a scalar and a polynomial of degree less than or equal to n is a polynomial less than or equal to n.

This tells us that $P_n(F)$ is a subspace of the space of all polynomials.

Example 6. Let $C(\mathbb{R})$ denote the set of all continuous real-valued functions defined on $\mathbb{R}$. We know that $C(\mathbb{R})$ is a subset of the vector space $\mathcal{F}(\mathbb{R}, \mathbb{R})$ defined in Example 3. We will show that $C(\mathbb{R})$ is a subspace of $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

1. Note that the zero function $f(x) = 0$ for all $x \in \mathbb{R}$ of $\mathcal{F}(\mathbb{R}, \mathbb{R})$ is a continuous real-valued function.
2. Let $f, g \in C(\mathbb{R})$. Since the sum of f and g is also continuous for all $x \in \mathbb{R}$, we have that $f + g \in C(\mathbb{R})$.
3. Let $c \in \mathbb{R}$ and $f \in C(\mathbb{R})$. We have that for any $x \in \mathbb{R}$, $(cf)(x) = cf(x)$ is a continuous function. Hence, property (c) is satisfied.

Hence, $C(\mathbb{R})$ is a subspace.

Definition 1.2.4 (Upper Triangular)**.** An $m \times n$ matrix A is called **upper triangular** if all its entries below the diagonal entries are zero; that is, if $A_{ij} = 0$ whenever $i > j$.

Example 7. Let B be an upper triangular 3×4 matrix

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \end{pmatrix}$$

Definition 1.2.5 (Diagonal Matrix)**.** An $n \times n$ matrix M is called a **diagonal matrix** if $M_{ij} = 0$ whenever $i \neq j$; that is, if all its non-diagonal entries are zero.

Example 8. Let A be diagonal 3×3 matrix.

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

Example 9. The set of diagonal matrices is a subspace of $M_{n\times n}(F)$. To see why, we have

1. Let O be the zero matrix of $M_{n\times n}(F)$. Since O is also a diagonal matrix, we know that O must be in the set of diagonal matrices.
2. Let A, B in the set of diagonal matrices. Let $i \neq j$ such that
$$(A + B)_{ij} = A_{ij} + B_{ij} = 0 + 0 = 0.$$
Hence, $A + B$ is a diagonal matrix and so addition is closed.
3. Let A be a diagonal matrix as before and let $c \in F$. Let $i \neq j$ again, and observe that
$$(cA)_{ij} = cA_{ij} = c \cdot 0 = 0.$$
Hence, cA is a diagonal matrix and so scalar multiplication is closed.

Since the set of diagonal matrices satisfies all properties of theorem 3, we conclude that it is indeed a subspace of $M_{n\times n}(F)$.

Definition 1.2.6 (Trace). The **trace** of an $n \times n$ matrix M, denoted $\mathrm{tr}(M)$, is the sun of the diagonal entries of M; that is,

$$\mathrm{tr}(M) = M_{11} + M_{22} + \cdots + M_{nn}.$$

The set of all $n \times n$ matrices that have a trace equal to zero is a subspace of $M_{n\times n}(F)$ (proved in Exercise 6).

Example 10 (Non-example). Denote V as the set of matrices in $M_{m\times n}(\mathbb{R})$ having non-negative entries. The subset V is not a subspace because it is not closed under scalar multiplication. We can see this by multiplying any matrix in V be a negative number and observe that the entries of said matrix are no longer all non-negative.

We can create subspaces out of other subspaces.

Theorem 1.2.2. Any intersection of subspaces of a vector space V is a subspace of V.

Proof. Let C be a collection of subspaces of V, and let W denote the intersection of the subspaces in C. Since every subspace contains the zero vector and the intersection $W \neq \emptyset$, the zero vector $0 \in W$. Let $a \in F$ and $x, y \in W$. Since each subspace of C is closed under addition and scalar multiplication, it follows that $x + y$ and ax are contained each subspace in C and hence W must be closed under addition and closed under scalar multiplication. Thus, W is a subspace of V by theorem 3. ■

A natural question to ask is whether or not the union of subspaces of a vector space V is a subspace of V too. We can see that indeed the union of subspaces of V satisfy the first first and third property of theorem 3. However, property 2 need not be satisfied all the time. In fact, the union can only be a subspace of V if and only if one the subspaces is a subset of the other.

1.3 Linear Combinations and Systems of Linear Equations

Definition 1.3.1 (Linear Combinations). Let V be a vector space and S a nonempty subset of V. A vector $v \in V$ is called **linear combination** of vectors of S if there exist a finite number of vectors $u_1, u_2, \ldots, u_n \in S$ and scalars $a_1, a_2, \ldots, a_n \in F$ such that

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n.$$

In this case, we say that v is a linear combination of $u_1, u_2, \ldots, u_n \in V$ and call $a_1, a_2, \ldots, a_n \in F$ the **coefficients** of the linear combination.

- An easy example of a vector expressed as a linear combination is the zero vector O where $0v = O$ for each $v \in V$.
- Determining whether a vector is a linear combination of other vectors often involves solving a system of linear equations.

Example 11. Suppose we wanted to express the vector $(2, 6, 8) \in \mathbb{R}^3$ as a linear combination of

$$u_1 = (1, 2, 1),\ u_2 = (-2, -4, -2),\ u_3 = (0, 2, 3),\ u_4 = (2, 0, -3), \text{ and } u_5 = (-3, 8, 16).$$

Our goal is to find scalars a_1, a_2, a_3, a_4 and a_5 such that

$$(2, 6, 8) = a_1u_1 + a_2u_2 + a_3u_3 + a_4u_4 + a_5u_5.$$

Doing a bit of algebra, we would need to solve the following system of linear equations

$$\begin{aligned} a_1 - 2a_2 \quad\; + 2a_4 - 3a_5 &= 2 \\ 2a_1 - 4a_2 + 2a_3 \quad\; + 8a_5 &= 6 \\ a_1 - 2a_2 + 3a_3 + 16a_5 &= 8 \end{aligned}$$

Solving the system of equations above involves three types of operations:

1. Interchanging the order of any two equations in the system;
2. multiplying any equation in the system by some non-zero constant;
3. adding a constant multiple of any equation to another equation in the system.

We will learn in a later section that the operations listed above do not change the set of solutions to the original system.

Example 12. We claim that

$$2x^3 - 2x^2 + 12x - 6$$

is a linear combination of

$$x^3 - 2x^2 - 5x - 3 \text{ and } 3x^3 - 5x^2 - 4x - 9$$

in $P_3(\mathbb{R})$, but that

$$3x^3 - 2x^2 + 7x + 8$$

is not. In the first case we wish to find scalars a and b such that

$$\begin{aligned} 2x^3 - 2x^2 + 12x - 6 &= a(x^3 - 2x^2 - 5x - 3) \\ &\quad + b(3x^3 - 5x^2 - 4x - 9) \\ &= (a + 3b)x^3 + (-2a - 5b)x^2 \\ &\quad + (-5a - 4b)x + (-3a - 9b). \end{aligned}$$

Thus, we have the following system of linear equations:

$$\begin{aligned} a + 3b &= 2 \\ -2a - 5b &= -2 \\ -5a - 4b &= 12 \\ -3a - 9b &= -6. \end{aligned}$$

Adding the appropriate multiples of the first equation to the others in order to eliminate a, we find that

$$\begin{aligned} a + 3b &= 2 \\ b &= 2 \\ 11b &= 22 \\ 0b &= 0. \end{aligned}$$

Then we get that $a = -4, b = 2, 0 = 0, 0 = 0$. Hence, we have

$$2x^3 - 2x^2 + 12x - 6 = -4(x^3 - 2x^2 - 5x - 3) + 2(3x^3 - 5x^2 - 4x - 9).$$

Looking at the second case now, we observe that using the preceding technique leads us to the following system of linear equations

$$\begin{aligned} a + 3b &= 3 \\ -2a - 5b &= -2 \\ -5a - 4b &= 7 \\ -3a - 9b &= 8. \end{aligned}$$

Eliminating a as before yields the following:

$$\begin{aligned} a + 3b &= 3 \\ b &= 4 \\ 11b &= 22 \\ 0 &= 17. \end{aligned}$$

The presence of the non-sensical result in the last equality tells us that there are no solutions to the system and that the second polynomial cannot be written as a linear combination of the two polynomials.

We can denote the set of all linear combinations of some set of vectors in the following:

Definition 1.3.2 (Span). Let S be a nonempty subset of a vector space V. The **span** of S, denoted $\operatorname{span}(S)$, is the set consisting of all linear combinations of the vectors in S. For convenience, we define $\operatorname{span}(\emptyset) = \{0\}$.

Some immediate examples of spans are:

Example 13. In $\mathbb{R}^3$, the span of the set $S = \{(1,0,0),(0,1,0)\}$ consist of all vectors in $\mathbb{R}^3$ such that for some scalars $a, b \in \mathbb{R}$, we have

$$a(1,0,0) + b(0,1,0) = (a,b,0).$$

This tells us that S contains all the points in the xy-plane. One can show that S is a subspace of $\mathbb{R}^3$.

Theorem 1.3.1. The span of any subset S of a vector space v is a subspace of V that contains S. Moreover, any subspace of V that contains S must also contain the span of S.

Proof. Suppose $S = \emptyset$. Then the span of S is just $\operatorname{span}(\emptyset) = \{0\}$ which is a subspace in which S is contained in. Moreover, $\operatorname{span}(\emptyset) = \{0\}$ is always contained in any subspace X of V because every subspace contains the zero vector 0. Suppose $S \neq \emptyset$, then S contains a vector z. We need to show that $\operatorname{span}(S)$ is a subspace of V.

(a) Since $S \neq \emptyset$, we know that S contains a vector v such that $0z = 0$. Hence, $0 \in \operatorname{span}(S)$.

(b) Let $x, y \in \operatorname{span}(S)$. We need to show that $x + y \in \operatorname{span}(S)$. If $x \in \operatorname{span}(S)$, then we can find $a_1, a_2, \ldots, a_n \in F$ and $x_1, x_2, \ldots, x_n \in S$ such that

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n.$$

Likewise, $y \in \operatorname{span}(S)$ implies that we can find scalars $b_1, b_2, \ldots, b_m$ and vectors $y_1, y_2, \ldots, y_m$ such that

$$b_1y_1 + b_2y_2 + \cdots + b_my_m.$$

Then

$$x + y = a_1x_1 + a_2x_2 + \cdots + a_nx_n + b_1y_1 + b_2y_2 + \cdots + b_my_m$$

is a linear combination and thus $x + y \in \text{span}(S)$.

(c) Note that $c \in F$ implies that

$$\begin{aligned} cx &= c(a_1x_1 + a_2x_2 + \cdots + a_nx_n) \\ &= (ca_1)x_1 + (ca_2)x_2 + \cdots + (ca_n)x_n \end{aligned}$$

is a linear combination and thus $cx \in \text{span}(S)$.

Hence, $\text{span}(S)$ is a subspace of V. Now we need to show that S is contained within $\text{span}(S)$; that is, $S \subseteq \text{span}(S)$. Let $v \in S$. Then using (VS 5), we can see that $1 \cdot v = v$ is a linear combination; so we have $v \in \text{span}(S)$ and hence, the span of S contains S.

Now let W be any subspace of V that contains S. We need to show that $\text{span}(S) \subseteq W$. Let $v \in \text{span}(S)$. Then we can find scalars $a_1, a_2, \ldots, a_n \in F$ and $x_1, x_2, \ldots, x_n \in S$ such that

$$v = a_1x_1 + a_2x_2 + \cdots + a_nx_n.$$

Since $S \subseteq W$, we know that $x_1, x_2, \ldots, x_n \in W$. Using exercise 20 from section 1.3 and using the same set of scalars $a_1, a_2, \ldots, a_n \in F$, we have $v = a_1x_1 + a_2x_2 + \cdots + a_nx_n \in W$. Hence, $\text{span}(S) \subseteq W$. ■

Definition 1.3.3. A subset S of a vector space V **generates** (or **spans**) V if $\text{span}(S) = V$. In this case, we also say that the vectors of S generates (or span) V.

Example 14 (Vectors in $\mathbb{R}^3$)**.** The vectors $(1, 1, 0), (1, 0, 1), (0, 1, 1)$ generate $\mathbb{R}^3$ since any given vector $v \in \mathbb{R}^3$ is a linear combination of the three given vectors. Furthermore, there exists scalars $r, s, t \in \mathbb{R}$ such that

$$r(1, 1, 0) + s(1, 0, 1) + t(0, 1, 1) = (a_1, a_2, a_3) = v$$

where

$$r = \tfrac{1}{2}(a_1 + a_2 - a_3), s = \tfrac{1}{2}(a_1 - a_2 + a_3) \text{ and } t = \tfrac{1}{2}(-a_1 + a_2 + a_3).$$

Example 15 (Polynomials)**.** The polynomials $x^2+3x-2, 2x^2+5x-3$, and $-x^2-4x+4$ generate $P_2(\mathbb{R})$ because each of the three given polynomials belongs to $P_2(\mathbb{R})$ and each polynomial $ax^2 + bx + c \in P_2(\mathbb{R})$ is a linear combination of these three. It can be shown that we can find $a, b, c \in \mathbb{R}$ such that

$$a(x^2 + 3x - 2) + b(2x^2 + 5x - 3) + c(-x^2 - 4x + 4) = ax^2 + bx + c.$$

Example 16 (Matrices)**.** Note that the matrices

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

generate $M_{2\times2}(\mathbb{R})$ because an every $A \in M_{2\times2}(\mathbb{R})$ can be expressed as a linear combination of the four given matrices found below where there exists scalars $a, b, c, d \in \mathbb{R}$ such that

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + b\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + c\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + d\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

where it can be show that

$$\begin{aligned}
a &= \frac{1}{3}a_{11} + \frac{1}{3}a_{12} + \frac{1}{3}a_{21} - \frac{2}{3}a_{22} \\
b &= \frac{1}{3}a_{11} + \frac{1}{3}a_{12} - \frac{2}{3}a_{21} + \frac{1}{3}a_{22} \\
c &= \frac{1}{3}a_{11} - \frac{2}{3}a_{12} + \frac{1}{3}a_{21} + \frac{1}{3}a_{22} \\
d &= -\frac{2}{3}a_{11} + \frac{1}{3}a_{12} + \frac{1}{3}a_{21} + \frac{1}{3}a_{22}.
\end{aligned}$$

A non-example set of matrices are

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

do not generate $M_{2\times 2}(\mathbb{R})$ since all of the matrices have equal diagonal entries. Thus, every the set of matrices above cannot generate every $A \in M_{2\times 2}(\mathbb{R})$.

1.4 Linear Dependence and Linear Independence

1.4.1 Motivation

Suppose V is a vector space over a field F and that $W \subseteq V$ is a subspace of V.

- Suppose we have a set S that will generate every vector in V as a linear combination of vectors in S.
- It is desirable to find a subset of S that is as small as possible or rather just enough to generate each vector in V.
- The reason for this is to reduce the amount of computations done to represent a vector in V.
- Finding this subset is a matter of expressing one of the vectors in the spanning set as a linear combination of the other vectors in S.
- This will naturally lead to a system of linear equations for which we have to solve for the coefficients.
- The answer to the question is whether this system of equations leads to a solution or not.
- An easier way to answer this question is to find coefficients (not all zero) for which the zero vector can be written as a linear combination of vectors in S.
- Doing this, guarantees that we can write any one of the vectors in S as a linear combination of the others.

Definition 1.4.1 (Linearly Dependence)**.** A subset S of a vector space V is called **linearly dependent** if there exists a finite number of distinct vectors in $u_1, u_2, \ldots, u_n \in S$ and scalars $a_1, a_2, \ldots, a_n$ not all zero, such that

$$a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0.$$

In this case, we also say that the vectors of S are linearly dependent.

- The **trivial representation** of 0 is a linear combination of scalars $a_1, a_2, \ldots, a_n \in F$ and distinct vectors $v_1, v_2, \ldots, v_n \in S$ where for all $1 \le i \le n$, we have $a_i = 0$.
- This tells us that our definition of **linear dependence** implies that 0 is a non-trivial linear combination.
- Any subset of V that contains the zero vector is subsequently linearly dependent since 0 can be written as non-trivial representation; that is, $1 \cdot 0 = 0$.

1.4.2 Examples of Linearly Dependent Sets

Example 17. Consider a subset in $\mathbb{R}^4$ defined by

$$S = \{(1,3,-4,2),(2,2,-4,0),(1,-3,2,-4),(-1,0,1,0)\}.$$

To show that S is a linear dependent set, we need to find scalars $a_1, a_2, a_3, a_4 \in \mathbb{R}$ such that a_i for all $1 \le i \le 4$ not all zero such that

$$a_1(1,3,-3,2) + a_2(2,2,-4,0) + a_3(1,-3,2,-4) + a_4(-1,0,1,0) = 0.$$

We can represent this as a system of linear equations with each equation being set equal to zero and solving for the scalars. Solving this system of equations leads to the coefficients $a_1 = 4, a_2 = -3, a_3 = 2$, and $a_4 = 0$. Thus, we have that S is linearly dependent subset of $\mathbb{R}^4$ and hence we can write any vector in S as a linear combination of the other vectors contained in S.

Example 18. Define a subset of $M_{2\times 2}(\mathbb{R})$

$$S = \left\{ \begin{pmatrix} 1 & -3 & 2 \\ -4 & 0 & 5 \end{pmatrix} \}, \begin{pmatrix} -3 & 7 & 4 \\ 6 & -2 & -7 \end{pmatrix}, \begin{pmatrix} -2 & 3 & 11 \\ -1 & -3 & 2 \end{pmatrix} \right\}.$$

This set is **linearly dependent** because we can find coefficients a_1, a_2, a_3 such that $a_1 = 5, a_2 = 3$, and $a_3 = -2$ where

$$5\begin{pmatrix} 1 & -3 & 2 \\ -4 & 0 & 5 \end{pmatrix} + 3\begin{pmatrix} -3 & 7 & 4 \\ 6 & -2 & -7 \end{pmatrix} - 2\begin{pmatrix} -2 & 3 & 11 \\ -1 & -3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Definition 1.4.2 (Linear Independence)**.** A subset S of a vector space that is not linearly dependent is called **linearly independent**. As before, we also say that the vectors of S are linearly independent.

Whereas linear dependence requires solutions that are not all zero, linear independence **requires** all the solutions to be zero. This provides us an easy way to determine if a finite set is linearly independent.

1.4.3 Examples of Linear Independent Sets

Some facts about linearly independent sets include:

- The empty set is linearly independent since we don't have any vectors to take linear combinations of.
- A set consisting of only one non-zero vector; that is, $\{v\}$ is linearly independent.
- If $\{v\}$ is linearly dependent, then it the singleton has to be the zero vector 0. This is because $au = 0$

$$u = 1 \cdot u = (a^{-1}a)u = a^{-1}(au) = a^{-1} \cdot 0 = 0.$$

Example 19 (A Set of Vectors in $\mathbb{R}^4$)**.** It can be shown that the set

$$S = \{(1,0,0,-1),(0,1,0,-1),(0,0,1,-1),(0,0,0,1)\}$$

is linearly independent by showing that all the coefficients of the linear combination of 0 represent the trivial representation; that is, we have scalars $a_1 = a_2 = a_3 = a_4 = 0$.

Example 20. For $k = 0, 1, \ldots, n$, let $p_k = x^k + x^{k+1} + \cdots + x^n$. The set

$$\{p_0(x), p_1(x), \ldots p_n(x)\}$$

is linearly independent in $P_n(F)$. It can be shown that for some scalars $a_0, a_1, \dots, a_n \in F$, the following equation

$$a_0 + (a_0 + a_1 + a_2)x^2 + \cdots + (a_0 + a_1 + \cdots + a_n)x^n = 0$$

has the trivial-representation; that is, $a_i = 0$ for all $1 \leq i \leq n$.

Theorem 1.4.1. Let V be a vector space, and let $S_1 \subseteq S_2 \subseteq V$. If S_1 is linearly dependent, then S_2 is linearly dependent.

Proof. Let V be a vector space. Suppose S_1 is linearly dependent. Then there exists a finite number of distinct vectors $v_1, v_2, \dots, v_n \in S_1$ and scalars $a_1, a_2, \dots, a_n \in F$ not all zero such that

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0.$$

Since $S_1 \subseteq S_2$, we must have $v_1, v_2, \dots, v_n \in S_2$ as well. If these vectors are linearly dependent, then these vectors are also linearly dependent in S_2; that is, we have scalars $a_1, a_2, \dots, a_n \in F$ not all zero such that

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0.$$

Hence, S_2 is linearly dependent. ■

Corollary. Let V be a vector space, and let $S_1 \subseteq S_2 \subseteq V$. If S_2 is linearly independent, then S_1 is linearly independent.

Proof. Note that this corollary is just the contrapositive of the theorem before it. Hence, S_1 is linearly independent. ■

- Determining whether there exists a minimal generating set for the span of S is related to solving the problem of finding whether some vector in S that can be written in terms of a linear combination of other vectors in S.
- We can see that in a given subset of $\mathbb{R}^3$ defined by

  $$S = \{u_1, u_2, u_3, u_4\}$$

  where $u_1 = (2, -1, 4), u_2 = (1, -1, 3), u_3 = (1, 1, -1)$, and $u_4 = (1, -2, 1)$. Note that this S is linearly dependent. This tells us that any vector in S can be written as a linear combination of the others. Suppose we pick u_3 and write as a linear combination of the vectors u_1, u_2, and u_4. We would find that the span of these three vectors (denote this set as S') generates the same spanning set S but with a cardinality one less than S; that is, $\text{span}(S) = \text{span}(S')$.
- If we find that there does not exist a proper subset that is equivalent to the span of S, then S must be a linearly independent set.

Theorem 1.4.2. Let S be a linearly independent subset of a vector space V, and let $v \in V$ but not in S. Then $S \cup \{v\}$ is linearly dependent if and only if $v \in \text{span}(S)$.

Proof. ($\Rightarrow$) Suppose $S \cup \{v\}$ is a linearly dependent set. Then there exists a finite number of scalars $a_1, a_2, \dots, a_n \in F$ and vectors $u_1, u_2, \dots, u_n \in S \cup \{v\}$ such that

$$a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0.$$

Since S is a linearly independent set, we must have $u_i = v$ for some $1 \leq i \leq n$. Choose $i = 1$ (any choice of i will do) such that

$$a_1v + a_2u_2 + \cdots + a_nu_n = 0.$$

Solving for v by subtracting a_1v on both sides, multiplying by a_1^{-1} on both sides of the equation, and distributing by a_1^{-1} yields the following:

$$v = (-a_1^{-1}a_2)u_2 + (-a_1^{-1}a_3)u_3 + \cdots + (-a_1^{-1}a_n)u_n.$$

Since v is a linear combination of vectors $u_1, u_2, \ldots, u_n \in S$, we know that $v \in \text{span}(S)$.

($\Leftarrow$) Conversely, suppose $v \in \text{span}(S)$. This implies that there exists a finite amount of scalars $a_1, a_2, \ldots, a_n \in F$ and $u_1, u_2, \ldots, u_n \in S$ such that

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n.$$

Subtracting v from both sides of the equation above gives us the following equation:

$$a_1u_1 + a_2u_2 + \cdots + a_nu_n - v = 0.$$

Note that v is not contained in S, so $v \neq u_i$ for all $1 \leq i \leq n$. Since S is a linearly independent set, we know that $a_i = 0$ for all $1 \leq i \leq n$. This implies that the only coefficient that is non-zero is with $-v = -1v$. Hence, the set of vectors $S' = \{u_1, u_2, \ldots, u_n, v\}$ is linearly dependent. Since $S' \subseteq S \cup \{v\}$, we know that $S \cup \{v\}$ is also linearly independent by Theorem 6. ■

1.5 Bases And Dimension

- Recall that S is a generating set for a subspace W and no proper subset of S is a generating set for W, then S must be linearly independent.
- Linearly independent sets possess the unique property that every vector that its spanning set generates is unique.
- This is property is what allows generating sets to be the building blocks of vector spaces.

Definition 1.5.1 (Basis). A **basis** β for a vector space V is linearly independent subset of V that generates V. If β is a basis for V, we also say that the vectors of β form a basis for V.

Example 21.

- Recall that the empty set $\emptyset$ is linearly independent and that $\text{span}(\emptyset) = \{0\}$. The empty set $\emptyset$ in this case is the basis for the zero vector space.
- Note that in F^n, the vectors $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0) \ldots, e_n = (0, 0, \ldots, 0, 1)$ form a basis for F^n.
- The basis for $M_{m\times n}(F)$ is the set of matrices E^{ij} such that the only nonzero entry is a 1 in the ith and jth column.
- As we have seen in the last section, the set $\{1, x, x^2, \ldots, x^n\}$ is a basis for $P_n(F)$.
- In $P(F)$, the set $\{1, x, xx^2, \ldots\}$ is a basis. *Bases are not limited to finite sets. They can be infinite.*

Theorem 1.5.1. Let V be a vector space and $u_1, u_2, \ldots, u_n$ be distinct vectors in V. Then $\beta = \{u_1, u_2, \ldots, u_n\}$ is a basis for V if and only if each $v \in V$ can be unique expressed as a linear combination of vectors in β, that is, expressed in the form

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$$

for unique scalars $a_1, a_2, \ldots, a_n$.

Proof. ($\Rightarrow$) Suppose $\beta = \{u_1, u_2, \ldots, u_n\}$ is a basis for V. Then $\text{span}(\beta) = V$. If $v \in V$, then $v \in \text{span}(\beta)$. Hence, we can write v as a linear combination of vectors in β such that choosing scalars $a_1, a_2, \dots, a_n \in F$ leads to

$$v = \sum_{i=1}^{n} a_i u_i.$$

Suppose there exists another representation of $v \in V$ such that

$$v = \sum_{i=1}^{n} b_i x_i$$

Hence, observe that

$$\begin{aligned} \sum_{i=1}^{n} a_i x_i &= \sum_{i=1}^{n} b_i y_i \\ \Rightarrow \sum_{i=1}^{n} (a_i - b_i) x_i &= 0. \end{aligned}$$

Since β is linearly independent, we know that $a_i - b_i = 0$ which implies $a_i = b_i$ for all $1 \leq i \leq n$. Hence, v can be expressed as a unique linear combination of vectors in β.

($\Leftarrow$) Let $v \in V$. Then v can be uniquely expressed as a linear combination of vectors in $\beta = \{u_1, u_2, \dots, u_n\}$ such that

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n \tag{1}$$

for unique a_i for all $1 \leq i \leq n$. Suppose for sake of contradiction that β is **NOT** a basis for V. Then either β is linearly dependent or $\operatorname{span}(\beta) \neq V$. Suppose $\operatorname{span}(\beta) \neq V$, then $v \in V$ cannot be written a linear combination of vectors in β which is a contradiction. Suppose β is linearly dependent. Then either $u_1 = 0$ or $u_{k+1} \in \operatorname{span}(\{u_1, u_2, \dots, u_k\})$ for some $1 \leq k < n$. If $u_1 = 0$, then (1) can be re-written as

$$v = a_1 0 + a_2 u_2 + \cdots + a_n u_n.$$

This implies that $a_1 \in F$ can be made arbitrary such that v can be written in another representation. But this is a contradiction since we assumed that v contains a unique linear combination. Now, suppose $u_{k+1} \in \operatorname{span}(\{u_1, u_2, \dots, u_k\})$. Then we have

$$u_{k+1} = \delta_1 u_1 + \delta_2 u_2 + \cdots + \delta_k u_k.$$

Substituting this equation for u_{k+1} in (1) produces the following equation:

$$\begin{aligned} v &= a_1 u_1 + a_2 u_2 + \cdots + a_k u_k + a_{k+1} u_{k+1} + \cdots a_n u_n \\ &= a_1 u_1 + a_2 u_2 + \cdots + a_k u_k + a_{k+1}(\delta_1 u_1 + \delta_2 u_2 + \cdots + \delta_k u_k) + \cdots + a_n u_n \\ &= (a_1 + a_{k+1}\delta_1) u_1 + (a_2 + a_{k+1}\delta_2) u_2 + \cdots + (a_k + a_{k+1}\delta_k) u_k + \cdots + a_n u_n. \end{aligned} \tag{2}$$

Since v has a unique representation, we must have

$$a_i = a_i + a_{k+1}\delta_i \Rightarrow a_{k+1}\delta_i = 0 \tag{3}$$

for all $1 \leq i \leq k$. Assuming that $a_{k+1} \neq 0$, the only way for $a_{k+1}\delta_i = 0$ is if $\delta_i = 0$ for all $1 \leq i \leq k$. But note that not all δ_i are zero since $\{u_1, u_2, \dots, u_k\} \subseteq \beta$ is linearly dependent by Exercise 16. Hence, there exists at least one i such that $a_{k+1}\delta_i \neq 0$ which is a contradiction. Hence, β must be a basis for V.

■

- Any vector $v \in V$ can be written as a linear combination of vectors from the basis containing $u_1, u_2, \dots, u_n \in V$.
- This determines a unique n-tuple of scalars $(a_1, a_2, \dots, a_n)$ and conversely, each n-tuple of scalars determines a unique vector $v \in V$ such that each coefficient from the linear combination of $u_1, u_2, \dots, u_n$ is an entry from said tuple.

- For example, in our vector space F^n, n is the number of vectors that should be in the basis for F^n which is indeed the case.
- In this book, we are only concerned with finite bases.

Theorem 1.5.2. If a vector space V is generated by a finite set S, then some subset of S is a basis for V. Hence, V has a finite basis.

Proof. ■