5 Vector Spaces

As suggested at the end of chapter 4, the vector spaces $\mathbb{R}^n$ are not the only vector spaces. We now give a general definition that includes $\mathbb{R}^n$ for all values of n, and $\mathbb{R}^S$ for all sets S, and more. This mathematical structure is applicable to a wide range of real-world problems and allows for tremendous economy of thought; the idea of a basis for a vector space will drive home the main idea of vector spaces; they are sets with very simple structure.

The two key properties of vectors are that they can be added together and multiplied by scalars. Thus, before giving a rigorous definition of vector spaces, we restate the main idea.

> A vector space is a set that is closed under addition and scalar multiplication.

Definition A **vector space** $(V, +, . , \mathbb{R})$ is a set V with two operations $+$ and $\cdot$ satisfying the following properties for all $u, v \in V$ and $c, d \in \mathbb{R}$:

(+i) (Additive Closure) $u + v \in V$. *Adding two vectors gives a vector.*

(+ii) (Additive Commutativity) $u + v = v + u$. *Order of addition does not matter.*

(+iii) (Additive Associativity) $(u + v) + w = u + (v + w)$. *Order of adding many vectors does not matter.*

(+iv) (Zero) There is a special vector $0_V \in V$ such that $u + 0_V = u$ for all u in V.

(+v) (Additive Inverse) For every $u \in V$ there exists $w \in V$ such that $u + w = 0_V$.

($\cdot$ i) (Multiplicative Closure) $c \cdot v \in V$. *Scalar times a vector is a vector.*

($\cdot$ ii) (Distributivity) $(c+d) \cdot v = c \cdot v + d \cdot v$. *Scalar multiplication distributes over addition of scalars.*

($\cdot$ iii) (Distributivity) $c \cdot (u+v) = c \cdot u + c \cdot v$. *Scalar multiplication distributes over addition of vectors.*

($\cdot$ iv) (Associativity) $(cd) \cdot v = c \cdot (d \cdot v)$.

($\cdot$ v) (Unity) $1 \cdot v = v$ for all $v \in V$.

`Examples of each rule`

Remark Rather than writing $(V, +, ., \mathbb{R})$, we will often say "let V be a vector space over $\mathbb{R}$". If it is obvious that the numbers used are real numbers, then "let V be a vector space" suffices. Also, don't confuse the scalar product $\cdot$ with the dot product $\bullet$. The scalar product is a function that takes as its two inputs one number and one vector and returns a vector as its output. This can be written

$$\cdot \colon \mathbb{R} \times V \to V .$$

Similarly

$$+ : V \times V \to V .$$

On the other hand, the dot product takes two vectors and returns a number. Succinctly: $\bullet \colon V \times V \to \mathbb{R}$. Once the properties of a vector space have been verified, we'll just write scalar multiplication with juxtaposition $cv = c \cdot v$, though, to keep our notation efficient.

5.1 Examples of Vector Spaces

One can find many interesting vector spaces, such as the following:

Example of a vector space

Example 58

$$\mathbb{R}^{\mathbb{N}} = \{f \mid f\colon \mathbb{N} \rightarrow \mathbb{R}\}$$

Here the vector space is the set of functions that take in a natural number n and return a real number. The addition is just addition of functions: $(f_1+f_2)(n) = f_1(n)+f_2(n)$. Scalar multiplication is just as simple: $c \cdot f(n) = cf(n)$.

We can think of these functions as infinitely large ordered lists of numbers: $f(1) = 1^3 = 1$ is the first component, $f(2) = 2^3 = 8$ is the second, and so on. Then for example the function $f(n) = n^3$ would look like this:

$$f = \begin{pmatrix} 1 \\ 8 \\ 27 \\ \vdots \\ n^3 \\ \vdots \end{pmatrix}.$$

Thinking this way, $\mathbb{R}^{\mathbb{N}}$ is the space of all infinite sequences. Because we can not write a list infinitely long (without infinite time and ink), one can not define an element of this space explicitly; definitions that are implicit, as above, or algebraic as in $f(n) = n^3$ (for all $n \in \mathbb{N}$) suffice.

Let's check some axioms.

(+i) (Additive Closure) $(f_1 + f_2)(n) = f_1(n) + f_2(n)$ is indeed a function $\mathbb{N} \rightarrow \mathbb{R}$, since the sum of two real numbers is a real number.

(+iv) (Zero) We need to propose a zero vector. The constant zero function $g(n) = 0$ works because then $f(n) + g(n) = f(n) + 0 = f(n)$.

The other axioms should also be checked. This can be done using properties of the real numbers.

Reading homework: problem 1

Example 59 The space of functions of one real variable.

$$\mathbb{R}^{\mathbb{R}} = \{f \mid f\colon \mathbb{R} \rightarrow \mathbb{R}\}$$

The addition is point-wise

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

as is scalar multiplication

$$c\cdot f(x)=cf(x)\,.$$

To check that $\mathbb{R}^{\mathbb{R}}$ is a vector space use the properties of addition of functions and scalar multiplication of functions as in the previous example.

We can not write out an explicit definition for one of these functions either, there are not only infinitely many components, but even infinitely many components between any two components! You are familiar with algebraic definitions like $f(x)=e^{x^2-x+5}$. However, most vectors in this vector space can not be defined algebraically. For example, the nowhere continuous function

$$f(x)=\begin{cases}1\,, & x\in\mathbb{Q}\\ 0\,, & x\notin\mathbb{Q}\end{cases}\,.$$

Example 60 $\mathbb{R}^{\{*,\star,\#\}}=\{f:\{*,\star,\#\}\to\mathbb{R}\}$. Again, the properties of addition and scalar multiplication of functions show that this is a vector space.

You can probably figure out how to show that $\mathbb{R}^S$ is vector space for any set S. This might lead you to guess that all vector spaces are of the form $\mathbb{R}^S$ for some set S. The following is a counterexample.

Example 61 Another very important example of a vector space is the space of all differentiable functions:

$$\left\{f:\mathbb{R}\to\mathbb{R}\,\middle|\,\frac{d}{dx}f\text{ exists}\right\}.$$

From calculus, we know that the sum of any two differentiable functions is differentiable, since the derivative distributes over addition. A scalar multiple of a function is also differentiable, since the derivative commutes with scalar multiplication ($\frac{d}{dx}(cf)=c\frac{d}{dx}f$). The zero function is just the function such that $0(x)=0$ for every x. The rest of the vector space properties are inherited from addition and scalar multiplication in $\mathbb{R}$.

Similarly, the set of functions with at least k derivatives is always a vector space, as is the space of functions with infinitely many derivatives. None of these examples can be written as $\mathbb{R}^S$ for some set S. Despite our emphasis on such examples, it is also not true that all vector spaces consist of functions. Examples are somewhat esoteric, so we omit them.

Another important class of examples is vector spaces that live inside $\mathbb{R}^n$ but are not themselves $\mathbb{R}^n$.

Example 62 (Solution set to a homogeneous linear equation.)
Let
$$M = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}.$$
The solution set to the homogeneous equation $Mx = 0$ is
$$\left\{ c_1 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \middle| c_1, c_2 \in \mathbb{R} \right\}.$$
This set is not equal to $\mathbb{R}^3$ since it does not contain, for example, $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. The sum of any two solutions is a solution, for example
$$\left[2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right] + \left[7 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + 5 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right] = 9 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + 8 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$
and any scalar multiple of a solution is a solution
$$4 \left[5 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - 3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right] = 20 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - 12 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$
This example is called a *subspace* because it gives a vector space inside another vector space. See chapter 9 for details. Indeed, because it is determined by the linear map given by the matrix M, it is called $\ker M$, or in words, the *kernel* of M, for this see chapter 16.

Similarly, the solution set to any homogeneous linear equation is a vector space: Additive and multiplicative closure follow from the following statement, made using linearity of matrix multiplication:

If $Mx_1 = 0$ and $Mx_2 = 0$ then $M(c_1x_1+c_2x_2) = c_1Mx_1+c_2Mx_2 = 0+0 = 0$.

A powerful result, called the subspace theorem (see chapter 9) guarantees, based on the closure properties alone, that homogeneous solution sets are vector spaces.

More generally, if V is any vector space, then any hyperplane through the origin of V is a vector space.

Example 63 Consider the functions $f(x) = e^x$ and $g(x) = e^{2x}$ in $\mathbb{R}^\mathbb{R}$. By taking combinations of these two vectors we can form the plane $\{c_1 f + c_2 g | c_1, c_2 \in \mathbb{R}\}$ inside of $\mathbb{R}^\mathbb{R}$. This is a vector space; some examples of vectors in it are $4e^x - 31e^{2x}$, $\pi e^{2x} - 4e^x$ and $\frac{1}{2}e^{2x}$.

A hyperplane which does not contain the origin cannot be a vector space because it fails condition (+iv).

It is also possible to build new vector spaces from old ones using the product of sets. Remember that if V and W are sets, then their product is the new set

$$V \times W = \{(v, w) | v \in V, w \in W\},$$

or in words, all ordered pairs of elements from V and W. In fact $V \times W$ is a vector space if V and W are. We have actually been using this fact already:

Example 64 The real numbers $\mathbb{R}$ form a vector space (over $\mathbb{R}$). The new vector space

$$\mathbb{R} \times \mathbb{R} = \{(x, y) | x \in \mathbb{R}, y \in \mathbb{R}\}$$

has addition and scalar multiplication defined by

$$(x, y) + (x', y') = (x + x', y + y') \text{ and } c.(x, y) = (cx, cy).$$

Of course, this is just the vector space $\mathbb{R}^2 = \mathbb{R}^{\{1,2\}}$.

5.1.1 Non-Examples

The solution set to a linear non-homogeneous equation is not a vector space because it does not contain the zero vector and therefore fails (iv).

Example 65 The solution set to

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

is $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} -1 \\ 1 \end{pmatrix} \middle| c \in \mathbb{R} \right\}$. The vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is not in this set.

Do notice that if just one of the vector space rules is broken, the example is not a vector space.

Most sets of n-vectors are not vector spaces.

Example 66 $P := \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \middle| \, a, b \geq 0 \right\}$ is not a vector space because the set fails ($\cdot$i) since $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \in P$ but $-2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \end{pmatrix} \notin P$.

Sets of functions other than those of the form $\mathbb{R}^S$ should be carefully checked for compliance with the definition of a vector space.

Example 67 The set of all functions which are nowhere zero

$$\{f \colon \mathbb{R} \to \mathbb{R} \mid f(x) \neq 0 \text{ for any } x \in \mathbb{R}\} ,$$

does not form a vector space because it does not satisfy (+i). The functions $f(x) = x^2+1$ and $g(x) = -5$ are in the set, but their sum $(f+g)(x) = x^2-4 = (x+2)(x-2)$ is not since $(f+g)(2) = 0$.

5.2 Other Fields

Above, we defined vector spaces over the real numbers. One can actually define vector spaces over any *field*. This is referred to as choosing a different *base field*. A field is a collection of "numbers" satisfying properties which are listed in appendix B. An example of a field is the complex numbers,

$$\mathbb{C} = \big\{ x + iy \mid i^2 = -1, x, y \in \mathbb{R} \big\} .$$

Example 68 In quantum physics, vector spaces over $\mathbb{C}$ describe all possible states a physical system can have. For example,

$$V = \left\{ \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \mid \lambda, \mu \in \mathbb{C} \right\}$$

is the set of possible states for an electron's spin. The vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ describe, respectively, an electron with spin "up" and "down" along a given direction. Other vectors, like $\begin{pmatrix} i \\ -i \end{pmatrix}$ are permissible, since the base field is the complex numbers. Such states represent a mixture of spin up and spin down for the given direction (a rather counterintuitive yet experimentally verifiable concept), but a given spin in some other direction.

Complex numbers are very useful because of a special property that they enjoy: every polynomial over the complex numbers factors into a product of linear polynomials. For example, the polynomial

$$x^2 + 1$$

doesn't factor over real numbers, but over complex numbers it factors into

$$(x + i)(x - i)\,.$$

In other words, there are *two* solutions to

$$x^2 = -1,$$

$x = i$ and $x = -i$. This property has far-reaching consequences: often in mathematics problems that are very difficult using only real numbers become relatively simple when working over the complex numbers. This phenomenon occurs when diagonalizing matrices, see chapter 13.

The rational numbers $\mathbb{Q}$ are also a field. This field is important in computer algebra: a real number given by an infinite string of numbers after the decimal point can't be stored by a computer. So instead rational approximations are used. Since the rationals are a field, the mathematics of vector spaces still apply to this special case.

Another very useful field is bits

$$B_2 = \mathbb{Z}_2 = \{0, 1\}\,,$$

with the addition and multiplication rules

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

These rules can be summarized by the relation $2 = 0$. For bits, it follows that $-1 = 1$!

The theory of fields is typically covered in a class on abstract algebra or Galois theory.

5.3 Review Problems

Webwork:

Reading problems	1
Addition and inverse	2

1. Check that $\left\{\begin{pmatrix}x\\y\end{pmatrix}\middle| x,y\in\mathbb{R}\right\}=\mathbb{R}^2$ (with the usual addition and scalar multiplication) satisfies all of the parts in the definition of a vector space.

2. (a) Check that the complex numbers $\mathbb{C}=\{x+iy\mid i^2=-1,x,y\in\mathbb{R}\}$, satisfy all of the parts in the definition of a vector space over $\mathbb{C}$. Make sure you state carefully what your rules for vector addition and scalar multiplication are.

   (b) What would happen if you used $\mathbb{R}$ as the base field (try comparing to problem 1).

3. (a) Consider the set of convergent sequences, with the same addition and scalar multiplication that we defined for the space of sequences:

$$V=\left\{f\mid f\colon\mathbb{N}\to\mathbb{R},\lim_{n\to\infty}f(n)\in\mathbb{R}\right\}\subset\mathbb{R}^{\mathbb{N}}.$$

   Is this still a vector space? Explain why or why not.

   (b) Now consider the set of divergent sequences, with the same addition and scalar multiplication as before:

$$V=\left\{f\mid f\colon\mathbb{N}\to\mathbb{R},\lim_{n\to\infty}f(n)\text{ does not exist or is }\pm\infty\right\}\subset\mathbb{R}^{\mathbb{N}}.$$

   Is this a vector space? Explain why or why not.

4. Consider the set of 2×4 matrices:

$$V=\left\{\begin{pmatrix}a&b&c&d\\e&f&g&h\end{pmatrix}\middle| a,b,c,d,e,f,g,h\in\mathbb{C}\right\}$$

Propose definitions for addition and scalar multiplication in V. Identify the zero vector in V, and check that every matrix in V has an additive inverse.

5. Let $P_3^{\mathbb{R}}$ be the set of polynomials with real coefficients of degree three or less.

   (a) Propose a definition of addition and scalar multiplication to make $P_3^{\mathbb{R}}$ a vector space.

   (b) Identify the zero vector, and find the additive inverse for the vector $-3-2x+x^2$.

   (c) Show that $P_3^{\mathbb{R}}$ is not a vector space over $\mathbb{C}$. Propose a small change to the definition of $P_3^{\mathbb{R}}$ to make it a vector space over $\mathbb{C}$. (Hint: Every little symbol in the the instructions for par (c) is important!)

Hint

6. Let $V=\{x\in\mathbb{R}|x>0\}=:\mathbb{R}_+$. For $x,y\in V$ and $\lambda\in\mathbb{R}$, define

$$x\oplus y=xy\,,\qquad \lambda\otimes x=x^\lambda\,.$$

   Show that $(V,\oplus,\otimes,\mathbb{R})$ is a vector space.

7. The component in the ith row and jth column of a matrix can be labeled m^i_j. In this sense a matrix is a function of a pair of integers. For what set S is the set of 2×2 matrices the same as the set $\mathbb{R}^S$? Generalize to other size matrices.

8. Show that any function in $\mathbb{R}^{\{*,\star,\#\}}$ can be written as a sum of multiples of the functions $e_*, e_\star, e_\#$ defined by

$$e_*(k)=\begin{cases}1\,, & k=*\\ 0\,, & k=\star\\ 0\,, & k=\#\end{cases}\,,\ e_\star(k)=\begin{cases}0\,, & k=*\\ 1\,, & k=\star\\ 0\,, & k=\#\end{cases}\,,\ e_\#(k)=\begin{cases}0\,, & k=*\\ 0\,, & k=\star\\ 1\,, & k=\#\end{cases}\,.$$

9. Let V be a vector space and S any set. Show that the set V^S of all functions $S\to V$ is a vector space. *Hint:* first decide upon a rule for adding functions whose outputs are vectors.

9

Subspaces and Spanning Sets

It is time to study vector spaces more carefully and return to some fundamental questions:

1. *Subspaces*: When is a subset of a vector space itself a vector space? (This is the notion of a *subspace*.)

2. *Linear Independence*: Given a collection of vectors, is there a way to tell whether they are independent, or if one is a "linear combination" of the others?

3. *Dimension*: Is there a consistent definition of how "big" a vector space is?

4. *Basis*: How do we label vectors? Can we write any vector as a sum of some basic set of vectors? How do we change our point of view from vectors labeled one way to vectors labeled in another way?

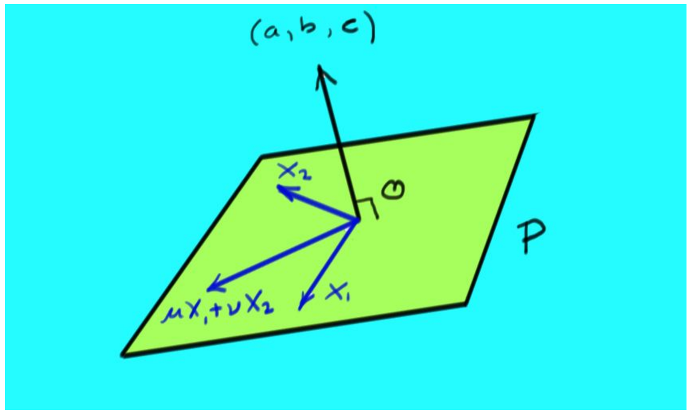
Let's start at the top!

9.1 Subspaces

Definition We say that a subset U of a vector space V is a **subspace** of V if U is a vector space under the inherited addition and scalar multiplication operations of V.

Example 109 Consider a plane P in $\mathbb{R}^3$ through the origin:

$$ax + by + cz = 0.$$



This equation can be expressed as the homogeneous system $\begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$, or $MX = 0$ with M the matrix $\begin{pmatrix} a & b & c \end{pmatrix}$. If X_1 and X_2 are both solutions to $MX = 0$, then, by linearity of matrix multiplication, so is $\mu X_1 + \nu X_2$:

$$M(\mu X_1 + \nu X_2) = \mu M X_1 + \nu M X_2 = 0.$$

So P is closed under addition and scalar multiplication. Additionally, P contains the origin (which can be derived from the above by setting $\mu = \nu = 0$). All other vector space requirements hold for P because they hold for all vectors in $\mathbb{R}^3$.

Theorem 9.1.1 (Subspace Theorem)**.** *Let U be a non-empty subset of a vector space V. Then U is a subspace if and only if $\mu u_1 + \nu u_2 \in U$ for arbitrary u_1, u_2 in U, and arbitrary constants μ, ν.*

Proof. One direction of this proof is easy: if U is a subspace, then it is a vector space, and so by the additive closure and multiplicative closure properties of vector spaces, it has to be true that $\mu u_1 + \nu u_2 \in U$ for all u_1, u_2 in U and all constants constants μ, ν.

The other direction is almost as easy: we need to show that if $\mu u_1 + \nu u_2 \in U$ for all u_1, u_2 in U and all constants μ, ν, then U is a vector space. That is, we need to show that the ten properties of vector spaces are satisfied. We already know that the additive closure and multiplicative closure properties are satisfied. Further, U has all of the other eight properties because V has them. □

Note that the requirements of the subspace theorem are often referred to as "closure".

We can use this theorem to check if a set is a vector space. That is, if we have some set U of vectors that come from some bigger vector space V, to check if U itself forms a smaller vector space we need check only two things:

1. If we add any two vectors in U, do we end up with a vector in U?

2. If we multiply any vector in U by any constant, do we end up with a vector in U?

If the answer to both of these questions is yes, then U is a vector space. If not, U is not a vector space.

Reading homework: problem 1

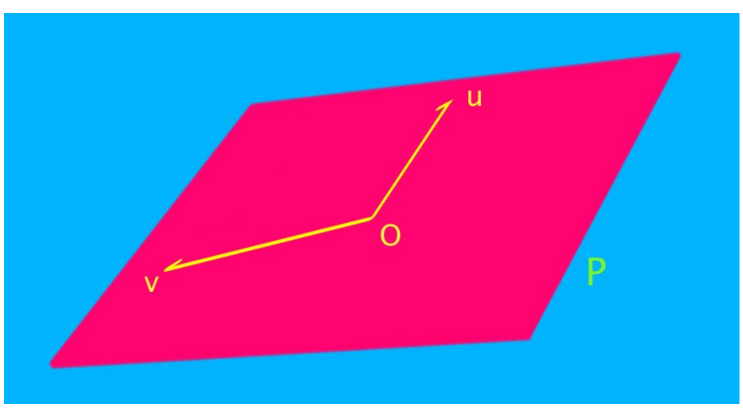
9.2 Building Subspaces

Consider the set

$$U = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \subset \mathbb{R}^3 .$$

Because U consists of only two vectors, it clear that U is *not* a vector space, since any constant multiple of these vectors should also be in U. For example, the 0-vector is not in U, nor is U closed under vector addition.

But we know that any two vectors define a plane:



In this case, the vectors in U define the xy-plane in $\mathbb{R}^3$. We can view the xy-plane as the set of all vectors that arise as a linear combination of the two vectors in U. We call this set of all linear combinations the *span* of U:

$$\operatorname{span}(U) = \left\{ x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \middle| x, y \in \mathbb{R} \right\}.$$

Notice that any vector in the xy-plane is of the form

$$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in \operatorname{span}(U).$$

Definition Let V be a vector space and $S = \{s_1, s_2, \ldots\} \subset V$ a subset of V. Then the **span of** S, denoted $\operatorname{span}(S)$, is the set

$$\operatorname{span}(S) := \{r^1 s_1 + r^2 s_2 + \cdots + r^N s_N \mid r^i \in \mathbb{R}, N \in \mathbb{N}\}.$$

That is, the span of S is the set of all finite linear combinations[1] of elements of S. Any *finite* sum of the form "a constant times s_1 plus a constant times s_2 plus a constant times s_3 and so on" is in the span of S.[2].

Example 110 Let $V = \mathbb{R}^3$ and $X \subset V$ be the x-axis. Let $P = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and set

$$S = X \cup \{P\}.$$

The vector $\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$ is in $\operatorname{span}(S)$, because $\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. Similarly, the vector $\begin{pmatrix} -12 \\ 17.5 \\ 0 \end{pmatrix}$ is in $\operatorname{span}(S)$, because $\begin{pmatrix} -12 \\ 17.5 \\ 0 \end{pmatrix} = \begin{pmatrix} -12 \\ 0 \\ 0 \end{pmatrix} + 17.5 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. Similarly, any vector

[1]Usually our vector spaces are defined over $\mathbb{R}$, but in general we can have vector spaces defined over different base fields such as $\mathbb{C}$ or $\mathbb{Z}_2$. The coefficients r^i should come from whatever our base field is (usually $\mathbb{R}$).

[2]It is important that we only allow finitely many terms in our linear combinations; in the definition above, N must be a finite number. It can be any finite number, but it must be finite. We can relax the requirement that $S = \{s_1, s_2, \ldots\}$ and just let S be any set of vectors. Then we shall write $\operatorname{span}(S) := \{r^1 s_1 + r^2 s_2 + \cdots + r^N s_N \mid r^i \in \mathbb{R}, s_i \in S, N \in \mathbb{N}, \}$

of the form

$$x\begin{pmatrix}1\\0\\0\end{pmatrix}+y\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}x\\y\\0\end{pmatrix}$$

is in $\operatorname{span}(S)$. On the other hand, any vector in $\operatorname{span}(S)$ must have a zero in the z-coordinate. (Why?) So $\operatorname{span}(S)$ is the xy-plane, which is a vector space. (Try drawing a picture to verify this!)

Reading homework: problem 2

Lemma 9.2.1. *For any subset $S \subset V$, $\operatorname{span}(S)$ is a subspace of V.*

Proof. We need to show that $\operatorname{span}(S)$ is a vector space.

It suffices to show that $\operatorname{span}(S)$ is closed under linear combinations. Let $u, v \in \operatorname{span}(S)$ and λ, μ be constants. By the definition of $\operatorname{span}(S)$, there are constants c^i and d^i (some of which could be zero) such that:

$$\begin{aligned}
u &= c^1 s_1 + c^2 s_2 + \cdots \\
v &= d^1 s_1 + d^2 s_2 + \cdots \\
\Rightarrow \lambda u + \mu v &= \lambda(c^1 s_1 + c^2 s_2 + \cdots) + \mu(d^1 s_1 + d^2 s_2 + \cdots) \\
&= (\lambda c^1 + \mu d^1)s_1 + (\lambda c^2 + \mu d^2)s_2 + \cdots
\end{aligned}$$

This last sum is a linear combination of elements of S, and is thus in $\operatorname{span}(S)$. Then $\operatorname{span}(S)$ is closed under linear combinations, and is thus a subspace of V. □

Note that this proof, like many proofs, consisted of little more than just writing out the definitions.

Example 111 For which values of a does

$$\operatorname{span}\left\{\begin{pmatrix}1\\0\\a\end{pmatrix},\begin{pmatrix}1\\2\\-3\end{pmatrix},\begin{pmatrix}a\\1\\0\end{pmatrix}\right\}=\mathbb{R}^3?$$

Given an arbitrary vector $\begin{pmatrix}x\\y\\z\end{pmatrix}$ in $\mathbb{R}^3$, we need to find constants r^1, r^2, r^3 such that

$$r^1\begin{pmatrix}1\\0\\a\end{pmatrix}+r^2\begin{pmatrix}1\\2\\-3\end{pmatrix}+r^3\begin{pmatrix}a\\1\\0\end{pmatrix}=\begin{pmatrix}x\\y\\z\end{pmatrix}.$$

We can write this as a linear system in the unknowns r^1, r^2, r^3 as follows:

$$\begin{pmatrix}1&1&a\\0&2&1\\a&-3&0\end{pmatrix}\begin{pmatrix}r^1\\r^2\\r^3\end{pmatrix}=\begin{pmatrix}x\\y\\z\end{pmatrix}.$$

If the matrix $M=\begin{pmatrix}1&1&a\\0&2&1\\a&-3&0\end{pmatrix}$ is invertible, then we can find a solution

$$M^{-1}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}r^1\\r^2\\r^3\end{pmatrix}$$

for *any* vector $\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3$.

Therefore we should choose a so that M is invertible:

$$i.e.,\ 0\neq\det M=-2a^2+3+a=-(2a-3)(a+1).$$

Then the span is $\mathbb{R}^3$ if and only if $a\neq -1, \frac{3}{2}$.

Linear systems as spanning sets

Some other very important ways of building subspaces are given in the following examples.

Example 112 (The kernel of a linear map).

Suppose $L: U\to V$ is a linear map between vector spaces. Then if

$$L(u)=0=L(u'),$$

linearity tells us that

$$L(\alpha u+\beta u')=\alpha L(u)+\beta L(u')=\alpha 0+\beta 0=0.$$

Hence, thanks to the subspace theorem, the set of all vectors in U that are mapped to the zero vector is a subspace of V. It is called the kernel of L:

$$\mathrm{ker}L := \{u \in U | L(u) = 0\} \subset U.$$

Note that finding a kernel means finding a solution to a homogeneous linear equation.

Example 113 (The image of a linear map).

Suppose $L : U \to V$ is a linear map between vector spaces. Then if

$$v = L(u) \text{ and } v' = L(u')\,,$$

linearity tells us that

$$\alpha v + \beta v' = \alpha L(u) + \beta L(u') = L(\alpha u + \beta u')\,.$$

Hence, calling once again on the subspace theorem, the set of all vectors in V that are obtained as outputs of the map L is a subspace. It is called the image of L:

$$\mathrm{im}L := \{L(u) \mid u \in U\} \subset V.$$

Example 114 (An eigenspace of a linear map).

Suppose $L : V \to V$ is a linear map and V is a vector space. Then if

$$L(u) = \lambda u \text{ and } L(v) = \lambda v\,,$$

linearity tells us that

$$L(\alpha u + \beta v) = \alpha L(u) + \beta L(v) = \alpha L(u) + \beta L(v) = \alpha \lambda u + \beta \lambda v = \lambda(\alpha u + \beta v)\,.$$

Hence, again by subspace theorem, the set of all vectors in V that obey the *eigenvector equation* $L(v) = \lambda v$ is a subspace of V. It is called an eigenspace

$$V_\lambda := \{v \in V | L(v) = \lambda v\}.$$

For most scalars λ, the only solution to $L(v) = \lambda v$ will be $v = 0$, which yields the trivial subspace $\{0\}$. When there are nontrivial solutions to $L(v) = \lambda v$, the number λ is called an eigenvalue, and carries essential information about the map L.

Kernels, images and eigenspaces are discussed in great depth in chapters 16 and 12.

9.3 Review Problems

Webwork:

Reading Problems	1, 2
Subspaces	3, 4, 5, 6
Spans	7, 8

1. Determine if $x - x^3 \in \text{span}\{x^2, 2x + x^2, x + x^3\}$.

2. Let U and W be subspaces of V. Are:

   (a) $U \cup W$

   (b) $U \cap W$

   also subspaces? Explain why or why not. Draw examples in $\mathbb{R}^3$.

Hint

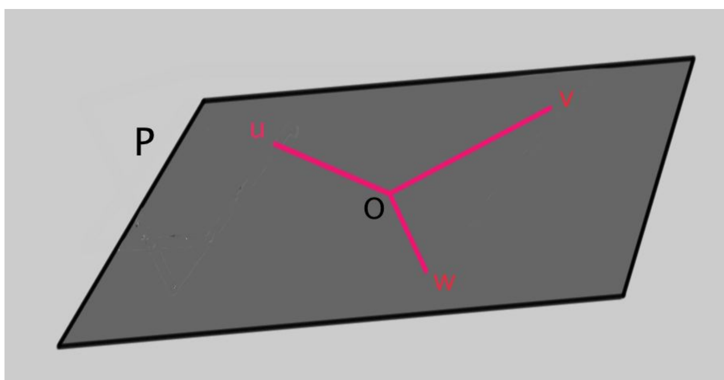
3. Let $L : \mathbb{R}^3 \to \mathbb{R}^3$ where

$$L(x, y, z) = (x + 2y + z, 2x + y + z, 0)\,.$$

   Find $\ker L$, $\text{im} L$ and the eigenspaces $\mathbb{R}^3_{-1}$, $\mathbb{R}^3_3$. Your answers should be subsets of $\mathbb{R}^3$. Express them using span notation.

10
Linear Independence

Consider a plane P that includes the origin in $\mathbb{R}^3$ and non-zero vectors $\{u, v, w\}$ in P.



If no two of u, v and w are parallel, then $P = \text{span}\{u, v, w\}$. But any two vectors determines a plane, so we should be able to span the plane using only two of the vectors u, v, w. Then we could choose two of the vectors in $\{u, v, w\}$ whose span is P, and express the other as a linear combination of those two. Suppose u and v span P. Then there exist constants d^1, d^2 (not both zero) such that $w = d^1u + d^2v$. Since w can be expressed in terms of u and v we say that it is not independent. More generally, the relationship

$$c^1u + c^2v + c^3w = 0 \qquad c^i \in \mathbb{R},\ \text{some } c^i \neq 0$$

expresses the fact that u, v, w are not all independent.

Definition We say that the vectors $v_1, v_2, \ldots, v_n$ are **linearly dependent** if there exist constants[1] $c^1, c^2, \ldots, c^n$ not all zero such that

$$c^1 v_1 + c^2 v_2 + \cdots + c^n v_n = 0.$$

Otherwise, the vectors $v_1, v_2, \ldots, v_n$ are **linearly independent.**

Remark The zero vector 0_V can *never* be on a list of independent vectors because $\alpha 0_V = 0_V$ for any scalar α.

Example 115 Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{pmatrix} 4 \\ -1 \\ 3 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} -3 \\ 7 \\ 4 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 5 \\ 12 \\ 17 \end{pmatrix}, \qquad v_4 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Are these vectors linearly independent?

No, since $3v_1 + 2v_2 - v_3 + v_4 = 0$, the vectors are linearly *dependent*.

`Worked Example`

10.1 Showing Linear Dependence

In the above example we were given the linear combination $3v_1 + 2v_2 - v_3 + v_4$ seemingly by magic. The next example shows how to find such a linear combination, if it exists.

Example 116 Consider the following vectors in $\mathbb{R}^3$:

$$v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Are they linearly independent?

We need to see whether the system

$$c^1 v_1 + c^2 v_2 + c^3 v_3 = 0$$

[1]Usually our vector spaces are defined over $\mathbb{R}$, but in general we can have vector spaces defined over different base fields such as $\mathbb{C}$ or $\mathbb{Z}_2$. The coefficients c^i should come from whatever our base field is (usually $\mathbb{R}$).

has any solutions for c^1, c^2, c^3. We can rewrite this as a homogeneous system by building a matrix whose columns are the vectors v_1, v_2 and v_3:

$$\begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} c^1 \\ c^2 \\ c^3 \end{pmatrix} = 0.$$

This system has solutions if and only if the matrix $M = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}$ is singular, so we should find the determinant of M:

$$\det M = \det \begin{pmatrix} 0 & 1 & 1 \\ 0 & 2 & 2 \\ 1 & 1 & 3 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = 0.$$

Therefore nontrivial solutions exist. At this point we know that the vectors are linearly dependent. If we need to, we can find coefficients that demonstrate linear dependence by solving

$$\left(\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 1 & 1 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

The solution set $\{\mu(-2, -1, 1) \mid \mu \in \mathbb{R}\}$ encodes the linear combinations equal to zero; any choice of μ will produce coefficients c^1, c^2, c^3 that satisfy the linear homogeneous equation. In particular, $\mu = 1$ corresponds to the equation

$$c^1 v_1 + c^2 v_2 + c^3 v_3 = 0 \Rightarrow -2v_1 - v_2 + v_3 = 0.$$

Reading homework: problem 1

Definition Any sum of vectors $v_1, \ldots, v_k$ multiplied by scalars $c^1, \ldots, c^k$, namely

$$c^1 v_1 + \cdots + c^k v_k \,,$$

is called a *linear combination* of $v_1, \ldots, v_k$.

Theorem 10.1.1 (Linear Dependence)**.** *An ordered set of non-zero vectors $(v_1, \ldots, v_n)$ is linearly dependent if and only if one of the vectors v_k is expressible as a linear combination of the preceding vectors.*

Proof. The theorem is an if and only if statement, so there are two things to show.

i. First, we show that if $v_k = c^1 v_1 + \cdots c^{k-1} v_{k-1}$ then the set is linearly dependent.

This is easy. We just rewrite the assumption:

$$c^1 v_1 + \cdots + c^{k-1} v_{k-1} - v_k + 0 v_{k+1} + \cdots + 0 v_n = 0.$$

This is a vanishing linear combination of the vectors $\{v_1, \ldots, v_n\}$ with not all coefficients equal to zero, so $\{v_1, \ldots, v_n\}$ is a linearly dependent set.

ii. Now we show that linear dependence implies that there exists k for which v_k is a linear combination of the vectors $\{v_1, \ldots, v_{k-1}\}$.

The assumption says that

$$c^1 v_1 + c^2 v_2 + \cdots + c^n v_n = 0.$$

Take k to be the largest number for which c_k is not equal to zero. So:

$$c^1 v_1 + c^2 v_2 + \cdots + c^{k-1} v_{k-1} + c^k v_k = 0.$$

(Note that $k > 1$, since otherwise we would have $c^1 v_1 = 0 \Rightarrow v_1 = 0$, contradicting the assumption that none of the v_i are the zero vector.)

So we can rearrange the equation:

$$\begin{array}{rcl} c^1 v_1 + c^2 v_2 + \cdots + c^{k-1} v_{k-1} & = & -c^k v_k \\ \Rightarrow \ -\dfrac{c^1}{c^k} v_1 - \dfrac{c^2}{c^k} v_2 - \cdots - \dfrac{c^{k-1}}{c^k} v_{k-1} & = & v_k . \end{array}$$

Therefore we have expressed v_k as a linear combination of the previous vectors, and we are done.

□

`Worked proof`

Example 117 Consider the vector space $P_2(t)$ of polynomials of degree less than or equal to 2. Set:

$$
\begin{aligned}
v_1 &= 1+t \\
v_2 &= 1+t^2 \\
v_3 &= t+t^2 \\
v_4 &= 2+t+t^2 \\
v_5 &= 1+t+t^2.
\end{aligned}
$$

The set $\{v_1, \ldots, v_5\}$ is linearly dependent, because $v_4 = v_1 + v_2$.

10.2 Showing Linear Independence

We have seen two different ways to show a set of vectors is linearly dependent: we can either find a linear combination of the vectors which is equal to zero, or we can express one of the vectors as a linear combination of the other vectors. On the other hand, to check that a set of vectors is linearly *independent*, we must check that every linear combination of our vectors with non-vanishing coefficients gives something other than the zero vector. Equivalently, to show that the set $v_1, v_2, \ldots, v_n$ is linearly independent, we must show that the equation $c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0$ has no solutions other than $c_1 = c_2 = \cdots = c_n = 0$.

Example 118 Consider the following vectors in $\mathbb{R}^3$:

$$
v_1 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}.
$$

Are they linearly independent?

We need to see whether the system

$$
c^1v_1 + c^2v_2 + c^3v_3 = 0
$$

has any solutions for c^1, c^2, c^3. We can rewrite this as a homogeneous system:

$$
\begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} c^1 \\ c^2 \\ c^3 \end{pmatrix} = 0.
$$

This system has solutions if and only if the matrix $M = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}$ is singular, so we should find the determinant of M:

$$\det M = \det \begin{pmatrix} 0 & 2 & 1 \\ 0 & 2 & 4 \\ 2 & 1 & 3 \end{pmatrix} = 2 \det \begin{pmatrix} 2 & 1 \\ 2 & 4 \end{pmatrix} = 12.$$

Since the matrix M has non-zero determinant, the only solution to the system of equations

$$\begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} c^1 \\ c^2 \\ c^3 \end{pmatrix} = 0$$

is $c_1 = c_2 = c_3 = 0$. So the vectors v_1, v_2, v_3 are linearly independent.

Here is another example with bits:

Example 119 Let $\mathbb{Z}_2^3$ be the space of 3×1 bit-valued matrices (i.e., column vectors). Is the following subset linearly independent?

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

If the set is linearly dependent, then we can find non-zero solutions to the system:

$$c^1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c^2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + c^3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 0,$$

which becomes the linear system

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} c^1 \\ c^2 \\ c^3 \end{pmatrix} = 0.$$

Solutions exist if and only if the determinant of the matrix is non-zero. But:

$$\det \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = 1 \det \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} - 1 \det \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = -1 - 1 = 1 + 1 = 0$$

Therefore non-trivial solutions exist, and the set is not linearly independent.

Reading homework: problem 2

10.3 From Dependent Independent

Now suppose vectors $v_1, \ldots, v_n$ are linearly dependent,

$$c^1 v_1 + c^2 v_2 + \cdots + c^n v_n = 0$$

with $c^1 \neq 0$. Then:

$$\operatorname{span}\{v_1, \ldots, v_n\} = \operatorname{span}\{v_2, \ldots, v_n\}$$

because any $x \in \operatorname{span}\{v_1, \ldots, v_n\}$ is given by

$$\begin{aligned} x &= a^1 v_1 + \cdots + a^n v_n \\ &= a^1 \left(-\frac{c^2}{c_1} v_2 - \cdots - \frac{c^n}{c_1} v_n \right) + a^2 v_2 + \cdots + a^n v_n \\ &= \left(a^2 - a^1 \frac{c^2}{c_1} \right) v_2 + \cdots + \left(a^n - a^1 \frac{c^n}{c_1} \right) v_n . \end{aligned}$$

Then x is in $\operatorname{span}\{v_2, \ldots, v_n\}$.

When we write a vector space as the span of a list of vectors, we would like that list to be as short as possible (this idea is explored further in chapter 11). This can be achieved by iterating the above procedure.

Example 120 In the above example, we found that $v_4 = v_1 + v_2$. In this case, any expression for a vector as a linear combination involving v_4 can be turned into a combination without v_4 by making the substitution $v_4 = v_1 + v_2$.

Then:

$$\begin{aligned} S &= \operatorname{span}\{1+t, 1+t^2, t+t^2, 2+t+t^2, 1+t+t^2\} \\ &= \operatorname{span}\{1+t, 1+t^2, t+t^2, 1+t+t^2\}. \end{aligned}$$

Now we notice that $1+t+t^2 = \frac{1}{2}(1+t) + \frac{1}{2}(1+t^2) + \frac{1}{2}(t+t^2)$. So the vector $1+t+t^2 = v_5$ is also extraneous, since it can be expressed as a linear combination of the remaining three vectors, v_1, v_2, v_3. Therefore

$$S = \operatorname{span}\{1+t, 1+t^2, t+t^2\}.$$

In fact, you can check that there are no (non-zero) solutions to the linear system

$$c^1(1+t) + c^2(1+t^2) + c^3(t+t^2) = 0.$$

Therefore the remaining vectors $\{1+t, 1+t^2, t+t^2\}$ are linearly independent, and span the vector space S. Then these vectors are a minimal spanning set, in the sense that no more vectors can be removed since the vectors are linearly independent. Such a set is called a *basis* for S.

10.4 Review Problems

Webwork:

Reading Problems	1, 2
Testing for linear independence	3, 4
Gaussian elimination	5
Spanning and linear independence	6

1. Let B^n be the space of $n \times 1$ bit-valued matrices (*i.e.*, column vectors) over the field $\mathbb{Z}_2$. Remember that this means that the coefficients in any linear combination can be only 0 or 1, with rules for adding and multiplying coefficients given here.

   (a) How many different vectors are there in B^n?

   (b) Find a collection S of vectors that span B^3 and are linearly independent. In other words, find a basis of B^3.

   (c) Write each other vector in B^3 as a linear combination of the vectors in the set S that you chose.

   (d) Would it be possible to span B^3 with only two vectors?

Hint

2. Let e_i be the vector in $\mathbb{R}^n$ with a 1 in the ith position and 0's in every other position. Let v be an arbitrary vector in $\mathbb{R}^n$.

   (a) Show that the collection $\{e_1, \ldots, e_n\}$ is linearly independent.

   (b) Demonstrate that $v = \sum_{i=1}^{n} (v \cdot e_i) e_i$.

   (c) The span$\{e_1, \ldots, e_n\}$ is the same as what vector space?

3. Consider the ordered set of vectors from $\mathbb{R}^3$

$$\left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix} \right)$$

   (a) Determine if the set is linearly independent by using the vectors as the columns of a matrix M and finding RREF(M).

(b) If possible, write each vector as a linear combination of the preceding ones.

(c) Remove the vectors which can be expressed as linear combinations of the preceding vectors to form a linearly independent ordered set. (Every vector in your set set should be from the given set.)

4. Gaussian elimination is a useful tool to figure out whether a set of vectors spans a vector space and if they are linearly independent. Consider a matrix M made from an ordered set of column vectors $(v_1, v_2, \ldots, v_m) \subset \mathbb{R}^n$ and the three cases listed below:

   (a) RREF(M) is the identity matrix.

   (b) RREF(M) has a row of zeros.

   (c) Neither case (a) or (b) apply.

   First give an explicit example for each case, state whether the column vectors you use are linearly independent or spanning in each case. Then, in general, determine whether $(v_1, v_2, \ldots, v_m)$ are linearly independent and/or spanning $\mathbb{R}^n$ in each of the three cases. If they are linearly dependent, does RREF(M) tell you which vectors could be removed to yield an independent set of vectors?