

Lecture Notes for Math 3410,  
with Computational Examples



# Lecture Notes for Math 3410, with Computational Examples

Sean Fitzpatrick  
University of Lethbridge

October 6, 2019



# Preface

Linear algebra is a mature, rich subject, full of both fascinating theory and useful applications. One of the things you might have taken away from a first course in the subject is that there's a lot of tedious calculation involved. This is true, if you're a human. But the algorithms you learn in a course like Math 1410 are easily implemented on a computer. If we want to be able to discuss any of the interesting applications of linear algebra, we're going to need to learn how to do linear algebra on a computer.

There are many good mathematical software products that can deal with linear algebra, like Maple, Mathematica, and MatLab. But all of these are proprietary, and expensive. Sage is a popular open source system for mathematics, and students considering further studies in mathematics would do well to learn Sage. Since most people in Math 3410 are probably not considering a career as a mathematician, we'll try to do everything in Python.

Python is a very popular programming language, partly because of its ease of use. Those of you enrolled in Education may find yourself teaching Python to your students one day. Also, if you do want to use Sage, you're in luck: Sage is an amalgamation of many different software tools, including Python. So any Python code you encounter in this course can also be run on Sage.

# Contents

<b>Preface</b>	<b>v</b>
<b>1 Computational Tools</b>	<b>1</b>
1.1 Jupyter . . . . .	1
1.2 Python basics . . . . .	1
1.3 SymPy for linear algebra . . . . .	3
<b>2 Vector spaces</b>	<b>5</b>
2.1 Abstract vector spaces . . . . .	5
2.2 Span . . . . .	5
2.3 Linear Independence . . . . .	7
2.4 Basis and dimension . . . . .	9
<b>3 Linear Transformations</b>	<b>15</b>
3.1 Definition and examples . . . . .	15
3.2 Kernel and Image . . . . .	20

# Chapter 1

## Computational Tools

### 1.1 Jupyter

The first thing you need to know about doing linear algebra in Python is how to access a Python environment. Fortunately, you do not need to install any software for this. The University of Lethbridge has access to the **Syzygy Jupyter Hub** service, provided by PIMS (the Pacific Institute for Mathematical Sciences), Cybera, and Compute Canada. To access Syzygy, go to [uleth.syzygy.ca](http://uleth.syzygy.ca) and log in with your ULeth credentials.

Note: if you click the login button and nothing happens, click the back button and try again. Sometimes there's a problem with our single sign-on service.

The primary type of document you'll encounter on Syzygy is the **Jupyter notebook**. Content in a Jupyter notebook is organized into **cells**. Some cells contain text, which can be in either HTML or **Markdown**. Markdown is a simple markup language. It's not as versatile as HTML, but it's easier to use. On Jupyter, markdown supports the LaTeX language for mathematical expressions. Use single dollar signs for inline math:  $\frac{d}{dx}\sin(x) = \cos(x)$  produces  $\frac{d}{dx}\sin(x) = \cos(x)$ , for example.

If you want “display math”, use double dollar signs. Unfortunately, entering matrices is a bit tedious. For example,  $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$  produces

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}.$$

Later we'll see how to enter things like matrices in Python.

It's also possible to use markdown to add *emphasis*, images, URLs, etc.. For details, see the following [Markdown cheatsheet](#), or this [quick reference](#) from [callysto.ca](http://callysto.ca).

What's cool about a Jupyter notebook is that in addition to markdown cells, which can present content and provide explanation, we can also include *code cells*. Jupyter supports many different programming languages, but we will stick mainly to Python.

### 1.2 Python basics

OK, so you've logged into Syzygy and you're ready to write some code. What does basic code look like in Python? The good news is that you don't need

to be a programmer to do linear algebra in Python. Python includes many different *libraries* that keep most of the code under the hood, so all you have to remember is what command you need to use to accomplish a task. That said, it won't hurt to learn a little bit of basic coding.

Basic arithmetic operations are understood, and you can simply type them in. Hit **shift+enter** in a code cell to execute the code and see the result.

```
3+4
```

```
3*4
```

```
3**4
```

OK, great. But sometimes we want to do calculations with more than one step. For that, we can assign variables.

```
a = 14
b = -9
c = a+b
print(a, b, c)
```

Sometimes you might need input that's a string, rather than a number. We can do that, too.

```
string_var = "Hey,_look_at_my_string!"
print(string_var)
```

Another basic construction is a list. Getting the hang of lists is useful, because in a sense, matrices are just really fancy lists.

```
empty_list = list()
this_too = []
list_of_zeros = [0]*7
print(list_of_zeros)
```

Once you have an empty list, you might want to add something to it. This can be done with the **append** command.

```
empty_list.append(3)
print(empty_list)
print(len(empty_list))
```

Go back and re-run the above code cell two or three more times. What happens? Probably you can guess what the **len** command is for. Now let's get really carried away and do some "for real" coding, like loops!

```
for i in range(10):
    empty_list.append(i)
print(empty_list)
```

Notice the indentation in the second line. This is how Python handles things like for loops, with indentation rather than bracketing. We could say more about lists but perhaps it's time to talk about matrices. For further reading, you can [start here](#).



## 1.3 SymPy for linear algebra

**SymPy** is a Python library for symbolic algebra. On its own, it's not as powerful as programs like Maple, but it handles a lot of basic manipulations in a fairly simple fashion, and when we need more power, it can interface with other Python libraries.

Another advantage of SymPy is sophisticated “pretty-printing”. In fact, we can enable MathJax within SymPy, so that output is rendered in the same way as when LaTeX is entered in a markdown cell.

### 1.3.1 SymPy basics

Running the following Sage cell will load the SymPy library and turn on MathJax.

```
from sympy import *
init_printing()
```

**Note:** if you are going to be working with multiple libraries, and more than one of them defines a certain command, instead of `from sympy import` all you can do `import sympy as sy`. If you do this, each SymPy command will need to be appended with `sy`; for example, you might write `sy.Matrix` instead of simply `Matrix`. Let's use SymPy to create a  $2 \times 3$  matrix.

```
A = Matrix(2,3,[1,2,3,4,5,6])
A
```

The `A` on the second line asks Python to print the matrix using SymPy's printing support. If we use Python's `print` command, we get something different:

```
print(A)
```

We'll have more on matrices in [Subsection 1.3.2](#). For now, let's look at some more basic constructions. One basic thing to be mindful of is the type of numbers we're working with. For example, if we enter `2/7` in a code cell, Python will interpret this as a floating point number (essentially, a division).

(If you are using Sage cells in HTML rather than Jupyter, this will automatically be interpreted as a fraction.)

```
2/7
```

But we often do linear algebra over the rational numbers, and so SymPy will let you specify this:

```
Rational(2,7)
```

You might not think to add the comma above, because you're used to writing fractions like `2/7`. Fortunately, the SymPy authors thought of that:

```
Rational(2/7)
```

Hmm... You might have got the output you expected in the cell above, but maybe not. If you got a much worse looking fraction, read on.

Another cool command is the `sympify` command, which can be called with the shortcut `S`. The input `2` is interpreted as an `int` by Python, but `S(2)` is a “SymPy Integer”:

```
S(2)/7
```

Of course, sometimes you *do* want to use floating point, and you can specify this, too:

```
2.5
```

```
Float(2.5)
```

```
Float(2.5e10)
```

One note of caution: `Float` is part of SymPy, and not the same as the core Python `float` command. You can also put decimals into the `Rational` command and get the corresponding fraction:

```
Rational(0.75)
```

The only thing to beware of is that computers convert from decimal to binary and then back again, and sometimes weird things can happen:

```
Rational(0.2)
```

Of course, there are workarounds. One way is to enter `0.2` as a string:

```
Rational('0.2')
```

Another is to limit the size of the denominator:

```
Rational(0.2).limit_denominator(10**12)
```

Try some other examples above. Some inputs to try are `1.23` and `23e-10`

We can also deal with repeating decimals. These are entered as strings, with square brackets around the repeating part. Then we can “sympify”:

```
S('0.[1]')
```

Finally, SymPy knows about mathematical constants like  $e$ ,  $\pi$ ,  $i$ , which you’ll need from time to time in linear algebra. If you ever need  $\infty$ , this is entered as `oo`.

```
I*I
```

```
I-sqrt(-1)
```

```
pi.is_irrational
```

### 1.3.2 Matrices in SymPy

(more to come)

# Chapter 2

## Vector spaces

### 2.1 Abstract vector spaces

(To Do)

### 2.2 Span

Recall that a **linear combination** of a set of vectors  $\vec{v}_1, \dots, \vec{v}_k$  is a vector expression of the form

$$\vec{w} = c_1 \vec{v}_1 + c_2 \vec{v}_2 + \dots + c_k \vec{v}_k,$$

where  $c_1, \dots, c_k$  are scalars.

The **span** of those same vectors is the set of all possible linear combinations:

$$\text{span}\{\vec{v}_1, \dots, \vec{v}_k\} = \{c_1 \vec{v}_1 + \dots + c_k \vec{v}_k \mid c_1, \dots, c_k \in \mathbb{F}\}.$$

Therefore, the questions “Is the vector  $\vec{w}$  in  $\text{span}\{\vec{v}_1, \dots, \vec{v}_k\}$ ?” is really asking, “Can  $\vec{w}$  be written as a linear combination of  $\vec{v}_1, \dots, \vec{v}_k$ ?”

With the appropriate setup, all such questions become questions about solving systems of equations. Here, we will look at a few such examples.

**Exercise 2.2.1** Determine whether the vector  $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$  is in the span of the vectors  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ .

This is really asking: are there scalars  $s, t$  such that

$$s \begin{bmatrix} 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}?$$

And this, in turn, is equivalent to the system

$$\begin{aligned} s - t &= 2 \\ s + 2t &= 3, \end{aligned}$$

which is the same as the matrix equation

$$\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Solving the system confirms that there is indeed a solution, so the answer to our original question is yes.

To confirm the above example (and see what the solution is), we can use the computer.

```
from sympy import *
init_printing()
```

```
A = Matrix(2,3,[1,-1,2,1,2,3])
A.rref()
```

The above code produces the reduced row-echelon form of the augmented matrix for our system. Do you remember how to get the answer from here? Here's another approach.

```
B = Matrix(2,2,[1,-1,1,2])
B
```

```
C = Matrix(2,1,[2,3])
X = (B**-1)*C
X
```

Our next example involves polynomials. At first this looks like a different problem, but it's essentially the same once we set it up.

**Exercise 2.2.2** Determine whether  $p(x) = 1 + x + 4x^2$  belongs to  $\text{span}\{1 + 2x - x^2, 3 + 5x + 2x^2\}$ .

We seek scalars  $s, t$  such that

$$s(1 + 2x - 2x^2) + t(3 + 5x + 2x^2) = 1 + x + 4x^2.$$

On the left-hand side, we expand and gather terms:

$$(s + 3t) + (2s + 5t)x + (-2s + 2t)x^2 = 1 + x + 4x^2.$$

These two polynomials are equal if and only if we can solve the system

$$\begin{aligned} s + 3t &= 1 \\ 2s + 5t &= 1 \\ -2s + 2t &= 4. \end{aligned}$$

We can solve this computationally using matrices again.

```
M = Matrix(3,3,[1,3,1,2,5,1,-2,2,4])
M.rref()
```

So, what's the answer? Is  $p(x)$  in the span? Can we determine what polynomials are in the span? Let's consider a general polynomial  $q(x) = a + bx + cx^2$ . A bit of thought tells us that the coefficients  $a, b, c$  should replace the constants 1, 1, 4 above.

```
a, b, c = symbols('a_b_c', real = True, constant = True)
N = Matrix(3,3,[1,3,a,2,5,b,-2,2,c])
N
```

Asking the computer to reduce this matrix won't produce the desired result. (Maybe someone can figure out how to define the constants in a way that works?) But we can always specify row operations.

```

N1 =
    N.elementary_row_operation(op='n->n+km', row1=1, row2=0, k=-2)
N1

```

Now we repeat. Here are two more empty cells to work with: We end this section with a few non-computational, but useful results, which will be left as exercises to be done in class, or by the reader.

**Exercise 2.2.3** Let  $V$  be a vector space, and let  $X, Y \subseteq V$ . Show that if  $X \subseteq Y$ , then  $\text{span } X \subseteq \text{span } Y$ .

**Exercise 2.2.4** Can  $\{(1, 2, 0), (1, 1, 1)\}$  span  $\{(a, b, 0) \mid a, b \in \mathbb{R}\}$ ?

**Theorem 2.2.5** Let  $V$  be a vector space, and let  $\vec{v}_1, \dots, \vec{v}_k \in V$ . If  $\vec{u} \in \text{span}\{\vec{v}_1, \dots, \vec{v}_k\}$ , then

$$\text{span}\{\vec{u}, \vec{v}_1, \dots, \vec{v}_k\} = \text{span}\{\vec{v}_1, \dots, \vec{v}_k\}.$$

The moral of [Theorem 2.2.5](#) is that one vector in a set is a linear combination of the others, we can remove it from the set without affecting the span. This suggests that we might want to look for the most “efficient” spanning sets – those in which no vector in the set can be written in terms of the others. Such sets are called **linearly independent**, and they are the subject of [Section 2.3](#).

## 2.3 Linear Independence

In any vector space  $V$ , we say that a set of vectors

$$\{\vec{v}_1, \dots, \vec{v}_k\}$$

is **linearly independent** if for any scalars  $c_1, \dots, c_k$

$$c_1\vec{v}_1 + \dots + c_k\vec{v}_k = \vec{0} \quad \Rightarrow \quad c_1 = \dots = c_k = 0.$$

This means that no vector in the set can be written as a linear combination of the other vectors in that set. We will soon see that when looking for vectors that span a subspace, it is especially useful to find a spanning set that is also linearly independent. The following lemma establishes some basic properties of independent sets.

**Lemma 2.3.1** In any vector space  $V$ :

1. If  $\vec{v} \neq \vec{0}$ , then  $\{\vec{v}\}$  is independent.
2. If  $S \subseteq V$  contains the zero vector, then  $S$  is dependent.

The definition of linear independence tells us that if  $\{\vec{v}_1, \dots, \vec{v}_k\}$  is an independent set of vectors, then there is only one way to write  $\vec{0}$  as a linear combination of these vectors; namely,

$$\vec{0} = 0\vec{v}_1 + 0\vec{v}_2 + \dots + 0\vec{v}_k.$$

In fact, more is true: *any* vector in the span of a linearly independent set can be written in only one way as a linear combination of those vectors.

Computationally, questions about linear independence are just questions about homogeneous systems of linear equations. For example, suppose we want to know if the vectors

$$\vec{u} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}, \vec{v} = \begin{bmatrix} 0 \\ 2 \\ -3 \end{bmatrix}, \vec{w} = \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix}$$

are linearly independent in  $\mathbb{R}^3$ . This question leads to the vector equation

$$x\vec{u} + y\vec{v} + z\vec{w} = \vec{0},$$

which becomes the matrix equation

$$\begin{bmatrix} 1 & 0 & 4 \\ -1 & 2 & 0 \\ 4 & -3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We now apply some basic theory from linear algebra. A unique (and therefore, trivial) solution to this system is guaranteed if the matrix  $A = \begin{bmatrix} 1 & 0 & 4 \\ -1 & 2 & 0 \\ 4 & -3 & -3 \end{bmatrix}$  is invertible, since in that case we have  $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}\vec{0} = \vec{0}$ .

This approach is problematic, however, since it won't work if we have 2 vectors, or 4. Instead, we look at the reduced row-echelon form. A unique solution corresponds to having a leading 1 in each column of  $A$ . Let's check this condition.

```
from sympy import *
init_printing()
```

```
A = Matrix(3,3,[1,0,4,-1,2,0,4,-3,-3])
A.rref()
```

One observation is useful here, and will lead to a better understanding of independence. First, it would be impossible to have 4 or more linearly independent vectors in  $\mathbb{R}^3$ . Why? (How many leading ones can you have in a  $3 \times 4$  matrix?) Second, having two or fewer vectors makes it more likely that the set is independent.

The largest set of linearly independent vectors possible in  $\mathbb{R}^3$  contains three vectors. You might have also observed that the smallest number of vectors needed to span  $\mathbb{R}^3$  is 3. Hmm. Seems like there's something interesting going on here. But first, some more computation.

**Exercise 2.3.2** Determine whether the set  $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 9 \end{bmatrix} \right\}$  is linearly independent in  $\mathbb{R}^3$ .

Again, we set up a matrix and reduce:

```
A = Matrix(3,3,[1,-1,-1,2,0,4,0,3,9])
A.rref()
```

Notice that this time we don't get a unique solution, so we can conclude that these vectors are *not* independent. Furthermore, you can probably deduce from the above that we have  $2\vec{v}_1 + 3\vec{v}_2 - \vec{v}_3 = \vec{0}$ . Now suppose that  $\vec{w} \in \text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ . In how many ways can we write  $\vec{w}$  as a linear combination of these vectors?

**Exercise 2.3.3** Which of the following subsets of  $P_2(\mathbb{R})$  are independent?

- (a)  $S_1 = \{x^2 + 1, x + 1, x\}$
- (b)  $S_2 = \{x^2 - x + 3, 2x^2 + x + 5, x^2 + 5x + 1\}$

In each case, we set up the defining equation for independence, collect terms, and then analyze the resulting system of equations. (If you work with

polynomials often enough, you can probably jump straight to the matrix. For now, let's work out the details.)

Suppose

$$r(x^2 + 1) + s(x + 1) + tx = 0.$$

Then  $rx^2 + (s + t)x + (r + s) = 0 = 0x^2 + 0x + 0$ , so

$$r = 0$$

$$s + t = 0$$

$$r + s = 0.$$

And in this case, we don't even need to ask the computer. The first equation gives  $r = 0$  right away, and putting that into the third equation gives  $s = 0$ , and the second equation then gives  $t = 0$ .

Since  $r = s = t = 0$  is the only solution, the set is independent.

Repeating for  $S_2$  leads to the equation

$$(r + 2s + t)x^2 + (-r + s + 5t)x + (3r + 5s + t)1 = 0.$$

This gives us:

```
A = Matrix(3,3,[1,2,1,-1,1,5,3,5,1])
A.rref()
```

**Exercise 2.3.4** Determine whether or not the set

$$\left\{ \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \right\}$$

is linearly independent in  $M_2(\mathbb{R})$ .

Again, we set a linear combination equal to the zero vector, and combine:

$$a \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + c \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + d \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} -a + b + c & -b + c - d \\ -b + c - d & -a + b + c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

We could proceed, but we might instead notice right away that equations 1 and 4 are identical, and so are equations 2 and 3. With only two distinct equations and 4 unknowns, we're certain to find nontrivial solutions.

## 2.4 Basis and dimension

Next, we begin with an important result, sometimes known as the “Fundamental Theorem”:

**Theorem 2.4.1 Fundamental Theorem (Steinitz Exchange Lemma).** Suppose  $V = \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$ . If  $\{\vec{w}_1, \dots, \vec{w}_m\}$  is a linearly independent set of vectors in  $V$ , then  $m \leq n$ .

If a set of vectors spans a vector space  $V$ , and it is not independent, we observed that it is possible to remove a vector from the set and still span  $V$ . This suggests that spanning sets that are also linearly independent are of particular importance, and indeed, they are important enough to have a name.

**Definition 2.4.2** Let  $V$  be a vector space. A set  $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$  is called a **basis** of  $V$  if  $\mathcal{B}$  is linearly independent, and  $\text{span } \mathcal{B} = V$ .  $\diamond$

The importance of a basis is that vector  $\vec{v} \in V$  can be written in terms of the basis, and this expression as a linear combination of basis vectors is *unique*. Another important fact is that every basis has the same number of elements.

**Theorem 2.4.3 Invariance Theorem.** *If  $\{\vec{e}_1, \dots, \vec{e}_n\}$  and  $\{\vec{f}_1, \dots, \vec{f}_m\}$  are both bases of a vector space  $V$ , then  $m = n$ .*

Suppose  $V = \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$ . If this set is not linearly independent, Theorem 2.2.5 tells us that we can remove a vector from the set, and still span  $V$ . We can repeat this procedure until we have a linearly independent set of vectors, which will then be a basis. These results let us make a definition.

**Definition 2.4.4** Let  $V$  be a vector space. If  $V$  can be spanned by a finite number of vectors, then we call  $V$  a **finite-dimensional** vector space. If  $V$  is finite-dimensional (and non-trivial), and  $\{\vec{e}_1, \dots, \vec{e}_n\}$  is a basis of  $V$ , we say that  $V$  has **dimension**  $n$ , and write

$$\dim V = n.$$

If  $V$  cannot be spanned by finitely many vectors, we say that  $V$  is **infinite-dimensional**.  $\diamond$

**Exercise 2.4.5** Find a basis for  $U = \{X \in M_{22} \mid XA = AX\}$ , if  $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

**Exercise 2.4.6** Show that the following are bases of  $\mathbb{R}^3$ :

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

```
from sympy import *
init_printing()
```

**Exercise 2.4.7** Show that the following is a basis of  $M_{22}$ :

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}.$$

**Exercise 2.4.8** Show that  $\{1+x, x+x^2, x^2+x^3, x^3\}$  is a basis for  $P_3$ .

**Exercise 2.4.9** Find a basis and dimension for the following subspaces of  $P_2$ :

- a  $U_1 = \{a(1+x) + b(x+x^2) \mid a, b \in \mathbb{R}\}$
- b  $U_2 = \{p(x) \in P_2 \mid p(1) = 0\}$
- c  $U_3 = \{p(x) \in P_2 \mid p(x) = p(-x)\}$

**Solution.**

- a By definition,  $U_1 = \text{span}\{1+x, x+x^2\}$ , and these vectors are independent, since neither is a scalar multiple of the other. Since there are two vectors in this basis,  $\dim U_1 = 2$ .
- b If  $p(1) = 0$ , then  $p(x) = (x-1)q(x)$  for some polynomial  $q$ . Since  $U_2$  is a subspace of  $P_2$ , the degree of  $q$  is at most 2. Therefore,  $q(x) = ax + b$  for some  $a, b \in \mathbb{R}$ , and

$$p(x) = (x-1)(ax+b) = a(x^2-x) + b(x-1).$$

Since  $p$  was arbitrary, this shows that  $U_2 = \text{span}\{x^2-x, x-1\}$ .



The set  $\{x^2 - x, x - 1\}$  is also independent, since neither vector is a scalar multiple of the other. Therefore, this set is a basis, and  $\dim U_2 = 2$ .

- c If  $p(x) = p(-x)$ , then  $p(x)$  is an even polynomial, and therefore  $p(x) = a + bx^2$  for  $a, b \in \mathbb{R}$ . (If you didn't know this it's easily verified: if

$$a + bx + cx^2 = a + b(-x) + c(-x)^2,$$

we can immediately cancel  $a$  from each side, and since  $(-x)^2 = x^2$ , we can cancel  $cx^2$  as well. This leaves  $bx = -bx$ , or  $2bx = 0$ , which implies that  $b = 0$ .)

It follows that the set  $\{1, x^2\}$  spans  $U_3$ , and since this is a subset of the standard basis  $\{1, x, x^2\}$  of  $P_2$ , it must be independent, and is therefore a basis of  $U_3$ , letting us conclude that  $\dim U_3 = 2$ .

We've noted a few times now that if  $\vec{w} \in \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$ , then

$$\text{span}\{\vec{w}, \vec{v}_1, \dots, \vec{v}_n\} = \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$$

If  $\vec{w}$  is not in the span, we can make another useful observation:

**Lemma 2.4.10 Independent Lemma.** *Suppose  $\{\vec{v}_1, \dots, \vec{v}_n\}$  is a linearly independent set of vectors in a vector space  $V$ . If  $\vec{u} \in V$  but  $\vec{u} \notin \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$ , then  $\{\vec{u}, \vec{v}_1, \dots, \vec{v}_n\}$  is independent.*

*Proof.* Suppose  $S = \{\vec{v}_1, \dots, \vec{v}_n\}$  is independent, and that  $\vec{u} \notin \text{span } S$ . Suppose we have

$$a\vec{u} + c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n = \vec{0}$$

for scalars  $a, c_1, \dots, c_n$ . We must have  $a = 0$ ; otherwise, we can multiply by  $\frac{1}{a}$  and rearrange to obtain

$$\vec{u} = -\frac{c_1}{a}\vec{v}_1 - \dots - \frac{c_n}{a}\vec{v}_n,$$

but this would mean that  $\vec{u} \in \text{span } S$ , contradicting our assumption.

With  $a = 0$  we're left with

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n = \vec{0},$$

and since we assumed that the set  $S$  is independent, we must have  $c_1 = c_2 = \dots = c_n = 0$ . Since we already showed  $a = 0$ , this shows that  $\{\vec{u}, \vec{v}_1, \dots, \vec{v}_n\}$  is independent. ■

This is, in fact, an “if and only if” result. If  $\vec{u} \in \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$ , then  $\{\vec{u}, \vec{v}_1, \dots, \vec{v}_n\}$  is not independent. Above, we argued that if  $V$  is finite dimensional, then any spanning set for  $V$  can be reduced to a basis. It probably won't surprise you that the following is also true.

**Lemma 2.4.11** *Let  $V$  be a finite-dimensional vector space, and let  $U$  be any subspace of  $V$ . Then any independent set of vectors  $\{\vec{u}_1, \dots, \vec{u}_k\}$  in  $U$  can be enlarged to a basis of  $U$ .*

*Proof.* This follows from [Lemma 2.4.10](#). If our independent set of vectors spans  $U$ , then it's a basis and we're done. If not, we can find some vector not in the span, and add it to our set to obtain a larger set that is still independent. We can continue adding vectors in this fashion until we obtain a spanning set.

Note that this process *must* terminate:  $V$  is finite-dimensional, so there is a finite spanning set for  $V$ , and therefore for  $U$ . By the Steinitz Exchange lemma, our independent set cannot get larger than this spanning set. ■

**Theorem 2.4.12** Any finite-dimensional vector space  $V$  has a basis. Moreover:

1. If  $V$  can be spanned by  $m$  vectors, then  $\dim V \leq m$ .
2. Given an independent set  $I$  in  $V$ , and a basis  $\mathcal{B}$  of  $V$ , we can enlarge  $I$  to a basis of  $V$  by adding elements of  $\mathcal{B}$ .

If  $U$  is a subspace of  $V$ , then:

1.  $U$  is finite-dimensional, and  $\dim U \leq \dim V$ .
2. If  $\dim U = \dim V$ , then  $U = V$ .

**Exercise 2.4.13** Find a basis of  $M_{22}(\mathbb{R})$  that contains the vectors

$$\vec{v} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \vec{w} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

**Solution.** By the previous theorem, we can form a basis by adding vectors from the standard basis

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

It's easy to check that  $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$  is not in the span of  $\{\vec{v}, \vec{w}\}$ . To get a basis, we need one more vector. Observe that all three of our vectors so far have a zero in the  $(2, 1)$ -entry. Thus,  $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$  cannot be in the span of the first three vectors, and adding it gives us our basis.

**Exercise 2.4.14** Extend the set  $\{1 + x, x + x^2, x - x^3\}$  to a basis of  $P_3(\mathbb{R})$ .

**Solution.** Again, we only need to add one vector from the standard basis  $\{1, x, x^2, x^3\}$ , and it's not too hard to check that any of them will do.

**Exercise 2.4.15** Give two examples of infinite-dimensional vector spaces. Support your answer.

Let's recap our results so far:

- A basis for a vector space  $V$  is an independent set of vectors that spans  $V$ .
- The number of vectors in any basis of  $V$  is a constant, called the dimension of  $V$ .
- The number of vectors in any independent set is always less than or equal to the number of vectors in a spanning set.
- In a finite-dimensional vector space, any independent set can be enlarged to a basis, and any spanning set can be cut down to a basis by deleting vectors that are in the span of the remaining vectors.

Another important aspect of dimension is that it reduces many problems, such as determining equality of subspaces, to counting problems.

**Theorem 2.4.16** Let  $U$  and  $W$  be subspaces of a finite-dimensional vector space  $V$ .

1. If  $U \subseteq W$ , then  $\dim U \leq \dim W$ .
2. If  $U \subseteq W$  and  $\dim U = \dim W$ , then  $U = W$ .

*Proof.*

1. Suppose  $U \subseteq W$ , and let  $B = \{\vec{u}_1, \dots, \vec{u}_k\}$  be a basis for  $U$ . Since  $B$  is a basis, it's independent. And since  $B \subseteq U$  and  $U \subseteq W$ ,  $B \subseteq W$ . Thus,  $B$  is an independent subset of  $W$ , and since any basis of  $W$  spans  $W$ , we know that  $\dim U = k \leq \dim W$ , by [Theorem 2.4.1](#).
2. Suppose  $U \subseteq W$  and  $\dim U = \dim W$ . Let  $B$  be a basis for  $U$ . As above,  $B$  is an independent subset of  $W$ . If  $W \neq U$ , then there is some  $\vec{w} \in W$  with  $\vec{w} \notin U$ . But  $U = \text{span } B$ , so that would mean that  $B \cup \{\vec{w}\}$  is a linearly independent set containing  $\dim U + 1$  vectors. This is impossible, since  $\dim W = \dim U$ , so no independent set can contain more than  $\dim U$  vectors.

■

An even more useful counting result is the following:

**Theorem 2.4.17** *Let  $V$  be an  $n$ -dimensional vector space. If the set  $S$  contains  $n$  vectors, then  $S$  is independent if and only if  $\text{span } S = V$ .*

*Proof.* If  $S$  is independent, then it can be extended to a basis  $B$  with  $S \subseteq B$ . But  $S$  and  $B$  both contain  $n$  vectors (since  $\dim V = n$ ), so we must have  $S = B$ .

If  $S$  spans  $V$ , then  $S$  must contain a basis  $B$ , and as above, since  $S$  and  $B$  contain the same number of vectors, they must be equal. ■

**New subspaces from old.** On your first assignment, you showed that if  $U$  and  $W$  are subspaces of a vector space  $V$ , then the intersection  $U \cap W$  is also a subspace of  $V$ . You also showed that the union  $U \cup W$  is generally not a subspace, unless one subspace is contained in the other (in which case the union is just the larger of the two subspaces we already have).

In class, we discussed the fact that the right way to define a subspace containing both  $U$  and  $W$  is using their **sum**: we define the sum  $U + W$  of two subspaces by

$$U + W = \{\vec{u} + \vec{w} \mid \vec{u} \in U \text{ and } \vec{w} \in W\}.$$

We proved that  $U + W$  is again a subspace of  $V$ .

If  $U \cap W = \{\vec{0}\}$ , we say that the sum is a **direct sum**, and write it as  $U \oplus W$ . The following theorem might help us understand why direct sums are singled out for special attention:

**Theorem 2.4.18** *Let  $U$  and  $W$  be subspaces of a finite-dimensional vector space  $V$ . Then  $U + W$  is finite-dimensional, and*

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

If the sum is direct, then we have simply  $\dim(U \oplus W) = \dim U + \dim W$ . The other reason why direct sums are preferable, is that any  $\vec{v} \in U \oplus W$  can be written *uniquely* as  $\vec{v} = \vec{u} + \vec{w}$  where  $\vec{u} \in U$  and  $\vec{w} \in W$ .

**Theorem 2.4.19** *For any subspaces  $U, W$  of a vector space  $V$ ,  $U \cap W = \{\vec{0}\}$  if and only if for every  $\vec{v} \in U + W$  there exist unique  $\vec{u} \in U, \vec{w} \in W$  such that  $\vec{v} = \vec{u} + \vec{w}$ .*

*Proof.* Suppose that  $U \cap W = \{\vec{0}\}$ , and suppose that we have  $\vec{v} = \vec{u}_1 + \vec{w}_1 = \vec{u}_2 + \vec{w}_2$ , for  $\vec{u}_1, \vec{u}_2 \in U, \vec{w}_1, \vec{w}_2 \in W$ . Then  $\vec{0} = (\vec{u}_1 - \vec{u}_2) + (\vec{w}_1 - \vec{w}_2)$ , which implies that

$$\vec{w}_1 - \vec{w}_2 = -(\vec{u}_1 - \vec{u}_2).$$

Now,  $\vec{u} = \vec{u}_1 - \vec{u}_2 \in U$ , since  $U$  is a subspace, and similarly,  $\vec{w} = \vec{w}_1 - \vec{w}_2 \in W$ . But we also have  $\vec{w} = -\vec{u}$ , which implies that  $\vec{w} \in U$ . Therefore,  $\vec{w} \in U \cap W$ , which implies that  $\vec{w} = \vec{0}$ , so  $\vec{w}_1 = \vec{w}_2$ . But we must also then have  $\vec{u} = \vec{0}$ , so

$$\vec{u}_1 = \vec{u}_2.$$

Conversely, suppose that every  $\vec{v} \in U + W$  can be written uniquely as  $\vec{v} = \vec{u} + \vec{w}$ , with  $\vec{u} \in U$  and  $\vec{w} \in W$ . Suppose that  $\vec{a} \in U \cap W$ . Then  $\vec{a} \in U$  and  $\vec{a} \in W$ , so we also have  $-\vec{a} \in W$ , since  $W$  is a subspace. But then  $\vec{0} = \vec{a} + (-\vec{a})$ , where  $\vec{a} \in U$  and  $-\vec{a} \in W$ . On the other hand,  $\vec{0} = \vec{0} + \vec{0}$ , and  $\vec{0}$  belongs to both  $U$  and  $W$ . It follows that  $\vec{a} = \vec{0}$ . Since  $\vec{a}$  was arbitrary,  $U \cap W = \{\vec{0}\}$ . ■

We end with one last application of the theory we've developed on the existence of a basis for a finite-dimensional vector space. As we continue on to later topics, we'll find that it is often useful to be able to decompose a vector space into a direct sum of subspaces. Using bases, we can show that this is always possible.

**Theorem 2.4.20** *Let  $V$  be a finite-dimensional vector space, and let  $U$  be any subspace of  $V$ . Then there exists a subspace  $W \subseteq V$  such that  $U \oplus W = V$ .*

*Proof.* Let  $\{\vec{u}_1, \dots, \vec{u}_m\}$  be a basis of  $U$ . Since  $U \subseteq V$ , the set  $\{\vec{u}_1, \dots, \vec{u}_m\}$  is a linearly independent subset of  $V$ . Since any linearly independent set can be extended to a basis of  $V$ , there exist vectors  $\vec{w}_1, \dots, \vec{w}_n$  such that

$$\{\vec{u}_1, \dots, \vec{u}_m, \vec{w}_1, \dots, \vec{w}_n\}$$

is a basis of  $V$ .

Now, let  $W = \text{span}\{\vec{w}_1, \dots, \vec{w}_n\}$ . Then  $W$  is a subspace, and  $\{\vec{w}_1, \dots, \vec{w}_n\}$  is a basis for  $W$ . (It spans, and must be independent since it's a subset of an independent set.)

Clearly,  $U + W = V$ , since  $U + W$  contains the basis for  $V$  we've constructed. To show the sum is direct, it suffices to show that  $U \cap W = \{\vec{0}\}$ . To that end, suppose that  $\vec{v} \in U \cap W$ . Since  $\vec{v} \in U$ , we have

$$\vec{v} = a_1\vec{u}_1 + \dots + a_m\vec{u}_m$$

for scalars  $a_1, \dots, a_m$ . Since  $\vec{v} \in W$ , we can write

$$\vec{v} = b_1\vec{w}_1 + \dots + b_n\vec{w}_n$$

for scalars  $b_1, \dots, b_n$ . But then

$$\vec{0} = \vec{v} - \vec{v} = a_1\vec{u}_1 + \dots + a_m\vec{u}_m - b_1\vec{w}_1 - \dots - b_n\vec{w}_n.$$

Since  $\{\vec{u}_1, \dots, \vec{u}_m, \vec{w}_1, \dots, \vec{w}_n\}$  is a basis for  $V$ , it's independent, and therefore, all of the  $a_i, b_j$  must be zero, and therefore,  $\vec{v} = \vec{0}$ . ■

The subspace  $W$  constructed in the theorem above is called a **complement** of  $U$ . It is not unique; indeed, it depends on the choice of basis vectors. For example, if  $U$  is a one-dimensional subspace of  $\mathbb{R}^2$ ; that is, a line, then any other non-parallel line through the origin provides a complement of  $U$ . Later we will see that an especially useful choice of complement is the **orthogonal complement**.

## Chapter 3

# Linear Transformations

At an elementary level, Linear Algebra is the study of vectors (in  $\mathbb{R}^n$ ) and matrices. Of course, much of that study revolves around systems of equations. Recall that if  $\vec{x}$  is a vector in  $\mathbb{R}^n$  (viewed as an  $n \times 1$  column matrix), and  $A$  is an  $m \times n$  matrix, then  $\vec{y} = A\vec{x}$  is a vector in  $\mathbb{R}^m$ . Thus, multiplication by  $A$  produces a function from  $\mathbb{R}^n$  to  $\mathbb{R}^m$ .

This example motivates the definition of a *linear transformation*, and as we'll see, provides the archetype for all linear transformations in the finite-dimensional setting. Many areas of mathematics can be viewed at some fundamental level as the study of sets with certain properties, and the functions between them. Linear algebra is no different. The sets in this context are, of course, vector spaces. Since we care about the linear algebraic structure of vector spaces, it should come as no surprise that we're most interested in functions that preserve this structure. That is precisely the idea behind linear transformations.

### 3.1 Definition and examples

Let  $V$  and  $W$  be vector spaces. At their most basic, all vector spaces are sets. Given any two sets, we can consider functions from one to the other. The functions of interest in linear algebra are those that respect the vector space structure of the sets.

**Definition 3.1.1** Let  $V$  and  $W$  be vector spaces. A function  $T : V \rightarrow W$  is called a **linear transformation** if:

1. For all  $\vec{v}_1, \vec{v}_2 \in V$ ,  $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$ .
2. For all  $\vec{v} \in V$  and scalars  $c$ ,  $T(c\vec{v}) = cT(\vec{v})$ .

We often use the term **linear operator** to refer to a linear transformation  $T : V \rightarrow V$  from a vector space to itself.  $\diamond$

Note on notation: it is common usage to drop the usual parentheses of function notation when working with linear transformations, as long as this does not cause confusion. That is, one might write  $T\vec{v}$  instead of  $T(\vec{v})$ , but one should never write  $T\vec{v} + \vec{w}$  in place of  $T(\vec{v} + \vec{w})$ , for the same reason that one should never write  $2x + y$  in place of  $2(x + y)$ . Mathematicians often think of linear transformations in terms of matrix multiplication, which probably explains this notation to some extent.

The properties of a linear transformation tell us that a linear map  $T$  *preserves* the operations of addition and scalar multiplication. (When the domain

and codomain are different vector spaces, we might say that  $T$  *intertwines* the operations of the two vector spaces.) In particular, any linear transformation  $T$  must preserve the zero vector, and respect linear combinations.

**Theorem 3.1.2** *Let  $T : V \rightarrow W$  be a linear transformation. Then*

1.  $T(\vec{0}_V) = \vec{0}_W$ , and
2. For any scalars  $c_1, \dots, c_n$  and vectors  $\vec{v}_1, \dots, \vec{v}_n \in V$ ,

$$T(c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_n\vec{v}_n) = c_1T(\vec{v}_1) + c_2T(\vec{v}_2) + \dots + c_nT(\vec{v}_n).$$

*Proof.*

1. Since  $\vec{0}_V + \vec{0}_V = \vec{0}_V$ , we have

$$T(\vec{0}_V) = T(\vec{0}_V + \vec{0}_V) = T(\vec{0}_V) + T(\vec{0}_V).$$

Adding  $-T(\vec{0}_V)$  to both sides of the above gives us  $\vec{0}_W = T(\vec{0}_V)$ .

2. The addition property of a linear transformation can be extended to sums of three or more vectors using associativity. Therefore, we have

$$\begin{aligned} T(c_1\vec{v}_1 + \dots + c_n\vec{v}_n) &= T(c_1\vec{v}_1) + \dots + T(c_n\vec{v}_n) \\ &= c_1T(\vec{v}_1) + \dots + c_nT(\vec{v}_n), \end{aligned}$$

where the second line follows from the scalar multiplication property. ■

**Example 3.1.3** Let  $V = \mathbb{R}^n$  and let  $W = \mathbb{R}^m$ . For any  $m \times n$  matrix  $A$ , the map  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  defined by

$$T_A(\vec{x}) = A\vec{x}$$

is a linear transformation. (This follows immediately from properties of matrix multiplication.)

Let  $B = \{\vec{e}_1, \dots, \vec{e}_n\}$  denote the standard basis of  $\mathbb{R}^n$ . Recall that  $A\vec{e}_i$  is equal to the  $i$ th column of  $A$ . Thus, if we know the value of a linear transformation  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  on each basis vector, we can immediately determine the matrix  $A$  such that  $T = T_A$ :

$$A = [T(\vec{e}_1) \quad T(\vec{e}_2) \quad \dots \quad T(\vec{e}_n)].$$

This is true because  $T$  and  $T_A$  agree on the standard basis: for each  $i = 1, 2, \dots, n$ ,

$$T_A(\vec{e}_i) = A\vec{e}_i = T(\vec{e}_i).$$

Moreover, if two linear transformations agree on a basis, they must be equal. Given any  $\vec{x} \in \mathbb{R}^n$ , we can write  $\vec{x}$  uniquely as a linear combination

$$\vec{x} = c_1\vec{e}_1 + c_2\vec{e}_2 + \dots + c_n\vec{e}_n.$$

If  $T(\vec{e}_i) = T_A(\vec{e}_i)$  for each  $i$ , then by [Theorem 3.1.2](#) we have

$$\begin{aligned} T(\vec{x}) &= T(c_1\vec{e}_1 + c_2\vec{e}_2 + \dots + c_n\vec{e}_n) \\ &= c_1T(\vec{e}_1) + c_2T(\vec{e}_2) + \dots + c_nT(\vec{e}_n) \\ &= c_1T_A(\vec{e}_1) + c_2T_A(\vec{e}_2) + \dots + c_nT_A(\vec{e}_n) \\ &= T_A(c_1\vec{e}_1 + c_2\vec{e}_2 + \dots + c_n\vec{e}_n) \\ &= T_A(\vec{x}). \end{aligned}$$

□

Let's look at some other examples of linear transformations.

- For any vector spaces  $V, W$  we can define the **zero transformation**  $0 : V \rightarrow W$  by  $0(\vec{v}) = \vec{0}$  for all  $\vec{v} \in V$ .
- On any vector space  $V$  we have the **identity transformation**  $1_V : V \rightarrow V$  defined by  $1_V(\vec{v}) = \vec{v}$  for all  $\vec{v} \in V$ .
- Let  $V = F[a, b]$  be the space of all functions  $f : [a, b] \rightarrow \mathbb{R}$ . For any  $c \in [a, b]$  we have the **evaluation map**  $E_a : V \rightarrow \mathbb{R}$  defined by  $E_a(f) = f(a)$ .

To see that this is linear, note that  $E_a(0) = 0(a) = 0$ , where  $0$  denotes the zero function; for any  $f, g \in V$ ,

$$E_a(f + g) = (f + g)(a) = f(a) + g(a) = E_a(f) + E_a(g),$$

and for any scalar  $c \in \mathbb{R}$ ,

$$E_a(cf) = (cf)(a) = c(f(a)) = cE_a(f).$$

Note that the evaluation map can similarly be defined as a linear transformation on any vector space of polynomials.

- On the vector space  $C[a, b]$  of all *continuous* functions on  $[a, b]$ , we have the integration map  $I : C[a, b] \rightarrow \mathbb{R}$  defined by  $I(f) = \int_a^b f(x) dx$ . The fact that this is a linear map follows from properties of integrals proved in a calculus class.
- On the vector space  $C^1(a, b)$  of continuously differentiable functions on  $(a, b)$ , we have the differentiation map  $D : C^1(a, b) \rightarrow C(a, b)$  defined by  $D(f) = f'$ . Again, linearity follows from properties of the derivative.
- Let  $\mathbb{R}^\infty$  denote the set of sequences  $(a_1, a_2, a_3, \dots)$  of real numbers, with term-by-term addition and scalar multiplication. The shift operators

$$S_L(a_1, a_2, a_3, \dots) = (a_2, a_3, a_4, \dots)$$

$$S_R(a_1, a_2, a_3, \dots) = (0, a_1, a_2, \dots)$$

are both linear.

- On the space  $M_{mn}(\mathbb{R})$  of  $m \times n$  matrices, the trace defines a linear map  $\text{tr} : M_{mn}(\mathbb{R}) \rightarrow \mathbb{R}$ , and the transpose defines a linear map  $T : M_{mn}(\mathbb{R}) \rightarrow M_{nm}(\mathbb{R})$ . The determinant and inverse operations on  $M_{nn}$  are *not* linear.

For finite-dimensional vector spaces, it is often convenient to work in terms of a basis. The properties of a linear transformation tell us that we can completely define any linear transformation by giving its values on a basis. In fact, it's enough to know the value of a transformation on a spanning set. The argument given in [Example 3.1.3](#) can be applied to any linear transformation, to obtain the following result.

**Theorem 3.1.4** *Let  $T : V \rightarrow W$  and  $S : V \rightarrow W$  be two linear transformations. If  $V = \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$  and  $T(\vec{v}_i) = S(\vec{v}_i)$  for each  $i = 1, 2, \dots, n$ , then  $T = S$ .*

If the above spanning set is not also independent, then one might be concerned about the fact that there will be more than one way to express a vector as linear combination of vectors in that set. If we define  $T$  by giving its values on a spanning set, will it be well-defined? Suppose that we have scalars  $a_1, \dots, a_n, b_1, \dots, b_n$  such that

$$\vec{v}_1 a_1 + \dots + \vec{v}_n a_n = \vec{v}_1 b_1 + \dots + \vec{v}_n b_n$$

$$b_1\vec{v}_1 + \cdots + b_n\vec{v}_n$$

We then have

$$\begin{aligned} a_1T(\vec{v}_1) + \cdots + a_nT(\vec{v}_n) &= T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n) \\ &= T(b_1\vec{v}_1 + \cdots + b_n\vec{v}_n) \\ &= b_1T(\vec{v}_1) + \cdots + b_nT(\vec{v}_n). \end{aligned}$$

The next theorem seems like an obvious consequence of the above, and indeed, one might wonder where the assumption of a basis is needed. The distinction here is that the vectors  $\vec{w}_1, \dots, \vec{w}_n \in W$  are chosen in advance, and then we set  $T(\vec{v}_i) = \vec{w}_i$ , rather than simply defining each  $\vec{w}_i$  as  $T(\vec{b}_i)$ .

**Theorem 3.1.5** *Let  $V, W$  be vector spaces. Let  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  be a basis of  $V$ , and let  $\vec{w}_1, \dots, \vec{w}_n$  be any vectors in  $W$ . (These vectors need not be distinct.) Then there exists a unique linear transformation  $T : V \rightarrow W$  such that  $T(\vec{b}_i) = \vec{w}_i$  for each  $i = 1, 2, \dots, n$ ; indeed, we can define  $T$  as follows: given  $\vec{v} \in V$ , write  $\vec{v} = c_1\vec{b}_1 + \cdots + c_n\vec{b}_n$ . Then*

$$T(\vec{v}) = T(c_1\vec{b}_1 + \cdots + c_n\vec{b}_n) = c_1\vec{w}_1 + \cdots + c_n\vec{w}_n.$$

With the basic theory out of the way, let's look at a few basic examples.

**Exercise 3.1.6** Suppose  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is a linear transformation. If  $T \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$  and  $T \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ , find  $T \begin{bmatrix} -2 \\ 4 \end{bmatrix}$ .

**Solution.** Since we know the value of  $T$  on the standard basis, we can use properties of linear transformations to immediately obtain the answer:

$$\begin{aligned} T \begin{bmatrix} -2 \\ 4 \end{bmatrix} &= T \left( -2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \\ &= -2T \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 4T \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= -2 \begin{bmatrix} 3 \\ -4 \end{bmatrix} + 4 \begin{bmatrix} 5 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 14 \\ 16 \end{bmatrix}. \end{aligned}$$

**Exercise 3.1.7** Suppose  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is a linear transformation. Given that  $T \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$  and  $T \begin{bmatrix} 2 \\ -5 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ , find  $T \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ .

**Solution.** At first, this example looks the same as the one above, and to some extent, it is. The difference is that this time, we're given the values of  $T$  on a basis that is not the standard one. This means we first have to do some work to determine how to write the given vector in terms of the given basis.

Suppose we have  $a \begin{bmatrix} 3 \\ 1 \end{bmatrix} + b \begin{bmatrix} 2 \\ -5 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$  for scalars  $a, b$ . This is equivalent to the matrix equation

$$\begin{bmatrix} 3 & 2 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

Solving (perhaps using the code cell below), we get  $a = \frac{26}{17}, b = -\frac{5}{17}$ . Therefore,

$$T \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \frac{26}{17} \begin{bmatrix} 1 \\ 4 \end{bmatrix} - \frac{5}{17} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 16/17 \\ 109/17 \end{bmatrix}.$$



```

from sympy import *
init_printing()
A = Matrix(2,2,[3,2,1,-5])
B = Matrix(2,1,[4,3])
(A**-1)*B

```

**Exercise 3.1.8** Suppose  $T : P_2(\mathbb{R}) \rightarrow \mathbb{R}$  is defined by

$$T(x+2) = 1, T(1) = 5, T(x^2+x) = 0.$$

Find  $T(2-x+3x^2)$ .

**Solution.** We need to find scalars  $a, b, c$  such that

$$2-x+3x^2 = a(x+2) + b(1) + c(x^2+x).$$

We could set up a system and solve, but this time it's easy enough to just work our way through. We must have  $c = 3$ , to get the correct coefficient for  $x^2$ . This gives

$$2-x+3x^2 = a(x+2) + b(1) + 3x^2 + 3x.$$

Now, we have to have  $3x + ax = -x$ , so  $a = -4$ . Putting this in, we get

$$2-x+3x^2 = -4x-8+b+3x^2+3x.$$

Simplifying this leaves us with  $b = 10$ . Finally, we find:

$$\begin{aligned} T(2-x+3x^2) &= T(-4(x+2) + 10(1) + 3(x^2+x)) \\ &= -4T(x+2) + 10T(1) + 3T(x^2+x) \\ &= -4(1) + 10(5) + 3(0) = 46. \end{aligned}$$

**Exercise 3.1.9** Find a linear transformation  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  such that

$$T(1, 2) = (1, 1, 0) \quad \text{and} \quad T(-1, 1) = (0, 2, -1).$$

Then, determine the value of  $T(3, 2)$ .

**Solution.** Since  $\{(1, 2), (-1, 1)\}$  forms a basis of  $\mathbb{R}^2$  (the vectors are not parallel and there are two of them), it suffices to determine how to write a general vector in terms of this basis. suppose

$$x(1, 2) + y(-1, 1) = (a, b)$$

for a general element  $(a, b) \in \mathbb{R}^2$ . This is equivalent to the matrix equation

$$\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}. \text{ We find:}$$

$$(a, b) = \frac{1}{3}(a+b)(1, 2) + \frac{1}{3}(-2a+b)(-1, 1).$$

Thus,

$$\begin{aligned} T(a, b) &= \frac{1}{3}(a+b)T(1, 2) + \frac{1}{3}(-2a+b)T(-1, 1) \\ &= \frac{1}{3}(a+b)(1, 1, 0) + \frac{1}{3}(-2a+b)(0, 2, -1) \\ &= \left( \frac{a+b}{3}, -a+b, \frac{2a-b}{3} \right). \end{aligned}$$

Therefore,

$$T(3, 2) = \left( \frac{5}{3}, -1, \frac{4}{3} \right).$$

```
a, b = symbols('a_b', real = True, constant = True)
A = Matrix(2, 2, [1, -1, 2, 1])
B = Matrix(2, 1, [a, b])
(A**-1)*B
```

## 3.2 Kernel and Image

Given any linear transformation  $T : V \rightarrow W$  we can associate two important subspaces: the **kernel** of  $T$  (also known as the **nullspace**), and the **image** of  $T$  (also known as the **range**).

**Definition 3.2.1** Let  $T : V \rightarrow W$  be a linear transformation. The **kernel** of  $T$ , denoted  $\ker T$ , is defined by

$$\ker T = \{\vec{v} \in V \mid T(\vec{v}) = \vec{0}\}.$$

The **image** of  $T$ , denoted  $\operatorname{im} T$ , is defined by

$$\operatorname{im} T = \{(T\vec{v}) \mid \vec{v} \in V\}.$$

◇

Note that the kernel of  $T$  is just the set of all vectors  $T$  sends to zero. The image of  $T$  is the range of  $T$  in the usual sense of the range of a function.