Linear Algebra

J. Hoste
Pitzer College

April 26, 2017

Contents

Dedication

To Edmund Deaton, who taught me Euclidean Geometry "Texas style."

Preface

0.1 For the Student

These notes cover most of what would typically be included in any first course in Linear Algebra. But unlike a textbook, these notes contain *very little* in the way of explanations, examples, or proofs of important propositions and theorems! Instead, they comprise a bare-bones outline of the subject, simply giving all the necessary definitions and then stating the important results as exercises which are left to the reader! The idea is for each student to develop the subject for herself, discovering proofs of theorems rather than simply reading them in a textbook. In this way, a much greater depth of understanding of the material should follow. Fortunately, the Instructor will be there to help! And, in addition to learning linear algebra, the Student will learn basic strategies of proof and how to speak and write about mathematics in a careful, rigorous, and ultimately, elegant way.

This philosophy of teaching is old but was championed in recent history by the American mathematician R. L. Moore (1882–1974). A favorite Chinese proverb of Moore's was "I hear, I forget. I see, I remember. I do, I understand."[1]. He has been quoted as saying, "The student is taught the best who is told the least."[2] Indeed, Moore would prohibit students from consulting books or even talking to each other! Instead, they were to work everything out for themselves. Of course, in class, Moore would teach student how to think about mathematics, how to approach problem solving, and how to speak and write about mathematics with style.

I was first introduced to the "Moore method" of teaching while still in high school when I attended a summer program in mathematics at San Diego State University. There I took a course in Euclidean Geometry taught by Professor Edmund Deaton in strict Moore fashion. Deaton earned his

[1] See Halmos, P.R. (1985) *I want to be a mathematician: an automathography.* Springer-Verlag: 258

[2] Parker, John, 2005. *R. L. Moore: Mathematician and Teacher*

Ph.D. in mathematics from the University of Texas, Austin, in 1960 while Moore was still on the faculty there and presumably experienced Moore's teaching first-hand. The experience that summer in San Diego was truly inspirational, starting me on my life-long path as a mathematician.

Jim Hoste
Claremont, CA 2017

0.2 Some Symbols and Definitions

1. $\mathbb{N}$, the set of *natural* numbers, 1, 2, 3, 4, $\ldots$.
2. $\mathbb{Z}$, the set of *integers*, $\ldots$, -2, -1, 0, 1, 2, $\ldots$
3. $\mathbb{Q}$, the set of *rational* numbers, that is, all fractions $\frac{n}{m}$ where n and m are integers and $m \neq 0$.
4. $\mathbb{R}$, the set of *real* numbers. Every real number can be expressed as a decimal, possibly and infinite decimal.
5. i, the imaginary number $\sqrt{-1}$
6. $\mathbb{C}$, the set of *complex* numbers, $\mathbb{C} = \{a + bi \,|\, a, b, \in \mathbb{R}\}$.
7. $A \subset B$, means A is a subset of B, that is, every element of A is an element of B.
8. $\mathbb{F}$, a *field*, usually the real or complex numbers.
9. $x \in S$, x is an element of the set S.
10. $f(x)$ is *even* if $f(-x) = f(x)$ for all x.
11. $f(x)$ is *odd* if $f(-x) = -f(x)$ for all x.

Chapter 1

Fields

1.1 Complex Numbers

Definition 1. *Let i be the "imaginary number" whose square is -1. The complex numbers are the set*

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

If $z = a + bi$, then a is called the real part of z, denoted Re z and b is called the imaginary part of z, denoted Im z. The complex numbers $a + bi$ and $c+di$ are equal if and only if their real and imaginary parts, respectively, are equal. That is, if and only if $a = c$ and $b = d$. Addition and multiplication of complex numbers are defined as

$$\begin{aligned}(a + bi) + (c + di) &= (a + b) + (c + d)i\\ (a + bi)(c + di) &= (ac - bd) + (ad + bc)i\end{aligned}$$

Exercise 1. *Show all the following properties for complex numbers.*

1. *Addition and multiplication are commutative, that is,*

$$u + v = v + u \quad \text{and} \quad uv = vu$$

   for all complex numbers u and v.

2. *Addition and multiplication are associative, that is,*

$$(u + v) + w = u + (v + w) \quad \text{and} \quad (uv)w = u(vw)$$

   for all complex numbers u and v.

3. *The numbers* 0 *and* 1 *are additive and multiplicative identities, respectively. That is,*

$$0 + u = u \quad \text{and} \quad 1u = u$$

for any complex number u.

4. *There exist additive and multiplicative inverses. That is*

   (a) *For every* $u \in \mathbb{C}$, *there exists a unique* $v \in \mathbb{C}$ *such that* $u + v = 0$.

   (b) *For every* $u \in \mathbb{C}$ *with* $u \neq 0$, *there exists a unique* $v \in \mathbb{C}$ *such that* $uv = 1$.

5. *The complex numbers satisfy the distributive property. That is,*

$$u(v + w) = uv + uw$$

for any complex numbers u, v, *and* w.

Definition 2.

1. *If* z *is any complex number, let* $-z$ *denote its additive inverse. Define subtraction as* $w - z = w + (-z)$.

2. *If* z *is any nonzero complex number, let* $1/z$ *denote its multiplicative inverse. Define division as* $w/z = w(1/z)$.

3. *If* $z = a + bi$ *is any complex number, its complex conjugate, denoted by* $\overline{z}$, *is defined as* $\overline{z} = a - bi$.

4. *The absolute value of* $z = a + bi$, *denoted* $|z|$, *is defined as* $|z| = \sqrt{a^2 + b^2}$.

Exercise 2. *Let* u, v, *and* w *be any complex numbers. Show that all the following are true.*

1. $\mathit{Re}(u + v) = \mathit{Re}\ u + \mathit{Re}\ v$

2. $\mathit{Im}(u + v) = \mathit{Im}\ u + \mathit{Im}\ v$

3. $u + \overline{u} = 2\mathit{Re}\ u$

4. $u - \overline{u} = 2(\mathit{Im}\ u)i$

5. $u\overline{u} = |u|^2$

6. $\overline{u + v} = \overline{u} + \overline{v}$

7. $\overline{uv} = \overline{u}\,\overline{v}$

8. $\overline{\overline{u}} = u$

9. $|uv| = |u|\,|v|$

The complex number $a + bi$ can be associated to the ordered pair of real numbers (a, b). This allows us to graph, or plot, a complex number as a point in the plane using Cartesian coordinates. The horizontal axis is called the *real* axis and the vertical axis is called the *imaginary* axis.

Exercise 3. *Plot the following complex numbers in the plane:*
$5, 0, -5, 1 + i, (-1 + \sqrt{3}i)/2, i, i^2, i^3, i^4, 4 - 3i$.

Exercise 4. *Given a point in the plane that represents the complex number z and given any real number a, describe what point in the plane represents the complex number az? Does it matter if a is positive, negative, or zero?*

Exercise 5. *Given two points in the plane that represent the complex numbers z and w, give a graphical explanation of how to find the point that corresponds to their sum $z + w$.*

Exercise 6 (*)**.** *Given two points in the plane that represent the complex numbers z and w, give a graphical explanation of how to find the point that corresponds to their product zw. This is beautiful, but not so easy to discover. Follow this outline: Given a point in the plane that represents the complex number z, consider the line ℓ through the origin and z and let θ be the angle this line makes with the real axis. Assume that θ is measured from the real axis to ℓ in a counterclockwise fashion*

1. *Show that $z = |z|(\cos\theta + i\sin\theta)$.*

2. *If $z = |z|(\cos\theta + i\sin\theta)$ and $w = |w|(\cos\phi + i\sin\phi)$, show that*

$$zw = |z|\,|w|(\cos(\theta + \phi) + i\sin(\theta + \phi)).$$

Combine these facts with the fact that $|wz| = |w|\,|z|$ to finish the exercise.

Exercise 7. *If a point in the plane represents the complex number $z \neq 0$, where are the points in the plane that represent the following:*

1. $-z$

2. $\overline{z}$

3. $1/z$

The rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$ are each examples of a more general mathematical object called a *field*. We don't really need to know what a field is in general, but for completeness, the definition is given below. Notice how the *definition* of a field is based exactly on the *properties* that are listed in Exercise 1.

Definition 3. *A field is a set $\mathbb{F}$ with two binary operations, called addition and multiplication, that satisfy the following:*

1. *Addition is commutative:*

$$x + y = y + x \text{ for all } x \text{ and } y \text{ in } \mathbb{F}.$$

2. *Addition is associative:*

$$x + (y + z) = (x + y) + z \text{ for all } x, y, \text{ and } z \text{ in } \mathbb{F}.$$

3. *There is a unique additive identity in $\mathbb{F}$ which is denoted 0:*

$$x + 0 = x \text{ for all } x \text{ in } \mathbb{F}.$$

4. *Every element has a unique additive inverse:*

$$\text{if } x \in \mathbb{F}, \text{ then there exists a unique element } y \in \mathbb{F} \text{ such that } x + y = 0.$$

   The additive inverse of x is denoted $-x$.

5. *Multiplication is commutative:*

$$xy = yx \text{ for all } x, y \in \mathbb{F}.$$

6. *Multiplication is associative:*

$$x(yz) = (yx)z \text{ for all } x, y, z \in \mathbb{F}.$$

7. *There is a unique, non-zero, multiplicative identity in $\mathbb{F}$, denoted by 1:*

$$1x = x \text{ for all } x \text{ in } \mathbb{F}.$$

8. *Every non-zero element $x \in \mathbb{F}$ has a unique multiplicative inverse, denoted by $1/x$ or x^{-1}:*

$$xx^{-1} = 1 \text{ for all } x \in \mathbb{F}.$$

9. *Multiplication distributes over addition:*

$$x(y+z) = xy + xz \text{ for all } x, y, z \in \mathbb{F}.$$

If we had started by giving the definition of a field, then Exercise 1 could have said: Show that the rational numbers, the real numbers, and the complex numbers are all fields. This would have been harder because we used the fact that $\mathbb{R}$ is a field to *show* that $\mathbb{C}$ is a field!

Exercise 8. *Show that the integers do NOT form a field.*

Chapter 2

Vector Spaces

2.1 Vector Spaces

For the remainder of this course $\mathbb{F}$ will stand for a *field*. We have three important examples of a field: either the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, or the complex numbers $\mathbb{C}$. These are perhaps the most important examples of fields, but they are not the only possible fields. In this course, we (probably) will not use any fields other than these. The numbers in $\mathbb{F}$ will often be called *scalars*.

Definition 4. *A vector space over the field* $\mathbb{F}$ *of scalars is a set* V *of objects, called vectors, such that*

1. *There is a rule, or operation, called vector addition, that associates to every pair of vectors* $u, v \in V$*, a unique element of* V *called their sum and denoted* $u + v$*, in such a way that*
   (a) *addition is commutative:* $u + v = v + u$ *for all* $u, v \in V$*,*
   (b) *addition is associative:* $u+(v+w) = (u+v)+w$ *for all* $u, v, w \in V$*,*
   (c) *there exists a unique element of* V*, called the additive identity and denoted by* 0*, such that* $0 + v = v$ *for all* $v \in V$*, and*
   (d) *every vector has a unique additive inverse: if* v *is any element of* V*, then there exists a unique element* u *of* V *such that* $v + u = 0$*. We denote the additive inverse of* v *by* $-v$*.*
2. *There is a rule, or operation, called scalar multiplication, that associates to every scalar* $a \in \mathbb{F}$ *and every vector* $v \in V$ *a vector in* V*, called the product of* a *and* v *and denoted by* av*, in such a way that*
   (a) *If* v *is any vector and* 1 *is the multiplicative identity in* $\mathbb{F}$*, then* $1v = v$*.*
   (b) *If* v *is any vector and* $a, b \in \mathbb{F}$*, then* $a(bv) = (ab)v$*.*
3. *Addition and scalar multiplication satisfy the following distributive laws:*
   (a) $a(u + v) = au + av$ *for all* $a \in \mathbb{F}$ *and for all* $u, v \in V$
   (b) $(a + b)v = av + bv$ *for all* $a, b \in \mathbb{F}$ *and for all* $v \in V$*.*

We define subtraction of vectors as $u - v = u + (-v)$*. If* $\mathbb{F} = \mathbb{Q}, \mathbb{R},$ *or* $\mathbb{C}$*, then* V *is called a rational, real, or complex vector space, respectively.*

Warning: We are going to use 0 to denote both the *number*, or *scalar*, in $\mathbb{F}$ that is the additive identity in $\mathbb{F}$ **and** the *vector* in V that is the additive identity in V. You are going to have to be aware of the context to know whether 0 means the number or the vector!

In general, a vector space is a totally abstract thing. So we need some concrete examples of vector spaces.

Definition 5. The vector space $\mathbb{F}^n$ *Let* n *be a natural number. An ordered set* $(x_1, x_2, \ldots, x_n)$ *of* n *elements of* $\mathbb{F}$ *is called an* n*-tuple. The set of all*

n-tuples is denoted $\mathbb{F}^n$.

We define addition on $\mathbb{F}^n$ by

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

If $a \in \mathbb{F}$ we define scalar multiplication by

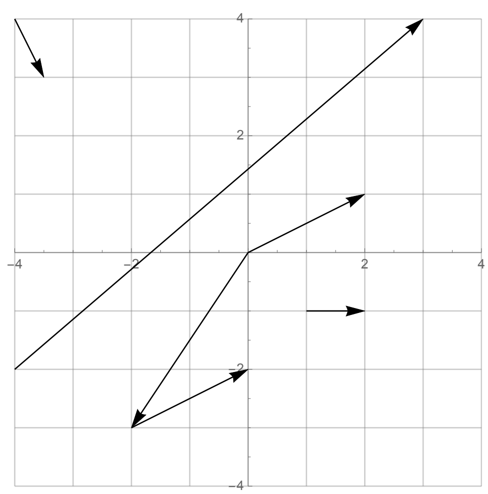
$$a(x_1, x_2, \ldots, x_n) = (ax_1, ax_2, \ldots, ax_n).$$

Exercise 9. *Show that for each n, the set $\mathbb{F}^n$ is a vector space over $\mathbb{F}$.*

Note that if $n = 1$, we have that $\mathbb{F}$ is a vector space over $\mathbb{F}$.

The vector spaces $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3, \ldots$ are perhaps our most important examples as we begin to learn about vector spaces. A nice thing about $\mathbb{R}, \mathbb{R}^2$, and $\mathbb{R}^3$ is that we can actually visualize them. A vector in $\mathbb{R}^2$ is an ordered pair of real numbers (a, b). These numbers, a and b, are called the *components* of the vector. We can think of a vector as a point in the plane using Cartesian coordinates, but we can also think of it as a "directed line segment" or an "arrow" in the plane with *any* initial point (x_0, y_0) and terminal point $(x_1, y_1) = (x_0 + a, y_0 + b)$. If the initial point is $(0, 0)$, then the terminal point is (a, b), the same as the vector. But if the initial point is not $(0, 0)$, then neither the initial point nor the terminal point need be equal to the vector (a, b).

Exercise 10. *A number of vectors (arrows) in $\mathbb{R}^2$ are shown below. For each one, determine the components of the vector. Are any of the vectors the same?*



Exercise 11. *If two vectors in $\mathbb{R}^2$ are given as arrows in the plane, show how to find the vector (an arrow) that corresponds to their sum.*

Exercise 12. *If a vector v in $\mathbb{R}^2$ is given as an arrow in the plane, explain how to find an arrow corresponding to the vector $2v$. More generally, how do you find an arrow representing av for any real number a?*

A vector in $\mathbb{F}^n$ is an n-tuple $(x_1, x_2, \dots, x_n)$. Sometimes we will write this as a *row* $(x_1, x_2, \dots, x_n)$ and sometimes we will write it as a *column* $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$.

Exercise 13. *Show that the set of all functions $f : \mathbb{F} \to \mathbb{F}$, with the sum of two functions defined as $(f + g)(x) = f(x) + g(x)$ and the product of a scalar and a function defined as $(af)(x) = af(x)$, is a vector space over $\mathbb{F}$.*

Definition 6. *A function $p : \mathbb{F} \to \mathbb{F}$ is called a polynomial with coefficients in $\mathbb{F}$ if either it is the zero function, that is, $p(z) = 0$ for all z, or if there exist numbers $a_0, a_1, \dots, a_m \in \mathbb{F}$ with $a_m \neq 0$ such that*

$$p(z) = a_0 + a_1 z + a_2 z^2 + \dots a_m z^m$$

for all $z \in \mathbb{F}$. In the latter case, the non-negative integer m is called the degree of the polynomial.

Exercise 14 (*)**.** *Show that the degree of a polynomial is well defined. That is, if $p : \mathbb{F} \to \mathbb{F}$ is a polynomial of degree m, then it cannot be a polynomial of degree ℓ with $\ell \neq m$. Equivalently, if p and q are polynomials of different degrees, then they are not the same function.*

Exercise 15.

1. *Show that the sum of two polynomials with coefficients in $\mathbb{F}$ is a polynomial with coefficients in $\mathbb{F}$.*

2. *If p is a polynomial with coefficients in $\mathbb{F}$ and $a \in F$, show that ap is a polynomial with coefficients in $\mathbb{F}$.*

3. *Show that $\mathcal{P}(\mathbb{F})$, the set of all polynomials with coefficients in $\mathbb{F}$, is a vector space over $\mathbb{F}$.*

Exercise 16. *Show that* $0v = 0$ *for all* $v \in V$*. (Warning: The "0" on the left side of the equality is the additive identity in* $\mathbb{F}$*, while the "0" on the right side of the equality is the additive identity in* V*. The first is a number, the second is a vector.)*

Exercise 17. *Show that* $a0 = 0$ *for all* $a \in \mathbb{F}$*. (Here both* 0*'s are the zero vector in* V*.)*

Exercise 18. *Assume* V *is a vector space over* $\mathbb{F}$*. Show that* $(-a)v = a(-v) = -(av)$ *for every* $v \in V$ *and every* $a \in \mathbb{F}$*. Because of this result, we will simply write* $-av$ *for any of* $(-a)v, a(-v)$*, or* $-(av)$*.*

Exercise 19. *Assume* V *is a vector space over* $\mathbb{F}$*. Show that* $-(-v) = v$ *for all* $v \in V$*.*

Exercise 20. *Assume* V *is a vector space over* $\mathbb{F}$ *with* $a \in \mathbb{F}$ *and* $v \in V$*. Show that if* $av = 0$*, then* $a = 0$ *or* $v = 0$*.*

2.2 Subspaces of Vector Spaces

Definition 7. *Suppose that V is a vector space over $\mathbb{F}$. A subset $W \subset V$ is called a subspace if all three of the following are true.*

1. *The zero vector 0 is in W. (W contains the zero vector.)*
2. *If u and v are in W, then so is $u + v$. (W is closed under addition.)*
3. *If u is in W and a is in $\mathbb{F}$, then $au \in W$. (W is closed under scalar multiplication.)*

Exercise 21. *Determine if each of the following subsets W of $\mathbb{R}^2$ is a subspace. Either prove that it is or show that it is not.*

1. $W = \{(0, 0)\}$.
2. *W is a finite set of two or more vectors.*
3. *W is a line containing the origin, that is, all scalar multiples of some vector v.*
4. *W is a line not containing the origin, that is, $W = \{w + tv\}$ where w and v are nonzero vectors, w is not a multiple of v, and $t \in \mathbb{R}$.*
5. *W is the set of all vectors $v = (a, b)$ with $\sqrt{a^2 + b^2} \leq 1$.*
6. $W = \mathbb{R}^2$.

Exercise 22 (*)**.** *Show that a subspace of $\mathbb{R}^2$ is either*

1. $\{(0, 0)\}$, *or*
2. *A line containing the origin, or*
3. *All of $\mathbb{R}^2$.*

Exercise 23. *Give an example of a nonempty subset W of $\mathbb{R}^3$ that is closed under addition but not under scalar multiplication (and hence is not a subspace).*

Exercise 24. *Give an example of a nonempty subset W of $\mathbb{R}^3$ that is closed under scalar multiplication but not under addition (and hence is not a subspace).*

Exercise 25. *If U and W are subspaces of V, show that $U \cap W$ is a subspace of V.*

Definition 8. *If $W_1, W_2, \ldots, W_k$ are subspaces of V, define their sum $W_1 + W_2 \cdots + W_k$ to be set of all possible sums of vectors from $W_1, W_2, \ldots, W_k$. That is,*

$$W_1 + W_2 \cdots + W_k = \{w_1 + w_2 \cdots + w_k \,|\, w_1 \in W_1, w_2 \in W_2, \ldots, w_k \in W_k\}.$$

Exercise 26. *If $W_1, W_2, \ldots, W_k$ are subspaces of V, show that $W_1 + W_2 \cdots + W_k$ is a subspace of V.*

Exercise 27. *Determine the sum $W_1 + W_2$ for each of the following examples.*

1. *$V = \mathbb{R}^3$, $W_1 = \{(x, 0, 0) \,|\, x \in \mathbb{R}\}$, and $W_1 = \{(0, y, 0) \,|\, y \in \mathbb{R}\}$.*
2. *$V = \mathcal{P}(\mathbb{F})$, W_1 is the subspace of even polynomials, and W_2 is the subspace of odd polynomials.*

Exercise 28. *Assume that $W_1, W_2, \ldots, W_k$ are subspaces of V. Show that $W_1 + W_2 \cdots + W_k$ is the smallest subspace of V containing each W_i. In other words, if W is a subspace of V and $W_i \subset W$ for every i, then $W_1 + W_2 \cdots + W_k \subset W$.*

Definition 9. *The vector space V is the direct sum of the subspaces $W_1, W_2, \ldots, W_k$, written*

$$V = W_1 \oplus W_2 \cdots \oplus W_k,$$

if each vector $v \in V$ can be written uniquely as $v = w_1 + w_2 \cdots + w_k$, where each $w_i \in W_i$.

Exercise 29. *Determine if the given vector space is the direct sum of the given subspaces.*

1. *$V = \mathbb{F}^3$*

$$\begin{aligned} W_1 &= \{(x, y, 0) \,|\, x, y, \in \mathbb{F}\} \\ W_2 &= \{(0, 0, z) \,|\, z \in \mathbb{F}\} \end{aligned}$$

2. *$V = \mathbb{F}^3$*

$$\begin{aligned} W_1 &= \{(x, y, 0) \,|\, x, y, \in \mathbb{F}\} \\ W_2 &= \{(0, 0, z) \,|\, z \in \mathbb{F}\} \\ W_3 &= \{(0, z, z) \,|\, z \in \mathbb{F}\} \end{aligned}$$

3. *$V = \mathcal{P}(\mathbb{F})$, W_1 is the subspace of even polynomials, and W_2 is the subspace of odd polynomials.*

Exercise 30. *Show that if $W_1, W_2, \ldots, W_k$ are subspaces of V, then $V = W_1 \oplus W_2 \cdots \oplus W_k$ if and only if both of the following conditions are true.*

1. *$V = W_1 + W_2 \cdots + W_k$*

2. *If $0 = w_1 + w_2 \cdots + w_k$, with each $w_i \in W_i$, then $w_i = 0$ for all i.*

Exercise 31. *Show that if W_1 and W_2 are subspaces of V, then $V = W_1 \oplus W_2$ if and only if both of the following conditions are true.*

1. *$V = W_1 + W_2$*

2. *$W_1 \cap W_2 = \{0\}$*

Exercise 32. *Prove that each of the following statements is true, or give a counterexample to show that it is false.*

1. *If W_1, W_2, and U are subspaces of V such that $W_1 + U = W_2 + U$, then $W_1 = W_2$.*

2. *If W_1, W_2, and U are subspaces of V such that $V = W_1 \oplus U$ and $V = W_2 \oplus U$, then $W_1 = W_2$.*

2.3 Span and Linear Independence

Definition 10. *A linear combination of a list of vectors* $v_1, v_2, \ldots, v_k$ *in a vectors space* V *over* $\mathbb{F}$ *is a vector of the form*

$$a_1 v_1 + a_2 + v_2 \cdots + a_k v_k$$

where $a_i \in \mathbb{F}$. *The linear combination is called trivial if* $a_i = 0$ *for all* i *and non trivial if at least one of the* a_i*'s is nonzero. The set of all such linear combinations is called the span of* $v_1, v_2, \ldots v_k$ *and is denoted* $span(v_1, v_2, \ldots, v_k)$.

Exercise 33. *Show that if* $v_1, v_2, \ldots, v_k$ *are vectors in* V, *then* $W = span(v_1, v_2, \ldots, v_k)$ *is a subspace of* V. *We say that the vectors* $v_1, v_2, \ldots, v_k$ *span* W.

Definition 11. *A list of vectors* $v_1, v_2, \ldots, v_k$ *in* V *is called linearly independent if* $a_1 v_1 + a_2 v_2 \cdots + a_k v_k = 0$ *implies that* $a_1 = a_2 \cdots = a_k = 0$. *In other words, a list of vectors is linearly independent if it is only their trivial linear combination that is zero. A list of vectors is called linearly dependent if it is not linearly independent. In other words, a list of vectors is linearly dependent if some nontrivial linear combination of them is zero.*

Exercise 34. *Show that a list of vectors in* V *is linearly dependent if and only if one of the vectors is a linear combination of the others.*

Exercise 35.

1. *Show that* $(2, 1, -1), (0, 3, 1)$, *and* $(5, 4, -2)$ *are linearly dependent in* $\mathbb{R}^3$.
2. *Show that* $1, z, z^2, \ldots, z^m$ *are linearly independent in* $\mathcal{P}(\mathbb{F})$.
3. *Show that a list of a single vector is linearly independent if and only if the vector is nonzero.*
4. *Show that a list of two vectors is linearly independent if and only if neither is a scalar multiple of the other.*
5. *Show that any sublist of a list of linearly independent vectors is linearly independent.*
6. *Show that any superlist of a list of linearly dependent vectors is linearly dependent.*

Definition 12. *A list of vectors* $v_1, v_2, \ldots v_k$ *is a basis for* V *if it spans* V *and is linearly independent.*

Exercise 36. *For* $1 \le i \le n$*, let* $e_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{F}^n$ *be the vector whose* i*-th component is 1 and all other components are 0. Show that* $e_1, e_2, \ldots, e_n$ *form a basis for* $\mathbb{F}^n$*. This is called the standard basis.*

Exercise 37. *Show that a list of vectors* $v_1, v_2, \ldots v_k$ *is a basis for* V *if and only if every vector in* V *can be written uniquely as a linear combination of* $v_1, v_2, \ldots v_k$*.*

Exercise 38 (**)**. Basis Extension Theorem** *Suppose that* $(v_1, v_2, \ldots, v_k)$ *are linearly independent vectors in* V*. Suppose further that* $(v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_j)$ *span* V*. Then there is some sublist* $(u_1, u_2, \ldots, u_\ell)$ *of* $(w_1, w_2, \ldots, w_j)$*, possibly empty, so that* $(v_1, v_2, \ldots, v_k, u_1, \ldots, u_\ell)$ *form a basis of* V*.*

Exercise 39 (**)**.** *If* $\mathcal{B}_1 = (v_1, v_2, \ldots, v_n)$ *and* $\mathcal{B}_2 = (w_1, w_2, \ldots, w_m)$ *are bases of* V*, then for each* v_i *there exists some* w_j *such that the list* $(v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n, w_j)$ *obtained from* $\mathcal{B}_1$ *by removing* v_i *and adding* w_j *is a basis for* V*.*

Exercise 40 (*)**.** *If* $(v_1, v_2, \ldots, v_n)$ *and* $(w_1, w_2, \ldots, w_m)$ *are bases of* V*, then* $n = m$*.*

Definition 13. *A vector space is called finite dimensional if it is spanned by a finite number of vectors. A vector space is called infinite dimensional if it is not finite dimensional.*

Exercise 41. *Every finite dimensional vector space has a basis.*

Definition 14. *If* V *has a basis* $(v_1, v_2, \ldots, v_n)$*, then* n *is called the dimension of* V*, denoted as dim*V*.*

Additional Exercise 1. *Suppose* W *is a subspace of the finite dimensional vector space* V*. Then there exists a basis* $(w_1, w_2, \ldots, w_i, v_{i+1}, v_{i+2}, \ldots, v_n)$ *of* V *such that* $(w_1, \ldots, w_i)$ *is a basis for* W*.*

Proof: Suppose $(w_1, w_2, \ldots, w_j)$ is any set of linearly independent vectors in W. If we adjoin to these vectors a basis for V, we will obtain a set of vectors that span V. Hence, by the Basis Extension Theorem, there is a basis for V that contains $w_1, w_2, \ldots, w_j$ as a subset. (One way to say this is that we can "extend" the list $(w_1, w_2, \ldots, w_j)$ to a basis of V. In particular, notice that $j \le \dim V$.

Now let $(w_1, w_2, \ldots, w_i)$ be a set of linearly independent vectors in W where i is as large as possible. We can get such a set as follows. Because W is a subspace, it must contain 0. If this is all it contains, then $i = 0$, and the largest set of linearly independent vectors in W is empty. Otherwise W contains a nonzero vector w_1. If w_1 does not span W, then W contains a vector w_2 not in the span of w_1. Thus (w_1, w_2) are linearly independent. If they do not span W, then there is a vector w_3 in W that is not in the span of w_1 and w_2. Hence (w_1, w_2, w_3) are linearly independent. Continuing in this way, if $(w_1, w_2, \ldots, w_j)$ are a set of linearly independent vectors in W that do not span W, then there must be a vector w_{j+1} in W that is not in the span of $w_1, w_2, \ldots, w_j$, in which case we add it to the list of w_i's. And, $(w_1, w_2, \ldots, w_{j+1})$ must be linearly independent. This process cannot go on indefinitely, because, as we just proved, the number of linearly independent vectors in W cannot exceed the dimension of V. Hence we must eventually reach a linearly independent set of vectors, $(w_1, w_2, \ldots, w_i)$, that span W. This set is therefore a basis for W.

Now, again using the Basis Extension Theorem, we can extend this basis to a basis of V. $\square$

Exercise 42. *If W is a subspace of the finite dimensional vector space V, then W is finite dimensional and $dim W \leq dim V$.*

Exercise 43.

1. *Show that $\mathbb{F}^n$ is finite dimensional and $dim \mathbb{F}^n = n$.*
2. *Let m be any nonnegative integer. Show that the set of all polynomials with degree less than or equal to m and coefficients in $\mathbb{F}$ is a finite dimensional subspace of $\mathcal{P}(\mathbb{F})$.*
3. *Show that $\mathcal{P}(\mathbb{F})$ is infinite dimensional.*

Exercise 44. *Suppose $v_1, v_2, \ldots, v_k$ are vectors in the finite dimensional vector space V. If $k > dim V$, then $(v_1, v_2, \ldots, v_k)$ is linearly dependent.*

Exercise 45 (**)**.** *Suppose W_1 and W_2 are finite dimensional subspaces of V. Then*

$$dim(W_1 \cap W_2) + dim(W_1 + W_2) = dim(W_1) + dim(W_2).$$

Chapter 3

Linear Transformations

3.1 Linear Functions and Matrices

Definition 15. *A linear transformation, or linear map, is a function $T : V \to W$ from one vector space to another that satisfies both the following properties:*

1. $T(v + u) = T(v) + T(u)$, *for all vectors* $v, u \in V$.
2. $T(av) = aT(v)$, *for all* $a \in F$ *and for all* $v \in V$

Exercise 46. *Show that each of the following is a linear map.*

1. *The Zero Map,* $0 : V \to W$ *defined by* $0v = 0$ *for all* $v \in V$. *The image of every vector in* V *is the zero vector in* W. *We will denote this function by* 0. *(This is now three different uses of the symbol* 0. *What are they?)*
2. *The Identity Map,* $I : V \to V$, *defined by* $Iv = v$ *for all* $v \in V$.
3. *Differentiation of polynomials,* $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$.
4. *Integration of polynomials:* $Int : \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ *given by*

$$Int(p) = \int_0^1 p(x)\, dx$$

5. *Multiplying a polynomial by* x: $T : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ *defined by* $T(p) = xp(x)$.

Exercise 47. *Show that the map $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $T(x, y, z) = (ax + by + cz, dx + ey + fz)$ where $a, b, c, d, e,$ and f are any real numbers, is a linear map.*

Exercise 48. *Show that any linear map from one vector space to another must take the zero vector to the zero vector, that is, $T(0) = 0$.*

Definition 16. *Let $\mathcal{L}(V, W)$ denote the set of all linear maps from V to W. If $T, S \in \mathcal{L}(V, W)$ and $a \in \mathbb{F}$, define the function $T + S$ as $(T + S)(v) = T(v) + S(v)$ and the function aT as $(aT)(v) = aT(v)$.*

Exercise 49. *Show that if $S, T \in \mathcal{L}(V, W)$ and $a \in \mathbb{F}$, then $S+T \in \mathcal{L}(V, W)$ and $aT \in \mathcal{L}(V, W)$.*

Exercise 50. *Show that if V and W are vector spaces over $\mathbb{F}$, then $\mathcal{L}(V, W)$ is a vector space over $\mathbb{F}$.*

Definition 17. *If $T : V \to W$ is a linear map, the null space of T, denoted null T is defined to be*

$$\textit{null}\ T = \{v \in V \mid Tv = 0\}.$$

Exercise 51. *Determine the null space of each of the linear maps in Exercise 46.*

Exercise 52. *If $T : V \to W$ is a linear map, show that null T is a subspace of V.*

Definition 18. *The function $f : X \to Y$ is called one-to-one if $f(a) = f(b)$ implies that $a = b$. Equivalently, f is one-to-one if $a \neq b$ implies $f(a) \neq f(b)$. The function f is called onto if given any $b \in Y$, there is some $a \in X$ such that $f(a) = b$.*

Exercise 53. *Show that the linear map $T : V \to W$ is one-to-one if and only if null $T = \{0\}$.*

Definition 19. *If $T : V \to W$ is a linear map, the image of T, denoted image T is defined to be*

$$\textit{image}\ T = \{Tv \mid v \in V\}.$$

Exercise 54. *If $T : V \to W$ is a linear map, show that image T is a subspace of W.*

Exercise 55. *If $T : V \to W$ is a linear map and V is finite dimensional, show that*

$$\dim V = \dim \textit{null}\, T + \dim \textit{image}\, T.$$

Exercise 56. *Suppose that $T : V \to W$ is a linear map and both V and W are finite dimensional.*

1. *Show that if dim $V >$ dim W, then T cannot be one-to-one.*

2. *Show that if dim $V <$ dim W, then T cannot be onto.*

Additional Exercise 2. *Suppose that both V and W are finite dimensional and that $T : V \to W$ is a linear map that is both one-to-one and onto. Then* $\dim V = \dim W$.

Suppose $T : V \to W$ is a linear map and that $(v_1, v_2, \ldots, v_n)$ is a basis for V and $(w_1, w_2, \ldots, w_m)$ is a basis for W. If v is any vector in V, then v is some linear combination of the basis element, that is, there exist $a_1, a_2, \ldots, a_n \in \mathbb{F}$ such that $v = a_1v_1 + a_2v_2 + \ldots a_nv_n$. Now, because T is linear, we have

$$Tv = a_1Tv_1 + a_2Tv_2 + \ldots a_nTv_n.$$

What this says is that the values of T on the bases vectors *determines* the value of T on *all* vectors. Each vector Tv_i is in W and so is a linear combination of the basis vectors of W. That is, for each $1 \le i \le n$, there exist numbers $a_{i1}, a_{i2}, \ldots, a_{im}$ such that

$$Tv_i = a_{i1}w_1 + a_{i2}w_2 + \ldots a_{im}w_m.$$

Thus the set of numbers a_{ij} for $1 \le i \le n$ and $1 \le j \le m$ completely determine the map T. We can arrange all of these numbers into the following rectangular array with m rows and n columns, which we call an *m by n matrix over* $\mathbb{F}$.

$$\begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \ldots & a_{nm} \end{pmatrix}$$

This matrix completely encodes the map T, knowing the bases $(v_1, v_2, \ldots, v_n)$ and $(w_1, w_2, \ldots, w_m)$.

Definition 20. *An m by n matrix over $\mathbb{F}$ is a rectangular array of numbers from $\mathbb{F}$ having m rows and n columns. The matrix is called square if $n = m$.*

NOTE: If $T : \mathbb{F}^n \to \mathbb{F}^m$, then we should assume that the bases in question are the *standard* bases, $(e_1, e_2, \ldots, e_n)$ for $\mathbb{F}^n$ and $(e_1, e_2, \ldots, e_m)$ for $\mathbb{F}^m$.

Exercise 57. *Write down the matrix that encodes each of the following linear maps with respect to the* **standard** *bases.*

1. $T : \mathbb{R}^2 \to \mathbb{R}^3$ *given by* $T(x, y) = (2x - y, 3y, 4x + 2y)$.
2. $T : \mathbb{C}^3 \to \mathbb{C}^2$ *given by* $T(x, y, z) = (2x - iz, 0)$.

IMPORTANT: A linear map $T : V \to W$ is encoded by a matrix! (Together with a choice of bases for V and W.)

3.2 Systems of Linear Equations

Consider a system of n linear equations in m unknowns (that is, m variables)

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 \cdots + a_{nm}x_m &= b_n \end{aligned}$$

where all the *coefficients* $a_{ij} \in \mathbb{F}$ and all the constants $b_i \in \mathbb{F}$.

Solving the system of equations means finding values (in $\mathbb{F}$) for each of the variables x_i that make all the equations *simultaneously* true. This set of values for the variables is called a *solution*.

Exercise 58. *Explain why a system of two linear equations in two unknowns, in the case where $\mathbb{F} = \mathbb{R}$, can be thought of as corresponding to two lines in $\mathbb{R}^2$. Use this to explain why there will either be no solutions, exactly one solution, or infinitely many different solutions. Give examples for each case.*

We can use the language of vectors to rewrite the system of equations as a single equation between vector quantities. First, rewrite the system of n equations as the following single vector equation.

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 \cdots + a_{nm}x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Here we are writing the vectors as columns. Next, let's rewrite the left hand side of this equation as the sum of m vectors

$$\begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{n1}x_1 \end{pmatrix} + \begin{pmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{n2}x_2 \end{pmatrix} \cdots + \begin{pmatrix} a_{1m}x_m \\ a_{2m}x_m \\ \vdots \\ a_{nm}x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

and then rewrite each of the m vectors on the left hand side of the equation as a scalar multiple of a vector

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} \cdots + x_m \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

We now have a single equation with m unknowns. To solve this single equation is to find values for all the x_i's so that the m vectors on the left add up to the single vector on the right. Notice that what we are trying to do is find some linear combination of the vectors on the left that is equal to the vector on the right! By using matrices, we can rewrite this in a nice way so as to finally think of having just *one equation in one unknown*!

We can now write our system of equations as

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{3.1}$$

where, in order to *make sense* of the left hand side of the equation, we *define* the product of a matrix and a vector as follows.

Definition 21. *Suppose A is the n by m matrix over $\mathbb{F}$*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

and v is the vector $v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$ with entries in $\mathbb{F}$. Then the product Av is defined as

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} \cdots + x_m \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix} \tag{3.2}$$

Exercise 59. *Use Definition 21 to perform the following multiplications. Write each answer as a single column vector.*

1. $\begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$

3. $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}$

4. $\begin{pmatrix} 1 & -12 & 3 \\ 4 & -1/2 & 0 \\ 2 & -2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ -5 \\ 2 \end{pmatrix}$

IMPORTANT: Definition 21 says that an n by m matrix A over $\mathbb{F}$ defines a function $f_A : \mathbb{F}^m \to \mathbb{F}^n$ by $f_A(v) = Av$.

Exercise 60. *Rewrite each of the following systems of equations in matrix form as in Equation 3.1.*

1. $\begin{array}{rcrcr} 2x & + & 3y & = & 4 \\ -15x & + & y & = & -2 \end{array}$

2. $\begin{array}{rcrcrcr} 2x & + & 3y & - & iz & = & 4+i \\ -15x & & & + & z & = & -2i \end{array}$

3. $\begin{array}{rcrcrcrcr} 2x & & & - & iz & + & w & = & 4+i \\ -15x & + & y & & & & & = & -2i \\ & & & & z & - & w & = & 0 \end{array}$

Exercise 61. *Explain why Equation 3.1 can now be thought of as a single equation with a single unknown. What is the single unknown? What kind of object is it? (A number? A vector?)*

Definition 22. *Suppose that A is an n by m matrix over $\mathbb{F}$ and that B is an m by ℓ matrix over $\mathbb{F}$. Then we can multiply A times B to obtain an n by ℓ matrix $C = AB$ over $\mathbb{F}$ by defining the k-th column of C to be the product of A times the k-th column of B.*

Exercise 62. *Perform the following multiplications.*

1. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 0 & -3 \\ 3 & -1 & -1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 4 \\ 6 & 7 \end{pmatrix}$

3. $\begin{pmatrix} 1 \\ 3 \\ 0 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \end{pmatrix}$

4. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

Definition 23. *If A is n by m matrix over $\mathbb{F}$, let $A[i, j]$ denote the entry of A that lies in the i-th row and j-th column.*

Exercise 63. *Suppose A is an n by m matrix and B is an m by ℓ matrix. Let $C = AB$. Find a formula for $C[i, j]$ in terms of the entries of A and B.*

Definition 24. *A matrix A is called diagonal if $A[i, j] = 0$ whenever $i \neq j$. The entries $A[i, i]$ are called diagonal entries and the entries $A[i, j]$ where $i \neq j$ are called off-diagonal. So, diagonal matrices are ones where all the off-diagonal entries are zero.*

Exercise 64. *Show that the product of two diagonal matrices is diagonal. In particular, show that if A and B are diagonal, then $AB[i, i] = A[i, i]B[i, i]$.*

Definition 25. *Let $M(n, m)(\mathbb{F})$ be the set of all n by m matrices over $\mathbb{F}$. If A and B are in $M(n, m)(\mathbb{F})$, define their sum $A + B$ to be the n by m matrix over $\mathbb{F}$ where $(A+B)[i, j] = A[i, j]+B[i, j]$. If $x \in \mathbb{F}$, define the scalar multiple of A, xA, to be the n by m matrix over $\mathbb{F}$ where $(xA)[i, j] = xA[i, j]$.*

Exercise 65. *Perform the following additions and scalar multiplications.*

1. $\begin{pmatrix} 2 & -3 & \sqrt{2} \\ 0 & i & 3 \end{pmatrix} + \begin{pmatrix} -2 & 0 & 3 \\ 5 & 6 & -1 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 4 & 3 & 2 \end{pmatrix}$

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

Exercise 66. *Show that $M(n, m)(\mathbb{F})$ with addition and scalar multiplication defined as in Defintion 25 forms a vector space over $\mathbb{F}$.*

3.3 Solving Systems of Linear Equations with Gaussian Elimination

Suppose we want to solve the system of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 \cdots + a_{nm}x_m &= b_n \end{aligned}$$

There is a general method called *Gaussian elimination* that can always be used to find all possible solutions (or to determine if none exist). The process will repeatedly use the following steps to alter the equations, but not their solution, until it is obvious what the solutions are.

Steps Used in Gaussian Elimination

1. Trade any two equations.
2. Multiply both sides of one of the equations by any non-zero element of $\mathbb{F}$.
3. Add a multiple of one equation to another equation.

Example 1. *Solve the system*

$$\begin{aligned} 3x - 4y + z &= 0 \\ x - 2y &= 6 \\ 2y - 3z &= 12 \end{aligned}$$

Step 1: Multiply the first equation by $1/3$*, so that the coefficient of* x *in the first equation is 1.*

$$\begin{aligned} x - \frac{4}{3}y + \frac{1}{3}z &= 0 \\ x - 2y &= 6 \\ 2y - 3z &= 12 \end{aligned}$$

Step 2. Subtract the first equation from the second, so as to eliminate x from the second equation. We obtain,

$$\begin{aligned} x - \frac{4}{3}y + \frac{1}{3}z &= 0 \\ -\frac{2}{3}y - \frac{1}{3}z &= 6 \\ 2y - 3z &= 12 \end{aligned}$$

Step 3. Multiply the second equation by $-\frac{3}{2}$, so that the coefficient of k in the second equation is 1.

$$\begin{aligned} x - \frac{4}{3}y + \frac{1}{3}z &= 0 \\ y + \frac{1}{2}z &= -9 \\ 2y - 3z &= 12 \end{aligned}$$

Step 4. Add -2 times the second equation to the third equation, so as to eliminate y from the last equation. This gives,

$$\begin{aligned} x - \frac{4}{3}y + \frac{1}{3}z &= 0 \\ y + \frac{1}{2}z &= -9 \\ -4z &= 30 \end{aligned}$$

Step 5. Multiply the last equation by $-1/4$, giving

$$\begin{aligned} x - \frac{4}{3}y + \frac{1}{3}z &= 0 \\ y + \frac{1}{2}z &= -9 \\ z &= -\frac{15}{2} \end{aligned}$$

We now know what z is.

Step 6. Add $-\frac{1}{2}$ times the last equation to the second equation, so as to eliminate z from the second equation.

$$\begin{aligned} x - \frac{4}{3}y + \frac{1}{3}z &= 0 \\ y &= -\frac{21}{4} \\ z &= -\frac{15}{2} \end{aligned}$$

We now know what y is.

Step 7. Add $-\frac{1}{3}$ times the last equation to the first equation, in order to eliminate z from the first equation.

$$\begin{aligned} x - \frac{4}{3}y &= \frac{5}{2} \\ y &= -\frac{21}{4} \\ z &= -\frac{15}{2} \end{aligned}$$

Step 8. Add $\frac{4}{3}$ times the second equation to the first equation, in order to eliminate y from the first equation.

$$\begin{aligned} x &= -\frac{9}{2} \\ y &= -\frac{21}{4} \\ z &= -\frac{15}{2} \end{aligned}$$

We now know what all of x, y, and z are.

Exercise 67. *Prove that each of the three operations used in Gaussian elimination change the set of equations, but do NOT change the set of solutions to the equations.*

In practice, it is easier to carry out all the above steps by using matrices and vectors. We can rewrite the initial set of equations as

$$\begin{pmatrix} 3 & -4 & 1 \\ 1 & -2 & 0 \\ 0 & 2 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ 12 \end{pmatrix}$$

We need only keep track of the matrix of coefficients and the constant vector which we combine into one matrix which we call the *augmented* matrix. In the augmented matrix, we use a vertical line to separate the coefficient matrix from the constant vector. The eight steps in Example 1 now become

$$
\left(\begin{array}{ccc|c} 3 & -4 & 1 & 0 \\ 1 & -2 & 0 & 6 \\ 0 & 2 & -3 & 12 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -\frac{4}{3} & \frac{1}{3} & 0 \\ 1 & -2 & 0 & 6 \\ 0 & 2 & -3 & 12 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -\frac{4}{3} & \frac{1}{3} & 0 \\ 0 & -\frac{2}{3} & -\frac{1}{3} & 6 \\ 0 & 2 & -3 & 12 \end{array}\right)
$$

$$
\to \left(\begin{array}{ccc|c} 1 & -\frac{4}{3} & \frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} & -9 \\ 0 & 2 & -3 & 12 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -\frac{4}{3} & \frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} & -9 \\ 0 & 0 & -4 & 30 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -\frac{4}{3} & \frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{2} & -9 \\ 0 & 0 & 1 & -\frac{15}{2} \end{array}\right)
$$

$$
\to \left(\begin{array}{ccc|c} 1 & -\frac{4}{3} & \frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{21}{4} \\ 0 & 0 & 1 & -\frac{15}{2} \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -\frac{4}{3} & 0 & \frac{5}{2} \\ 0 & 1 & 0 & -\frac{21}{4} \\ 0 & 0 & 1 & -\frac{15}{2} \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{9}{2} \\ 0 & 1 & 0 & -\frac{21}{4} \\ 0 & 0 & 1 & -\frac{15}{2} \end{array}\right)
$$

In the last augmented matrix, the solution appears as the last column.

The three allowable steps, when applied to the augmented matrix, are now called *elementary row operations*.

Definition 26. *The elementary row operations that may be applied to any matrix with entries in $\mathbb{F}$ are*

1. *Trade any two rows.*
2. *Multiply any row by a non-zero element of $\mathbb{F}$.*
3. *Add a multiple of one row to another row*

Definition 27. *A matrix is in reduced echelon form if all of the following are true*

1. *In every row, the left-most non-zero entry is 1.*
2. *The left-most non-zero entry of every row is always to the right of the left-most non-zero entry of the row above.*
3. *The left-most non-zero entry of each row is the only non-zero entry in its column.*

Exercise 68. *Use elementary row operations to transform each of the following matrices to reduced echelon form.*

1. $\begin{pmatrix} 2 & 3 \\ 1 & -4 \end{pmatrix}$
2. $\begin{pmatrix} -1 & 0 & 3 & -1 \\ 0 & 1 & 5 & 3 \\ -2 & 3 & 6 & 2 \end{pmatrix}$

3. $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & -4 \\ -1 & 1 & -3 & 0 \end{pmatrix}$

4. $\begin{pmatrix} 2 & 3 \\ 1 & -4 \\ 3 & 3 \\ 6 & -2 \end{pmatrix}$

Exercise 69. *What are all possible 2 by 2, 2 by 3, and 2 by 4 matrices in reduced echelon form?*

Exercise 70. *What are all possible 3 by 2 and 4 by 2 matrices in reduced echelon form?*

Exercise 71. *What are all possible 3 by 3 matrices in reduced echelon from?*

Exercise 72. *Describe how to use elementary row operations to transform any matrix to one in reduced echelon form.*

Exercise 73. *Solve the following systems of linear equations by forming the augmented matrix and then using elementary row operations to tranform the augmented matrix into a matrix where all the but the last column is in reduced echelon form.*

1.

$$\begin{aligned} 2x + 2y + 10z &= -1 \\ 2x + 5y + 16z &= -4 \\ -4x - y - 14z &= 2 \end{aligned}$$

2.

$$\begin{aligned} 15x - 4y &= 53 \\ 5x - 2y &= 19 \end{aligned}$$

3.

$$\begin{aligned} 2x + 6y &= 8 \\ x + 3y + z &= 6 \\ 3x + 9y + z &= 14 \end{aligned}$$

Exercise 74. *Describe how any system of linear equations can be solved by using the elementary row operations to transform the augmented matrix into a matrix where all but the last column is in reduced echelon form. Describe how the augmented matrix, once it is in this form, will look if*

1. *The system has no solutions.*
2. *The system has a unique solution.*
3. *The system has infinitely many different solutions.*

3.4 More on Multiplication of Matrices

Let's explore to what extant multiplication of matrices behaves like multiplication of real or complex numbers. Recall that we can only multiply the matrix A times the matrix B if the number of columns of A equals the number of rows of B.

Exercise 75. *How many rows and columns must two matrices A and B each have in order for both AB and BA to be defined?*

Exercise 76. *How many rows and columns must a matrix A have in order for $A^2 = AA$ to be defined?*

Because of the above, it is often nice to focus only on square matrices of a given size. Then we can multiply them in any order we wish. But don't forget, non-square matrices arise naturally in systems of linear equations.

Exercise 77. *Let 0 be the matrix consisting of all zeroes. (Note that we have one zero matrix for every size matrix.) Show that $0A = A0 = 0$ for all n by m matrices A, where the zero matrices in these equations have to have the correct sizes. (What sizes are those?)*

Exercise 78. *With real or complex numbers, if the product of two numbers is zero, then at least one of the numbers must be zero. Show this is false for matrices.*

Exercise 79. *Unlike multiplication of real or complex numbers, multiplication of matrices is NOT commutative. Find an example of two 2 by 2 matrices A and B so that $AB \neq BA$.*

Exercise 80. *Show that matrix multiplication is associative. Don't forget that in order for $A(BC)$ and $(AB)C$ to be defined, the three matrices will have to be the correct sizes! Explain what sizes the matrices must be for this to make sense.*

Exercise 81. *Show that if A, B and C are matrices of the correct sizes, then the following distributive laws hold.*

1. $A(B + C) = AB + AC$
2. $(A + B)C = AC + BC$

Explain what sizes the matrices must be for this to make sense.

Definition 28. *The n by n identity matrix is the n by n matrix I_n defined by*

$$I[i,j] = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

There is one identity matrix for each natural number n.

Exercise 82. *Let A be an n by m matrix over $\mathbb{F}$. Show that*

1. *$I_n A = A$, and*
2. *$AI_m = A$*

Thus multiplicative identities exist for multiplication on the right and for multiplication on the left.

Non-zero real or complex numbers have multiplicative inverses. The next exercise explores the analogy for matrices.

Definition 29. *Suppose A is and n by m matrix. The m by n matrix B is called a left (multiplicative) inverse of A if $BA = I_m$. The m by n matrix C is called a right (multiplicative) inverse of A if $AC = I_n$. Usually when it is obvious we are talking about multiplication, we will just talk about the right or left inverse, not the right or left multiplicative inverse.*

Exercise 83. *Given any n and m, show there is an n by m matrix that does not have either a right or left inverse. (Don't forget that not all real or complex numbers have multiplicative inverses!)*

Exercise 84. *Suppose that $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and that $ad - bc \neq 0$. Let*

$$B = \begin{pmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Show that $AB = BA = I_2$. Thus B is both the left and right inverse of A.

Exercise 85. *Suppose that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} j & k \\ m & n \end{pmatrix}$. Show that*

$$(ad - bc)(eh - fg) = jn - km.$$

Use this to show that a 2 by 2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has a left and right inverse if and only if $ad - bc \neq 0$.

Exercise 86. *Suppose that A is an n by m matrix that has both left and right inverses. Show that the two inverses are equal. That is, if $AB = I_n$ and $B'A = I_m$ then $B = B'$.*

Exercise 87. *Suppose that A is an n by m matrix that has both a left and right inverse (which must be the same by Exercise 86). Show that the inverse is unique. That is, if $BA = I_n$ and $AB = I_m$ and also $B'A = I_n$ and $AB' = I_m$, then $B = B'$.*

Definition 30. *A matrix is called invertible if it has both a left and right inverse. In this case, we denote the inverse of A by A^{-1}.*

Exercise 88. *If A and B are invertible matrices, then show that AB is also invertible and that $(AB)^{-1} = B^{-1}A^{-1}$. (Here A and B must have the correct sizes for this to make sense. What sizes?)*

How to find the inverse of a matrix, if it has one.

Suppose that A is an n by m matrix and we want to see if there is an m by n matrix B so that $AB = I_n$. Think of B as made up of its columns $b_1, b_2, \ldots, b_n$. That is, each b_i is a vector which is a column of B. Now Definition 22 says that the columns of AB are A times the columns of B. So to find the columns of B, we need to solve all of the following equations.

$$Ab_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, Ab_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, Ab_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

This means we want to use elementary row operations to transform each of the following augmented matrices

$$\left(A| \begin{matrix} 1 \\ 0 \\ \vdots \\ 0 \end{matrix} \right), \left(A| \begin{matrix} 0 \\ 1 \\ \vdots \\ 0 \end{matrix} \right), \ldots, \left(A| \begin{matrix} 0 \\ 0 \\ \vdots \\ 1 \end{matrix} \right)$$

into matrices where all but the last column is in reduced echelon form. In each case we will do the same elementary row operations! Hence we can do these all at once by starting with the augmented matrix

$$(A|I_n) = \left(A| \begin{matrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \vdots & 1 \end{matrix} \right)$$

Example 2. *Determine if* $A = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & -1 \end{pmatrix}$ *has a right inverse, and if so find one. We start with the augmented matrix*

$$\left(\begin{array}{ccc|cc} 1 & 2 & 5 & 1 & 0 \\ 3 & 4 & -1 & 0 & 1 \end{array}\right)$$

Using elementary row operations (check this!) we can transform this to

$$\left(\begin{array}{ccc|cc} 1 & 0 & -11 & -2 & 1 \\ 0 & 1 & 8 & 3/2 & -1/2 \end{array}\right)$$

This means that $A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ *has infinitely many solutions, namely* $\begin{pmatrix} -2+11z \\ 3/2-8z \\ z \end{pmatrix}$ *for any number* z*. Similarly, the solutions to the equation* $A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ *are* $\begin{pmatrix} 1+11w \\ -1/2-8w \\ w \end{pmatrix}$ *where* w *is any number. Thus there are infinitely many right inverses to* A*, namely all matrices of the form*

$$B = \begin{pmatrix} -2+11z & 1+11w \\ 3/2-8z & -1/2-8w \\ z & w \end{pmatrix}.$$

The reader should check that no matter what z *and* w *are,* $AB = I_2$*.*

Exercise 89. *In Example 2, the matrix* A *has infinitely many right inverses, one for each value of* z *and* w*. Are any of the right inverses of* A *a left inverse of* A*? Does* A *have a left inverse?*

Exercise 90. *Use the method of Example 2 to find the right inverses of the following matrices, if they exist.*

1. $\begin{pmatrix} 3 & 6 \\ -15 & -30 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$

3. $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

Exercise 91. *Use the methodology of Example 2 to give another proof that* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *has a right inverse if an only if* $ad - bc \neq 0$*. Furthermore, when it does have a right inverse, it is uniquely given by the formula in Exercise 84.*

3.5 Elementary Row Operations

Our goal in this section is to see that each elementary row operation correspond to multiplying on the left by a certain kind of square matrix.

Definition 31. *Let δ_{ij}^n be the n by n square matrix all of whose entries are zero except for the entry in row i and column j which is equal to 1.*

Exercise 92. *Let A be an n by m matrix. Show that $(I_n + a\delta_{ij}^n)A$ is the matrix obtained from A by adding a times the j-th row of A to the i-th row of A.*

Exercise 93.

1. *Show that $I_n + (a-1)\delta_{ii}^n$ is the matrix obtained from the identity matrix by replacing the 1 in the i-th row and i-th column by a.*

2. *Let A be an n by m matrix. Show that $(I_n + (a-1)\delta_{ii}^n)A$ is the matrix obtained from A by multiplying the i-th row of A by a.*

Definition 32. *Let P_{ij}^n be the n by n square matrix obtained from I_n by trading the i-th and j-th rows.*

Exercise 94. *Show that $P_{ij}^n = I_n - \delta_{ii}^n - \delta_{jj}^n + \delta_{ij}^n + \delta_{ji}^n$.*

Exercise 95. *Let A be an n by m matrix. Show that $P_{ij}^n A$ is the matrix obtained from A by trading the i-th and j-th rows of A.*

Definition 33. *An n by n elementary matrix E is an n by n matrix such that given any n by m matrix A, the matrix EA is obtained from A by an elementary row operation. From the above exercises we have that every elementary matrix is of the form $I_n + a\delta_{ij}$ or P_{ij}^n.*

Exercise 96.

1. *Show that* $\delta_{ij}^n \delta_{rs}^n = \begin{cases} 0 & \text{if } j \neq r \\ \delta_{is} & \text{if } j = r \end{cases}$

2. *Show that $(P_{ij}^n)^2 = I_n$.*

3. *Show that $(I_n + a\delta_{ij})(I_n - a\delta_{ij}) = I_n$*

Exercise 97. *Show that every elementary matrix is invertible and that the inverse is also an elementary matrix. For each of the three types of elementary matrices, describe its inverse.*

Let's take a look at Gaussian elimination again, this time from the point of view of multiplying on the left by elementary matrices. When we want to solve the single equation in one unknown $5x = 6$ we can multiply both sides of the equation on the left by the multiplicative inverse of 5, namely $1/5$. This gives the solution $x = 6/5$. Suppose now we want to solve the vector equation $Ax = b$. If the matrix A has a left inverse B so that $BA = I$, then if we multiply on the left by B we obtain $BAx = Bb$ which is $Ix = Bb$ or finally, $x = Bb$. Thus, multiplying on the left by the left inverse of A solves the equation for x. But what if A does not not have a left inverse? There may still be a solution, or infinitely many solutions! We can solve the problem by using Gaussian elimination.

Let's repeat Example 1. But now, let's think of each elementary row operation as multiplying on the left by an elementary matrix.

Example 3. *Solve the equation*

$$\begin{pmatrix} 3 & -4 & 1 \\ 1 & -2 & 0 \\ 0 & 2 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ 12 \end{pmatrix}$$

Step 1: Multiply the first equation by $1/3$*, so that the left-most entry in the first row is 1. To do this we multiply on the left by the elementary matrix* $E_1 = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. *This gives*

$$\begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -4 & 1 \\ 1 & -2 & 0 \\ 0 & 2 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 6 \\ 12 \end{pmatrix}$$

which simplifies to

$$\begin{pmatrix} 1 & -4/3 & 1/3 \\ 1 & -2 & 0 \\ 0 & 2 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ 12 \end{pmatrix}$$

Step 2. Subtract the first row from the second row, so as to eliminate the left-most entry from the second row. To do this we multiply on the left by the elementary matrix $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. *This gives*

$$\begin{pmatrix}1&0&0\\-1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&-4/3&1/3\\1&-2&0\\0&2&-3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1&0&0\\-1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}0\\6\\12\end{pmatrix}$$

which simplifies to

$$\begin{pmatrix}1&-4/3&1/3\\0&-2/3&-1/3\\0&2&-3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\6\\12\end{pmatrix}$$

Step 3. Multiply the second row by $-\frac{3}{2}$*, so that the left-most entry in the second row is 1. To do this we multiply on the left by the elementary matrix* $E_3=\begin{pmatrix}1&0&0\\0&-3/2&0\\0&0&1\end{pmatrix}$. *This gives*

$$\begin{pmatrix}1&0&0\\0&-3/2&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&-4/3&1/3\\0&-2/3&-1/3\\0&2&-3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1&0&0\\0&-3/2&0\\0&0&1\end{pmatrix}\begin{pmatrix}0\\6\\12\end{pmatrix}$$

which simplifies to

$$\begin{pmatrix}1&-4/3&1/3\\0&1&1/2\\0&2&-3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\-9\\12\end{pmatrix}$$

Step 4. Add -2 *times the second row to the third row, so as to eliminate left-most entry from the third row. To do this we multiply on the left by the elementary matrix* $E_4=\begin{pmatrix}1&0&0\\0&1&0\\0&-2&1\end{pmatrix}$. *This gives*

$$\begin{pmatrix}1&0&0\\0&1&0\\0&-2&1\end{pmatrix}\begin{pmatrix}1&-4/3&1/3\\0&1&1/2\\0&2&-3\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&-2&1\end{pmatrix}\begin{pmatrix}0\\-9\\12\end{pmatrix}$$

which simplifies to

$$\begin{pmatrix}1&-4/3&1/3\\0&1&1/2\\0&0&-4\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\-9\\30\end{pmatrix}$$

Step 5. Multiply the last row by $-1/4$. To do this we multiply on the left by the elementary matrix $E_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/4 \end{pmatrix}$. *This gives*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/4 \end{pmatrix} \begin{pmatrix} 1 & -4/3 & 1/3 \\ 0 & 1 & 1/2 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/4 \end{pmatrix} \begin{pmatrix} 0 \\ -9 \\ 30 \end{pmatrix}$$

which simplifies to

$$\begin{pmatrix} 1 & -4/3 & 1/3 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -9 \\ -15/2 \end{pmatrix}$$

Step 6. Add $-\frac{1}{2}$ times the last row to the second row. To do this we multiply on the left by the elementary matrix $E_6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{pmatrix}$. *This gives*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -4/3 & 1/3 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -9 \\ -15/2 \end{pmatrix}$$

which simplifies to

$$\begin{pmatrix} 1 & -4/3 & 1/3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -21/4 \\ -15/2 \end{pmatrix}$$

Step 7. Add $-\frac{1}{3}$ times the last row to the first row.To do this we multiply on the left by the elementary matrix $E_7 = \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. *This gives*

$$\begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -4/3 & 1/3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -21/4 \\ -15/2 \end{pmatrix}$$

which simplifies to

$$\begin{pmatrix} 1 & -4/3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5/2 \\ -21/4 \\ -15/2 \end{pmatrix}$$

Step 8. Add $\frac{4}{3}$ times the second row to the first row. To do this we multiply on the left by the elementary matrix $E_8 = \begin{pmatrix} 1 & 4/3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. *This gives*

$$\begin{pmatrix} 1 & 4/3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -4/3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 4/3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5/2 \\ -21/4 \\ -15/2 \end{pmatrix}$$

which simplifies to

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -9/2 \\ -21/4 \\ -15/2 \end{pmatrix}$$

or

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -9/2 \\ -21/4 \\ -15/2 \end{pmatrix}$$

Starting with $Ax = b$ we have repeatedly multiplied on the left by elementary matrices obtaining the final equation of

$$E_8E_7E_6E_5E_4E_3E_2E_1Ax = E_8E_7E_6E_5E_4E_3E_2E_1b$$

which is the same as

$$Ix = E_8E_7E_6E_5E_4E_3E_2E_1b$$

because the elementary row operations take A to the identity matrix. Notice that the product of all the elementary matrices

$$B = E_8E_7E_6E_5E_4E_3E_2E_1$$

is a left inverse of A so that, in the end, we are multiplying both sides of the original equation on the left by B. Exercise 97 says that each of the elementary matrices is invertible. If we multiply both sides of

$$E_8E_7E_6E_5E_4E_3E_2E_1A = I$$

on the left by E_8^{-1} we obtain

$$E_7E_6E_5E_4E_3E_2E_1A = E_8^{-1}.$$

If now multiply both sides on the right by E_8^{-1} we obtain

$$E_7E_6E_5E_4E_3E_2E_1AE_8 = I.$$

Similarly, we can move E_7 from the beginning to the end of the left hand side of the equation to obtain

$$E_6E_5E_4E_3E_2E_1AE_8E_7 = I.$$

Continuing in this way, we obtain

$$AE_8E_7E_6E_5E_4E_3E_2E_1 = I.$$

Thus A has a right inverse as well as left inverse and hence is invertible.

Exercise 98. *Show that if the reduced echelon form of the square matrix A is I_n, then A is invertible and A^{-1} is the product of the elementary matrices used to take A to reduced echelon form.*

Definition 34. *Suppose $f : X \to Y$ and $g : Y \to Z$ are functions. Their composition is the function $g \circ f : X \to Z$ defined as $(g \circ f)(x) = g(f(x))$.*

Exercise 99. *Suppose $f : X \to Y$ and $g : Y \to Z$ are functions.*

1. *Show that if f and g are both one-to-one, then their composition is one-to-one.*
2. *Show that if f and g are both onto, then their composition is onto.*
3. *Show that if their composition is one-to-one, then f is one-to-one.*
4. *Show that if their composition is onto, then g is onto.*

Exercise 100. *Let A be an n by m matrix. Then A defines a function $T_A : \mathbb{F}^m \to \mathbb{F}^n$ by $T_A(v) = Av$.*

1. *Show that T_A is a linear map.*
2. *If B is an ℓ by n matrix, show that $T_B \circ T_A = T_{BA}$.*

Exercise 101. *Let A be an n by m matrix.*

1. *Show that if A has a left inverse, then T_A is one-to-one.*
2. *Show that if A has a right inverse, then T_A is onto.*

Additional Exercise 3. *Show that if the matrix A is invertible, then it is square.*

Proof: Suppose that A is an n by m matrix that is invertible. This means that it has both a left and right inverse. By Exercise 101, it follows that $T_A : \mathbb{F}^m \to \mathbb{F}^n$ is both one-to-one and onto. Now by Additional Exercise 2 we have that $n = m$. $\square$

Exercise 102. *Let A be an n by m matrix. Show that $T_A : \mathbb{F}^m \to \mathbb{F}^n$ is one-to-one if and only if $Av = 0$ implies $v = 0$.*

Exercise 103. *Show that if A is an invertible matrix, then the reduced echelon form of A is I.*

Combine previous exercises to sum up what we know in the following exercise.

Additional Exercise 4. *Suppose A is a matrix with entries in $\mathbb{F}$. The following are equivalent.*

1. *A is invertible*
2. *The reduced echelon form of A is the identity matrix I.*
3. *T_A is one-to-one and onto.*

Definition 35. *A square matrix is called singular if it is not invertible, and non-singular if it is invertible.*

3.6 The Rank of a Matrix

Definition 36. *Suppose A is an n by m matrix with entries in $\mathbb{F}$. The row space of A is defined to be the subspace of $\mathbb{F}^m$ spanned by the rows of A. The row rank of A is the dimension of the row space. The column space of A is defined to be the subspace of $\mathbb{F}^n$ spanned by the columns of A. The column rank of A is the dimension of the column space.*

Exercise 104. *Show that the row rank of a matrix A is the maximal number of linearly independent rows in A. Show that the column rank of A is the maximal number of linearly independent columns in A.*

Exercise 105. *Suppose that A is an n by m matrix and that E is an n by n elementary matrix. Show that A and EA have the same row space and hence the same row rank.*

Exercise 106. *Let A be a matrix and R its reduced echelon form. Show that A and R have the same row space and same row rank and that the row rank is the number of non-zero rows in R.*

Exercise 107. *Let A be a matrix and R its reduced echelon form. Show that the null space of T_A and the null space of T_R are the same.*

Exercise 108. *Show that the column space of a matrix A is the same as the image of T_A.*

Example 4. *If we change a matrix A by a row operation, the row space of A and null space of T_A do not change. BUT the column space of A and image of T_A CAN change. Here is an example: Let $A = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$. The column space is spanned by $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ -2 \end{pmatrix}$ which has $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ as a basis. This is a 1-dimensional subspace of $\mathbb{R}^2$ and is the same as the image of T_A. Now let's change A to row reduced echelon form, $R = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$. The column space is now spanned by $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. So, we see that the column space of A and the image of T_A have changed! However, notice that the dimension of the column space of A has not changed.*

Exercise 109. *Let A be an n by m matrix and B an invertible matrix.*

1. *Show that if any set of columns in A are linearly independent, then the same set of columns in BA are also linearly independent. (Hint: Don't forget that Av is as linear combination of the columns of A.)*

2. *Show that the column rank of A and BA are the same because they both have the same maximal number of linearly independent columns.*

3. *Show that if R is the reduced echelon form of A, then A and R have the same column rank.*

Exercise 110. *Show that for any matrix A, the row rank and the column rank are equal. (Hint: Consider taking A to reduced echelon form and use earlier exercises.)*

Definition 37. *The rank of a matrix is defined to be its row rank (or column rank—the two are equal).*

Definition 38. *Let A be an n by m matrix. We say that A has full rank if its rank is as large as possible, which is the minimum of n and m.*

To find the rank of a matrix, put it in reduced echelon form and count the number of non-zero rows. Thus a square matrix is nonsingular, or invertible, if and only if it is of full rank. Combing this with Additional Exercise 4 we now have

Theorem 1. *Suppose A is a matrix with entries in $\mathbb{F}$. The following are equivalent.*

1. *A is invertible.*

2. *The reduced echelon form of A is the identity matrix I.*

3. *A has full rank.*

4. *T_A is one-to-one and onto.*

3.7 Isomorphisms

Definition 39. *A linear map $T : V \to W$ between vector spaces is called invertible if there exists a linear map $S : W \to V$ such that $S \circ T : V \to V$ is the identity map on V and the map $T \circ S : W \to W$ is the identity map on W. The maps S and T are called inverses of each other.*

Exercise 111. *Show that if $T : V \to W$ is an invertible map, then its inverse is unique.*

Exercise 112. *Show that a linear map $T : V \to W$ is invertible if and only if it is one-to-one and onto.*

Definition 40. *A linear map $T : V \to W$ that is one-to-one and onto is called an isomorphism. Two vector spaces are called isomorphic if there is an isomorphism between them.*

Exercise 113. *Show that two finite dimensional vector spaces are isomorphic if and only if they have the same dimension.*

Exercise 114. *Suppose V is a finite dimensional vector space and $T : V \to V$ is a linear map. Show that the following three statements are equivalent.*

1. *T is invertible.*
2. *T is one-to-one.*
3. *T is onto.*

Remember that linear maps correspond to matrices! BUT, don't forget, *this correspondence depends on a choice of bases* for the vector spaces. Suppose that V is a finite dimensional vector space with basis $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$. The vector space is totally abstract—we may have no idea what the vectors are in V. But by using the basis, the vectors in V correspond to vectors in $\mathbb{F}^m$. Namely, if $v \in V$, then v is some linear combination of the basis vectors and we have

$$v = a_1v_1 + a_2v_2 + \cdots + a_mv_m.$$

Hence we can record the m-tuple of numbers $(a_1, a_2, \ldots, a_m)$ to represent v and this is a vector in $\mathbb{F}^m$. This defines a map

$$\Phi_{\mathcal{B}} : V \to \mathbb{F}^m$$

by $\Phi_{\mathcal{B}}(a_1v_1 + a_2v_2 + \cdots + a_mv_m) = (a_1, a_2, \ldots, a_m)$.

Exercise 115. *The map $\Phi_{\mathcal{B}} : V \to \mathbb{F}^m$ is an isomorhpism. (Don't forget to show that $\Phi_{\mathcal{B}}$ is a linear map before showing that it is one-to-one and onto.)*

We can now see how a linear map between vector spaces can be described by a matrix once bases have been chosen for each space. Suppose that $T : V \to W$ is a linear map, that $\mathcal{B}_1 = \{v_1, v_2, \ldots, v_m\}$ and $\mathcal{B}_2 = \{w_1, w_2, \ldots, w_n\}$ are bases for V and W, respectively. Now we have a chain of maps

$$\mathbb{F}^m \xrightarrow{\Phi_{\mathcal{B}_1}^{-1}} V \xrightarrow{T} W \xrightarrow{\Phi_{\mathcal{B}_2}} \mathbb{F}^n$$

which is a composition of linear maps and hence a linear map. But we have already seen that a linear map from $\mathbb{F}^m$ to $\mathbb{F}^n$ is represented by a matrix, namely it is T_A for some n by m matrix A.

Exercise 116. *Continuing with the discussion above, assume that $T(v_i) = a_{1i}w_1 + a_{2i}w_2 + \cdots + a_{ni}w_n$. Let A be the n by m matrix defined by $A[i,j] = a_{ij}$. Show that*

$$\Phi_{\mathcal{B}_2} \circ T \circ \Phi_{\mathcal{B}_1}^{-1} = T_A. \tag{3.3}$$

Warning: If we use different bases for V or W, or both, the matrix that represents T will change!

Definition 41. *We say that the linear map $T : V \to W$ is represented by the matrix A with respect to the bases $\mathcal{B}_1 = \{v_1, v_2, \ldots, v_m\}$ and $\mathcal{B}_2 = \{w_1, w_2, \ldots, w_n\}$, of V and W respectively, if Equation 3.3 it true.*

Because linear maps are represented by matrices, with respect to choices of bases, we have all kinds of parallel statements for linear maps and matrices. Here is a nice one.

Exercise 117. *Suppose that $T : V \to W$ is a linear map, represented by the n by m matrix A with respect to the bases $\mathcal{B}_1$ and $\mathcal{B}_2$. Show that T is invertible if and only if A is invertible. Moreover, if T is invertible, its inverse is represented by A^{-1}*

Chapter 4

Determinants

4.1 Definition of the Determinant

Associated to any square matrix A with entries in $\mathbb{F}$ is a number in $\mathbb{F}$ called its *determinant* and denoted $\det A$. We will prove the following theorem in a series of exercises and use it to find a formula for the determinant.

Theorem 2. *Let $M(n, \mathbb{F})$ be the set of all n by n matrices with entries in $\mathbb{F}$. There is a unique function $\det : M(n, \mathbb{F}) \to \mathbb{F}$, called the determinant, that satisfies all the following properties:*

1. $\det I_n = 1$
2. *If the rank of A is less than n, then $\det A = 0$.*
3. *The $\det$ is linear in each row. This means that for each i,*

$$\det \begin{pmatrix} r_1 \\ \vdots \\ r_{i-1} \\ av + bw \\ r_{i+1} \\ \vdots \\ r_n \end{pmatrix} = a \det \begin{pmatrix} r_1 \\ \vdots \\ r_{i-1} \\ v \\ r_{i+1} \\ \vdots \\ r_n \end{pmatrix} + b \det \begin{pmatrix} r_1 \\ \vdots \\ r_{i-1} \\ w \\ r_{i+1} \\ \vdots \\ r_n \end{pmatrix}$$

where each r_j is a row of the matrix, v and w are rows, and a and b are scalars.

Exercise 118. *Suppose that* $\det : M(n, \mathbb{F}) \to \mathbb{F}$ *is a function that satisfies the three properties of Theorem 2. Show that the following are true:*

1. *If* A' *is obtained from* A *by multiplying some row by the scalar* a*, then* $\det A' = a \det A$*.*

2. *If* A' *is obtained from* A *by adding a multiple of one row to another, then* $\det A' = \det A$*.*

3. *If* A' *is obtained from* A *by trading two rows, then* $\det A' = -\det A$*.*

Exercise 119. *Suppose* $\det$ *and* $\det'$ *are two functions satisfying the properties of Theorem 2. Suppose that* B *is obtained from* A *by elementary row operations. Then* $\det A = \det' A$ *if and only if* $\det B = \det' B$*.*

Exercise 120. *Show that if* $\det$ *is a function satisfying the properties of Theorem 2, then it is unique. (Hint: First consider a matrix with less than full rank. Then consider a matrix that has full rank, and therefore can be taken to the identity matrix by elementary row operations.)*

So, we now know that if a function satisfying the properties of Theorem 2 exists, then it is unique. But we still need to show such a function exists! We will do this by induction on n.

Exercise 121. *If* $A = (a)$ *is a 1 by 1 matrix, define the determinant as* $\det A = a$*. Show that this function satisfies the properties of Theorem 2.*

Definition 42. *Let* A *be an* n *by* n *matrix. Define* A_{ij} *to be the* $n-1$ *by* $n-1$ *matrix obtained from* A *by removing the* i*-th row and* j*-th column.*

Exercise 122 (**)**.** *Assume that* $\det : M(n, \mathbb{F}) \to \mathbb{F}$ *is a function that satisfies the three properties of Theorem 2. Fix any* i *with* $1 \leq i \leq n+1$ *and define* $\det : M(n+1, \mathbb{F}) \to \mathbb{F}$ *as*

$$\det A = \sum_{j=1}^{n+1} (-1)^{i+j} A[j, i] \det A_{ji}. \tag{4.1}$$

Show that $\det : M(n+1, \mathbb{F}) \to \mathbb{F}$ *is a function that satisfies the three properties of Theorem 2.*

Computing the determinant of a matrix by using Equation 4.1 is called *expanding the determinant* along the i-th column. Because the determinant function is unique, it doesn't matter which column we expand on! This is useful in practice: choosing a column that has one or more zeroes shortens the computation.

Before proving Exercise 122, let's look at what it says for a 2 by 2 or 3 by 3 matrix. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then expanding on the first column gives

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\det(d) - c\det(b) = ad - cb.$$

If $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}$ then expanding on the second column gives

$$\begin{aligned} \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} &= -b\det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + e\det \begin{pmatrix} a & c \\ g & i \end{pmatrix} - h\det \begin{pmatrix} a & c \\ d & f \end{pmatrix} \\ &= -b(di - fg) + e(ai - cg) - h(af - cd). \end{aligned}$$

Exercise 123. *Compute the determinant of the following matrices.*

1. $\begin{pmatrix} -3 & 4 \\ 1 & -2 \end{pmatrix}$
2. $\begin{pmatrix} 1 & 0 & -1 \\ 2 & 3 & -1 \\ 1 & -2 & 4 \end{pmatrix}$
3. $\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 3 \\ 1 & -2 & 5 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$

Proof of Exercise 122: Supose that A is an $n+1$ by $n+1$ matrix and fix i, the number of the column that we will expand on. We need to show that the function defined by Equation 4.1 satisfies the three properties of Theorem 2, *assuming* that the determinant function applied to each of the n by n submatrices does satisfy the properties of Theorem 2.

If $A = I_{n+1}$, then each term of the sum in Equation 4.1 where $j \neq i$ will be zero because $A[j,i] = 0$ if $j \neq i$. The term with $j = i$ will be 1 because $A_{ii} = I_n$ and we are assuming that for this smaller size matrix, $\det I_n = 1$. Thus $\det I_{n+1} = 1$.

Next we will show that Equation 4.1 gives a function that is linear in each row. This is going to be really messy to write out! Suppose

$$A = \begin{pmatrix} r_1 \\ \vdots \\ r_{j-1} \\ av + bw \\ r_{j+1} \\ \vdots \\ r_{n+1} \end{pmatrix}$$

where each $r_k = (r_{k,1}, r_{k,2}, \ldots, r_{k,n+1})$ is a row vector, $v = (v_1, v_2, \ldots, v_{n+1})$ and $w = (w_1, w_2, \ldots, w_{n+1})$ are row vectors, and a and b are scalars. Define the $n+1$ by $n+1$ matrices B and C as

$$B = \begin{pmatrix} r_1 \\ \vdots \\ r_{j-1} \\ v \\ r_{j+1} \\ \vdots \\ r_{n+1} \end{pmatrix} \qquad C = \begin{pmatrix} r_1 \\ \vdots \\ r_{j-1} \\ w \\ r_{j+1} \\ \vdots \\ r_{n+1} \end{pmatrix}.$$

Out goal is to show that $\det A = a \det B + b \det C$. Notice that v and w appear in the j-th row of B and C and that $av + bw$ appears in the j-th row of A. So, staying away from the j-th row, we have the same entries. In particular,

$$A[k,i] = B[k,i] = C[k,i], \text{if } k \neq j.$$

Moreover, if we strike the j-th row and i-th column out of A, B, and C, we get the same n by n matrix. That is

$$A_{ji} = B_{ji} = C_{ji}.$$

But if we strike out the k-th row and i-th column from each matrix with $k \neq j$ we get smaller matrices and, because we are assuming the properties of Theorem 2, we have

$$\det A_{ki} = a \det B_{ki} + b \det C_{ki}, \text{ if } k \neq j.$$

Now, let's apply Equation 4.1 to A, splitting the sum into the first $j-1$ terms, the j-th term, and then the last $n+1-j$ terms. Regrouping terms and using the above facts, we get,

$$\begin{aligned}
\det A = &\sum_{k=1}^{j-1}(-1)^{k+i}A[k,i]\det A_{ki} \\
&+ (-1)^{j+i}A[j,i]\det A_{ji} \\
&\quad + \sum_{k=j+1}^{n+1}(-1)^{k+i}A[k,i]\det A_{ki} \\
= &\sum_{k=1}^{j-1}(-1)^{k+i}A[k,i](a\det B_{ki} + b\det C_{ki}) \\
&+ (-1)^{j+i}(av_j + bw_j)\det A_{ji} \\
&\quad + \sum_{k=j+1}^{n+1}(-1)^{k+i}A[k,i](a\det B_{ki} + b\det C_{ki}) \\
= &\, a\left[\sum_{k=1}^{j-1}(-1)^{k+i}A[k,i]\det B_{ki} + (-1)^{j+i}v_j\det A_{ji} + \sum_{k=j+1}^{n+1}(-1)^{k+i}A[k,i]\det B_{ki}\right] \\
&+ b\left[\sum_{k=1}^{j-1}(-1)^{k+i}A[k,i]\det C_{ki} + (-1)^{j+i}w_j\det A_{ji} + \sum_{k=j+1}^{n+1}(-1)^{k+i}A[k,i]\det C_{ki}\right] \\
= &\, a\left[\sum_{k=1}^{j-1}(-1)^{k+i}B[k,i]\det B_{ki} + (-1)^{j+i}v_j\det B_{ji} + \sum_{k=j+1}^{n+1}(-1)^{k+i}B[k,i]\det B_{ki}\right] \\
&+ b\left[\sum_{k=1}^{j-1}(-1)^{k+i}C[k,i]\det C_{ki} + (-1)^{j+i}w_j\det C_{ji} + \sum_{k=j+1}^{n+1}(-1)^{k+i}C[k,i]\det C_{ki}\right] \\
= &\, a\det B + b\det C.
\end{aligned}$$

It remains to show that if A does not have full rank, then $\det A = 0$. So, let's suppose that A does not have full rank. Then some row is a linear combination of the other rows and hence this row can be made to be the zero vector by using elementary row operations of the type where we add a multiple of one row to another. If an n by n matrix has a row of zeroes, then it does not have full rank and hence has determinant equal to zero because we are assuming that the determinant function satisfies the properties of Theorem 2 for n by n matrices. Therefore, if A has a row of zeroes, Equation 4.1 will produce zero. Hence it suffices to show that the function defined by Equation 4.1 is unchanged by adding a multiple of one row to another. Suppose A' is obtained from A by adding b times the k-th row to the j-th row. Because we have already shown that Equation 4.1 defines a function that is linear in each row, we get that $\det A' = \det A + b \det B$, where B is the matrix obtained from A be replacing row j with row k. Thus, B has two rows that are the same: row j and row k. So, finally, it remains to show that if A is a matrix with two rows that are the same, then the function defined by Equation 4.1 produces zero.

Therefore, suppose that A has two rows that are the same, say $r_s = r_t$. Each term in Equation 4.1 with k not equal to either s or t will give zero because the smaller matrix A_{ki} will have two rows that are the same, and hence will not be of full rank. Hence its determinant will be zero. Thus two terms in the sum are left to analyze: the terms where $k = s$ and $k = t$. The formula gives

$$\begin{aligned} \det A &= (-1)^{s+i} A[s,i] \det A_{si} + (-1)^{t+i} A[t,i] \det A_{ti} && (4.2) \\ &= (-1)^{s+i} A[s,i] (\det A_{si} + (-1)^{t-s} \det A_{ti}) && (4.3) \end{aligned}$$

since the two rows are the same and so $A[s,i] = A[t,i]$. If the two rows are adjacent, then $t - s = \pm 1$ and $A_{si} = A_{ti}$, so we get zero. What if the rows are separated by one row? That is, $|s - t| = 2$. In this case, A_{si} and A_{ti} differ by swapping two adjacent rows and hence $\det A_{si} = -\det A_{ti}$ By Exercise 118. More generally, if $|s - t| = p$, then the claim is that A_{si} and A_{ti} differ by swapping a pair of rows $p - 1$ times. Thus if p is even, the signs of the two terms in (4.3) are the same, but the two determinants are opposite and cancel. Or, if p is odd, the determinants have the same sign, but the signs of the terms are different and again the terms cancel. $\square$

Exercise 124. *If* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, *show that its determinant is the number* $\det A = ad - bc$.

4.2 Properties of the Determinant

Definition 43. *A matrix is called upper triangular if all entries below the diagonal are zero. A matrix is called lower triangular if all entries above the diagonal are zero. A matrix is called triangular if it is either upper or lower triangular.*

Exercise 125. *Show that the determinant of a triangular matrix is the product of the diagonal elements. (Hint: Induct on the size of the matrix.)*

Exercise 126. *There are three types of elementary matrices. What is the determinant of each kind?*

Exercise 127. *Suppose that* $E_1, E_2, \ldots, E_k$ *are elementary matrices.*

1. *Show that* $\det(E_1 A) = \det E_1 \det A$.
2. *Use induction on* k *to show that*

$$\det(E_k E_{k-1} \ldots E_1 A) = \det E_k \det E_{k-1} \ldots \det E_1 \det A.$$

Exercise 128. *Show that* $\det(AB) = \det A \det B$. *(Hint: Show that* $A = E_1 E_2 \ldots E_k A'$ *where each* E_i *is an elementary matrix and* A' *is either the identity or a matrix whose last row is all zeroes. Now use Exercise 127.)*

Exercise 129. *If* A *is invertible, show that* $\det(A^{-1}) = (\det A)^{-1}$.

Exercise 130. *Show that a square matrix is invertible if and only if its determinant is not zero.*

Definition 44. *Let* A *be an* n *by* m *matrix. The transpose of* A*, denoted* A^T *is the* m *by* n *matrix defined as* $A^T[i,j] = A[j,i]$. *The rows of* A *are the columns of* A^T *and vice versa.*

Exercise 131. *Show that* $(AB)^T = B^T A^T$.

Exercise 132. *Show that* $(A+B)^T = A^T + B^T$ *and that* $(cA)^T = cA^T$.

Definition 45. *An elementary column operation on a matrix* A *is one of the following:*

1. *Swap two columns.*
2. *Multiply a column by a nonzero scalar.*

3. *Add a multiple of one column to another column.*

Exercise 133. *Let E be an elementary matrix.*

1. *Show that E^T is an elementary matrix.*

2. *Show that* $\det E = \det E^T$.

Exercise 134. *Show that an elementary column operation on A can be achieved by multiplying A on the right by an elementary matrix.*

Exercise 135. *Show that* $\det A = \det A^T$. *Hint: Split this into two cases: A does not have full rank, and A does have full rank. In the latter case, recall that A is a product of elementary matrices.*

Exercise 136. *Show that the determinant of any matrix can also be computed by "expanding on any row."*

4.3 A Formula for the Inverse of a Matrix

Definition 46. *Given a square matrix A, its adjugate matrix, denoted $\widetilde{A}$, is defined as* $\widetilde{A}[i,j] = (-1)^{i+j} \det A_{ji}$.

Exercise 137 (*). *Show that if A is an n by n matrix, then* $A\widetilde{A} = (\det A) I_n$. *Therefore, if* $\det A \neq 0$ *a formula for the inverse of A is*

$$A^{-1} = \frac{1}{\det A}\widetilde{A}. \tag{4.4}$$

Exercise 138. *Use Equation 4.4 to find the inverses of these matrices.*

1. $\begin{pmatrix} 3 & 5 \\ -1 & 2 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 0 & 3 \\ -1 & 2 & 1 \\ 2 & 0 & 1 \end{pmatrix}$

Chapter 5

Inner Product Spaces

NOTE: In this ENTIRE chapter we will only consider the field $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

5.1 Inner Products

Definition 47. *Let V be a vector space over $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. An inner product on V is a function from $V \times V$ to $\mathbb{F}$, where the image of the pair of vectors (v, w) is denoted by $\langle v, w \rangle$, such that*

1. $\langle v, v \rangle \geq 0$ *for all v and $\langle v, v \rangle = 0$ if and only if $v = 0$. (positive definiteness)*
2. $\langle v, w \rangle = \overline{\langle w, v \rangle}$ *for all $v, w \in V$. (conjugate symmetry)*
3. $\langle av + bw, u \rangle = a\langle v, u \rangle + b\langle w, u \rangle$ *all $a, b \in \mathbb{F}$. (linearity in the first slot)*

A vector space equipped with an inner product is called an inner product space.

If $\mathbb{F} = \mathbb{C}$, then $\langle v, w \rangle$ is a complex number, and so perhaps $\langle v, v \rangle$ is a complex number. But the first property says that $\langle v, v \rangle \geq 0$ for all vectors v. This means that $\langle v, v \rangle$ is a *real* number and moreover that it is nonnegative. Recall that the conjugate of a complex number w is denoted $\overline{w}$ and defined as $\overline{a + bi} = a - bi$. In the second property we see that the inner product is not exactly symmetric—instead it is conjugate symmetric. However, if we are using $\mathbb{F} = \mathbb{R}$, then the conjugate of a real number is itself and we do have symmetry. The last property says the inner product is linear in the first position. Combing this with the second property we have the following result.

Exercise 139. *If V is an inner product space, then*

$$\langle u, av + bw\rangle = \overline{a}\langle u, v\rangle + \overline{b}\langle u, w\rangle$$

Thus, if $\mathbb{F} = \mathbb{R}$, we have that the inner product is linear in both slots.

Exercise 140. *If V is an inner product space, show that $\langle v, 0\rangle = \langle 0, v\rangle = 0$ for all $v \in V$.*

The definition of an inner product is motivated by the "dot" product in $\mathbb{F}^n$.

Exercise 141. *If $v = (v_1, v_2, \ldots, v_n)$ and $(w_1, w_2, \ldots, w_n)$ are vectors in $\mathbb{F}^n$, show that the function*

$$v \cdot w = v_1\overline{w_1} + v_2\overline{w_2} + \cdots + v_n\overline{w_n}$$

defines an inner product. We call this the standard inner product on $\mathbb{F}^n$. If $\mathbb{F} = \mathbb{R}$, then we do not need the conjugation. This gives the classic dot product on $\mathbb{R}^n$, which we call the standard inner product on $\mathbb{R}^n$.

Notice that if we think of vectors in $\mathbb{F}^n$ as columns, then $v \cdot w$ is simply the matrix multiplication $v^T\overline{w}$. NOTE: Throughout the rest of this chapter we will think of all vectors as columns so that v^T is a row.

Exercise 142. *Let A be an invertible n by n matrix with entries in $\mathbb{F}$. Show that $\langle v, w\rangle = Av \cdot Aw = (Av)^T(\overline{Aw}) = v^T A^T\, \overline{A}\overline{w}$ defines an inner product on $\mathbb{F}^n$. Thus there are infinitely many different possible inner products on $\mathbb{F}^n$.*

Exercise 143. *Suppose that V is a finite dimensional inner product space with basis $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$. Suppose that $\langle\ ,\ \rangle$ is some inner product on V. Show that if we know the value of the inner product on all pairs of basis elements, then we know its value on any pair of vectors. In particular, let B be the matrix where $B[i, j] = \langle v_i, v_j\rangle$. Show that*

$$\langle x, y\rangle = (\Phi_{\mathcal{B}}(x))^T B\overline{\Phi_{\mathcal{B}}(y)}$$

for all $x, y \in V$. If $V = \mathbb{F}^n$, the basis is the standard basis, and the inner product is the dot product, show that $B = I_n$.

Exercise 144. *For which values of $a, b, c,$ and d will $v^T \begin{pmatrix} a & b \\ c & d \end{pmatrix} \overline{w}$ define an inner product on $\mathbb{F}^2$?*

Exercise 145. *Let V be the vector space of all continuous functions from $[a,b]$ to $\mathbb{F}$. Show that*

$$\langle f,g\rangle = \int_a^b f(x)\overline{g(x)}\,dx$$

defines an inner product on V.

Definition 48. *If V is an inner space, the norm of a vector v is defined to be $||v|| = \sqrt{\langle v,v\rangle}$. The distance between two vectors v and w is defined to be $||v-w||$.*

Exercise 146 (Cauchy-Schwartz Inequality)**.** *If V is an inner product space, show that*

$$|\langle v,w\rangle| \leq ||v||\,||w||$$

for all vectors v and w. Hint: If $w \neq 0$, start with

$$0 \leq \langle v - \langle v,w\rangle/||w||^2, v - \langle v,w\rangle/||w||^2\rangle$$

and then expand the right side.

Exercise 147. *If V is an inner product space, show that all the following are true:*

1. $||v|| \geq 0$ *for all* $v \in V$.
2. $||v|| = 0$ *if and only if* $v = 0$.
3. $||av|| = |a|\,||v||$ *for all* $v \in V, a \in \mathbb{F}$.
4. $||v+w|| \leq ||v|| + ||w||$.

The last property is called the triangle inequality. Why?

Additional Exercise 5. *Suppose V in an inner product space.*

1. *If $\mathbb{F} = \mathbb{R}$, show that*

$$\langle u,v\rangle = (||u+v||^2 - ||u-v||^2)/4.$$

2. *If $\mathbb{F} = \mathbb{C}$, show that*

$$\langle u,v\rangle = (||u+v||^2 - ||u-v||^2 + i||u+iv||^2 - i||u-iv||^2)/4.$$

Definition 49. *Two vectors v and w are called perpendicular, or orthogonal to each other, denoted as $v \perp w$, if $\langle v,w\rangle = 0$. If W is a set of vectors in V, we will write $u \perp W$ to mean that u is orthogonal to every vector in W.*

Exercise 148 (Pythagorean Theorem)**.** *If V is an inner product space and u and v are orthogonal vectors, then*

$$||u+v||^2 = ||u||^2 + ||v||^2. \tag{5.1}$$

Conversely, if V is a **real** *inner product space ($\mathbb{F} = \mathbb{R}$), show that Equation 5.1 implies that u and v are othogonal.*

Definition 50. *If W is a subset of the inner product space V, then the orthogonal complement of W is defined as*

$$W^{\perp} = \{v \in V \mid v \perp w \textit{ for all } w \in W\}.$$

Exercise 149. *If W is any subset of an inner product space V, show that $W^{\perp}$ is a subspace of V.*

Definition 51. *A set of vectors $\{v_1, v_2, \dots . v_k\}$ in an inner product space is called an orthonormal system if*

$$\langle v_i, v_j \rangle = \begin{cases} 1 & \textit{if } i = j \\ 0 & \textit{if } i \neq j \end{cases}$$

Exercise 150. *Show that the standard basis on $\mathbb{F}^n$ is an orthonormal system with respect to the standard inner product.*

Exercise 151. *Show that an orthonormal system of vectors is always linearly independent.*

Exercise 152. *Suppose $\{v_1, v_2, \dots, v_n\}$ is an orthonormal basis for V. Show that for any vector $v \in V$,*

$$v = \sum_{i=1}^{n} \langle v, v_i \rangle v_i$$

and

$$||v||^2 = |\langle v, v_1 \rangle|^2 + |\langle v, v_2 \rangle|^2 + \cdots + |\langle v, v_n \rangle|^2.$$

Exercise 153. *Suppose $\{w_1, w_2, \dots, w_k\}$ is an orthonormal system in V and let W be the subspace spanned by $\{w_1, w_2, \dots, w_k\}$. Show that for any vector $v \in V$, we can write v uniquely as $v = w + u$ where $w \in W$ and $u \in W^{\perp}$. Show that*

$$w = \sum_{i=1}^{k} \langle v, w_i \rangle w_i$$

and $u = v - w$,

Exercise 154 (**Gram-Schmidt Orthonormalization Process).** *Let V be an inner product space and $\{v_1, v_2, \ldots, v_k\}$ a set of linearly independent vectors in V. Let $v_1' = v_1/||v_1||$ and inductively define v_{i+1}' as*

$$v_{i+1}' = w_{i+1}/||w_{i+1}|| \text{ where } w_{i+1} = v_{i+1} - \sum_{j=1}^{i} \langle v_{i+1}, v_j' \rangle v_j'.$$

Show that for each $1 \leq i \leq k$ we have span$(v_1, v_2, \ldots, v_i) =$ span$(v_1', v_2', \ldots, v_i')$ and that $\{v_1', v_2', \ldots, v_k'\}$ is an orthonormal set.

Exercise 155. *Apply the Gram-Schmidt process to the basis*

$$\{(1, 0, -1), (0, 1, 3), (0, 0, 1)\}$$

of $\mathbb{R}^3$ (using the standard inner product on $\mathbb{R}^3$).

Exercise 156. *Consider the real vector space of polynomials with real coefficients and of degree less than or equal to 5 with domain $[-\pi, \pi]$ and the inner product defined in Exercise 145. Apply the Gram-Schmidt process to the basis $\{1, x, x^2, x^3, x^4, x^5\}$.*

Exercise 157. *Show that every finite dimensional subspace of an inner product space has an orthonormal basis. (Notice that this implies that every finite dimensional inner product space has an orthonormal basis.)*

Exercise 158. *Let W be a finite dimensional subspace of an inner product space V. Show there exists a unique linear map $\pi_W : V \to W$ such that $\pi_W(v) = v$ for all $v \in W$, Null$(\pi_W) = W^\perp$, and π_W is onto. The map π_W is called the orthogonal projection of V onto W. Given a vector $v \in V$, we might sometimes write proj$_W v$ for $\pi_W(v)$.*

NOTE: In the previous exercise, the fact that $\pi_W(v) = v$ for all $v \in W$ can be described by saying that "π_W restricts to the identity map on W."

Exercise 159. *Suppose V is an inner product space and W is a finite dimensional subspace. Show that $\pi_W \circ \pi_W = \pi_W$.*

Exercise 160. *Suppose V is an inner product space and W is a finite dimensional subspace.*

1. *Show that $(W^\perp)^\perp = W$.*

2. *Show that $V = W \oplus W^\perp$.*

5.2 Minimization Problems

Given an inner product space V, a subspace W and a vector $v \in V$, we often want to find the closest vector in W to v. The next exercise states that the closest vector of W to v is the orthogonal projection of v into W.

Exercise 161. *Suppose that W is a finite dimensional subspace of an inner product space V and $v \in V$. Then the closest vector in W to v is $\pi_W(v)$. Equivalently,*

$$||v - \pi_W(v)|| \leq ||v - w||$$

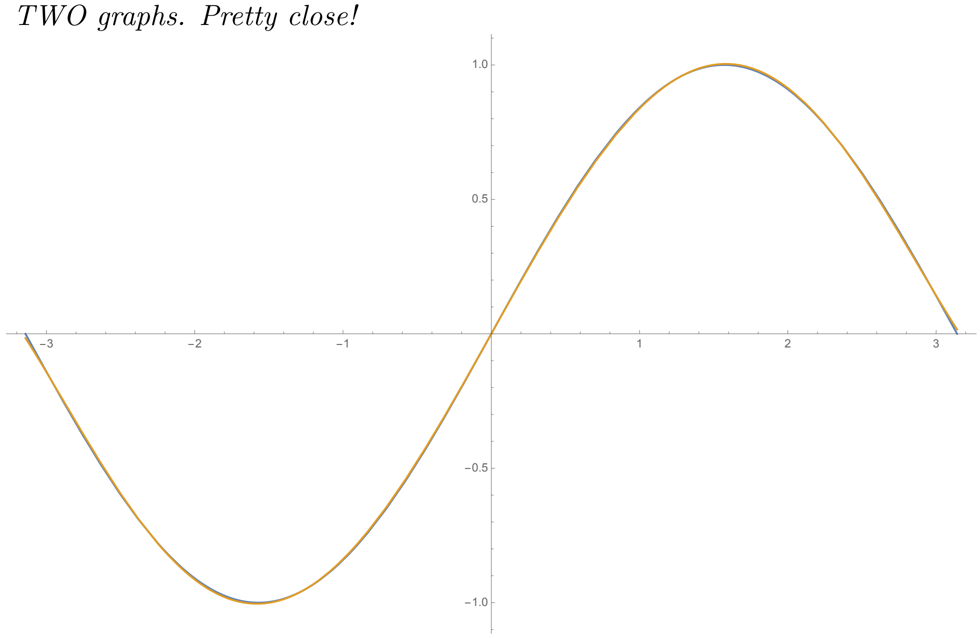
for all $w \in W$.

Example 5. *Consider the vector space of continuous functions from $[-\pi, \pi]$ to $\mathbb{R}$ with the inner product defined in Exercise 145. It is hard to compute values of $\sin x$ and easy to compute values of polynomials. So, it is useful to approximate $\sin x$ with a polynomial. Let $P_5(x)$ be the subspace of real polynomial functions of degree at most five. Let's find the closest element of $P_5(x)$ to $\sin x$ by orthogonally projecting $\sin x$ into $P_5(x)$.*

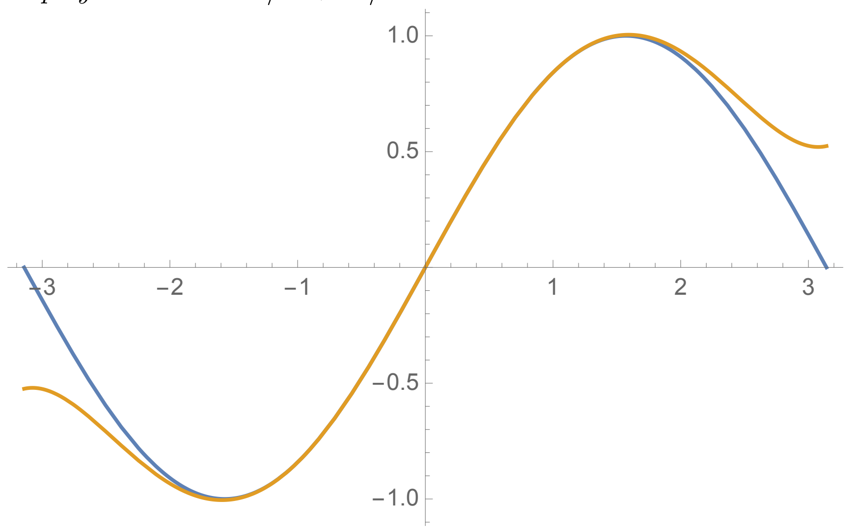
Exercise 162. *Use the orthonormal basis for $P_5(x)$ found in Exercise 156 to show that*

$$\pi_{P_5(x)}(\sin x) = 0.987862x - 0.155271x^3 + 0.00564312x^5.$$

Plotting BOTH $\sin x$ and the polynomial above, we get the following TWO graphs. Pretty close!



Compare this with the following plot of both $\sin x$ *and the fifth degree Taylor polynomial* $1 - x^3/3! + x^5/5!$.



Example 6. *Suppose we want to solve a system of linear equations* $Ax = b$. *There will only be a solution if* b *lies in the image of the map given by multiplication by* A. *In other words, if* b *lies in the column space Col* A *of* A. *Suppose that* b *does not lie in Col* A. *Then we can't solve the system. Instead, let's try to approximate the solution—let's find the vector* x *so that* Ax *is as close as possible to* b. *Since* Ax *is always in the column space of* A, *the closest vector* Ax *to* b *is the orthogonal projection of* b *into the column space. Call this vector* b'. *Now we can solve* $Ax = b'$ *and this gives a vector* x *whose image* Ax *is as close as possible to* b.

Exercise 163 (Method of Least Squares). *Suppose we want to find the "best fitting line" to a set of data points* $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$. *Suppose the line is given by* $y = mx + b$. *Our task is to find the best choices for the variables* m *and* b *so that the line is as "close as possible" to the data points. We want to solve the system*

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_n & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Of course, if $n > 2$ *it is unlikely that a line contains all the points. That is, it is unlikely that the vector* $y = (y_1, y_2, \ldots, y_n)$ *is in the column space* *of the the coefficient matrix. Therefore, we will orthogonally project y into the column space to obtain the vector y' and then solve the system with y replaced by y'.*

1. *Starting with the basis $w_1 = (1, 1, \ldots, 1)$ and $x = (x_1, x_2 \ldots, x_n)$ for the column space, use Gram-Schmidt to find an orthonormal basis $\{w'_1, w'_2\}$ for the column space.*

2. *Next, project y into the column space, writing its image y' as a linear combination of w'_1 and w'_2.*

3. *Now find m and b. It will be helpful to introduce the following notation for various average values:*

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$
$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \qquad \overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i$$

   You will finally give formulae for m and b in terms of $\overline{x}, \overline{y}, \overline{x^2}$, and $\overline{xy}$.

4. *This method of fitting a line to the data is also called the method of least squares because the answer minimizes the function*

$$E(m, b) = \sum_{i=1}^{n} (y_i - mx_i - b)^2$$

   which is the sum of the squares of the vertical distances between each data point and the line $y = mx + b$. Solve the problem again by using calculus to find the minimum value of E. Do this by finding the partial derivatives of E with respect to both m and b and finding out what values of m and b make both partial derivatives equal to zero.

Example 7 (Fourier Series)**.** *There are many cases when it is useful to approximate a periodic function with sines and cosines. A Fourier approximation of order n to a continuous function $f(x)$ on the interval $[0, 2\pi]$ is*

$$a_0 + a_1 \cos x + a_2 \cos 2x + \cdots + a_n \cos nx + b_1 \sin x + b_2 \sin 2x + \cdots + b_n \sin nx.$$

Exercise 164. *Show that the functions*

$$1, \cos x, \cos 2x, \ldots, \cos nx, \sin x, \sin 2x, \ldots, \sin nx$$

are an orthonormal set with respect to the inner product given in Exercise 145.

Exercise 165. *Find the Fourier approximation of order 4 to the function $f(x) = x$ on the interval $[0, 2\pi]$ by projecting $f(x)$ into the subspace spanned by*

$$1, \cos x, \cos 2x, \ldots, \cos 4x, \sin x, \sin 2x, \ldots, \sin 4x.$$

Using a graphing utility, plot both $f(x)$ and its Fourier approximation.

Exercise 166. *Let $f(x)$ be 1 on the interval $[0, \pi]$ and -1 on the interval $(\pi, 1]$. Find the Fourier approximation of order 4 to the function $f(x)$ on the interval $[0, 2\pi]$. Using a graphing utility, plot both $f(x)$ and its Fourier approximation.*

5.3 Isometries

Definition 52. *Suppose $T : V \to W$ is a linear map between inner product spaces. The map T is called an isometry if $||Tv||_W = ||v||_V$ for all $v \in V$. Here the subscripts of W and V have been added to emphasize that the norm of Tv is taken using the norm in W while the norm of v is taken using the norm in V.*

Exercise 167. *Show that an isometry between inner product spaces is injective.*

Exercise 168. *Suppose that V is a finite dimensional inner product space and $T : V \to V$ is an isometry. Show that T is invertible.*

Exercise 169. *Suppose $T : V \to W$ is a linear map between inner product spaces. Show that the following are equivalent.*

1. *T is an isometry.*
2. *$\langle Tv, Tu\rangle_W = \langle v, u\rangle_V$ for all $u, v \in V$.*
3. *T takes every orthonormal set of vectors to an orthonormal set of vectors.*

Hint: Use Additional Exercise 5

5.4 Orthogonal Matrices

Definition 53. *An n by n matrix with entries in $\mathbb{F}$ is called orthogonal if its columns are an orthonormal set of vectors in $\mathbb{F}^n$ with respect to the standard inner product.*

Exercise 170. *Suppose A is an n by n matrix with entries in $\mathbb{F}$. Show that the following are all equivalent.*

1. *A is orthogonal.*
2. *The columns of A form an orthonormal system with respect to the standard inner product on $\mathbb{F}^n$.*
3. $A^T\overline{A} = \overline{A}^T A = I_n$.
4. *A is invertible and $A^{-1} = \overline{A}^T$.*
5. $A\overline{A}^T = \overline{A}A^T = I_n$.
6. *The rows of A form an orthonormal system with respect to the standard inner product on $\mathbb{F}^n$.*

Exercise 171. *Show that the determinant of an orthogonal matrix must be ± 1.*

Exercise 172. *Show that the product of two orthogonal matrices must be orthogonal.*

Exercise 173. *Suppose that A is an n by n orthogonal matrix with entries in $\mathbb{F}$.*

1. *Show that $v \cdot w = Av \cdot Aw$ for all vectors $v, w \in \mathbb{F}^n$.*
2. *Show that $||Av|| = ||v||$ for all $v \in \mathbb{F}^n$. (Here the norm is defined with respect to the standard inner product.*

Exercise 174. *Let $V = \mathbb{R}^2$ with the standard metric.*

1. *Show that every orthogonal matrix must be of the form*

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \quad \textit{or} \quad \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}$$

2. *If $A = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$, show that $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is rotation through the angle t.*
3. *Let $A = \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}$. Show that $A = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and hence that $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is reflection across the x axis followed by rotaton through the angle t.*

Chapter 6

Eigenvalues and Eigenvectors

6.1 Change of Bases

Don't Forget:

$\boxed{\textbf{A linear map between vector spaces corresponds to a matrix.}}$

Let's recall how this works. The simplest case is:

> If $T : \mathbb{F}^m \to \mathbb{F}^n$ is a linear map, then their exists an n by m matrix A such that $T(v) = Av$ for every vector $v \in \mathbb{F}^m$. In otherwords, the image of v is given by multiplying v by the matrix A. The first column of A is the image of the vector $e_1 = (1, 0, \ldots, 0)$, the second column of A is the image of the vector $e_2 = (0, 1, \ldots, 0)$, and so on.

More generally:

> If V is an m-dimensional vector space and W is an n-dimensional vector space, then a linear map $T : V \to W$ *corresponds to* multiplication by an n by m matrix *once we have chosen bases for both V and W*. Now is a GREAT time to review Exercise 116!

A particularly simple case is when $V = W$ and we consider a linear map $T : V \to V$. A particularly simple case of this is when T is the identity. We are now talking about "a change of baisis for V." For example, suppose that V is the space of real-valued polynomials of degree less than or equal to three. One basis for V is $\mathcal{B}_1 = \{1, x, x^2, x^3\}$ and another basis is the orthonormal basis $\mathcal{B}_2$ obtained from $\mathcal{B}_1$ by the Gram-Schmidt process. in Exercise 156 we found that

$$\mathcal{B}_2 = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\sqrt{\frac{3}{2}}x}{\pi^{3/2}}, -\frac{\sqrt{\frac{5}{2}}\left(\pi^2 - 3x^2\right)}{2\pi^{5/2}}, \frac{5\sqrt{\frac{7}{2}}\left(x^3 - \frac{3\pi^2 x}{5}\right)}{2\pi^{7/2}} \right\}.$$

Suppose v is a vector in V and we express it with respect to the first basis. For example, suppose that $v = 3x^3 - 2x + 4$. With respect to the first basis, this is

$$v = 4 \cdot 1 - 2 \cdot x + 0 \cdot x^2 + 3 \cdot x^3$$

or

$$v = (4, -2, 0, 3)$$

if we just write the components of v. Question: What are the component of v, *the very same vector*, with respect to the basis $\mathcal{B}_2$? Of course, Exercise 152 gives the answer because $\mathcal{B}_2$ is an orthonormal basis. Namely we get

$$v = \langle v, w_1 \rangle w_1 + \langle v, w_2 \rangle w_2 + \langle v, w_3 \rangle w_3 + \langle v, w_4 \rangle w_4$$

where $\mathcal{B}_2 = \{w_1, w_2, w_3, w_4\}$.

Exercise 175. *Continuing with the above example, suppose v has components $\mathbf{a} = (a_1, a_2, a_3, a_4)$ with respect to the basis*

$$\mathcal{B}_2 = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\sqrt{\frac{3}{2}}x}{\pi^{3/2}}, -\frac{\sqrt{\frac{5}{2}}\left(\pi^2 - 3x^2\right)}{2\pi^{5/2}}, \frac{5\sqrt{\frac{7}{2}}\left(x^3 - \frac{3\pi^2 x}{5}\right)}{2\pi^{7/2}} \right\}.$$

Show that the components of v with respect to the basis $\mathcal{B}_1$ are given by $A\mathbf{a}^T$, where

$$A = \begin{pmatrix} \frac{1}{\sqrt{2\pi}} & 0 & -\frac{1}{2}\sqrt{\frac{5}{2\pi}} & 0 \\ 0 & \frac{\sqrt{\frac{3}{2}}}{\pi^{3/2}} & 0 & -\frac{3\sqrt{\frac{7}{2}}}{2\pi^{3/2}} \\ 0 & 0 & \frac{3\sqrt{\frac{5}{2}}}{2\pi^{5/2}} & 0 \\ 0 & 0 & 0 & \frac{5\sqrt{\frac{7}{2}}}{2\pi^{7/2}} \end{pmatrix}.$$

Thus to go from a linear combination of the basis elements in $\mathcal{B}_2$ to a linear combination of the basis elements in $\mathcal{B}_1$, we multiply the 4-tuple of components with respect to the basis $\mathcal{B}_2$ by the matrix A to get the 4-tuple of *components with respect to the basis $\mathcal{B}_1$. The matrix A is called the "change of bases matrix."*

To go the other way, from $\mathcal{B}_1$ to $\mathcal{B}_2$, we must multiply by the inverse of A. Find A^{-1}.

NOTE: Just to make the previous exercise a bit more clear, consider the polynomial

$$f(x) = -\frac{\sqrt{\frac{5}{2}}\left(\pi^2 - 3x^2\right)}{2\pi^{5/2}}.$$

This polynomial has components $(0,0,1,0)$ with respect to $\mathcal{B}_2$ and it obviously has components

$$\left(-\frac{\sqrt{\frac{5}{2}}\pi^2}{2\pi^{5/2}}, 0, \frac{3\sqrt{\frac{5}{2}}}{2\pi^{5/2}}, 0\right)$$

with respect to $\mathcal{B}_1$, which is what we get if we multiply A times $(0,0,1,0)^T$.

Exercise 176. *Let $\mathcal{B}_1 = \{(1,0,0),(0,1,0),(0,0,1)\}$ and $\mathcal{B}_2 = \{(1,-1,2),(3,1,-1),(1,1,1)\}$. Find the change of bases matrix to go from $\mathcal{B}_1$ to $\mathcal{B}_2$.*

Exercise 177. *Suppose that $T : V \to V$ is a linear map and that $\mathcal{B}_1$ and $\mathcal{B}_2$ are two bases for V. Suppose that T corresponds to the matrix A with respect to the first basis and the matrix B with respect to the second basis. Show that*

$$B = C^{-1}AC$$

where C is the change of bases matrix to go from $\mathcal{B}_2$ to $\mathcal{B}_1$.

Definition 54. *Two square matrices A and B are called similar if there is an invertible matrix P such that $B = P^{-1}AP$.*

Thus we see that changing the basis for a vector space V will change the matrix that represents a linear transformation $T : V \to V$ to a similar matrix.

Exercise 178. *Show that being similar is an equivalence relation on the set of square matrices. (See Appendix for more on equivalence relations.)*

So, a linear transformation $T : V \to V$ corresponds to an equivalence class of similar matrices.

Big Question: Given a linear transformation $T : V \to V$ how simple can we make the matrix that represents T by simply choosing a different basis for V? Equivalently, given a square matrix A, how simple can we make it by changing it to $P^{-1}AP$? Or, saying it yet another way, given a square matrix A, what is the simplest matrix that is similar to A? If our goal is to understand what T is like, or what T does, it will be simpler if we choose a basis such that the matrix representing T with respect to this basis is as simple as possible. Here we should think of the identity matrix, or a diagonal matrix, as REALLY simple, and a triangular matrix as next simple. A randomly chosen matrix is probably not so simple.

6.2 Eigenvectors

Definition 55. *Let $T : V \to V$ be a linear transformation. An eigenvector for T is a non-zero vector v such that $T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$. The scalar λ is called an eigenvalue with associated eigenvector v.*

Exercise 179. *Let $T : V \to V$ be a linear transformation and assume that V is finite dimensional. Then the matrix A that represents T with respect to the basis $\mathcal{B}$ is diagonal if and only if $\mathcal{B}$ consists entirely of eigenvectors. In the case that A is diagonal, then the diagonal entries are all eigenvalues.*

Definition 56. *Assume V is finite dimensional. A linear map $T : V \to V$ is called diagonalizable if there is a basis for V such that the matrix that represents T with respect to this basis is diagonal. Thus, the previous exercise says that T is diagonalizable if and only if there exists a basis of eigenvectors.*

Exercise 180. *Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be as follows. Determine if there is a basis of eigenvectors. If there is, find the corresponding diagonal matrix that describes T with respect to this basis.*

1. *T is reflection across the line $y = mx$.*
2. *T is rotation around the origin through the angle θ.*
3. *T is a "shear" map. An example of this is the map $T(x, y) = (ax, y)$, where a is some nonzero constant.*

Exercise 181. *Let $T : V \to V$ be a linear transformation.*

1. *Show that v is an eigenvector for T with associated eigenvalue λ if and only if v is in the null space of $T - \lambda Id$.*
2. *If $V = \mathbb{R}^n$ and T is given by multiplication by the matrix A, show that v is an eigenvector of T with associated eigenvalue λ if and only if $(A - \lambda I_n)v = 0$.*

Exercise 182. *Let A be an n by n matrix over $\mathbb{F}$. Show that* $\det(A - \lambda I_n)$ *is a degree n polynomial in the variable λ with coefficients in $\mathbb{F}$.*

Definition 57. *Let A be an n by n matrix over $\mathbb{F}$. The polynomial* $\det(A - \lambda I_n)$ *is called the characteristic polynomial of A.*

Exercise 183. *Show that if two matrices are similar, then they have the same characteristic polynomial.*

Exercise 184. *Let $T : V \to V$ be a linear transformation with V finite dimensional and $\mathcal{B}$ a basis for V. Suppose that with respect to this basis, T corresponds to multiplication by the matrix $A_{\mathcal{B}}$. Show that the characteristic polynomial of A does not depend on the choice of basis $\mathcal{B}$. In other words, if we change the basis and get a different matrix we will still have the same characteristic polynomial. Thus we may define the characteristic polynomial of T to be the characteristic polynomial of $A_{\mathcal{B}}$ for any basis $\mathcal{B}$ of V.*

Exercise 185. *Let $T : V \to V$ be a linear transformation with V finite dimensional. Show that there exists an eigenvector for T with associated eigenvalue λ if and only if λ is a root of the characteristic polynomial of T.*

Exercise 186. *Let $T : V \to V$ be a linear transformation with V finite dimensional. Assume that λ is a root of the characteristic polynomial of T. Show that the set E_λ of all eigenvectors with associated eigenvalue λ forms a subspace of V. This subspace is called the eigenspace associated to λ.*

Exercise 187. *For each of the following matrices find the characteristic polynomial and all of its roots. These are the eigenvalues associated to the matrix. For each eigenvalue, find a basis for the associated eigenspace.*

1. $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

2. $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$

3. $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$

Exercise 188. *Let $T : V \to V$ be a linear transformation with eigenvectors $\{v_1, v_2, \ldots, v_k\}$ and associated eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$, respectively. Show that if all of the eigenvalues are distinct, then the eigenvectors are linearly independent.*

We will assume the following BIG theorem and its corollaries.

Theorem 3 (Fundamental Theorem of Algebra)**.** *If*

$$P(z) = c_0 + c_1 z + \cdots + c_n z^n$$

is a polynomial with complex coefficients, $n \geq 1$, and $c_n \neq 0$, then $P(z)$ has at least one complex root.

Corollary 4. *Every polynomial with complex coefficients can be factored into linear factors. In other words, if*

$$P(z) = c_0 + c_1 z + \cdots + c_n z^n$$

is a polynomial with complex coefficients, $n \geq 1$, and $c_n \neq 0$, then we can factor $P(z)$ as

$$P(z) = c_n (z - z_1)^{n_1} (z - z_1)^{n_2} \cdots (z - z_k)^{n_k}$$

where

$$n_1 + n_2 + \cdots + n_k = n.$$

Here c_n and each z_i is a complex number. The natural number n_i is called the algebraic multiplicity of the root z_i.

Exercise 189. *Suppose V is n-dimensional and $T : V \to V$ is a linear transformation.*

1. *Show that T has at most n distinct eigenvalues.*

2. *Show that if T has n distinct eigenvalues, then T is diagonalizable.*

Definition 58. *Suppose V is n-dimensional and $T : V \to V$ is a linear transformation. The algebraic multiplicity of an eigenvalue λ is the algebraic multiplicity of λ as a root of the characteristic equation. The geometric multiplicity of λ is the dimension of the associated eigenspace E_λ.*

Exercise 190. *Suppose V is n-dimensional and $T : V \to V$ is a linear transformation. If λ is an eigenvalue, show that the geometric multiplicity of λ is always less than or equal to the algebraic multiplicity of λ. Hint: Extend a basis for E_λ to a basis for all of V and consider the matrix associated to T defined by this basis.*

6.3 Application: Fibonacci Numbers

The Fibonacci numbers are the sequence

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, \ldots$$

where the first two numbers are 0 and 1, and then each number thereafter is the sum of the previous two. Thus $f_0 = 0, f_1 = 1$ and define $f_n = f_{n-1} + f_{n-2}$ for each $n > 1$. We can use linear algebra to find a formula for f_n! (That just an exciting statement, not f_n *factorial*.)

Exercise 191. *Show that*

$$\begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Exercise 192. *Suppose that the n by n matrix has n linearly independent eigenvectors $v_1, v_2, \ldots, v_n$, with associated eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Let P be the matrix that has v_i as its i-th column. Show that $D = P^{-1}AP$ is diagonal with $D[i, i] = \lambda_i$. (In doing this, ask yourself, "Why is P invertible?")*

Exercise 193. *Find the eigenvalues and eigenvectors of*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Use this to write A as $A = PDP^{-1}$ where D is diagonal. What are P and D?

Exercise 194. *Suppose that A is a square matrix.*

1. *Show that if A is diagonal, then $A^k[i, i] = A[i, i]^k$.*
2. *If $A = PDP^{-1}$, show that $A^k = PD^kP^{-1}$.*

Exercise 195. *Combine the above exercises to show that the n-Fibonacci number is given by*

$$f_n = \frac{\left(\frac{1}{2}\left(1 + \sqrt{5}\right)\right)^n - \left(\frac{1}{2}\left(1 - \sqrt{5}\right)\right)^n}{\sqrt{5}}.$$

Cool! We have a formula for the n-th Fibonacci number! But it is a pretty weird formula. It is not even clear at all that this formula will produce integers at output! What is the easiest way to find f_{100}? Should we use the formula, or just start with $0, 1, 2, 3, 5, 8$ and keep going? We are going to need a computer for both, although we could, at least in theory, work out the second approach by hand. Could we even hope to compute f_{100} using the formula *by hand*? The answer is

$$f_{100} = 354224848179261915075.$$

The key to finding this formula was to diagonalize the matrix A which we could do because it has a complete set of eigenvectors. This method can be used to find formulae in all sorts of situations that involve the power of a matrix.

Appendix A

Appendix

A.1 Proof by Induction

Suppose we want to prove not just one statement, but infinitely many statements, $S_1, S_2, S_3, \ldots$. An important method to do such a thing is called *proof by induction*. The method is based on the following theorem.

Theorem 5. *Let $S_1, S_2, S_3, \ldots$ be a sequence of statements. If both of the following are true:*

1. *S_1 is true, and*

2. *For every i, if S_i is true, then S_{i+1} is true,*

then all the statements are true.

Proof: Assume that S_1 is true and that the truth of any one statement S_i implies the truth of the next statement, S_{i+1}. If all the statements are not true, then there must be some that are false. Moreover, there must be a first statement in the list that is false, say S_k. Now this first false statement is not S_1, because we are assuming that S_1 is true. Hence $k > 1$. Because S_k is the first false statement in the list, all the statements before it are true and, again, there are statements before it because $k > 1$. Therefore, the statement S_{k-1} is a true statement right before S_k. But now we have that S_k is true because we are assuming that if any one of the statements is true, then the next one is. We have arrived at a contradiction: the statement S_k is both true and false. Hence, our assumption that not all the statements are true must be wrong. In fact, they are all true! $\square$

Example 8. *Prove that $2^n > n$ for every positive integer n.*

Solution: We first need to recognize that this is in fact a sequence of statements, not just one statement. The statements are

$$\begin{aligned}
S_1 &: 2^1 > 1 \\
S_2 &: 2^2 > 2 \\
S_3 &: 2^3 > 3 \\
&\vdots
\end{aligned}$$

We will prove that all the statements are true by the method of induction. The first thing we need to do is prove the first statement, or *base case*, $2^1 > 1$. Proving the first statement is called "starting the induction." Since $2^1 = 2$ and $2 > 1$, the first statement is true. Next we ned to show that if any one of the statements is true, then the next statement is true. Doing this is called "proving the inductive step." So, we assume that for some $n > 1$, $2^n > 2$. This is our assumption, and using this assumption, we need to prove that $2^{n+1} > n + 1$. Notice that

$$\begin{aligned}
2^{n+1} &= 2 \cdot 2^n \\
&> 2n, \text{ because we are assuming that } 2^n > n \\
&= n + n \\
&> n + 1, \text{ because we are assuming that } n > 1
\end{aligned}$$

This completes the proof of the inductive step and hence the entire proof that all the statements are true.

Example 9. *Show that $2^n \geq n^2$ for all $n \geq 4$.*

Solution: We will prove this by induction on n. If $n = 4$, the statement is $2^4 \geq 4^2$ which is true because $2^4 = 16 = 4^2$. Now assume that for some $n > 4, 2^n \geq n^2$ and consider 2^{n+1}. We have,

$$\begin{aligned}
2^{n+1} &= 2 \cdot 2^n \\
&\geq 2 \cdot n^2, \text{ by assumption} \\
&\geq n^2 + n^2 \\
&\geq n^2 + n \cdot n \\
&\geq n^2 + 4n, \text{ since } n > 4 \\
&\geq n^2 + 2n + 2n \\
&\geq n^2 + 2n + 1, \text{ since } n > 4 \\
&\geq (n+1)^2
\end{aligned}$$

This completes the inductive step. Hence $2^n \geq n^2$ for all $n \geq 4$. $\square$

Exercise 196. *Use the method of induction to prove each of the following.*

1. $n! > 2^n$ *for all* $n \geq 4$
2. $(1+x)^n \geq 1 + nx$ *for all* $x > 0$ *and* $n \geq 1$
3. *Prove that if* $x_1, x_2, \ldots, x_n$ *are all odd integers, then* $x_1 x_2 \ldots x_n$ *is odd.*
4. $\displaystyle\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$

A.2 Equivalence Relations

Definition 59. *Suppose that* S *is a set. Let* R *be a subset of* $S \times S$*. In other words,* R *is a set of ordered pairs of elements from* S*. If* $(x, y) \in R$*, we will write* $x \sim y$ *and say that "*x *is related to* y*." We will refer to either the subset* R *or the symbol* $\sim$ *as "the relation." The relation* R*, or said differently, the relation* $\sim$*, is called an equivalence relation if the following three properties all hold.*

1. *(reflexive)* $x \sim x$ *for every* $x \in S$
2. *(symmetric) If* $x \sim y$ *then* $y \sim x$
3. *(transitive) If* $x \sim y$ *and* $y \sim z$*, then* $x \sim z$

Exercise 197. *Determine if the following relations on the given set are equivalence relations.*

1. $S = \mathbb{R}$ *and* $x \sim y$ *if* $x \leq y$*.*
2. $S = \mathbb{R}$ *and* $x \sim y$ *if* $x < y$*.*
3. S *is any set and* $x \sim y$ *if* $x = y$*.*
4. S *is any set and we declare* $x \sim y$ *if* $x, y \in S$*. In otherwords, every element is related to every element.*
5. *Let* n *be a fixed natural number.* S *is the set of integers and* $x \sim y$ *if* $x - y$ *is a multiple of* n*.*
6. S *is the set of people in the world and* $x \sim y$ *if* x *loves* y*.*

7. *S is the set of all people who have ever lived and $x \sim y$ if x and y share a common ancestor.*

Definition 60. *Suppose that $\sim$ is an equivalence relation on the set S. Then the equivalence class of $a \in S$ is defined as*

$$[a] = \{b \in S \mid b \sim a\}.$$

Exercise 198. *For each of the relations in Exercise 197 that are equivalence relations, describe the equivalence classes.*