Introduction

What is linear algebra, and more importantly, why do we study it? Well, believe it or not, you have wanted to ‘do’ linear algebra your whole life. Doubtless at some point in your mathematical life, you have wrongly assumed that \( \sin(x + y) = \sin(x) + \sin(y) \), or \( \log(3x) = 3 \log(x) \). The area of mathematics which does behave this nicely is linear algebra. Besides your innate desire, there are at least four important reasons for anyone who will be working in math/science/technology to study this subject.

1. You will learn some interesting and useful mathematics. The well known mathematician I.N. Herstein wrote, “Linear Algebra is real mathematics, interesting and exciting on its own, yet it is probably that part of mathematics which finds the widest application - in physics, chemistry, economics, in fact in almost every science and pseudo science.”

2. In studying linear algebra you will form a mental filing system to categorize the mathematics that you already know. For example, you will discover and explore the connections and common ground between such varied topics as matrices, sequences, differentiable functions, and polynomials. You’ll see how ideas like orthogonality and length can be generalized. This will allow you to differentiate functions by multiplying matrices, decompose continuous functions into a collection of trigonometric functions (the basis of all digital sound), and even to draw a square circle.

   The power and beauty of working with the “common ground” is that everything we prove will apply to lots of mathematics. It is similar to anthropologists, who after studying many cultures, sit back, reflect, and discover commonalities to all cultures. These commonalities then give insight on the common ground - human nature. It is also not unlike chemists and physicists who discover the common ground among a host of physical phenomena - they are called the Laws of Nature. For example, a great moment in history occurred when Isaac Newton discovered the common ground between a falling apple and the motion of planets.

3. Linear algebra is a relatively easy, common sense subject which provides you with a great opportunity to learn how to organize and write a grammatically correct mathematical proof. Why is this important? Anyone in a profession which uses mathematics must also know how to communicate in mathematics. Indeed, often it is the ability to communicate with managers and clients which distinguishes the person to be promoted to broader responsibilities. As an added bonus, learning precise writing and careful, critical thinking in mathematics carries over into clear, concise communication of all forms. Indeed, Abraham Lincoln studied mathematics (Euclid’s Elements) just to sharpen his mind and his debating skills. These are certainly goals of a liberal arts education.

4. A final skill you will begin developing is the ability to work and think with symbols and ideas which are hard to “taste, touch, and feel.” You will first encounter this when we define a vector space. In fact, this tendency to generalize and work with the
abstract is what distinguished mathematics from all other disciplines. It is also what makes mathematics so powerful.

How to do well...

1. Think of the theorems as problems with solutions. That is, when you read a theorem, try to think of the proof (or better yet - write it down!) before reading the proof. In this way, each theorem becomes a “problem with solution.” Don’t waste the opportunity to think creatively. Even if you don’t solve it, by struggling with it, you have carved out a little hollow in your mind so that you better appreciate and absorb the solution when you read it. Also notice the writing style of the proofs. Even memorizing the proofs of key theorems is a good idea because you will i) learn the theorem well, ii) learn techniques for solving other problems, and iii) get practice writing correctly.

2. Restate the theorem in your own words and apply the theorem to specific cases. Pretend that you are explaining the theorem to someone who doesn’t understand it. As you restate it, you will be forced to think more carefully about what it means. By applying the theorem to well known situations, the truth of the theorem may become more obvious. Better yet, draw a picture to illustrate the theorem. In linear algebra, this is often the best way to get a feeling for what is going on.

3. Think of writing a proof as solving a puzzle. Some puzzles are solved through straightforward mechanical means; others require a burst of insight. Therefore, much time is spent thinking, fiddling, exploring, and getting frustrated. As soon as you solve a problem, you will write it down and then go on to get frustrated on the next one. Sound discouraging? You do it all the time for ‘fun’ if you play crossword puzzles, video games, or learn any athletic skill. For most people though, developing a skill is more fun if done with a partner or group. That is a good reason to work together. You can build on each other’s ideas and also catch flaws in each others’ arguments.

4. Finally, proving a mathematical result requires two distinct stages much like that of a prosecutor in a criminal case. The prosecutor must first convince him/herself of the guilt of the defendant by cleverly piecing together the truth. Then, after being convinced oneself, the prosecutor must put together an airtight argument that convinces the jury. Similarly, the first stage in problem solving is a wild, creative, throw-out-any idea that comes to your mind mode. In linear algebra, this often involves drawing pictures and convincing oneself from geometric arguments. Then when you think you have pieced together the solution, you must put on the critical, examining, check-every-step-carefully-for-logical-errors mode. This is best done when you are writing the solution using mathematical symbols and careful logic. Thus the writing stage is an important part of the solution process.
0 Preliminaries

0.1 Definitions

Much of mathematical writing is done using the following words and symbols:

Definition 0.1.

1. Definition

2. Theorem - a major proved result (important to know and use).

3. Lemma - a result established in order to prove a theorem.

4. Corollary - a quick and easy result from a theorem.

5. \( \forall \): for all

6. \( \exists \): there exists

7. s.t. : such that

8. WLOG : without loss of generality

9. \( \Rightarrow \): implies

10. \( \Leftrightarrow \), iff : if and only if

11. \( \Rightarrow \Leftarrow \): contradiction

12. Statement: If in Detroit, then in Michigan.

13. Converse: If in Michigan, then in Detroit.

14. Inverse: If not in Detroit, then not in Michigan.

15. Contrapositive: If not in Michigan, then not in Detroit.

Notice that the contrapositive of a statement is true iff the statement is true, whereas the truth of the converse doesn’t follow from the truth of the statement.

Definition 0.2. A set is a well defined collection of objects. We specify a set by listing its elements: \( \{0, 1, 2\} \), or by giving a rule for set membership: \( \{x: x > 0\} \) (read: \( x \) such that \( x > 0 \)). We denote the empty set, the set with no elements, as \( \emptyset \).

Definition 0.3. The following symbols represent set operations and relations:

1. \( \in \): is an element of: \( x \in X \), e.g. \( 1 \in \{0, 1, 2\} \).

2. \( \subset \): is a subset of: \( Y \subset X \), e.g. \( \{1, 2\} \subset \{1, 2, 3\} \).

3. \( A = B \) iff \( A \subset B \) and \( B \subset A \).
4. \(\cup:\) union: \(A \cup B = \{a : a \in A \text{ or } a \in B\}\), i.e. all elements in either \(A\) or \(B\) (or both).
5. \(\cap:\) intersection: \(A \cap B = \{a : a \in A \text{ and } a \in B\}\), i.e. all elements in both \(A\) and \(B\).
6. \(\setminus:\) complementation: \(B \setminus A = \{a : a \in B \text{ and } a \notin A\}\), i.e. all elements in \(B\) but not in \(A\).

**Definition 0.4.** We also make frequent use of some predefined sets:

1. \(\mathbb{Z}:\) integers
2. \(\mathbb{R}:\) real numbers
3. \(\mathbb{C}:\) complex numbers
4. \(\mathbb{F}:\) will represent either \(\mathbb{R}\) or \(\mathbb{C}\) (both of which are fields)
5. \(\mathcal{C}[X]:\) continuous functions \(f : X \to \mathbb{R}\)
6. \(\mathcal{C}^\infty:\) infinitely differentiable functions \(f : \mathbb{R} \to \mathbb{R}\).
7. \(\mathcal{P}^n(X):\) polynomials of degree less than or equal to \(n\) from \(X\) into \(\mathbb{R}\).

### 0.2 Methods of Proof

We employ four basic techniques of proof, that is, four ways of establishing the truth of a theorem: direct proof, contraposition, proof by contradiction, and induction. We illustrate each proof technique with an example:

In a direct proof, we proceed logically from the hypothesis until we reach the desired conclusions.

**Theorem 0.5 (Direct Proof).** If \(n \in \mathbb{Z}\) is odd, \(n^2\) is also odd.

**Proof.** We begin with our hypothesis: \(n\) is odd. Thus we can write \(n = 2m + 1\) for some integer \(m\). Now \(n^2 = (2m + 1)^2 = 4m^2 + 4m + 1 = 2(2m^2 + 2m) + 1\), so \(n^2\) is odd.

Contraposition involves proving the contrapositive of a proposition. As noted above, the contrapositive of a statement is equivalent to the statement itself.

**Theorem 0.6 (Contraposition).** If \(n^2\) is even, \(n\) is also even.

**Proof.** We negate the conclusion: \(n\) is odd, and we negate the hypothesis: \(n^2\) is odd. Now forming the contrapositive of the statement, we get: If \(n\) is odd, \(n^2\) is odd, which we have just proven.

In proof by contradiction, we assume the opposite of the result we wish to prove and derive a contradiction, that is, we arrive at an impossible situation. If our logic throughout the proof was correct, the fact that we end up with an absurdity implies that the initial assumption was false.

Some of the most beautiful and important mathematical results have been established by proofs of contradiction. For example, Euclid showed there is an infinite number of primes and Cantor showed that there are more real numbers than rational numbers.
Theorem 0.7 (Contradiction). \( \sqrt{2} \) is irrational.

Proof. Here we start by assuming the opposite, \( \sqrt{2} \) is rational, that is, we can write \( \sqrt{2} = m/n \) where \( m/n \) is a fraction in lowest terms. Now, \( \sqrt{2} = m/n \Rightarrow n\sqrt{2} = m \Rightarrow 2n^2 = m^2 \), that is, \( m^2 \) is even, so by the last theorem, \( m \) must also be even, so we can write \( m = 2r \) for some \( r \in \mathbb{Z} \). Substituting back in, we get \( 2n^2 = (2r)^2 = 4r^2 \Rightarrow n^2 = 2r^2 \), so now \( n \) is even, so \( n = 2q \) for some \( q \in \mathbb{Z} \). But now \( m/n = 2r/2q \), which is not in lowest terms, so we have a contradiction. We conclude that \( \sqrt{2} \) is irrational.

Induction is a “domino effect” proof technique. If we have a row of dominoes to topple, we must first tip the first domino \((n = 1)\), and the falling of a domino down the line \((k)\) must cause the next domino \((k + 1)\) to fall. Analagously, if we are trying to prove that a result holds “for all \( n \)”, it suffices to show that the result holds for \( n = 1 \), and that if the result holds for some \( k \), it also holds for \( k + 1 \).

Theorem 0.8 (Induction). \( 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \)

Proof. First, we must show that the formula holds when \( n = 1 \): \( \frac{1(1+1)}{2} = 1 \).

Now, we suppose that for some \( k \), \( 1 + 2 + \cdots + k = \frac{k(k+1)}{2} \). We must show that our formula holds for \( k + 1 \), that is, \( 1 + \cdots + k + (k + 1) = \frac{(k+1)(k+1+1)}{2} \). Start by adding \( k + 1 \) to both sides of our formula for \( 1 + \cdots + k \):

\[
1 + 2 + \cdots + k + (k + 1) = \frac{k(k + 1)}{2} + (k + 1).
\]

Now get a common denominator and expand the right hand side:

\[
1 + 2 + \cdots + k + (k + 1) = \frac{k^2 + k + (2k + 2)}{2} = \frac{k^2 + 3k + 2}{2}.
\]

When we factor the numerator on the right, we obtain: \( 1 + 2 + \cdots + k + (k + 1) = \frac{(k+1)(k+2)}{2} \).

Exercise 0.9. Show that \( 1^3 + 2^3 + \cdots + n^3 = (1 + 2 + \cdots + n)^2 \).
1 Matrices and Systems of Equations

Definition 1.1. An \( m \times n \) matrix is a rectangular array of numbers with \( m \) rows and \( n \) columns. \( \mathbb{F}^m_n \) is the collection of all such matrices with entries in \( \mathbb{F} \). When the context is clear, we shall use \( \mathbb{F}^n \) to stand for either \( \mathbb{F}_n^n \) or \( \mathbb{F}_1^n \). If \( m = n \), the matrix is called square. We write \( A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \). The matrix \( 0 = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \) with all 0 entries is called the zero matrix.

Definition 1.2. Let \( A \) be a square matrix. The diagonal of \( A \) is the set of elements \( a_{kk} \) which proceed diagonally down and to the right in the matrix. \( A \) is diagonal if every non-diagonal entry is zero. \( A \) is called upper triangular if every entry below the diagonal is zero, and lower triangular if every entry above the diagonal is zero.

Definition 1.3. If \( A, B \in \mathbb{F}^m_n \), \( c \in \mathbb{F} \),

1. \( A = B \) iff \( a_{ij} = b_{ij} \ \forall i, j \)

2. \( A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \)

3. \( cA = \begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & \ddots & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{pmatrix} \)

Definition 1.4. The transpose \( A^t \) of an \( m \times n \) matrix \( A \) is the \( n \times m \) matrix obtained by making the \( i \)th row of \( A \) the \( i \)th column of \( A^t \). That is, if \( B = A^t \), then \( b_{ij} = a_{ji} \).

Definition 1.5. If \( A \in \mathbb{F}^m_n \) and \( B \in \mathbb{F}^n_p \), then we form the product \( AB \in \mathbb{F}^{m}_p \) by multiplying as follows:

\[
(ab)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.
\]

This definition is perhaps best illustrated by an example.

Exercise 1.6. Multiply \( \begin{pmatrix} 1 & -3 & 0 & 6 \\ 0 & 2 & 1 & 0 \\ 2 & 0 & 4 & 4 \end{pmatrix} \) \( \begin{pmatrix} 9 & -3 \\ 0 & -2 \\ 0 & 5 \\ 2 & 0 \end{pmatrix} \). We see that the matrices can be multiplied and that their product will be a \( 3 \times 2 \) matrix:

\[
\begin{pmatrix}
1(9) + 3(0) + 0(0) + 6(2) & 1(3) + 3(2) + 0(5) + 6(0) \\
0(9) + 2(0) + 1(0) + 0(2) & 0(-3) + 2(-2) + 1(5) + 0(0) \\
2(9) + 0(0) + 4(0) + -4(2) & 2(-3) + 0(-2) + 4(5) + -4(0)
\end{pmatrix} = \begin{pmatrix} 21 & 3 \\ 0 & 1 \\ 10 & 14 \end{pmatrix}.
\]
What properties does matrix multiplication share with the multiplication of numbers? We now consider ways in which matrix multiplication is both like and unlike the multiplication of numbers.

**Question 1.7.** Is $AB = BA$?

1. No - $AB$ and $BA$ are not both defined unless $m = n = p$.

2. Even if $m = n = p$, $A$ and $B$ need not commute:

$$\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ -6 & 8 \end{pmatrix}$$

$$\begin{pmatrix} -2 & 3 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ -7 & -1 \end{pmatrix}$$

**Question 1.8.** If $AB = 0$, must $A$ or $B$ be 0? No:

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

**Theorem 1.9.** When the following products are defined,

1. $A(BC) = (AB)C$

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

3. $A(kB) = k(AB) = (kA)B$

4. $k(A + B) = kA + kB$.

**Proof.** Exercise.

We now turn our attention to systems of equations. Consider the following $2 \times 2$ systems of equations:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) $2x - y = 1$</td>
<td>(2) $2x - y = 1$</td>
<td>(1) $2x - y = 1$</td>
</tr>
<tr>
<td>(2) $x + y = 2$</td>
<td>(2) $2x - y = 2$</td>
<td>(2) $4x - 2y = 2$</td>
</tr>
</tbody>
</table>

**Graphical Solution:**

(one solution) (no solution) (inf. many solns.)
Algebraic Solution:
A: Add (1) and (2) to obtain $3x = 3$, then substitute $x = 1$ into one of the equations.
B: Adding (1) and (2) yields the impossibility $0 = 1$.
C: Subtracting (1) from (2) gives us (1). (1) and (2) are essentially the same equation.

We now use matrices to find algebraic solutions to systems of equations.

**Definition 1.10.** Given an $m \times n$ matrix $A$ and a column vector $b$ (i.e. an $m \times 1$ matrix), we form the augmented matrix $(A|b)$ by adjoining $b$ to $A$:

$$
\begin{pmatrix}
  a_{11} & \cdots & a_{1n} & b_1 \\
  \vdots & \ddots & \vdots & \vdots \\
  a_{m1} & \cdots & a_{mn} & b_m
\end{pmatrix}
$$

You can do the same things with rows of a matrix that can be done with a system of equations. Namely:

1. Multiply a row by a non-zero constant
2. Interchange two rows
3. Add a multiple of one row to another.

These operations are called **elementary row operations (e.r.o.)**.

We can now write the system of equations

$$
\begin{align*}
  a_{11}x + a_{12}y &= b_1 \\
  a_{21}x + a_{22}y &= b_2
\end{align*}
$$

as a matrix equation:

$$
\begin{pmatrix}
  a_{11} & a_{12} \\
  a_{21} & a_{22}
\end{pmatrix}
\begin{pmatrix}
  x \\
  y
\end{pmatrix}
= 
\begin{pmatrix}
  b_1 \\
  b_2
\end{pmatrix}
$$

and we can write it as an augmented matrix

$$
\begin{pmatrix}
  a_{11} & a_{12} & b_1 \\
  a_{21} & a_{22} & b_2
\end{pmatrix}
$$

and solve it using elementary row operations. This process is known as **Gaussian Elimination**.

**Definition 1.11.** The first non-zero entry in any row of a matrix is called the **leading entry** of the row. If the leading entries of a matrix proceed diagonally down and to the right (i.e. the entries below a leading entry are all zeros and leading entries in rows nearer the bottom of a matrix are nearer the right), and all leading entries are equal to one, the matrix is in **row echelon form**. If, in addition, the entries above a leading entry are all zeros, then the matrix in is in **row-reduced echelon form (rref)**.

We now provide an algorithm for transforming matrices into rref using elementary row operations.
Gaussian Elimination

1. Move any rows containing only zeros to the bottom of the matrix.

2. Get a 1 in the upperleftmost position possible, either by switching the top row with a row whose leading entry is a 1 or by dividing the first row by its leading entry.

3. Add multiples of the first row to each row below it in order to get all zeros below the leading entry of the first row.

4. Repeat the procedure on the \((n - 1) \times (n - 1)\) matrix formed by deleting the first row and first column of the matrix being reduced.

5. The matrix is now in row echelon form. To get the matrix into rref, add multiples of each non-zero row to the rows above it to get zeros above all the leading entries.

The process is probably clarified by an example.

**Exercise 1.12.** Solve the system of equations

\[
\begin{align*}
2x_1 + 3x_2 + x_3 + 4x_4 - 9x_5 &= 17 \\
x_1 + x_2 + x_3 + x_4 - 3x_5 &= 6 \\
x_1 + x_2 + x_3 + 2x_4 - 5x_5 &= 8 \\
2x_1 + 2x_2 + 2x_3 + 3x_4 - 8x_5 &= 14.
\end{align*}
\]

\[
\left(\begin{array}{cccc|c}
2 & 3 & 1 & 4 & 9 \\
1 & 1 & 1 & 1 & -3 \\
1 & 1 & 1 & 2 & 8 \\
2 & 2 & 2 & 3 & 8
\end{array}\right) \rightarrow \left(\begin{array}{cccc|c}
1 & 1 & 1 & 1 & 3 \\
2 & 3 & 1 & 4 & -9 \\
1 & 1 & 1 & 2 & -5 \\
2 & 2 & 2 & 3 & 8
\end{array}\right) \rightarrow
\]

\[
\left(\begin{array}{cccc|c}
1 & 1 & 1 & 1 & -3 \\
0 & 1 & -1 & 2 & -3 \\
0 & 0 & 0 & 1 & 2 \\
0 & 0 & 1 & 2 & 2
\end{array}\right) \rightarrow \left(\begin{array}{cccc|c}
1 & 1 & 1 & 1 & -3 \\
0 & 1 & -1 & 2 & -3 \\
0 & 0 & 0 & 1 & 2 \\
0 & 0 & 0 & 0 & 0
\end{array}\right) \rightarrow
\]

\[
\left(\begin{array}{cccc|c}
1 & 1 & 1 & 0 & -1 \\
0 & 1 & -1 & 0 & 1 \\
0 & 0 & 0 & 1 & -2 \\
0 & 0 & 0 & 0 & 0
\end{array}\right) \rightarrow \left(\begin{array}{cccc|c}
1 & 0 & 2 & 0 & -2 \\
0 & 1 & -1 & 0 & 1 \\
0 & 0 & 0 & 1 & -2 \\
0 & 0 & 0 & 0 & 0
\end{array}\right)
\]

Now rewrite the original system as the reduced system:

\[
\begin{align*}
x_1 + 2x_3 & \quad 2x_5 = 3 \\
x_2 - x_3 + x_5 &= 1 \\
x_4 & \quad 2x_5 = 2.
\end{align*}
\]
Variables corresponding to leading entries are called leading variables \((x_1, x_2, x_4\) in this case). The other variables \((x_3, x_5)\) are called nonleading variables. Set \(x_3 = s, x_5 = t.\) Solving for each variable in terms of \(s\) and \(t\) we arrive at:

\[
x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2s + 2t + 3 \\ s - t + 1 \\ 2s + 2t \\ 2t + 2 \\ t \end{pmatrix} = s \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{pmatrix}.
\]

**Note 1.13.** Given the system of equations \(A\mathbf{x} = b\), we denote the set of solutions of \(A\mathbf{x} = b\) by \(X_g\). For example, in the previous exercise, \(X_g = \{s() + t() + () : s, t \in \mathbb{R}\}\). We often write \(x_g = s() + t() + ()\) and call \(x_g\) the general solution of \(A\mathbf{x} = b\). We obtain a particular solution (denoted \(x_p\)) of \(A\mathbf{x} = b\) by assigning particular numerical values to \(s\) and \(t\). For example, taking \(s = 0\) and \(t = 1\), \(x_p = \begin{pmatrix} 5 \\ 0 \\ 0 \\ 4 \\ 1 \end{pmatrix}\) is a particular solution to \(A\mathbf{x} = b\).

We now revisit our earlier example.

**Matrix Solution:**

<table>
<thead>
<tr>
<th>(A)</th>
<th>(B)</th>
<th>(C)</th>
</tr>
</thead>
</table>
| \[
\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},
\]
\(\Rightarrow\) \[
\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},
\]
\(\Rightarrow\) \[
\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},
\]
\(\Rightarrow\) \[
\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},
\]
| \[
\begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},
\]
\(\Rightarrow\) \[
\begin{pmatrix} 2 & -1 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},
\]
| \[
\begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},
\]
| \[
\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},
\]
| \[
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},
\]
| \[
\Rightarrow x = 1,
\]
| \|\|y = 1,\|
| \[
\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \end{pmatrix},
\]
| \[
\Rightarrow y = 2t - 1,
\]
| \[
\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ 2t - 1 \end{pmatrix},
\]
| \[
= t \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix},
\]

**Question 1.14.** What are the possible solution sets for a \(3 \times 3\) system? Each equation determines a plane; the possibilities for the intersection of three planes are:

1. no solution
2. 1 point
3. a line of solutions

4. a plane of solutions

**Definition 1.15.** The **rank** of a matrix is the number of non-zero rows of the matrix when it is in row echelon form.

**Theorem 1.16.** Given the system $Ax = b$ (1), where $A$ is an $n \times n$ matrix:

1. $\text{rank}(A) < \text{rank}(A|b) \Leftrightarrow (1)$ has no solution

2. $\text{rank}(A) = \text{rank}(A|b) = n \Leftrightarrow (1)$ has a unique solution

3. $\text{rank}(A) = \text{rank}(A|b) < n \Leftrightarrow (1)$ has infinitely many solutions (There are $n - \text{rank}(A)$ variables).

**Proof.** These three properties are mutually exclusive, and are also all of the possibilities for a system, since $\text{rank}(A) \leq \text{rank}(A|b)$ (the rank of a matrix cannot be decreased by augmenting it). Hence if 1 3 are proved in one direction, their converses hold as well.

1. If $\text{rank}(A) < \text{rank}(A|b)$, then a row of $A$ is all zeros, but the corresponding entry of $b$ is nonzero, so $Ax = b$ has no solution— one of the linear equations reduces to $0 = b$. 

2. If $\text{rank}(A) = \text{rank}(A|b) = n$ there is a leading entry in each column of $A$, so $A$ can be row-reduced to the identity matrix. After reducing $(A|b)$ so $A$ is the identity, the solution to the equation is the reduced form of $b$.

3. If $\text{rank}(A) = \text{rank}(A|b) < n$, then some row of $(A|b)$ is all zeros, so $A$ has less than $n$ leading entries. Now each nonleading entry becomes a parameter which can vary, so $Ax = b$ has infinitely many solutions.

**Definition 1.17.** The **nullity** of a matrix is the number of non-leading variables.

**Theorem 1.18.** Given $A \in \mathbb{F}^n_m$, $\text{rank}(A) + \text{null}(A) = n$.

**Proof.** It suffices to show that $\text{rank}(A)$ is the number of leading variables. $\text{rank}(A)$ is the number of non-zero rows of $A$ in row echelon form, so since each non-zero row has a leading entry and each entry corresponds to a different variable in row echelon form, $\text{rank}(A)$ is equal to the number of leading variables.

**Definition 1.19.** The system $Ax = b$ is called **homogeneous** if $b = 0$ and **nonhomogeneous** if $b \neq 0$.

**Theorem 1.20.** The system $Ax = 0$ has either the unique solution $x = 0$ or has infinitely many solutions.

**Proof.** Exercise.
Theorem 1.21. A linear combination of solutions to $Ax = 0$ is also a solution.

Proof. Let $x_1, \ldots, x_n$ be solutions to $Ax = 0$. Let $x = a_1x_1 + \cdots + a_nx_n$. Then $Ax = A(a_1x_1 + \cdots + a_nx_n) = A(a_1x_1) + \cdots + A(a_nx_n) = a_1A(x_1) + \cdots + a_nA(x_n) = 0 + \cdots + 0 = 0$.

Theorem 1.22. Let $X_g$ be the set of solutions to $Ax = b$ and $X_h$ be the set of solutions of the corresponding homogeneous system $Ax = 0$. Then $X_g = x_p + X_h$, where $x_p$ be a particular solution of $Ax = b$. That is,

1. if $x_g \in X_g$, then $x_g = x_p + x_h$, where $x_h \in X_h$, and

2. if $x = x_p + x_h$, then $x \in X_g$.

Proof. 1. Given $x_g \in X_g$, $A(x_g - x_p) = Ax_g - Ax_p = b - b = 0$. Thus $\exists x_h$ s.t. $x_g - x_p = x_h$, so $x_g = x_p + x_h$.

2. Given $x_h \in X_h$, $A(x_p + x_h) = Ax_p + Ax_h = b + 0 = b$, so $x_p + x_h \in X_g$.

Definition 1.23. A square matrix of the form

\[
\begin{pmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{pmatrix}
\]

is called an identity matrix - denoted $I$. $I$ has the property that $IA = AI = A$ (as long as the dimensions agree).

(Prove!)

Definition 1.24. $A \in \mathbb{F}_n$ is said to be invertible if $\exists B \in \mathbb{F}_n$ s.t. $AB = BA = I$. $B$ is said to be an inverse of $A$.

Theorem 1.25. If $A$ is invertible, its inverse is unique (denoted $A^{-1}$).

Proof. Suppose $B$ and $C$ are both inverses of $A$. Then

\[B = BI = B(AC) = (BA)C = IC = C.\]

Theorem 1.26. $(A^{-1})^{-1} = A$.

Proof. Exercise.

An important use of inverses is to find solutions of the system of equations $Ax = b$. If $A$ is invertible, then $A^{-1}Ax = A^{-1}b \Rightarrow x = A^{-1}b$. Thus, we will want to be able to find inverses of matrices. We will address the general question later, but the $2 \times 2$ case is straightforward:

\[
\begin{pmatrix}
a & b \\
c & d
\end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix}
d & b \\
c & a
\end{pmatrix}
\]

Proof. Check.

Exercise 1.27. Use inverses to solve the system of equations

\[
\begin{align*}
2x + 5y &= 12 \\
3x + 8y &= 7
\end{align*}
\]
Question 1.28. Do all matrices have inverses? No, for example, the zero matrix has no inverse. However, many other matrices also do not have inverses. We now characterize the matrices that do have inverses.

Lemma 1.29. Let $e(A)$ denote the matrix obtained by performing an elementary row operation on $A$. Then $e(A) = e(I)A$.

Proof.

1. Multiply a row of $A$ by a constant $k$:

\[
\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \cdots & k & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 1
\end{pmatrix}
\begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn} \\
\vdots & \ddots & \vdots \\
a_{n1} & \cdots & a_{nn}
\end{pmatrix}
= 
\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
\vdots & \vdots & \ddots & \vdots \\
ka_{m1} & ka_{m2} & \cdots & ka_{mn} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{pmatrix}
\]

2. Interchange two rows:

\[
\begin{pmatrix}
0 & \cdots & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 1
\end{pmatrix}
\begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn} \\
\vdots & \ddots & \vdots \\
a_{n1} & \cdots & a_{nn}
\end{pmatrix}
= 
\begin{pmatrix}
a_{m1} & a_{m2} & \cdots & a_{mn} \\
\vdots & \vdots & \ddots & \vdots \\
a_{11} & a_{12} & \cdots & a_{1n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{pmatrix}
\]

3. Add a multiple of one row to another

\[
\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
k & \cdots & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 1
\end{pmatrix}
\begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{m1} & \cdots & a_{mn} \\
\vdots & \ddots & \vdots \\
a_{n1} & \cdots & a_{nn}
\end{pmatrix}
= 
\begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
ka_{11} + ka_{m1} & \cdots & ka_{1n} + ka_{mn} \\
\vdots & \ddots & \vdots \\
a_{n1} & \cdots & a_{nn}
\end{pmatrix}
\]

Definition 1.30. Let $E$ be a matrix obtained by performing an e.r.o. on $I$. $E$ is called an elementary matrix.

Corollary 1.31. Let $R$ be the rref of $A$. Then there exist elementary matrices $E_1, \ldots, E_n$ s.t. $E_nE_{n-1}\cdots E_1A = R$.

Proof. Each step of the row-reduction of $A$ is an elementary row operation. By 1.29 each e.r.o. can be represented as multiplication by an elementary matrix, e.g. $e_1(e_2(A)) = e_1(e_2(I)A) = e_1(E_2A) = e_1(I)E_2A = E_1E_2A$, etc. where each $e_i$ is a step taken to reduce $A$ to rref and $E_i$ is the corresponding elementary matrix.
Lemma 1.32. Elementary matrices are invertible. (So \( A = E_1^{-1}E_2^{-1} \cdots E_{n-1}^{-1}E_n^{-1}R \).)

Proof. The following examples illustrate the result.

1. \[
\begin{pmatrix}
1 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
0 & \cdots & k & \cdots & 0 \\
0 & \cdots & 0 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
0 & \cdots & 1/k & \cdots & 0 \\
0 & \cdots & 0 & 1
\end{pmatrix} = I.
\]

2. \[
\begin{pmatrix}
1 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
0 & \cdots & 0 & 1 \\
0 & \cdots & 1 & 0
\end{pmatrix}
\begin{pmatrix}
1 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
0 & \cdots & 1 & 0 \\
0 & \cdots & 0 & 1
\end{pmatrix} = I.
\]

3. \[
\begin{pmatrix}
1 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
k & \cdots & 1 & \cdots & 0 \\
0 & \cdots & 1 & 0
\end{pmatrix}
\begin{pmatrix}
1 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
-k & \cdots & 1 & \cdots & 0 \\
0 & \cdots & 0 & 1
\end{pmatrix} = I.
\]

If \( \text{rref}(A) = I \), \( A = E_1^{-1}E_2^{-1} \cdots E_{n-1}^{-1}I \), so \( A^{-1} = E_nE_{n-1} \cdots E_1 = e_n(e_{n-1}(\cdots(e_1(I)))) \) where the \( e_i \)'s were used to transform \( A \) to \( I \). Thus, to obtain \( A^{-1} \), we just turn \( A \) into \( I \) and do the same e.r.o.s to \( I \) to get \( A^{-1} \).

Exercise 1.33. Find the inverse of \( A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 5 \\ 0 & 1 & 2 \end{pmatrix} \).

\[
(A|I) = \begin{pmatrix} 1 & 2 & 3 & | & 1 & 0 & 0 \\ 2 & 3 & 5 & | & 0 & 1 & 0 \\ 0 & 1 & 2 & | & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & | & -1 & 1 & -1 \\ 0 & 1 & 0 & | & 4 & 2 & 1 \\ 0 & 0 & 1 & | & -2 & 1 & 1 \end{pmatrix}.
\]

Check that \( A^{-1}A = I \).
Theorem 1.34. Given \( A \in \mathbb{F}_n \), the following are equivalent. (If one is true, then all are true; if one is false, then all are false.)

1. \( \text{rank}(A) = n \).
2. \( \text{null}(A) = 0 \).
3. \( \text{rref}(A) = I \).
4. \( A \) can be written as the product of elementary matrices.
5. \( A \) is invertible.
6. \( Ax = 0 \) has only the trivial solution \( x = 0 \).
7. \( Ax = b \) has a unique solution for all \( b \in \mathbb{F}^n \).
8. \( \det(A) \neq 0 \). (We’ll prove this later.)

Proof. \( 1 \Rightarrow 2 \) \( \text{rank}(A) = n \Rightarrow A \) has \( n \) leading entries \( \Rightarrow A \) has \( 0 \) non leading entries \( \Rightarrow \text{null}(A) = 0 \).

\( 2 \Rightarrow 3 \) \( \text{null}A = 0 \Rightarrow A \) has \( n \) leading entries \( \Rightarrow A \) can reduce to \( I \).

\( 3 \Rightarrow 4 \) \( \text{rref}(A) = I \Rightarrow A = E_1^{-1} \cdots E_{n-1}^{-1} E_n^{-1} I \).

\( 4 \Rightarrow 5 \) Since elementary matrices are invertible, \( A \) is the product of elementary matrices, so it is invertible.

\( 5 \Rightarrow 6 \) \( A \) invertible \( \Rightarrow A^{-1} Ax = A^{-1} 0 \Rightarrow x = 0 \).

\( 6 \Rightarrow 7 \) By 1.22, the general solution of \( Ax = b \) is \( x = x_h + x_f = 0 + x_f \) is unique.

\( 7 \Rightarrow 1 \) By 1.16.2.

Definition 1.35. Given \( A \in \mathbb{F}_n \), the minor of \( A_{jk} \) (denoted \( \hat{A}_{jk} \)) is the \( (n - 1) \times (n - 1) \) matrix obtained by deleting the \( j \)th row and \( k \)th column of \( A \).

Definition 1.36. Given \( A \in \mathbb{F}_n \), the determinant of \( A \), denoted \( \det(A) \) or \( |A| \) is defined as follows:

1. If \( A = (a) \), \( \det(A) = a \).
2. If \( A = (a \ b) \), \( \det(A) = ad - bc \).
3. If \( A \in \mathbb{F}_n \), \( n > 2 \), \( \det(A) = \sum_{k=1}^{n} A_{jk}(-1)^{j+k} \det(\hat{A}_{jk}) \). (Expand over \( j \)th row.)

Exercise 1.37. Calculate \( \det \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix} \).

Expanding over the first row:

\[
\begin{vmatrix} 1 & -2 & -2 \\ 3 & 1 & 2 \\ -2 & 2 & 1 \end{vmatrix} = 1 \begin{vmatrix} 1 & 2 \\ -2 & 1 \end{vmatrix} + 2 \begin{vmatrix} 3 & 2 \\ -2 & 1 \end{vmatrix} - 2 \begin{vmatrix} 3 & 1 \\ -2 & 2 \end{vmatrix} = 1(1 - 4) + 2(3 + 4) - 2(6 + 2) = -5.
\]
Although, it would take a supercomputer (operating at 1 trillion operations per second) 500,000 years to compute a $25 \times 25$ determinant using the above definition, matrices of size $100 \times 100$ or greater often arise in physical applications. The following theorems show how to speed computation of determinants using elementary row operations. Since their proofs are messy and computational, and not especially illuminating, we omit them.

**Theorem 1.38.** Given any column of $A$, say the $k$th column,

$$\det(A) = \sum_{j=1}^{n} A_{jk}(-1)^{j+k} \det(\tilde{A}_{jk}).$$

**Note 1.39.** One use of determinants is in solving systems of equations. If $Ax = b$ has a unique solution, where $A \in \mathbb{F}_n$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, then $x_i = \frac{\det(A')}{\det(A)}$, where $A'$ is the matrix obtained by replacing the $i$th column of $A$ with $b$. This method is called Cramer’s Rule.

**Exercise 1.40.** Solve the system of equations:

$$\begin{align*}
2x - 3y &= 7 \\
x + 4y &= 9
\end{align*}$$

$$x = \begin{vmatrix} 7 & -3 \\ 9 & 4 \\ 2 & 3 \\ 1 & 4 \end{vmatrix} = \frac{55}{11} = 5.$$  

$$y = \begin{vmatrix} 2 & 7 \\ 1 & 9 \\ 2 & -3 \\ 1 & 4 \end{vmatrix} = \frac{11}{11} = 1.$$  

**Theorem 1.41.** Let $A \in \mathbb{F}_n$. If $B$ is obtained from $A$ by

1. interchanging two rows or columns, then $\det(B) = -\det(A)$;  
2. multiplying each entry of a row or column by $k \in \mathbb{C}$, then $\det(B) = k \det(A)$;  
3. adding a multiple of a row (column) to a different row (column), then $\det(B) = \det(A)$.

**Theorem 1.42.** Let $A \in \mathbb{F}_n$.

1. $\det(I) = 1$.  
2. If $A$ has a row (column) of zeros, then $\det(A) = 0$. (Why? - interchange two rows.)  
3. If $A$ is an upper triangular (lower triangular, diagonal) matrix, $\det(A) = a_{11}a_{22} \cdots a_{nn}$.  
4. If $A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$ ($B, C, D$ square), then $\det(A) = \det(B) \det(D)$.  

16
5. \( \det(A^t) = \det(A) \).

6. \( \det(AB) = \det(A) \det(B) \).

7. \( \det(A^{-1}) = 1/\det(A) \). (Why are we not dividing by zero?)

We now prove the last part of 1.34.

**Theorem 1.43.** \( A \in \mathbb{F}^n \) is invertible iff \( \det(A) \neq 0 \).

**Proof.**  \( \Rightarrow \) \( A \) invertible \( \Rightarrow \) \( \text{rref}(A) = I \Rightarrow \det(A) = k \det(I) = k \neq 0 \).

\( \Leftarrow \) \( A \) not invertible \( \Rightarrow \) \( \text{rank}(A) < n \Rightarrow \det(A) = k \det(B) \) where \( B \) has a row of zeros \( \Rightarrow \det(A) = 0 \).

**Definition 1.44.** The trace of a square matrix \( A \) is \( a_{11} + a_{22} + \cdots + a_{nn} \).
2 Vector Spaces

**Definition 2.1.** A vector space over $\mathbb{F}$ (R or C) is a set of objects, $V$, together with two
operations (addition and scalar multiplication) which satisfy the following:

1. $\forall x, y \in V, \ x + y \in V.$
2. $\forall x \in V, \text{ and } \forall a \in \mathbb{F}, \ ax \in V.$
3. $\forall x, y \in V, \ x + y = y + x.$
4. $\forall x, y, z \in V, \ (x + y) + z = x + (y + z).$
5. $\exists \theta \in V \text{ s.t. } x + \theta = x, \ \forall x \in V. \ (\theta \text{ is called the zero vector}).$
6. For each $x \in V, \ \exists y \in V \text{ s.t. } x + y = \theta. \ (y \text{ is called the additive inverse of } x.)$
7. $\forall x \in V, \ 1x = x.$
8. $\forall a, b \in \mathbb{F}, \ \text{ and } \forall x \in V, \ (ab)x = a(bx).$
9. $\forall a \in \mathbb{F} \text{ and } \forall x, y \in V, \ a(x + y) = ax + ay.$
10. $\forall a, b \in \mathbb{F} \text{ and } \forall x \in V, \ (a + b)x = ax + bx.$

Giving a definition is like giving the rules of a game. Now we explore what else we can
find out about the game. (Ex: Given the rules of chess, can a knight land on every square?)

**Theorem 2.2.** Let $V$ be a vector space, $x, y, z \in V, \text{ and } a \in \mathbb{F}.$

1. The zero vector ($\theta$) is unique.
2. Given $x \in V, \text{ its additive inverse is unique (and is denoted } -x).$
3. If $x + z = y + z, \text{ then } x = y.$
4. $0x = \theta.$
5. $a\theta = \theta.$
6. $-1x = -x.$

**Proof.**

1. Suppose $\theta_1$ and $\theta_2$ each have the property that $x + \theta_i = x. \text{ Then } \theta_1 = \theta_1 + \theta_2 = \theta_2 + \theta_1 = \theta_2.$
2. Suppose $y_1$ and $y_2$ each have the property that $x + y_i = \theta. \text{ Then } y_1 = y_1 + \theta = y_1 + (x + y_2) = (y_1 + x) + y_2 = (x + y_1) + y_2 = \theta + y_2 = y_2 + \theta = y_2.$
3. $x = x + \theta = x + (z + -z) = (x + z) + -z = (y + z) + -z = y + (-z + z) = y + 0 = y.$
4. \( 0x = (0 + 0)x = 0x + 0x \Rightarrow 0x + \theta = 0x + 0x \overset{3}{\Rightarrow} \theta = 0x. \)

5. \( a\theta = a(\theta + \theta) = a\theta + a\theta \Rightarrow a\theta + \theta = a\theta + a\theta \overset{3}{\Rightarrow} \theta = a\theta. \)

6. \( x + \ 1x = 1x + 1x = (1 \ 1)x = 0x = \theta \overset{2}{\Rightarrow} 1x = x. \)

**Exercise 2.3.** Show that the following are vector spaces:

1. \( \mathbb{R}^n_m \): all \( m \times n \) matrices with real entries (includes \( \mathbb{R}^3, \mathbb{R}^n \))
2. \( \mathcal{F}(X) = \{ \text{all functions } f : X \to \mathbb{R} \} \)
3. \( \mathcal{P}(X) = \{ \text{all polynomials } p : X \to \mathbb{R} \} \)

**Exercise 2.4.** Determine why the following are not vector spaces:

1. \( V = \{ x : x \in [-1, 1] \}; x + y = xy, ax = 0 \)
2. \( V = \{ (x, y) : x, y \in \mathbb{R} \}; (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1y_2), a(x, y) = (ax, y) \)
3. \( V = \{ (x, y) : x, y \in \mathbb{R} \}; (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), a(x, y) = (x, 0). \)

**Definition 2.5.** Given the vector space \( V \) and \( W \subset V, \) \( W \) is a subspace of \( V \) if \( W \) is also a vector space under the operations on \( V. \)

**Theorem 2.6.** Given the vector space \( V \) and \( W \subset V, \) \( W \) is a subspace of \( V \) if and only if

1. \( \theta \in W, \) and
2. \( \forall x, y \in W, \ x + y \in W, \) and
3. \( \forall x \in W \) and \( \forall a \in \mathbf{F}, \ ax \in W. \)

**Proof.** Exercise.

**Exercise 2.7.** Determine whether or not the following are subspaces:

1. diagonal matrices \( \subset \mathbb{R}^n_n \)
2. symmetric matrices \( \subset \mathbb{R}^n_n \)
3. matrices with trace \( = 0 \subset \mathbb{R}^n_n \)
4. matrices with trace \( = 1 \subset \mathbb{R}^n_n \)
5. matrices with \( \det = 0 \subset \mathbb{R}^n_n \)
6. Continuous functions \( f : \mathbb{R} \to \mathbb{R} \subset \mathcal{F}(\mathbb{R}) \)
7. differentiable functions \( f : \mathbb{R} \to \mathbb{R} \subset \mathcal{F}(\mathbb{R}) \)
8. non differentiable functions \( f : \mathbb{R} \to \mathbb{R} \subset \mathcal{F}(\mathbb{R}) \)
9. even (odd) functions $f : \mathbb{R} \to \mathbb{R} \subset \mathcal{F}(\mathbb{R})$

10. functions with $f(5) = 0 \subset \mathcal{F}(\mathbb{R})$

11. functions with $f(5) = 3 \subset \mathcal{F}(\mathbb{R})$

12. $\mathcal{P}^n(\mathbb{R})$ (polynomials of degree $\leq n$) $\subset \mathcal{P}(\mathbb{R})$

**Exercise 2.8.** Describe 4 different subspaces of $\mathbb{R}^3$.

**Exercise 2.9.** Prove true or false: If $W_1$ and $W_2$ are subspaces of $V$, then

1. $W_1 \cup W_2$ is a subspace of $V$.

2. $W_1 \cap W_2$ is a subspace of $V$. 
3 Linear Independence, Spanning Sets and Bases

Definition 3.1. Given $x_1, x_2, \ldots, x_n \in V$ and $a_1, \ldots, a_n \in \mathbb{R}$, the vector $x = a_1 x_1 + \cdots + a_n x_n$ is said to be a linear combination of $x_1, \ldots, x_n$.

Exercise 3.2.

1. Is $(1,2,3)$ a linear combination of $\{(1,0,0), (0,1,0), (3,5,0)\}$?

2. Is $(3,5,7)$ a linear combination of $\{(1,1,1), (2,1,3), (1,1,4)\}$? $[5(1,1,1) - 2(2,1,3) + 2(1,1,4)]$

3. Show that any vector in $\mathbb{R}^3$ is a linear combinations of $(1, 0, 2), (3, -1, 5), (7, 2, 4)$.
   Procedure: $a(1, 0, 2) + b(3, -1, 5) + c(7, 2, 4) = (x, y, z) \Rightarrow$

   \[
   \begin{align*}
   a + 3b + 7c &= x \\
   b + 2c &= y \\
   2a + 5b + 4c &= z
   \end{align*}
   \]

   Now show that the determinant of the $3 \times 3$ matrix is nonzero and use 1.34.

Definition 3.3. Given a nonempty subset, $S$, of $V$, the collection of all linear combinations of vectors in $S$ is called the span of $S$ and is written $\text{span}(S)$.

Theorem 3.4. $\text{span}(S)$ is a subspace of $V$ and is the smallest subspace of $V$ containing $S$. (That is, if $W$ is a subspace containing $S$, then $\text{span}(S) \subset W$.)

Proof. We use 2.6. Given $x \in S$, $0x = \theta \in S$; etc.
If $W$ is a subspace of $V$ and $S \subset W$, then all linear combinations of elements of $S$ belong to $W$. Thus $\text{span}(S) \subset W$.

Definition 3.5. A set $S$ of vectors such that $\text{span}(S) = V$ is called a spanning set for $V$. $S$ is said to generate $V$.

Exercise 3.6. Determine which of the following are spanning sets for $\mathbb{R}^3$. (Use geometry.) If not $\mathbb{R}^3$, what subspace is spanned?

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

2. $\{(1,0,0), (0,1,0), (0,0,1), (2,5,7)\}$

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

4. $\{(2,7,1), (1,5,3)\}$

5. $\{(2,5,0), (1,2,0)\}$

6. $\{(2,1,3), (6,3,9)\}$
Theorem 3.7. If $S_1 \subseteq S_2$, then $\text{span}(S_1) \subseteq \text{span}(S_2)$.

Proof. Let $y \in \text{span}(S_1)$. Then $y = a_1 x_1 + \cdots + a_n x_n$ where $x_1, \ldots, x_n \in S_1$. Since $S_1 \subseteq S_2, x_1, \ldots, x_n \in S_2$. So $y \in \text{span}(S_2)$.

Theorem 3.8. If $W$ is a subspace of $V$, then $W = \text{span}(W)$.

Proof. \[ x \in W \Rightarrow 1x \in \text{span}(W) \Rightarrow x \in \text{span}(W) \Rightarrow W \subseteq \text{span}(W). \]
\[ \square \]
\[ x \in \text{span}(W) \Rightarrow x = a_1 x_1 + \cdots + a_n x_n \text{ where } x_1, \ldots, x_n \in W. \text{ But then } x \in W, \]
since $W$ is a subspace. (Why?)

Corollary 3.9. $\text{span}(S) = \text{span}(\text{span}(S))$.

Proof. By 3.4, $\text{span}(S)$ is a subspace. So by 3.8 we get the result.

Exercise 3.10. Show $\text{span}(S_1 \cap S_2) \subseteq \text{span}(S_1) \cap \text{span}(S_2).$ ($S_i \neq \emptyset$.) Give examples of equality and strict containment.

Proof. Let $x \in \text{span}(S_1 \cap S_2)$. Then $x = a_1 x_1 + \cdots + a_n x_n$ where $x_i \in S_1 \cap S_2 \forall i \in \{1, \ldots, n\}$. Since $x_i \in S_1 \forall i \in \{1, \ldots, n\}, x \in \text{span}(S_1)$. Since $x_i \in S_2 \forall i \in \{1, \ldots, n\}, x \in \text{span}(S_2)$. Thus, $x \in \text{span}(S_1) \cap \text{span}(S_2)$.

- $S_1 = \{(1,0,0), (0,1,0)\}, S_2 = \{(1,0,0), (0,2,0)\}$.
- $S_1 = \{(1,0,0), (0,1,0)\}, S_2 = \{(1,0,0), (0,0,1)\}$.

Exercise 3.11. Fill in the blank with $=$, $\subseteq$, or $\supseteq$: $\text{span}(S_1 \cup S_2) \ldots \text{span}(S_1) \cup \text{span}(S_2)$. Prove your conclusion.

Definition 3.12. A subset $S$ of $V$ is linearly dependent if there exists a finite number of distinct vectors $x_1, \ldots, x_n \in S$ and scalars $a_1, \ldots, a_n \in \mathbb{F}$ not all zero such $a_1 x_1 + \cdots + a_n x_n = \theta$. If $S$ is not linearly dependent, then $S$ is linearly independent. That is, $S$ is linearly independent if given any finite number of distinct vectors $x_1, \ldots, x_n \subseteq S$, $a_1 x_1 + \cdots + a_n x_n = \theta \Rightarrow a_1, \ldots, a_n = 0$.

Exercise 3.13. Determine whether the following sets are linearly dependent or linearly independent.

1. $\{(1,0,0), (0,1,0), (5,2,0)\}$
2. $\{(1,0,0), (0,1,0), (5,2,0) \text{ and anything else in } \mathbb{R}^3 \}$
3. $\{(0,0,0), (5,7,9), (8,7,12)\}$
4. $\{(1,0,0), (0,1,0), (0,0,1), (2,5,7)\}$
5. $\{(2,1,3), (6,3,9)\}$
6. $\{(1,0,2), (3, -1, 5), (7,2,4)\}$ (from 3.2.3)
Theorem 3.14. $S$ is linearly dependent iff $\exists x \in S$ s.t. $x \in \text{span}(S \setminus \{x\})$.

Proof. $\Rightarrow$ Exercise.
$\Leftarrow$ Exercise.

Theorem 3.15. Let $S_1 \subset S_2 \subset V$. If $S_1$ is linearly dependent, then $S_2$ is linearly dependent.

Proof. Exercise.

Corollary 3.16. Let $S_1 \subset S_2 \subset V$. If $S_2$ is linearly independent, then $S_1$ is linearly independent.

When you go on a backpacking trip, you want to take all you need - but no more. We now apply that common desire to spanning sets of vector spaces.

Exercise 3.17. Consider $\mathbb{R}^3$ or $\mathbb{R}^n$. Find a set, $S$, of vectors - as small as possible - s.t. $V = \text{span}(S)$. What do you notice about linear independence?

Definition 3.18. $\beta \subset V$ is called a basis of $V$ if

1. $\beta$ is a linearly independent set, and
2. $V = \text{span}(\beta)$.

Exercise 3.19. Find bases for

1. $\mathbb{R}^5$
2. $\mathbb{R}^2$
3. $\mathcal{P}^n(X)$
4. $\mathcal{P}(X)$.

Theorem 3.20. $\beta = \{x_1, \ldots, x_n\}$ is a basis for $V$ iff each element of $V$ can be uniquely expressed as a linear combination of elements of $\beta$.

Proof. $\Rightarrow$ Since $\beta$ is a basis, it spans (generates) $V$. Thus given $y \in V$, $\exists a_1, \ldots, a_n \in \mathbb{F}$ s.t. $y = a_1 x_1 + \cdots + a_n x_n$. If we can also write $y = b_1 x_1 + \cdots + b_n x_n$, then $0 = y - y = (a_1 - b_1)x_1 + \cdots + (a_n - b_n)x_n$. Since $\beta$ is linearly independent, $a_i - b_i = 0$, so $a_i = b_i \forall i = 1, \ldots, n$.

$\Leftarrow$ Since each vector in $V$ is a linear combination of elements of $\beta$, $V = \text{span}(\beta)$. Now suppose $a_1 x_1 + \cdots + a_n x_n = \theta$. We know that $0 x_1 + \cdots + 0 x_n = \theta$, so since $\theta$ is uniquely expressed, $a_i = 0 \forall i = 1, \ldots, n$. Thus $\beta$ is linearly independent.

Notice that $\{(1,0,0),(0,1,0),(0,0,1)\}$ and $\{(1,0,2),(3,1,5),(7,2,4)\}$ are both bases for $\mathbb{R}^3$. What do you notice? We now prove your observation (3.24).
Theorem 3.21. Let $V$ have a basis with $n$ elements. Let $S = \{y_1, \ldots, y_m\}$ be a linearly independent subset of $V$ where $m \leq n$. Then there exists a subset $S_1$ of $\beta$ containing exactly $n - m$ elements s.t. $S \cup S_1$ generates $V$.

Proof. By induction on $m$:

Step 1: The result holds for $m = 0$; i.e., $S = \emptyset$.

Now assume the statement is true for some $m < n$. We must show it is true for $m + 1$. So let $S = \{y_1, \ldots, y_m, y_{m+1}\}$ be a linearly independent subset of $V$. Then $\{y_1, \ldots, y_m\}$ is linearly independent by 3.16, so since the theorem is true in this case we know there exists a subset $\{x_1, \ldots, x_{n-m}\}$ of $\beta$ s.t. $\{y_1, \ldots, y_m\} \cup \{x_1, \ldots, x_{n-m}\}$ generates $V$. Hence $\exists a_1, \ldots, a_m, b_1, \ldots, b_{n-m}$ s.t. $y_{m+1} = a_1 y_1 + \ldots + a_m y_m + b_1 x_1 + \ldots + b_{n-m} x_{n-m}$.

Now at least one of the $b_i \neq 0$ (since $\{y_1, \ldots, y_{m+1}\}$ is a linearly independent set). Without loss of generality, assume $b_1 \neq 0$. Then $x_1$ can be written as a linear combination of $Q = \{y_1, \ldots, y_{m+1}, x_2, \ldots, x_{n-m}\}$. In particular $x_1 = \frac{a_1}{b_1} y_1 + \ldots + \frac{a_m}{b_1} y_m + \frac{b_1}{b_{n-m}} x_{n-m}$.

[This is the crux of the proof: putting in a $y_{m+1}$ and pulling out a $x_1$.] Hence $x_1 \in \text{span}(Q)$. Since $y_1, \ldots, y_m, x_2, \ldots, x_{n-m} \in \text{span}(Q)$, we have $\{y_1, \ldots, y_m, x_1, \ldots, x_{n-m}\} \subset \text{span}(Q)$. Thus $\text{span}(\{y_1, \ldots, y_m, x_1, \ldots, x_{n-m}\}) = V$, $\text{span}(Q) = V$. Thus the result holds for $m + 1$.

Corollary 3.22. Suppose $V$ has a basis $\beta$ with exactly $n$ elements. Then any linearly independent subset of $V$ containing exactly $n$ elements is a basis for $V$.

Proof. Apply 3.21 to the case when $S = \{y_1, \ldots, y_n\}$ has $n$ elements. Then $S_1$ has no elements, so $S$ generates $V$. Since $S$ is a linearly independent set, $S$ is a basis for $V$.

Corollary 3.23. If $V$ has a basis with exactly $n$ elements, then any subset of $V$ with more than $n$ elements is linearly dependent.

Proof. (By Contradiction.) Suppose $S_1 = \{x_1, \ldots, x_n, x_{n+1}, \ldots, x_m\}$ is a linearly independent set with $m > n$. Since $\{x_1, \ldots, x_n\}$ is linearly independent by 3.16, it is a basis by 3.22. Thus $x_m \in \text{span}(\{x_1, \ldots, x_n\})$. But then $\{x_1, \ldots, x_n, x_m\}$ is linearly dependent by 3.14 $\Rightarrow \{x_1, \ldots, x_n, x_m\}$ is linearly independent by 3.16.

Corollary 3.24. If $V$ has a basis $\beta$ with exactly $n$ elements, then every basis for $V$ has exactly $n$ elements.

Proof. Let $\beta_1$ be a basis for $V$. Then $\beta_1$ can have no more than $n$ elements by 3.23 since $\beta_1$ must be linearly independent. Also, if $\beta_1$ had fewer than $n$ elements, then by 3.23 $\beta_1$ would not be a basis. Hence $\beta_1$ has exactly $n$ elements.

Definition 3.25. If $V$ has a basis with exactly $n$ elements, then we say $V$ has dimension $n$ and we write $\text{dim}(V) = n$. Since $n$ is a finite number, we also say $V$ is finite-dimensional. $V$ is called infinite-dimensional if it is not finite dimensional. The subspace of $V$ consisting of the identity alone is finite dimensional and we define $\text{dim}(\emptyset) = 0$.

Theorem 3.26. If $V$ is generated by $S = \{x_1, \ldots, x_n\}$ then a subset of $S$, is a basis for $V$.

Proof. Exercise.
Theorem 3.27. Let $\beta = \{x_1, \ldots, x_n\}$ be a basis for $V$ and let $S = \{y_1, \ldots, y_m\}$ be a linearly independent subset of $V$. Then there exists a subset of $\beta$, $S_1$, s.t. $S \cup S_1$ is a basis for $V$.

Proof. By 3.23, $m \leq n$. So by 3.21 $\exists S_1 \subset \beta$ with $n - m$ elements s.t. $S \cup S_1$ generates $V$. Thus by 3.26, a subset of $S \cup S_1$ is a basis for $V$. But $S \cup S_1$ has $n$ elements, so by 3.24 all of $S \cup S_1$ is a basis for $V$.

Note 3.28. Summary: If $\beta = \{x_1, \ldots, x_n\}$ is a basis for $V$, then every basis for $V$ has exactly $n$ elements and we say $\dim(V) = n$. A set, $S_1$, with more than $n$ elements must be linearly dependent. If $\text{span}(S_1) = V$, $S_1$ can be reduced to a basis. A set $S_2$ with fewer than $n$ elements cannot span $V$. If $S_2$ is linearly independent, it can be extended to form a basis for $V$.

Exercise 3.29. What is the dimension of $\mathbb{R}^n$, $\mathbb{R}^m$, $\mathcal{P}(X)$, $\mathcal{P}(X)$? (Find a basis for each.)

Exercise 3.30. Find all subspaces of dimension 1, 2 and 3 of $\mathbb{R}^3$. Characterize them geometrically.

Exercise 3.31. Suppose for a backpacker a necessary and sufficient pack weighs 50 lbs. What can you conclude if the pack is 40 pounds? 70 pounds? If $\dim(V) = n$ choose the correct response for each statement.

1. Any basis for $V$ has (at least, exactly, no more than) $n$ elements.
2. A set with less than $n$ elements (must be, maybe, cannot be) linearly independent.
3. A set with less than $n$ elements (must, may, cannot) span $V$.
4. A set with more than $n$ elements (must be, maybe, cannot be) linearly independent.
5. A set with more than $n$ elements (must, may, cannot) span $V$.
6. A set with exactly $n$ elements (must be, maybe, cannot be) linearly independent.
7. A set with exactly $n$ elements (must, may, cannot) span $V$.
8. If $S \subset T$ and $\text{span}(S) = V$, then $\text{span}(T)$ (must, may, cannot) be $V$.
9. If $S \subset T$ and $S$ is linearly independent, then $T$ (must be, maybe, cannot be) linearly independent.
10. If $S$ is linearly independent and spans $V$, then $S$ has (at least, exactly, no more than) $n$ elements.
11. If $S$ spans $V$, then $S$ has (at least, exactly, no more than) $n$ elements.
12. If $S$ is linearly independent then $S$ has (at least, exactly, no more than) $n$ elements.
13. T or F: Every set with more than $n$ elements spans $V$.
14. T or F: Every set with less than $n$ elements is linearly independent.

Exercise 3.32. The rows of $A \in \mathbb{F}^m_n$ are elements of $\mathbb{F}^m_n$. Explain why the rows of $A$ are linearly independent iff $\text{rank}(A) = m$. 

25
4 Linear Transformations (Preserve Structure of Vector Spaces)

Definition 4.1. A function $T: V \rightarrow W$ ($V, W$ vector spaces) is called a linear transformation if

1. $T(x + y) = T(x) + T(y)$, and
2. $T(ax) = aT(x)$.

Theorem 4.2. $T(0) = 0$. (i.e., $T(\theta_V) = \theta_W$)

Proof. $T(0) = T(0 + 0) = T(0) + T(0) \Rightarrow T(0) = 0$.

Exercise 4.3. Determine whether each of the following is a linear transformation.

1. $V = W = C^\infty; T(f) = f'$
2. $V = W = C^\infty; T(f) = f'''$
3. $V = C(\mathbb{R}), W = \mathbb{R}; T(f) = \int_1^3 f(x)dx$
4. $T_x: \mathbb{R}^3 \rightarrow \mathbb{R}^3; T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$. (projection)
5. $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3; T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 3 \\ 4 & 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ Recall $Tx = Ax$ is linear!
6. $T_\theta: \mathbb{R}^2 \rightarrow \mathbb{R}^2$: $T$ rotates vectors ccw through the angle $\theta$

Proof. Find the matrix representation for $T_\theta$:

\[
x_1 = r \cos \alpha \\
x_2 = r \sin \alpha \\
y_1 = r \cos(\alpha + \theta) = r \cos \alpha \cos \theta - r \sin \alpha \sin \theta = x_1 \cos \theta - x_2 \sin \theta \\
y_2 = r \sin(\alpha + \theta) = r \sin \alpha \cos \theta + r \cos \alpha \sin \theta = x_2 \cos \theta + x_1 \sin \theta
\]

So $T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 \cos \theta - x_2 \sin \theta, x_2 \cos \theta + x_1 \sin \theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Thus $T$ is a linear transformation by 5.

Definition 4.4. The nullspace or kernel of $T: V \rightarrow W$ is the set $\ker(T) = \{x \in V : T(x) = 0\}$. The image or range of $T$ is the set $\text{ran}(T) = \{y \in W : \exists x \in V \text{ s.t. } Tx = y\}$.

Exercise 4.5. What are the kernel and range of $T_x: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ (4.3.4)? What do you notice?
Theorem 4.6. \( \ker(T) \) and \( \operatorname{ran}(T) \) are subspaces of \( V \) and \( W \) respectively.

**Proof.** 1. Let \( x, y \in \ker(T) \). Then \( T(x + y) = T(x) + T(y) = T(x) + T(y) = 0 + 0 = 0 \), so \( x + y \in \ker(T) \). What else must be checked?

2. If \( y_1, y_2 \in \operatorname{ran}(T) \), then \( \exists x_1, x_2 \in V \) s.t. \( T(x_1) = y_1, T(x_2) = y_2 \). Then \( T(x_1 + x_2) = T(x_1) + T(x_2) = y_1 + y_2 \), so \( y_1 + y_2 \in \operatorname{ran}(T) \) etc.

**Exercise 4.7.** Notice in 4.5 \( \dim(\mathbb{R}^3) = 3, \ \null(T_\theta) = 1, \ \operatorname{rank}(T_\theta) = 2 \). For \( T_\phi, \ \dim(\mathbb{R}^2) = 2, \ \null(T_\phi) = 0, \ \operatorname{rank}(T_\phi) = 2 \). For \( T : \mathcal{P}^7(X) \to \mathcal{P}^7(X) \) given by \( T(f) = f''' \), \( \dim(\mathcal{P}^7) = 8, \ \null(T) = 3, \ \operatorname{rank}(T) = 5 \).

What do you notice?

**Exercise 4.8.**

1. If \( \ker(T) \) contains more than one element, how many elements does it contain?

2. If \( \ker(T) \neq \{ \theta \} \), show that \( T \) sends an infinite number of elements in the domain to each element in \( \operatorname{ran}(T) \). Given \( y_0 \neq \theta \), does \( X = \{ x \in V : T(x) = y_0 \} \) have any particular structure? Keep reading.

**Definition 4.9.** A subset \( W \) of a vector space \( V \) is an **affine subspace** if there exists a vector \( v \in V \) and a subspace \( Y \) of \( V \) such that \( W = v + Y \), that is, every vector in \( W \) is the sum of the vector \( v \) and some vector in \( Y \). Note that an affine subspace is **not** in general a subspace.

We have already encountered the concept of affine subspace in 1.22. In that theorem, the solution set \( W \) of a system of equations is an affine subspace of \( \mathbb{F}^n \); \( W = x_p + \ker(A) \), where \( x_p \) is a particular solution of the system \( Ax = b \) and \( \ker(A) \) is the subspace of solutions to the homogeneous equation \( Ax = 0 \). This is a particular case of a more general result.

**Theorem 4.10.** Let \( T : V \to W \) be a linear transformation. The solution set of the equation \( T(x) = y \) is an affine subspace.

**Proof.** Let \( X \) be the solution set of \( T(x) = y \). Let \( x_p \) be a particular solution so \( T(x_p) = y \). Now if \( x \in X \), then \( T(x - x_p) = T(x) - T(x_p) = y - y = 0 \), so \( x - x_p \in \ker(T) \Rightarrow x \in x_p + \ker(T) \). Conversely, if \( x = x_p + x_h \) where \( x_h \in \ker(T) \), then \( T(x) = T(x_p + x_h) = T(x_p) + T(x_h) = y + 0 = y \). Hence the solution set of \( T(x) = y \) is the affine subspace \( x_p + \ker(T) \).

**Definition 4.11.** The **nullity** of \( T \), denoted \( \null(T) \), is the \( \dim(\ker(T)) \). The **rank** of \( T \), denoted \( \operatorname{rank}(T) \), is the \( \dim(\operatorname{ran}(T)) \).
Theorem 4.12. Given $T : V \to W$ where $\beta = \{x_1, \ldots, x_n\}$ is a basis for $V$, $\text{ran}(T) = \text{span}\{T(x_1), \ldots, T(x_n)\}$.

Proof. Since $\{T(x_1), \ldots, T(x_n)\} \subset \text{ran}(T)$, $\text{span}\{T(x_1), \ldots, T(x_n)\} \subset \text{span}(\text{ran}(T))$ by 3.7. Since $\text{ran}(T)$ is a subspace, $\text{span}(\text{ran}(T)) = \text{ran}(T)$ by 3.8. Thus $\text{span}\{T(x_1), \ldots, T(x_n)\} = \text{ran}(T)$.

Suppose $y \in \text{ran}(T)$. Then $\exists x \in V$ s.t. $y = T(x)$.

Since $\beta$ is a basis for $V$, $x = a_1 x_1 + \cdots + a_n x_n$, so $y = T(x) = T(a_1 x_1 + \cdots + a_n x_n) = a_1 T(x_1) + \cdots + a_n T(x_n) \Rightarrow y \in \text{span}\{T(x_1), \ldots, T(x_n)\}$.

Theorem 4.13. Given $T : V \to W$ where $V$ is finite dimensional, $\text{dim}(V) = \text{null}(T) + \text{rank}(T)$.

Proof. Suppose $\text{dim}(V) = n$ and $\{x_1, \ldots, x_k\}$ is a basis for $\text{ker}(T)$. By 3.27 we can extend $\{x_1, \ldots, x_k\}$ to $\beta = \{x_1, \ldots, x_k, x_{k+1}, \ldots, x_n\}$ which is a basis for $V$. We now show that $\{T(x_{k+1}), \ldots, T(x_n)\}$ is a basis for $\text{ran}(T)$.

First of all, by 4.12 $\text{ran}(T) = \text{span}\{T(x_1), \ldots, T(x_n)\} = \text{span}\{T(x_{k+1}), \ldots, T(x_n)\}$ since $T(x_1) = T(x_2) = \cdots = T(x_k) = \theta$.

Secondly, we must show the set $\{T(x_{k+1}), \ldots, T(x_n)\}$ is linearly independent. Suppose $b_{k+1} T(x_{k+1}) + \cdots + b_n T(x_n) = 0$. Then $T(b_{k+1} x_{k+1} + \cdots + b_n x_n) = 0$ so $b_{k+1} x_{k+1} + \cdots + b_n x_n \in \text{ker}(T)$. Since $\{x_1, \ldots, x_k\}$ is a basis for $\text{ker}(T)$, $\exists c_1, \ldots, c_k$ s.t. $c_1 x_1 + \cdots + c_k x_k = b_{k+1} x_{k+1} + \cdots + b_n x_n$. Hence $-c_1 x_1 - \cdots - c_k x_k + b_{k+1} x_{k+1} + \cdots + b_n x_n = \theta$. Since $\{x_1, \ldots, x_n\}$ is a basis for $V$, all the coefficients (in particular, all the $b_i$'s) must be zero.

Question 4.14. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be given by $T(x) = Ax$ where $A \in \mathbb{F}_n^m$. Then is $\text{rank}(T) = \text{rank}(A)$ and $\text{null}(T) = \text{null}(A)$? (Recall 1.15 and 1.17.) We now proceed to answer this question 4.18.

Theorem 4.15. If $\tilde{A}$ is gained from $A$ by elementary row operations, then $\text{rank}(\tilde{A}) = \text{rank}(A)$.

Proof. Thinking of the rows of $A$ and $\tilde{A}$ as vectors, e.r.o.s produce linear combinations of the original rows. Also, inverses of e.r.o.s are also e.r.o.s. So the rows of $\tilde{A}$ are linear combinations of the rows of $A$ and the rows of $A$ are linear combinations of the rows of $\tilde{A}$. Thus they span the same space called the row-space of $A$ which is the space generated by the (linearly independent) non-zero rows of the rref of $A$.

Theorem 4.16. Given $T : \mathbb{F}^n \to \mathbb{F}^m$ by $Tx = Ax$, $\text{ran}(T) = \text{span}(\text{columns of } A)$.

Proof. Let $e_i = (0, \ldots, 0, 1_{i \text{th}}, 0, \ldots, 0)$, so $\{e_i\}_{i=1,\ldots,n}$ is a basis for $\mathbb{F}^n$. Then note that the $i$th column of $A$ is just $Ae_i$. Thus by 4.12 $\text{ran}(T) = \text{span}\{T(e_1), \ldots, T(e_n)\} = \text{span}\{\text{columns of } A\}$, which is called the column-space of $A$. 

28
Theorem 4.17. \( \text{rank}(A) = \text{rank}(A^t) \) where \( A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \).

Proof. Let \( R_1, R_2, \ldots, R_m \) denote the rows of \( A \) where \( R_i = (a_{1i}, a_{2i}, \ldots, a_{ni}) \). Suppose the rank of \( A \) is \( r \) and the following \( r \) vectors form a basis for the row-space: \( S_1 = (b_{11}, b_{12}, \ldots, b_{1n}) \), \( S_2 = (b_{21}, \ldots, b_{2n}) \), \( \ldots, S_r = (b_{r1}, \ldots, b_{rn}) \). Then each of the row vectors is a linear combination of the \( S_i \):

\[
R_1 = k_{11}S_1 + \cdots + k_{1r}S_r \\
\vdots \\
R_m = k_{m1}S_1 + \cdots + k_{mr}S_r.
\]

Setting the \( i \)-th components equal to each other, we get for each \( i = 1, \ldots, n \):

\[
\begin{align*}
a_{1i} &= k_{11}b_{1i} + k_{12}b_{2i} + \cdots + k_{1r}b_{ri} \\
a_{2i} &= k_{21}b_{1i} + k_{22}b_{2i} + \cdots + k_{2r}b_{ri} \\
\vdots &= \vdots \\
a_{mi} &= k_{m1}b_{1i} + k_{m2}b_{2i} + \cdots + k_{mr}b_{ri}.
\end{align*}
\]

Thus for \( i = 1, \ldots, n \):

\[
\begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} = b_{1i} \begin{pmatrix} k_{11} \\ k_{21} \\ \vdots \\ k_{m1} \end{pmatrix} + b_{2i} \begin{pmatrix} k_{12} \\ k_{22} \\ \vdots \\ k_{m2} \end{pmatrix} + \cdots + b_{ri} \begin{pmatrix} k_{1r} \\ k_{2r} \\ \vdots \\ k_{mr} \end{pmatrix}.
\]

That is, each of the columns of \( A \) is a linear combination of the \( r \) vectors, so \( \text{rank}(A^t) = \dim(\text{column-space of } A) \leq r = \text{rank}(A) \). Reversing the argument, we get \( \text{rank}(A) \leq \text{rank}(A^t) \), so \( \text{rank}(A) = \text{rank}(A^t) \).

Theorem 4.18. If \( T : \mathbb{R}^n \to \mathbb{R}^m \) is given by \( Tx = Ax \), then \( \text{rank}(T) = \text{rank}(A) \).

Proof. \( \text{rank}(T) = \dim(\text{ran}(T)) = \dim(\text{span(columns of } A)) = \dim(\text{span(rows of } A^t)) = \text{number of linearly independent rows of } A^t = \text{rank}(A^t) = \text{rank}(A) \).

Question 4.19. Is it possible to find a matrix with exactly 2 linearly independent rows and 3 linearly independent columns?

Definition 4.20. A function \( f : A \to B \) is said to be

1. surjective (onto) if for each \( b \in B, \exists a \in A \) s.t. \( f(a) = b \);
2. injective (1-1) if \( a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2) \); (What’s the contrapositive?)
3. bijective if \( f \) is both injective and surjective.
4. Furthermore, \( I : A \to A \) given by \( I(x) = x \) is called the identity function.
5. If \( \exists g : B \to A \) s.t. \( g \circ f = I_A \) and \( f \circ g = I_B \), \( f \) is said to be invertible, and \( g \) is called an inverse of \( f \).

The results in the following theorem concerning general functions should already be familiar to the reader.

**Theorem 4.21.** Let \( f \) and \( g \) be functions.

1. If \( f \) is invertible, the inverse of \( f \) is unique (and is denoted \( f^{-1} \)).
2. \( (f \circ g)^{-1} = g^{-1} \circ f^{-1} \).
3. \( (f^{-1})^{-1} = f \).
4. \( f \) is invertible iff \( f \) is bijective.

**Proof.** Exercise. Use a functional diagram to motivate.

**Theorem 4.22.** \( T : V \to W \) is injective iff \( \ker(T) = \{0\} \).

**Proof.** \( \implies \) Suppose \( T(x) = 0 \). Then \( T(-x) = -T(x) = 0 \). Since \( T \) is injective, \( x = -x \Rightarrow 2x = 0 \Rightarrow x = 0 \).

\( \Leftarrow \) Suppose \( T(x) = T(y) \). Then \( T(x) - T(y) = 0 \), so \( T(x - y) = 0 \). Since \( \ker(T) = 0 \), \( x - y = 0 \), so \( x = y \).

**Theorem 4.23.** Given \( T : V \to W \) where \( \dim(V) = n = \dim(W) \), then \( T \) is injective iff \( T \) is surjective (\( \iff T \) is bijective \( \iff T \) is invertible).

**Proof.** \( \iff \) \( T \) injective \( \iff \) \( \ker(T) = \{0\} \) \( \iff \) \( \text{null}(T) = 0 \) \( \iff \) \( \text{rank}(T) = n \) \( \iff \) \( \text{ran}(T) = W \) \( \iff \) \( T \) is surjective.

**Theorem 4.24.** Let \( U, T : V \to W \) and \( \beta = \{x_1, \ldots, x_n\} \) be a basis for \( V \). If \( U(x_i) = T(x_i) \forall i = 1, \ldots, n \), then \( U = T \).

**Proof.** Given \( x \in V, x = a_1x_1 + \cdots + a_nx_n \). \( T(x) = \cdots \), etc.

**Theorem 4.25.** If \( T : V \to W \) is bijective and \( \beta = \{x_1, \ldots, x_n\} \) is a basis for \( V \), then \( \{T(x_1), \ldots, T(x_n)\} \) is a basis for \( W \).

**Proof.** First, by 4.13 \( \dim(W) = n \). Suppose \( a_1T(x_1) + \cdots + a_nT(x_n) = 0 \). Then \( T(a_1x_1 + \cdots + a_nx_n) = 0 \). Since \( T \) is injective, \( a_1x_1 + \cdots + a_nx_n = 0 \) by 4.22. Since \( \{x_1, \ldots, x_n\} \) is a basis for \( V \) (linearly independent), \( a_i = 0 \forall i = 1, \ldots, n \). Thus \( \{T(x_1), \ldots, T(x_n)\} \) is linearly independent. By 3.22 it is a basis for \( W \).

**Theorem 4.26.** Let \( T_1, T_2 : V \to W \) and \( U : W \to Z \) be linear transformations. Then

1. \( (T_1 + T_2): V \to W \) defined by \( (T_1 + T_2)(x) = T_1(x) + T_2(x) \) is a linear transformation.
2. \( (UT) : V \to Z \) defined by \( (UT)(x) = U(T(x)) \) is a linear transformation.
3. If \( T \) is invertible, \( T^{-1} \) is a linear transformation.
Proof. (of 3) Let $y_1, y_2 \in W$. Since $T$ is a surjection, $\exists x_1, x_2 \in V$ s.t. $T(x_1) = y_1$, $T(x_2) = y_2$. Since $T$ is an injection, $x_1$ and $x_2$ are unique. Thus $x_1 = T^{-1}(y_1)$ and $x_2 = T^{-1}(y_2)$. Hence $T^{-1}(y_1 + y_2) = T^{-1}(T(x_1) + T(x_2)) = T^{-1}(T(x_1 + x_2)) = x_1 + x_2 = T^{-1}(y_1) + T^{-1}(y_2)$. Also, $T^{-1}(c y_1) = T^{-1}(c T(x_1)) = T^{-1}(c x_1) = c x_1 = c T^{-1}(y_1)$. 1 and 2 are left as exercises.

Definition 4.27. $V$ is said to be isomorphic to $W$ if there exists a bijective linear transformation $T : V \to W$. ($T$ is called an isomorphism.)

Theorem 4.28. Let $\sim$ stand for “is isomorphic to”. Then

1. $V \sim V$.
2. If $V \sim W$, then $W \sim V$.
3. If $V \sim W$ and $W \sim Z$, then $V \sim Z$.

Proof. Exercise.

Definition 4.29. Any mathematical “verb” which satisfies the three conditions of 4.28 is said to be an equivalence relation.

Theorem 4.30. Let $V, W$ be finite dimensional vector spaces. Then $V$ is isomorphic to $W$ iff $\dim(V) = \dim(W)$.

Proof. $\Rightarrow$ $V \sim W$ $\Rightarrow$ $\exists$ a bijective linear transformation $T : V \to W$. Now $\dim(V) = \dim(\ker(T)) + \dim(\text{ran}(T)) \Rightarrow \dim(V) = 0 + \dim(W)$.

$\Leftarrow$ Let $\beta = \{x_1, \ldots, x_n\}$ be a basis for $V$ and $\{y_1, \ldots, y_n\}$ a basis for $W$.

Define $T : V \to W$ by $T(x_i) = y_i$. $T$ is well defined by 4.24. $T$ is surjective since given $y = \alpha_1 y_1 + \cdots + \alpha_n y_n$, $T(\alpha_1 x_1 + \cdots + \alpha_n x_n) = y$. Now suppose $T(x) = 0$. Write $x = a_1 x_1 + \cdots + a_n x_n$. Then $0 = T(x) = T(a_1 x_1 + \cdots + a_n x_n) = a_1 T(x_1) + \cdots + a_n T(x_n) = a_1 y_1 + \cdots + a_n y_n$. Since $\{y_1, \ldots, y_n\}$ is a basis for $W$, $a_i = 0 \ \forall i = 1, \ldots, n$. Thus $x = 0$, so $\ker(T) = 0$, so $T$ is injective by 4.22. Thus $T : V \to W$ is a bijection.
5 Matrix Representations of Linear Transformations

Definition 5.1. Given the vector space $V$ with basis $\beta = \{x_1, \ldots, x_k\}$ and given $x = a_1x_1 + \cdots + a_kx_k$, we write $[x]_\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}$. $\beta$ is called an ordered basis. (Since the order of elements is important) and $[x]_\beta$ is called the vector representation of $x$ relative to $\beta$.

Definition 5.2. Given $T : V \to W$ where $\beta = \{x_1, \ldots, x_m\}$ is an ordered basis for $V$ and $\gamma = \{y_1, \ldots, y_n\}$ is an ordered basis for $W$, we can write:

$$T(x_1) = a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n$$
$$T(x_2) = a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n$$
$$\vdots$$
$$T(x_m) = a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n.$$ 

Then $[T]_\beta^\gamma = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & \cdots & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & \cdots & \cdots & a_{mn} \end{pmatrix}$ (Notice rows $\to$ columns, columns $\to$ rows) is called the matrix representation of $T$ relative to $\beta$ and $\gamma$.

The following theorem says that a linear transformation can be expressed as a matrix with corresponding properties. ($V_{(\beta)}$ denotes the vector space $V$ with ordered basis $\beta$.)

Theorem 5.3. Let $T_1, T_2 : V_{(\beta)} \to W_{(\gamma)}$ and $U : W_{(\gamma)} \to Z_{(\delta)}$. Then

1. $[T(x)]_\gamma = [T]_\beta^\gamma[x]_\beta$;
2. $[T_1 + T_2]_\gamma^\gamma = [T_1]_\beta^\gamma + [T_2]_\beta^\gamma$;
3. $[kT]_\beta^\gamma = k[T]_\beta^\gamma$;
4. $[UT]_\beta^\gamma = [U]_\delta^\gamma[T]_\beta^\gamma$;
5. $T$ is invertible iff $[T]_\beta^\gamma$ is invertible and $(|[T]_\beta^\gamma|)^{-1} = [T^{-1}]_\gamma^\gamma$.

Proof. All the proofs are similar - we just prove 1. Notice that both $[T(x)]_\gamma$ and $[T]_\beta^\gamma[x]_\beta$ are $n \times 1$ vectors. On the LHS, $[T(x)]_\gamma = [T(b_1x_1 + \cdots + b_mx_m)]_\gamma = [b_1T(x_1) + \cdots + b_mT(x_m)]_\gamma$. So the entry in the $k$th place will be $b_1a_{1k} + b_2a_{2k} + \cdots + b_ma_{mk}$.

On the RHS we have the product: $(k$th row) $\to \begin{pmatrix} a_{1k} & a_{2k} & \cdots & a_{mk} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$. So the $k$th place will be $b_1a_{1k} + \cdots + b_ma_{mk}$.
Exercise 5.4. Let $T : \mathcal{P}^4(\mathbb{R}) \to \mathcal{P}^2(\mathbb{R})$ be given by $T(f) = f''$.

Take $\beta = \{1, x, x^2, x^3, x^4\}$, $\gamma = \{1, x, x^2\}$.

\[
T(1) = 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2
\]

\[
T(x) = 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2
\]

So,

\[
T(x^2) = 2 = 2 \cdot 1 + 0 \cdot x + 0 \cdot x^2
\]

\[
T(x^3) = 6x = 0 \cdot 1 + 6 \cdot x + 0 \cdot x^2
\]

\[
T(x^4) = 12x^2 = 0 \cdot 1 + 0 \cdot x + 12 \cdot x^2
\]

So $[T]_{\beta}^{\gamma} = \begin{pmatrix} 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 12 \end{pmatrix}$.

Given $f = 2x^4 - 5x^3 + 6x + 8$, find $T(f)$.

\[
[f]_{\beta} = \begin{pmatrix} 8 \\ 6 \\ 0 \\ -5 \\ 2 \end{pmatrix}, \quad \text{so} \quad [T]_{\beta}^{\gamma}[f]_{\beta} = \begin{pmatrix} 0 & 30 \\ 24 \end{pmatrix} = [T(f)]_{\gamma}.
\]

Thus, $T(f) = 0 \cdot 1 - 30x + 24x^2$.

Exercise 5.5. Repeat the previous exercise with the same $T$ and $f = 5x^4 - 7x^3 + 13x^2 + 41x + 17$.

One use of 5.3 is to write vectors in terms of a new basis. That is, given a vector space $V$ of dimension $n$, we consider $I : V(\beta) \to V(\gamma)$ and find the matrix $[I]_{\beta}^{\gamma}$. This is called a transition matrix.

Exercise 5.6. Consider $\mathbb{R}^3$ with two different bases: $\beta = \{(1, 1, 1), (1, 2, 0), (0, 1, 0)\}$ and $\gamma = \{(1, 0, 0), (0, 1, 1), (0, 0, 2)\}$ (Verify that these are both bases.)

Then

\[
I(1, 1, 1) = (1, 1, 1) = 1(1, 0, 0) + 1(0, 1, 1) + 0(0, 0, 2).
\]

\[
I(1, 2, 0) = (1, 2, 0) = 1(1, 0, 0) + 2(0, 1, 1) - 1(0, 0, 2).
\]

\[
I(0, 1, 0) = (0, 1, 0) = 0(1, 0, 0) + 1(0, 1, 1) \frac{1}{2}(0, 0, 2).
\]

So, $[I]_{\beta}^{\gamma} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & \frac{1}{2} \end{pmatrix}$.

Check: Take $[x]_{\beta} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}_{\beta}$

Then

\[
[I]_{\beta}^{\gamma}[x]_{\beta} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & -1 & -\frac{1}{2} \end{pmatrix}_{\beta} \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}_{\beta} = \begin{pmatrix} 5 \\ 12 \\ -5 \end{pmatrix}_{\gamma}.
\]

Check that $\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}_{\beta} = \begin{pmatrix} 5 \\ 12 \\ 5 \end{pmatrix}_{\gamma}$. 

33
Since $\beta$ and $\gamma$ generate the same space, $[I]^\gamma_\beta$ has an inverse. (Why? By 4.12 and 4.18, $\text{rank}([I]^\gamma_\beta) = n$. Thus $[I]^\gamma_\beta$ has an inverse by 1.34). Thus we have $[x]_\beta = ([I]^\gamma_\beta)^{-1}[x]_\gamma$ \cite{5,3,5} $= [I^{-1}]^\gamma_\beta[x]_\gamma = [I]^\gamma_\beta[x]_\gamma$ (as you’d expect).

**Question 5.7.** Given $T : V \rightarrow V$ where $\beta, \gamma$ are ordered bases for $V$, what is the relationship between $[T]^\beta_\beta$ and $[T]^\gamma_\gamma$?

Let $Q = [I]^\gamma_\beta$. Then

**Theorem 5.8.** $[T]^\gamma_\gamma = Q[T]^\beta_\beta Q^{-1}$, $[T]^\gamma_\beta = Q^{-1}[T]^\gamma_\gamma Q$

**Proof.** $Q[T]^\beta_\beta = [I]^\gamma_\beta[T]^\beta_\beta = [IT]^\gamma_\beta = [T][I]^\gamma_\beta = [T]^\gamma_\gamma Q \Rightarrow [T]^\gamma_\gamma = Q[T]^\beta_\beta Q^{-1}$ or $[T]^\gamma_\beta = Q^{-1}[T]^\gamma_\gamma Q$. This motivates:

**Definition 5.9.** If $A, B \in \mathbb{F}^n$ and $\exists Q \in \mathbb{F}^n$ s.t. $B = Q^{-1}AQ$, we say that $B$ is similar to $A$, and write $B \sim A$.

**Theorem 5.10.** Similarity of matrices is an equivalence relation (See 4.28)

**Proof.** Exercise.
6 Eigenvectors

**Definition 6.1.** Given \( T : V \to V \) if \( x \neq \theta \) satisfies \( T(x) = \lambda x \) for some scalar \( \lambda \), then \( x \) is called an eigenvector of \( T \) and \( \lambda \) is called an eigenvalue of \( T \) corresponding to \( x \). An eigenpair is a pairing \((x, \lambda)\) of an eigenvalue \( \lambda \) with a corresponding eigenvector \( x \).

**Question 6.2.** Since the action of linear transformations can be expressed in terms of matrices, we also say \( x \) is an eigenvector of \( A \in \mathbb{F}_n \) if \( Ax = \lambda x \). (\( \lambda \) is the eigenvalue.) What is this saying geometrically?

**Exercise 6.3.** Use geometrical considerations to find the eigenvectors and eigenvalues of the following linear transformation.

1. \( T_{\text{Ref}} : \mathbb{R}^2 \to \mathbb{R}^2 \) defined by \( T(x, y) = (x, -y) \)
2. \( T_{\text{pro}} : \mathbb{R}^2 \to \mathbb{R}^2 \) defined by \( T(x, y) = (x, 0) \)
3. \( T_{\theta=\pi/2} : \mathbb{R}^2 \to \mathbb{R}^2 \) where \( T \) rotates vectors in \( \mathbb{R}^2 \) by \( \pi/2 \) radians
4. \( T_{\theta=\pi} : \mathbb{R}^2 \to \mathbb{R}^2 \) where \( T \) rotates vectors in \( \mathbb{R}^2 \) by \( \pi \) radians

(answers: 1. \( x\)-axis\(\setminus\{0\} \) (\( \lambda = 1 \)), \( y\)-axis\(\setminus\{0\} \) (\( \lambda = -1 \)) 2. \( x\)-axis\(\setminus\{0\} \) (\( \lambda = 1 \)), \( y\)-axis\(\setminus\{0\} \) (\( \lambda = 0 \))
3. none 4. \( \mathbb{R}^2\setminus\{0\} \) (\( \lambda = -1 \))

Do you notice anything about the set of eigenvectors corresponding to an eigenvalue?

**Theorem 6.4.** Given \( \lambda, E_\lambda = \{ x : Tx = \lambda x \} \) is a subspace of \( V \).

**Proof.** Let \( x, y, \in E_\lambda \), then \( T(x + y) = Tx + Ty = \lambda x + \lambda y = \lambda(x + y) \Rightarrow x + y \in E_\lambda \), etc.

**Definition 6.5.** The space \( E_\lambda = \{ x : Tx = \lambda x \} \) of the above theorem is called the eigenspace of \( \lambda \). Note that \( \theta \in E_\lambda \) even though \( \theta \) is not considered an eigenvector.

**Exercise 6.6.** Let \( T : \mathbb{R}^2 \to \mathbb{R}^2 \) by given by \( T(x) = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} x \).

1. Determine whether \( x_1 = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \) and \( x_2 = \begin{pmatrix} 6 \\ 8 \end{pmatrix} \) are eigenvectors.

2. Can you find another eigenvector of \( T \)?

The following theorem, though simple to prove, provides a very useful means of finding eigenvalues.

**Theorem 6.7.** Given \( A \in \mathbb{F}_n \), there exists an eigenpair \((x, \lambda)\) \( \iff \) \( \det(A - \lambda I) = 0 \).

**Proof.** \( Ax = \lambda x, (x \neq \theta) \iff Ax = \lambdaIx \iff Ax = \lambda x \iff (A - \lambda I)(x) = 0 \iff \lambda \neq \theta \iff \det(A - \lambda I) = 0 \).

\( \det(A - \lambda I) \) is called the characteristic polynomial of \( A \) (in \( \lambda \)) and \( \det(A - \lambda I) = 0 \) is the characteristic equation.

**Exercise 6.8.** Find the eigenvalues and eigenvectors for \( A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \).
\[
\det(A - \lambda I) = \begin{vmatrix} 1 - \lambda & 3 \\ 2 & -2 - \lambda \end{vmatrix} = \lambda^2 - 3\lambda - 10 = (\lambda - 5)(\lambda + 2) = 0 \implies \lambda = 5, -2.
\]

Find the eigenspace for \( \lambda = 5 \):
\[
\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 5 \begin{pmatrix} x \\ y \end{pmatrix} \implies x + 3y = 5x \implies 4x + 2y = 5y \implies 4x - 3y = 0 \implies y = t \implies x = \frac{3}{4} t
\]
\[
\implies \begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} \frac{3}{4} \\ 1 \end{pmatrix} = t \begin{pmatrix} 3 \\ 4 \end{pmatrix}.
\]

So \( E_{\lambda=5} = \{t \begin{pmatrix} 3 \\ 4 \end{pmatrix} : t \in \mathbb{R}\} \).

Find the eigenspace for \( \lambda = -2 \):
\[
\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -2 \begin{pmatrix} x \\ y \end{pmatrix} \implies 4x + 4y = 0 \implies (\frac{y}{x = -t}) \implies \begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 1 \\ 1 \end{pmatrix} \implies E_{\lambda=-2} = \{t \begin{pmatrix} 1 \\ 1 \end{pmatrix} : t \in \mathbb{R}\}.
\]

**Exercise 6.9.** Given \( T : C^\infty \to C^\infty \) defined by \( T(f) = f' \), find all eigenvalues and eigenvectors. (Each \( \lambda \in \mathbb{R} \) is an eigenvalue and \( \{f : f = Ae^{\lambda t}\} \) is the corresponding eigenspace.)

**Theorem 6.10.** \( A \in \mathbb{R}^n \) has at most \( n \) eigenvalues. \( A \in \mathbb{C}^n \) has exactly \( n \) (not necessarily distinct) eigenvalues.

**Proof.** Observe the form of the characteristic equation (and note the Fundamental Theorem of Algebra).

**Note 6.11.** The above theorem reveals the difference between working in \( \mathbb{R} \) or \( \mathbb{C} \). Although not all polynomials factor in \( \mathbb{R} \) (e.g. \( \lambda^2 + 1 \)), all polynomials factor completely in \( \mathbb{C} \). If a polynomial does factor completely, we say the polynomial splits. You will see where this is important in Thm 6.20.

**Definition 6.12.** If \( \det(A - \lambda I) = 0 \) has exactly \( n \) factors of \( (\lambda - \lambda_0) \), then we say \( \lambda_0 \) has a multiplicity of \( n \).

**Exercise 6.13.** Find the eigenvalues and eigenvectors of \( A = \begin{pmatrix} 2 & 2 & 3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix} \).

\[
0 = \det(A - \lambda I) = \lambda^3 + \lambda^2 - 21\lambda - 45 = (\lambda - 5)(\lambda + 3)(\lambda + 3).
\]

Find \( E_{\lambda=5} \):
\[
\begin{pmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} 1 & 2 & 5 \\ 2 & -4 & -6 \\ -7 & 2 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} 1 & 2 & 5 \\ 0 & 10 & 50 \\ 0 & 16 & 32 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} y + 2z = 0 \\ z = t, y = 2t \end{pmatrix} \implies x = -t.
\]

\[
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix} \text{ so, } E_{\lambda=5} = \left\{ t \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} : t \in \mathbb{R} \right\}. \quad (\dim(E_{\lambda=5}) = 1.)
\]
Find $E_{\lambda=-3}$:
\[
\begin{pmatrix}
1 & 2 & -3 \\
2 & 4 & -6 \\
-1 & -2 & 3
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
\Rightarrow
\begin{cases}
x + 2y - 3z = 0 \\
z = s \\
y = t
\end{cases}
\Rightarrow
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= \begin{pmatrix}
-2t + 3s \\
t \\
s
\end{pmatrix}
\Rightarrow
E_{\lambda=-3} = \left\{ t \begin{pmatrix}
-2 \\
1 \\
0
\end{pmatrix} + s \begin{pmatrix}
3 \\
0 \\
1
\end{pmatrix} : s, t \in \mathbb{R} \right\}.

Theorem 6.14. $1 \leq \dim(E_{\lambda}) \leq \text{mult}(\lambda)$.

Proof. Let $\{x_1, \ldots, x_p\}$ be a basis for $E_{\lambda}$. Extend it to $\beta = \{x_1, \ldots, x_p, x_{p+1}, \ldots, x_n\}$ which is a basis for $V$. Since $T(x_i) = \lambda x_i$ for each $i = 1, \ldots, p$, $[T]_{\beta} = \begin{pmatrix}
I_p & B \\
O & C - tI
\end{pmatrix}$. Let $A = [T]_{\beta}$. Then
\[\det(A - tI_n) = \det \begin{pmatrix}
(\lambda - t)I_p & B \\
O & C - tI
\end{pmatrix} = \det((\lambda - t)I_p) \det(C - tI) = (-1)^p(\lambda - t)^p g(t).\]
Thus the multiplicity of $\lambda$ is at least $p$.

Exercise 6.15. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $\det(A - \lambda I) = \lambda^2 = 0 \Rightarrow \lambda = 0$. (So $\text{mult}(\lambda) = 2$). Find $E_{\lambda=0}$:
\[
\begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 0
\end{pmatrix}
\Rightarrow
0x + y = 0 \Rightarrow
y = 0
\Rightarrow
\begin{pmatrix}
x \\
y
\end{pmatrix}
= \begin{pmatrix}
t \\
0
\end{pmatrix}
\Rightarrow
E_{\lambda=0} = \left\{ t \begin{pmatrix}
1 \\
0
\end{pmatrix} : t \in \mathbb{R} \right\} \text{ so } \dim(E_{\lambda=0}) = 1.

Exercise 6.16. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then $\det(A - \lambda I) = \lambda^2 + 1 = 0 \Rightarrow \lambda = \pm i$.

Find $E_{\lambda=i}$:
\[
\begin{pmatrix}
-i & 1 & 0 \\
-1 & -i & 0
\end{pmatrix}
\Rightarrow
-ix + y = 0 \Rightarrow
y = it
\Rightarrow
\begin{pmatrix}
x \\
y
\end{pmatrix}
= t \begin{pmatrix}
1 \\
i
\end{pmatrix}
\Rightarrow
E_{\lambda=i} = \left\{ t \begin{pmatrix}
1 \\
i
\end{pmatrix} : t \in \mathbb{R} \right\}.

Find $E_{\lambda=-i}$:
\[
\begin{pmatrix}
i & 1 & 0 \\
1 & i & 0
\end{pmatrix}
\Rightarrow
ix + y = 0 \Rightarrow
y = -it
\Rightarrow
\begin{pmatrix}
x \\
y
\end{pmatrix}
= t \begin{pmatrix}
1 \\
i
\end{pmatrix}
\Rightarrow
E_{\lambda=-i} = \left\{ t \begin{pmatrix}
1 \\
i
\end{pmatrix} : t \in \mathbb{R} \right\}.

Theorem 6.17. If $A \in \mathbb{R}^n$ and $\lambda_+ = a + ib$ is an eigenvalue with eigenvector $x$, then $\lambda_- = a - ib$ is an eigenvalue with eigenvector $x$.

Proof. $\lambda_+$ and $\lambda_-$ are paired by an algebraic theorem (roots of polynomials). If $Ax = \lambda_+ x$ and $Ay = \lambda y$, then $Ay = \bar{\lambda} y \Rightarrow Ay = \lambda_+ y \Rightarrow y = x \Rightarrow y = y = x$.

What is the relationship between eigenvectors from different eigenvalues?

Theorem 6.18. If $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues of $A$, and if $x_i$ is an eigenvector corresponding to $\lambda_i$, then $\{x_1, \ldots, x_n\}$ is a linearly independent set.
Proof. (By induction.) True for \( n = 1 \) since \( \{ x_1 \} \) is a linearly independent set. Now assume true for \( n = k \). We must show true for \( n = k+1 \). So suppose (*)\( a_1 x_1 + \ldots + a_k x_k + a_{k+1} x_{k+1} = 0 \). Then

\[
0 = (A - \lambda_{k+1} I)(a_1 x_1 + \ldots + a_{k+1} x_{k+1})
\]

\[
= a_1(Ax_1 - \lambda_{k+1} x_1) + a_2(Ax_2 - \lambda_{k+1} x_2) + \ldots + a_k(Ax_k - \lambda_{k+1} x_k) + a_{k+1}(Ax_{k+1} - \lambda_{k+1} x_{k+1})
\]

\[
= a_1(\lambda_1 x_1 - \lambda_{k+1} x_1) + a_2(\lambda_2 x_2 - \lambda_{k+1} x_2) + \ldots + a_k(\lambda_k x_k - \lambda_{k+1} x_k) + a_{k+1}(\lambda_{k+1} x_{k+1} - \lambda_{k+1} x_{k+1})
\]

\[
= a_1(\lambda_1 - \lambda_{k+1}) x_1 + a_2(\lambda_2 - \lambda_{k+1}) x_2 + \ldots + a_k(\lambda_k - \lambda_{k+1}) x_k.
\]

Since \( \{ x_1, \ldots, x_k \} \) are linearly independent (by assumption), each coefficient \( a_i(\lambda_i - \lambda_{k+1}) \) = 0.

Since the \( \lambda_i \)'s are distinct, \( a_i = 0 \) \( \forall i = 1, \ldots, n \). Thus (*) becomes \( a_{k+1} x_{k+1} = 0 \).

Since \( x_{k+1} \neq 0 \), \( a_{k+1} = 0 \). Hence \( a_i = 0 \) \( \forall i = 1, \ldots, k + 1 \), so \( \{ x_1, \ldots, x_{k+1} \} \) is a linearly independent set.

**Definition 6.19.** If \( A \in \mathbb{F}_n^\mathbb{F} \) has \( n \) linearly independent eigenvectors, \( A \) is said to be **simple**. (Otherwise \( A \) is said to be **defective**)

**Theorem 6.20.** Let \( A \in \mathbb{F}_n^\mathbb{F} \).

1. If \( A \) is simple, then \( \dim(E_{\lambda}) = \text{mult}(\lambda) \) for all eigenvalues.

2. If the characteristic polynomial of \( A \) splits, then the converse also holds.

**Proof.** 1. \( A \in \mathbb{F}_n^\mathbb{F} \) has \( n \) linearly independent eigenvectors \( \Rightarrow \sum \dim(E_{\lambda_i}) = n \Rightarrow \sum \text{mult}(\lambda_i) = n \Rightarrow \dim(E_{\lambda_i}) = \text{mult}(\lambda_i) \forall i \).

2. If the characteristic polynomial of \( A \) splits, then \( n = \sum \text{mult}(\lambda_i) \). Thus \( \dim(E_{\lambda_i}) = \text{mult}(\lambda_i) \forall i \Rightarrow n = \sum \text{mult}(\lambda_i) = \sum \dim(E_{\lambda_i}) \Rightarrow A \) has \( n \) linearly independent eigenvectors.

**Theorem 6.21.** If \( A \sim B \ (A = Q \ 1 B Q) \), then

1. \( \det(A - \lambda I) = \det(B - \lambda I) \)

2. If \( (\lambda, x_0) \) is an eigenpair of \( A \), \( (\lambda, Q x_0) \) is an eigenpair of \( B \).

**Proof.**

1. \( \det(A - \lambda I) = \det(Q^{-1} B Q - \lambda I) = \det(Q^{-1} B Q - \lambda Q^{-1} I Q) = \det(Q^{-1} (B - \lambda I) Q) \)

\[ = \det(Q^{-1}) \det(B - \lambda I) \det(Q) = \frac{1}{\det(Q)} \det(B - \lambda I) \det(Q) = \det(B - \lambda I). \]

2. By 1) \( \det(A - \lambda I) = 0 \Leftrightarrow \det(B - \lambda I) = 0 \). That is \( A \) and \( B \) have the same eigenvalues. Now suppose \( A x_0 = \lambda x_0 \). Then, \( (Q^{-1} B Q)(x_0) = \lambda x_0 \Rightarrow Q^{-1} (B(Q x_0)) = \lambda x_0 \Rightarrow B(Q x_0) = \lambda Q x_0 \Rightarrow (\lambda, Q x_0) \) is an eigenpair of \( B \).
Theorem 6.22. If $D$ is a diagonal matrix, then the diagonal entries of $D$ are exactly the eigenvalues of $D$.

Proof. Either multiply $D$ by an appropriate vector $(e_i)$, or form the characteristic polynomial.

Corollary 6.23. If $A \sim D$, then the eigenvalues of $A$ are the diagonal entries of $D$.

Proof. From 6.21 and 6.22

Definition 6.24. $A \in \mathbb{F}_n$ is diagonalizable if $A \sim D$ for some diagonal matrix $D$.

Theorem 6.25. $A$ is diagonalizable $\iff A$ is simple.

Proof. $\implies$ Suppose $D = Q^{-1}AQ$. Then $QD = AQ$ where $D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$. Write $Q = (x_1, \ldots, x_n)$ where $x_i$ is a column. Then $A(x_1, \ldots, x_n) = (x_1, \ldots, x_n)D$. Multiplying this out, we see that $Ax_1 = \lambda_1 x_1, \ldots, Ax_n = \lambda_n x_n$.

Since the columns of $Q$ are linearly independent (Why? $Q$ invertible $\implies$ rank($Q$) = $n$ $\implies$ rank($Q^t$) = $n$ $\implies$ $Q^t$ has $n$ linearly independent rows $\implies$ $Q$ has $n$ linearly independent columns), $A$ has $n$ linearly independent eigenvectors.

$\impliedby$ Let $x_1, \ldots, x_n$ be linearly independent eigenvectors of $A$. Define $Q = (x_1, x_2, \ldots, x_n)$. Then $Q$ is invertible since rank($Q$) = rank($Q^t$) = $n$.

Let $e_k = (0, \ldots, 0, 1_{kth}, 0, \ldots, 0)^t$ Then

$$(e_1, \ldots, e_n) = I = Q^{-t}Q = Q^{-1}(x_1, \ldots, x_n) = (Q^{-1}x_1, Q^{-1}x_2, \ldots, Q^{-1}x_n).$$

Thus $Q^{-1}x_i = e_i$. Hence

$$Q^{-1}AQ = Q^{-1}A(x_1, x_2, \ldots, x_n) = Q^{-1}(Ax_1, Ax_2, \ldots, Ax_n) = Q^{-1}x_1, \lambda_2 x_2, \ldots, \lambda_n x_n

= (\lambda_1 Q^{-1}x_1, \lambda_2 Q^{-1}x_2, \ldots, \lambda_n Q^{-1}x_n) = (\lambda_1 e_1 \ldots \lambda_n e_n) = \begin{pmatrix} \lambda_1 & 0 \\ \vdots & \ddots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

Corollary 6.26. If $A$ is diagonalizable and $Q = (x_1, \ldots, x_n)$ where $x_i$ are eigenvectors, then $D = Q^{-1}AQ$.

Proof. Contained in the proof of 6.25

Corollary 6.27. If $A$ has $n$ distinct eigenvalues, then $A$ is diagonalizable.

Proof. Since each eigenvalue has associated with it an eigenvector, and since eigenvectors coming from different eigenvalues are linearly independent by 6.18, $A$ has $n$ linearly independent eigenvectors. Thus $A$ is simple; hence diagonalizable.

Question 6.28. Under what conditions is it true that: $A$ diagonalizable $\iff$ mult($\lambda_i$) = dim($E_{\lambda_i}$) $\forall i$?
Exercise 6.29. Diagonalize (if possible) $A_1 = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$, $A_2 = \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ 1 & 2 & 0 \end{pmatrix}$, $A_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $A_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (See 6.8, 6.13, 6.15, and 6.16.)

Eigenvectors for $A_1$ are $(\frac{3}{4})$ and $(\frac{1}{4})$, so $Q = (\frac{3}{4} \frac{1}{-4})$, $Q^{-1} = \frac{1}{3} (\frac{1}{4} \frac{-1}{3})$, $D = Q^{-1} A Q$.

Exercise 6.30. Use $D = Q^{-1} A Q$ to find $A^{57}$ for the matrices above.

Exercise 6.31. Find a $2 \times 2$ matrix $A$ which has eigenvalues 3 and 5 and associated eigenvectors $(\frac{1}{2})$ and $(\frac{1}{4})$ respectively.

We have been concentrating on eigenvalues and eigenvectors of matrices. We now return to linear transformations. We will assume the vector space $V$ is finite dimensional.

Definition 6.32. Given $T : V \rightarrow V$, the determinant of $T$, denoted $\det(T)$, is given by $\det(T) = \det([T]_{\beta}^\gamma)$ where $\beta$ is any ordered basis for $\gamma$.

Note 6.33. Definition 6.32 is well defined. That is, if $\beta$ and $\gamma$ are two ordered bases for $V$, then $\det([T]_{\beta}^\gamma) = \det([T]_{\gamma}^\beta)$.

Proof. By 5.8, $[T]_{\gamma}^\gamma = Q [T]_{\beta}^\beta Q^{-1}$. Thus,

$$\det([T]_{\gamma}^\gamma) = \det(Q [T]_{\beta}^\beta Q^{-1}) \overset{\text{142}}{=} \det(Q) \det([T]_{\gamma}^\gamma) \det(Q^{-1}) \overset{\text{142}}{=} \det(Q) \det([T]_{\gamma}^\gamma) \frac{1}{\det(Q)}.$$  

Theorem 6.34. Given $T : V \rightarrow V$, $\lambda$ is an eigenvalue of $T$ iff $\det(T - \lambda I) = 0$.

Proof. $\lambda$ is an eigenvalue of $T$ $\iff$ $\exists x \neq 0$ s.t. $T(x) = \lambda x \iff (T - \lambda I)(x) = 0 \iff T - \lambda I$ is not injective $\overset{\text{433}}{\iff} T - \lambda I$ is not invertible $\overset{\text{53}}{\iff} [T - \lambda I]_{\beta}^\beta$ is not invertible $\overset{\text{134}}{\iff} \det[T - \lambda I]_{\beta}^\beta = 0 \overset{\text{632}}{\iff} \det |T - \lambda I| = 0$.

Theorem 6.35. Given $T : V \rightarrow V$, $\lambda$ is an eigenvalue of $T$ iff $\lambda$ is an eigenvalue of $[T]_{\beta}$ and $(x, \lambda)$ is a eigenpair for $T$ iff $(|x| \beta, \lambda)$ is an eigenpair of $[T]_{\beta}$.

Proof.

1. $\lambda$ is an eigenvalue of $T$ $\overset{\text{634}}{\iff} \det(T - \lambda I) = 0 \overset{\text{632}}{\iff} \det[T - \lambda I]_{\beta}^\beta = 0 \overset{\text{53}}{\iff} \det([T]_{\beta}^\beta \lambda[I]_{\beta}^\beta) = 0 \overset{\text{67}}{\iff} \lambda$ is an eigenvalue for $[T]_{\beta}^\beta$.

2. $x$ is an eigenvector of $T$ corresponding to $\lambda$ $\iff T(x) = \lambda x \iff T(x) = \lambda I(x) \iff T(x) - \lambda I(x) = 0 \iff (T - \lambda I)x = 0 \iff [(T - \lambda I)(x)]_{\beta} = [0]_{\beta} = 0 \overset{\text{53}}{\iff} [T - \lambda I]_{\beta}^\beta [x]_{\beta} = 0 \overset{\text{163}}{\iff} [T]_{\beta}^\beta [x]_{\beta} - \lambda [I]_{\beta}^\beta [x]_{\beta} = 0 \iff [T]_{\beta}^\beta [x]_{\beta} = \lambda [x]_{\beta} \iff [x]_{\beta}$ is an eigenvector of $[T]_{\beta}$ corresponding to $\lambda$.  

40
7 Inner Product Spaces

Our definition of a vector space came from generalizing properties of $\mathbb{R}^2$ (and $\mathbb{R}^n$). However, $\mathbb{R}^3$ has other properties which we have not (as yet) generalized to a vector space. In particular, in $\mathbb{R}^3$, vectors have length and certain vectors are perpendicular (orthogonal) to other vectors. We now generalize these notions.

**Definition 7.1.** A vector space is a **normed linear space (NLS)** if there is a function $\| \cdot \| : V \to [0, \infty)$ satisfying

1. $\|x\| = 0$ iff $x = \theta$
2. $\|cx\| = |c|\|x\|$ for all $c \in \mathbb{F}$
3. $\|x + y\| \leq \|x\| + \|y\|.$

**Exercise 7.2.** The following are important norms on $\mathbb{R}^n$ where $x = (x_1, x_2, \ldots, x_n)$.

1. $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$
2. $\|x\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p} \quad 1 \leq p < \infty$
3. $\|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$

Show that 1 and 3 are norms.

**Exercise 7.3.** Is $\|x\|_2 = \min\{|x_1|, |x_2|\}$ a norm on $\mathbb{R}^2$? Prove or disprove.

**Exercise 7.4.** When is a circle a square? A circle is a set of points equidistant from a given point. A square is a 4 equal-sided figure with perpendicular adjacent sides.

Sketch the following “circles” in $\mathbb{R}^2$ (and label them).

1. $C_1 = \{x \in \mathbb{R}^2 : \|x\|_1 = 1\}$
2. $C_2 = \{x \in \mathbb{R}^2 : \|x\|_2 = 1\}$
3. $C_3 = \{x \in \mathbb{R}^2 : \|x\|_\infty = 1\}$

We now generalize orthogonality.

**Definition 7.5.** A vector space in an **inner product space (IPS)** if there is a function denoted $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ (or $\mathbb{R}$) called an inner product satisfying

1. $\langle x, x \rangle \geq 0$
2. $\langle x, x \rangle = 0 \iff x = \theta$
3. $\langle y, x \rangle = \overline{\langle x, y \rangle}$ ($\langle x, y \rangle = \langle y, x \rangle$ in $\mathbb{R}$)
4. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
5. $\langle cx, y \rangle = c\langle x, y \rangle, c \in \mathbb{F}.$
Theorem 7.6. For $x, y, z$ in an inner product space and $c \in \mathbb{F}$,

1. $\langle x, cy \rangle = c \langle x, y \rangle$, and
2. $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$.

**Proof.** Exercise

**Exercise 7.7.**

1. Given $x, y \in \mathbb{F}^n$, $\langle x, y \rangle = \sum_{i=1}^n x_i \overline{y_i}$ is an inner product.
2. Given $f, g \in C[0, 1]$, $\langle f, g \rangle = \int_0^1 f(x) \overline{g(x)} \, dx$ is an inner product.

**Definition 7.8.** If $V$ is an IPS, $x, y \in V$ are **orthogonal** if $\langle x, y \rangle = 0$. We write $x \perp y$. A subset $S \subset V$ is an orthogonal set if given $x, y \in S, x \neq y, x \perp y$.

**Exercise 7.9.** Find an orthogonal set of $\mathbb{R}^3$.

**Note 7.10.** Is there a connection between these two kinds of spaces? Yes. Given an IPS, $\sqrt{\langle x, x \rangle}$ is a norm. Check:

1. $\sqrt{\langle x, x \rangle} = 0 \iff \langle x, x \rangle = 0 \iff x = \theta$.
2. $\sqrt{\langle cx, cx \rangle} = \sqrt{c \langle x, x \rangle} = \sqrt{c^2 \langle x, x \rangle} = |c| \sqrt{\langle x, x \rangle}$.

To show that the third condition for being a norm is satisfied, we first need to prove the

**Theorem 7.11 (Cauchy-Schwartz Inequality).** Given $x, y \in V$ (an IPS),

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}.$$  

**Proof.** Case 1: $y = 0$. Both sides are equal to 0.

Case 2: $y \neq 0$. Then

$$0 \leq \langle x + ay, x + ay \rangle = \langle x, x \rangle + a \langle x, y \rangle + a \langle y, x \rangle + |a|^2 \langle y, y \rangle$$

for any $a$. In particular, for $a = \frac{\langle x, y \rangle}{\langle y, y \rangle}$, we get

$$0 \leq \langle x, x \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle y, x \rangle + \left| \frac{\langle x, y \rangle}{\langle y, y \rangle} \right|^2 \langle y, y \rangle.$$  

Multiplying by $\langle y, y \rangle$:

$$0 \leq \langle x, x \rangle \langle y, y \rangle - \langle x, y \rangle \langle x, y \rangle - \langle x, y \rangle \langle y, x \rangle + |\langle x, y \rangle|^2 \langle y, y \rangle$$  

$$\Rightarrow 0 \leq \langle x, x \rangle \langle y, y \rangle - \langle x, y \rangle \langle x, y \rangle - \langle x, y \rangle \langle y, x \rangle + |\langle x, y \rangle|^2$$  

$$\Rightarrow 0 \leq \langle x, x \rangle \langle y, y \rangle - 2|\langle x, y \rangle|^2 + |\langle x, y \rangle|^2 \Rightarrow 0 \leq \langle x, x \rangle \langle y, y \rangle - |\langle x, y \rangle|^2$$  

$$\Rightarrow |\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle \Rightarrow |\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}.$$
We now finish showing that $\sqrt{\langle x, x \rangle}$ is a norm:

3. 

\[
\langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\
= \langle x, x \rangle + \langle x, y \rangle + \langle y, y \rangle + 2\text{Re} \langle x, y \rangle \\
= \langle x, x \rangle + 2\langle x, y \rangle \\
\leq \langle x, x \rangle + 2\|x\|\|y\| \\
\leq \langle x, x \rangle + 2\sqrt{\langle x, x \rangle \sqrt{\langle y, y \rangle}} + \langle y, y \rangle \\
= (\sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle})^2.
\]

Thus $\sqrt{\langle x + y, x + y \rangle} \leq \sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle}$.

Thus we have shown that every IPS is a normed linear space with $\|x\| = \sqrt{\langle x, x \rangle}$.

**Note 7.12.** Is every NLS an IPS? That is, does every norm come from an inner product? No. A norm comes from an inner product iff the norm obeys the parallelogram law:

\[
\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).
\]

**Exercise 7.13.** Show that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ from 7.2 do not come from inner products.

**Definition 7.14.** An orthogonal set of vectors such that the norm of each vector is 1 is called an orthnormal set.

**Exercise 7.15.** Find (if possible) an orthnormal set of $\mathbb{R}^3$ containing

1. 2 vectors
2. 3 vectors
3. 4 vectors

**Theorem 7.16.** An orthnormal set of non-zero vectors $\{x_1, x_2, \ldots, x_n\}$ is linearly independent.

**Proof.** Suppose $a_1x_1 + \cdots + a_nx_n = 0$ (must show each $a_i = 0$). Now

\[
0 = \langle 0, x_i \rangle = a_1x_1 + \cdots + a_nx_n, x_i \rangle \\
= a_1\langle x_1, x_i \rangle + \cdots + a_i\langle x_i, x_i \rangle + \cdots + a_n\langle x_n, x_i \rangle \\
= a_1(0) + \cdots + a_i\|x_i\|^2 + \cdots + a_n(0).
\]

Since $\|x_i\| \neq 0$, $a_i = 0$.

**Remark 7.17.** Thus, in a vector space $V$ of dimension $n$, an orthnormal set of $n$ non-zero vectors is a basis of $V$ by 3.22. Are there any advantages for a basis to be orthnormal? You bet! In general, given $x \in V$ and a basis $\{x_1, \ldots, x_n\}$ we know we can find $a_1, \ldots, a_n$ such that $x = a_1x_1 + \cdots + a_nx_n$. However, doing so involves solving an $n \times n$ system of equations. If the basis is an orthnormal basis, the process is much easier.
Theorem 7.18. Let \( \dim(V) = n \). If \( \beta = \{x_1, \ldots, x_n\} \) is an orthogonal set of non-zero vectors, then \( \beta \) is a basis and for all \( x \in V \),

\[
x = \sum_{i=1}^{n} \frac{\langle x, x_i \rangle}{\langle x_i, x_i \rangle} x_i \quad (= \sum_{i=1}^{n} \langle x, x_i \rangle x_i \text{ if } \beta \text{ is orthonormal})
\]

Proof. By the above remark, \( \beta \) is a basis. Thus we can write \( x = a_1 x_1 + \cdots + a_n x_n \). Then

\[
\langle x, x_i \rangle = \langle a_1 x_1 + \cdots + a_n x_n, x_i \rangle = a_1 \langle x_1, x_i \rangle + \cdots + a_n \langle x_n, x_i \rangle \\
= a_i(0) + \cdots + a_i \langle x_i, x_i \rangle + \cdots + 0.
\]

Hence \( a_i = \frac{\langle x, x_i \rangle}{\langle x_i, x_i \rangle} \).

An important geometric interpretation: Recall that \( \frac{\langle x, x_i \rangle}{\|x_i\|} x_i \) is just the projection of \( x \) onto \( x_i \). So, if \( \{x_1, \ldots, x_n\} \) is an orthogonal set, then \( x \) is the sum of its projections on the basis vectors. If \( \{x_1, \ldots, x_n\} \) is not an orthogonal set, this need not be the case (see diagrams below). Thus with an orthogonal basis, as the projections of \( x \) onto the basis vectors are summed up, the sum will keep getting “closer and closer” to \( x \).

\[\text{Exercise 7.19. Show that } \{\sin(nx), \cos(mx) : m, n \in \mathbb{Z}\} \text{ is an orthogonal set of } C[-\pi, \pi] \text{ with inner product } \int_{-\pi}^{\pi} f(x)g(x)dx. \text{ (Note the similarity to } \sum x_iy_i).\]

\[\text{Note 7.20. Similar to the above, given } f \in C[-\pi, \pi],
\]

\[
f(x) = \sum_{n=-\infty}^{\infty} \langle f(x), \sin(nx) \rangle \frac{\sin(nx)}{\|\sin(nx)\|} + \sum_{n=-\infty}^{\infty} \langle f(x), \cos(nx) \rangle \frac{\cos(nx)}{\|\cos(nx)\|}
\]

The expansion given above is called the **Fourier series** of \( f \) and the inner products \( \langle f(x), \sin(nx) \rangle \) and \( \langle f(x), \cos(nx) \rangle \) are the **Fourier coefficients**.
**Note 7.21.** Gram-Schmidt Orthonormalization Process Now that we see the advantages of orthonormal bases, can we construct them? We can use 7.18 to obtain an orthonormal basis from any basis as follows: Let \( \{x_1, \ldots, x_n\} \) be a basis. We will construct an orthonormal basis \( \{w_1, \ldots, w_n\} \) as follows.

1. Take \( w_1 = x_1 \).

2. \( \frac{(x_2, w_1)}{\langle w_1, w_1 \rangle} w_1 \) is the projection of \( x_2 \) onto \( w_1 \) (see diagram), and \( w_2 = x_2 - \frac{(x_2, w_1)}{\langle w_1, w_1 \rangle} w_1 \perp w_1 \) (Prove!)

3. Similarly, set \( w_3 = x_3 - \frac{(x_3, w_1)}{\langle w_1, w_1 \rangle} w_1 - \frac{(x_3, w_2)}{\langle w_2, w_2 \rangle} w_2 \). Then \( w_3 \perp w_1 \) and \( w_3 \perp w_2 \).

4. Continue to get an orthogonal basis.

5. Normalize each vector \( \left( \frac{w_i}{\|w_i\|} \right) \) to get an orthonormal basis.

![Diagram](image)

**Exercise 7.22.** Start with your own basis for \( \mathbb{R}^3 \) (prove it is a basis), and use the Gram-Schmidt Process to get an orthonormal basis. (Prove it is orthogonal.) Write the vector \( (10, 12, 15) \) as a linear combination of your basis elements, then use 7.18 to write it as a linear combination of the elements of the orthonormal basis you formed with Gram-Schmidt.

**Exercise 7.23.** Starting with the basis \( \{1, x, x^2\} \) for \( \mathcal{P}^2 \), and using the inner product \( \int_{-1}^{1} f(x)g(x)\,dx \), use the Gram-Schmidt process to get an orthonormal basis. The polynomials obtained are the first three **Legendre polynomials**, which are useful in the solution of certain differential equations.
A Complex Numbers

Definition A.1. A complex number is a number of the form \( z = a + bi \), where \( a, b \in \mathbb{R} \), and \( i = \sqrt{-1} \). \( a \) is called the real part of \( z \) and \( b \) is called the imaginary part of \( z \) (denoted \( \text{Re} \ z \) and \( \text{Im} \ z \), respectively). The set of all complex numbers is written \( \mathbb{C} \).

The complex numbers can be identified with elements of \( \mathbb{R}^2 \) by viewing \( a \) as the \( x\)-coordinate and \( b \) as the \( y\)-coordinate of a vector. We thus view \( i \) as the vector \((0, 1)\) in \( \mathbb{R}^2 \).
We define addition and multiplication of complex numbers as follows, treating \( i \) as \( \sqrt{-1} \).

Definition A.2. If \( a + bi, c + di \in \mathbb{C} \),

1. \( a + bi = c + di \) iff \( a = b, c = d \)
2. \((a + bi) + (c + di) = (a + c) + (b + d)i \)
3. \((a + bi)(c + di) = (ac - bd) + (ad + bc)i \).

It is not immediately clear how to divide complex numbers. As an aid, we introduce another operation on complex numbers.

Definition A.3. The conjugate of \( z = a + bi \) (written \( \bar{z} \)) is the complex number \( a - bi \).

Notice \((a + bi)(a - bi) = (a + bi)(a - b)i = a^2 + b^2 \in \mathbb{R} \). Now we can do complex division by multiplying numerator and denominator by the conjugate of the denominator:

\[
\frac{a + bi}{c + di} = \frac{a + bi}{c + di} \frac{c - di}{c - di} = \frac{(a + bi)(c - di)}{c^2 + d^2}.
\]

Definition A.4. The modulus of a complex number \( z = a + bi \) is the length of the corresponding vector in \( \mathbb{R}^2, \sqrt{a^2 + b^2} \). We write the modulus of \( z \) as \( |z| \). Note that \( zz = |z|^2 \).

Just as in \( \mathbb{R}^2 \), we can express the complex number \( z \) in polar coordinates \( r \) and \( \theta \), where \( r \) is the distance from a point to the origin in the complex plane (which is the modulus we just defined) and \( \theta \) (called the argument of \( z \)) is the angle the complex vector makes with the positive real axis. Now every point \( z = a + ib \) can be written \( z = r \cos \theta + ir \sin \theta \) where \( r = |z| = \sqrt{a^2 + b^2} \) and \( \theta = \tan^{-1}(b/a) \). (Check that this agrees with geometric intuition.)

Using Euler’s formula

\[
e^{i\theta} = \cos \theta + i \sin \theta
\]

which is obtained by formally substituting \( i\theta \) into the power series expansion for \( e^z \), we can write

\[
z = r \cos \theta + ir \sin \theta = re^{i\theta}
\]

which is the exponential form of \( z \).

The exponential form gives us a geometric interpretation of multiplication. Let \( z_1 = r_1e^{i\theta_1} \) and \( z_2 = r_2e^{i\theta_2} \). Then

\[
z_1z_2 = r_1r_2e^{i(\theta_1 + \theta_2)}.
\]

That is, when two complex numbers are multiplied, the modulus of their product is the product of their moduli, and the argument of their product is the sum of their arguments.
Exercise A.5. Let \( z_1 = 2 + i \) and \( z_2 = 1 - 2i \). Find the modulus and argument of \( z_1, z_2, \) and \( z_1z_2 \) and verify the above statement about complex multiplication.

![Complex Addition](image1)

![Complex Multiplication](image2)

Note A.6. If we substitute the value \( \pi \) in for \( \theta \) in DeMoivre’s formula, we obtain the identity

\[
e^{i\pi} + 1 = 0
\]

a remarkable equality which includes the five basic constants in mathematics, as well as the three arithmetic operations of addition, multiplication, and exponentiation.

Definition A.7. A function \( f : \mathbb{R} \to \mathbb{C} \) is defined by \( f(x) = f_1(x) + if_2(x) \) where \( f_i : \mathbb{R} \to \mathbb{R} \). As with complex numbers, \( \overline{f(x)} = f_1(x) - if_2(x) \) and so \( f(x)\overline{f(x)} = |f(x)|^2 \).
B Generalized Eigenvectors and Jordan Normal Form

As noted previously, matrices which have repeated eigenvalues (eigenvalues with \( \text{mult}(\lambda) > 1 \)) fail to be diagonalizable if \( \text{mult}(\lambda) > \dim(E_\lambda) \) by 6.25. Such matrices were called defective. In this appendix we attempt to sidestep this defect and complete our set of eigenvectors with generalized eigenvectors. Although defective matrices are not diagonalizable, we find another useful form, the Jordan normal form of a matrix.

Diagonalization of a matrix \( A = Q^{-1}DQ \) can be interpreted in terms of change of basis. The matrix \( Q \) is a transition matrix from the standard basis of \( \mathbb{C}^n \) to the basis given by the eigenvectors of \( A \). Now \( Ax = Q^{-1}DQx \), that is, to imitate the action of \( A \) on \( x \), we first change the basis of \( x \) to the eigenvector basis using the matrix \( Q \), then we act on \( x \) in the new basis with \( D \), and finally use \( Q^{-1} \) to return to the standard basis for \( \mathbb{C}^n \). If \( A \) is simple, then our new basis consists entirely of eigenvectors and \( D \) is diagonal. If \( A \) is defective, then we do not have enough eigenvectors to form a basis. Although there are many ways of extending a linearly independent set to a basis, we would like to complete our basis in such a way that \( D \)'s replacement is as close to a diagonal matrix as possible. This motivates the

Definition B.1. Let \( A \in \mathbb{C}^n \) and let \((\lambda, x)\) be an eigenpair of \( A \). \( z \) is a generalized eigenvector of \( A \) corresponding to \( \lambda \) if \( \exists k > 0 \) such that \((A - \lambda I)^kz = 0\). \( G_\lambda = \{z \in \mathbb{C}^n : (A - \lambda I)^kz = 0\} \) is called the generalized eigenspace corresponding to \( \lambda \).

By letting \( k = 1 \) in the above definition, we see that every eigenvector is a generalized eigenvector.

Exercise B.2. Show that \( G_\lambda \) is a indeed subspace of \( \mathbb{C}^n \).

Definition B.3. The order of a generalized eigenvector \( z \) is the smallest \( k \) such that \((A - \lambda I)^kz = 0\). Using an generalized eigenvector \( y_k \) of order \( k \), we can generate a chain of generalized eigenvectors:

\[
\begin{align*}
y_k &= z \\
y_{k-1} &= (A - \lambda I)z \\
y_{k-2} &= (A - \lambda I)^2z \\
&\vdots \\
y_1 &= (A - \lambda I)^{k-1}z.
\end{align*}
\]

Note that in a chain of generalized eigenvectors, \( y_{i-1} = (A - \lambda I)y_i \). Also, \((A - \lambda I)y_1 = 0\) since \((A - \lambda I)^ky_k = 0\) because we chose \( y_k \) to have order \( k \). Hence \( y_1 \) (the end of a chain) is always an actual eigenvector.

Theorem B.4. A chain of generalized eigenvectors is a linearly independent set.

Proof. Let \( \{y_1, \ldots, y_k\} \) be a chain of generalized eigenvectors corresponding to the eigenvalue \( \lambda \) of order \( k \). Note that \( y_i \neq 0 \) for every \( i = 1, \ldots, k \). Suppose \( a_1y_1 + \cdots + a_ky_k = 0 \). Multiply both sides of the equation by the matrix \((A - \lambda I)^{k-1} \):

\[
\begin{align*}
a_1(A - \lambda I)^{k-1}y_1 + \cdots + a_k(A - \lambda I)^{k-1}y_k &= (A - \lambda I)^{k-1}0 \\
\Rightarrow a_1(0) + a_2(0) + \cdots + a_ky_1 &= 0
\end{align*}
\]

48
so \( a_k = 0 \), since \( y_1 \neq 0 \). Similarly, multiplying by \((\lambda I - A)^{k-2}\) gives us \( a_{k-1} = 0 \). Continuing likewise, we see that all of the \( a_i = 0 \), so \( \{y_1, \ldots, y_k\} \) is linearly independent.

We obtain a chain of generalized eigenvectors for each eigenvalue; note that some eigenvalues may have multiple chains (see example below). The union of these chains forms a basis of \( \mathbb{C}^n \), just as the eigenvectors of a diagonalizable matrix do. We omit the lengthy proof. Referring back to B.3, we can rewrite the equalities in a chain of generalized eigenvectors as:

\[
(A - \lambda I)y_{k-1} = y_{k-2} \\
\vdots \\
(A - \lambda I)y_2 = y_1 \\
(A - \lambda I)y_1 = 0
\]

which can be further manipulated to obtain

\[
Ay_{k-1} = \lambda y_{k-1} + y_{k-2} \\
\vdots \\
Ay_2 = \lambda y_2 + y_1 \\
Ay_1 = \lambda y_1
\]

so we can represent the action of \( A \) on the span of the chain \( \{y_{k-1}, \ldots, y_1\} \) by the **Jordan block**:

\[
A \left( \begin{array}{c} y_{k-1} \\ \vdots \\ y_1 \end{array} \right) = \left( \begin{array}{cccc} \lambda & 1 & 0 & \cdots \\ 0 & \lambda & 1 & \vdots \\ \vdots & \ddots & \ddots & \lambda \\ 0 & \cdots & \cdots & \lambda \end{array} \right) \left( \begin{array}{c} y_k \\ 1 \\ \vdots \\ y_1 \end{array} \right).
\]

**Definition B.5.** The **Jordan normal form** of a matrix is the matrix formed by arranging the Jordan blocks associated with each generalized eigenspace along the diagonal of the matrix:

\[
J = \begin{pmatrix}
J_1 & 0 & \cdots & 0 \\
0 & J_2 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & J_m
\end{pmatrix}
\]

**Exercise B.6.** Find the Jordan normal form of the matrix \( A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} \).

The only eigenvalue of \( A \) is 2. \( (A - 2I) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \), so \( (1,0,0) \) and \( (0,0,1) \) (and hence also \( (1,0,1) \)) are eigenvectors of \( A \), but they do not span \( \mathbb{C}^3 \). We turn to \( (A - 2I)^2 = 0, \)
so we can choose any vector from \( \mathbb{C}^3 \) which is not linearly dependent on \((1, 0, 0)\) and \((0, 0, 1)\) to complete our basis. For simplicity’s sake, we choose \((0, 1, 0)\). Note that \((A - 2I)(0, 1, 0) = (1, 0, 1)\), so \((0, 1, 0)\) and \((1, 0, 1)\) form a chain of generalized eigenvectors, and \((1, 0, 0)\) (or \((0, 0, 1)\) if you want) forms another chain by itself. Hence we have two Jordan blocks:

\[
J_1 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \quad J_2 = (2)
\]

so the Jordan normal form of our matrix is \[
\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.
\]

Every matrix is similar to a matrix in Jordan normal form. A specific case of this is diagonalization of simple matrices. In the same way, the matrix \( Q \) such that \( A = Q^{-1}JQ \) where \( J \) is the Jordan form of \( A \) is attained by placing the generalized eigenvectors as the columns. As before, this serves to change the basis from the standard basis for \( \mathbb{C}^n \) to the basis given by the generalized eigenvectors.

**Exercise B.7.** Find the matrix \( Q \) for the previous example.

We arrange the generalized eigenvectors by Jordan blocks. The first Jordan block corresponds to the chain of length 2 consisting of \((1, 0, 1)\) and \((0, 1, 0)\). The generalized eigenvectors are put in order left to right from the generalized eigenvector of order 2 to the actual eigenvector. The next Jordan block corresponds to the “chain” of the single eigenvector \((1, 0, 0)\). We attain

\[
\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}
\]

**Exercise B.8.** Find the Jordan normal form and transition matrix \( Q \) of the matrix \( A =
\[
\begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}
\]

\[
(\text{Answer: } A = \begin{pmatrix} 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}).
\]
C  Transition to Differential Equations

Definition C.1. The differential operator $D : V \to W$, where $V$ is a space of sufficiently differentiable functions takes $f \in V$ to its derivative $f'$. The $n$ fold composition $D^n$ of the operator takes $f$ to its $n$th derivative $f^{(n)}$.

Exercise C.2. Show that $D$ is a linear transformation.

Definition C.3. The $n$th order linear differential equation (DE) $p_n(t)y^{(n)} + \cdots + p_0(t)y(t) = g(t)$ can be written as $L(y) = g(t)$ where $L : V \to V$ is the linear transformation $L = p_n(t)D^n + \cdots + p_1(t)D + p_0(t)$ and $V$ is the space of sufficiently differentiable functions. $L$ is sometimes called an nth order linear operator or differential operator.

We state the following result from analysis without proof. The theorem is usually proved in an advanced calculus or analysis course.

Theorem C.4. Let $p_i \in C(a, b)$ for all $i = 0, \ldots, n - 1$. If $L = D^n + p_{n-1}D^{n-1} + \cdots + p_0$, then the equation $L(y) = 0$ has a unique solution satisfying a set of initial conditions $f(x_0) = c_0, f'(x_0) = x_1, \ldots, f^{(n-1)}(x_0) = c_{n-1}$.

We apply this theorem to find the number of linearly independent solutions of an nth order linear differential equation.

Theorem C.5. If $L$ is an nth order differential operator, then any basis for $\ker(L)$ has exactly $n$ elements. So, for example, if \{u_1(t), \ldots, u_n(t)\} is a linearly independent set of solutions, any solution to $L(y) = 0$ can be written as $y_0 = a_1u_1(t) + \cdots + a_nu_n(t)$.

Proof. Define a surjective linear transformation $T : \ker(L) \to \mathbb{R}^n$ which maps each solution $f$ of $L(y) = 0$ onto its value at the initial state, i.e. $T(f) = (f(x_0), f'(x_0), \ldots, f^{(n-1)}(x_0))$. Now the equation with all initial conditions prescribed to be $0$ has a unique solution by the previous theorem, that is, $T(f) = 0 \Rightarrow f = 0$, so $T$ is injective by 4.22. Now $T$ is bijective, so by 4.30 $\dim(\ker(L)) = \dim(\mathbb{R}^n) = n$.

Question C.6. How can we know if we have linearly independent solutions to a DE? For example, $\sqrt{t}$ and $\frac{1}{t}$ are solutions to $y'' + \frac{3}{2}y' - \frac{1}{2t^2}y = 0$. Are they linearly independent?

Definition C.7. Given the set of functions \{u_1(t), \ldots, u_n(t)\}, the Wronskian of the set is the determinant:

$$W(u_1(t), \ldots, u_n(t)) = \begin{vmatrix} u_1(t) & \cdots & u_n(t) \\ u_1'(t) & \cdots & u_n'(t) \\ \vdots & \ddots & \vdots \\ u_1^{(n)}(t) & \cdots & u_n^{(n)}(t) \end{vmatrix}.$$

Theorem C.8. If $W(u_1, \ldots, u_n) \neq 0$, \{u_1, \ldots, u_n\} is linearly independent.

Proof. Suppose $a_1u_1(t) + \cdots + a_nu_n(t) = 0$. (We must show $a_i = 0 \forall i$.) Then

$$a_1u_1' + \cdots + a_nu_n' = 0 \quad \vdots \quad a_1u_1^{(n)} + \cdots + a_nu_n^{(n)} = 0.$$

51
We then have
\[
\begin{pmatrix}
  u_1(t) & \cdots & u_n(t) \\
  u'_1(t) & \cdots & u'_n(t) \\
  \vdots & \ddots & \vdots \\
  u^{(n)}_1(t) & \cdots & u^{(n)}_n(t)
\end{pmatrix}
\begin{pmatrix}
a_1 \\
  \vdots \\
a_n
\end{pmatrix} = \begin{pmatrix} 0 \\
  \vdots \\
  0 \end{pmatrix}.
\]
(1)

Since the Wronskian is nonzero by assumption, the system of equations has the unique solution \(a_i = 0 \forall i\) by 1.34.

**Note C.9.** Since the general solution of an \(n\)th order DE has \(n\) arbitrary constants, the unique solution to the initial value problem can be found by using the \(n\) initial conditions to determine the constants.

**Proof.** See the previous theorem. Since the homogeneous equation 1 has a unique solution, the nonhomogeneous equation obtained by replacing the zero vector on the right hand side with the column vector of initial conditions has a unique solution by 1.34.

**Note C.10.** Since a linear differential equation is a linear equation, the solution to a system of differential equations is an affine subspace by 4.10. That is, if \(y_i\) is a particular solution to the nonhomogeneous DE \(L(y) = g(t)\), then \(y_y = y_i + y_p\) is the general solution, where \(y_i\) is the general solution of the homogeneous equation \(L(y) = 0\), for \(L(y_i + y_p) = L(y_i) + L(y_p) = 0 + g(t) = g(t)\).

**Exercise C.11.** Consider the DE \((\ddot{y})\ 9y = 0.\)

1. What is the order of the DE?
2. Write the DE in the form \(L(y) = 0.\) What is \(L?\)
3. Show that \(L\) is a linear transformation.
4. Is the DE homogeneous?
5. How many elements are an the basis of \(\ker(L)?\)
6. By inspection (guessing), find two different solutions of \((\ddot{y}).\)
7. Using the Wronskian, determine if these two solutions are linearly independent. If they are not, keep trying until you get two linearly independent solutions.
8. Explain why these two solutions form a basis for \(\ker(L).\)
9. Write the general solution to \((\ddot{y})\) and explain why it is the general solution.
10. By inspection, find a particular solution to
    \[
    \begin{align*}
    & (a) \ y'' - 9y = -18t \\
    & (b) \ y'' - 9y = 45 \\
    & (c) \ y'' - 9y = -1/t^2 - 9\ln t
    \end{align*}
    \]
11. What are the general solutions of a), b) and c) ?

12. What is a particular solution to \( y'' - 9y = -18t + 45 - 1/t^2 - 9\ln t \) ?

13. Write down a statement that generalizes 12 and prove it.

14. What is the general solution of the equation in 12?
Additional Exercises

1 Matrices and Systems of Equations

Perform the following multiplications:

1. \[
\begin{pmatrix}
1 & -2 & 0 \\
6 & 3 & -1 \\
2 & 2 & 4 \\
\end{pmatrix}
\begin{pmatrix}
3 & 3 & 4 \\
2 & -1 & 0 \\
11 & 4 & 5 \\
\end{pmatrix}
\]

2. \[
\begin{pmatrix}
3 & 3 & 4 \\
2 & -1 & 0 \\
11 & 4 & 5 \\
\end{pmatrix}
\begin{pmatrix}
1 & -2 & 0 \\
6 & 3 & -1 \\
2 & 2 & 4 \\
\end{pmatrix}
\]

Solve the following linear systems:

3. \[
\begin{align*}
x_1 + x_2 - x_3 + 7x_4 &= -8 \\
x_1 + x_3 + 4x_4 &= 1 \\
3x_1 + 2x_2 + 7x_3 + 6x_4 &= 21 \\
x_1 + x_2 - 3x_3 - x_4 &= -10 \\
\end{align*}
\]

4. \[
\begin{align*}
3x_1 + x_2 + 4x_3 &= 0 \\
9x_1 + 5x_2 + 16x_3 &= 0 \\
21x_1 + 11x_2 + 36x_3 &= 0 \\
\end{align*}
\]

5. \[
\begin{align*}
x_1 + 3x_2 + x_3 + 2x_4 &= 1 \\
-2x_1 - 5x_2 - x_3 - 5x_4 &= 1 \\
\end{align*}
\]

\[
\begin{align*}
a. \quad 2x_1 + 3x_2 &= 1 \\
b. \quad 2x_1 + 3x_2 &= 3 \\
c. \quad x_1 + 4x_2 &= 2 \\
d. \quad 2x_1 + 3x_2 &= 4 \\
\end{align*}
\]

Hint: Work smarter, not harder.

6. \[
\begin{align*}
a. \quad 2x_1 + 3x_2 &= 5 \\
b. \quad 2x_1 + 3x_2 &= 7 \\
c. \quad x_1 + 4x_2 &= 6 \\
d. \quad x_1 + 4x_2 &= 8 \\
\end{align*}
\]

7. Solve the following systems of equations using Cramer’s Rule:

a. \[
\begin{align*}
2x + y &= 7 \\
3x - 2y &= 0 \\
\end{align*}
\]

b. \[
\begin{align*}
x + y &= 1 \\
-x + 2y &= -4 \\
\end{align*}
\]

c. \[
\begin{align*}
2x + 2y + 4z &= 2 \\
2x - 2y + 8z &= 6 \\
\end{align*}
\]

Find the determinant of the following matrices, then find their inverses, if possible:

8. \[
\begin{pmatrix}
11 & 3 \\
9 & -2 \\
\end{pmatrix}
\]

9. \[
\begin{pmatrix}
0 & 1 \\
2 & 4 \\
\end{pmatrix}
\]

10. \[
\begin{pmatrix}
1 & 5 \\
-3 & -12 \\
\end{pmatrix}
\]

11. \[
\begin{pmatrix}
2 & -1 & 3 \\
1 & 1 & 2 \\
1 & 0 & 2 \\
\end{pmatrix}
\]

12. \[
\begin{pmatrix}
3 & 3 & 4 \\
6 & 3 & 2 \\
3 & -2 & -1 \\
\end{pmatrix}
\]

13. \[
\begin{pmatrix}
3 & 1 & 4 \\
5 & 2 & 7 \\
4 & 2 & -2 \\
\end{pmatrix}
\]

14. \[
\begin{pmatrix}
1 & 4 & 3 & 0 \\
1 & 5 & 5 & 4 \\
1 & 6 & -2 & -3 \\
\end{pmatrix}
\]

15. \[
\begin{pmatrix}
0 & 3 & 1 & 0 \\
2 & -6 & -2 & 4 \\
1 & 1 & 0 & 2 \\
\end{pmatrix}
\]

16. \[
\begin{pmatrix}
13 & 7 & -7 & -11 \\
3 & 4 & 0 & 2 \\
0 & 1 & 1 & 4 \\
\end{pmatrix}
\]

54
17. (a) Show that if $A \in \mathbb{F}_n$, $(A \quad I)(A + I) = A^2 \quad I$.
   
   (b) If $B \in \mathbb{R}_n$, is $(A - B)(A + B) = A^2 - B^2$?

18. Show that:
   
   (a) $(A^t)^t = A$
   
   (b) $(A + B)^t = A^t + B^t$
   
   (c) $(AB)^t = B^t A^t$
   
   (d) $(A^{-1})^t = (A^t)^{-1}$

19. Given $A \in \mathbb{F}_n$, $A$ is said to be **symmetric** if $A^t = A$ and we call $A$ **skew-symmetric** if $A^t = -A$.
   
   (a) Show that $A + A^t$ is symmetric and $A - A^t$ is skew-symmetric.
   
   (b) Write $A$ as the sum of a symmetric matrix and a skew symmetric matrix.

20. Show that if $A, B \in \mathbb{F}_n$ are invertible, then $AB$ is invertible with inverse $B^{-1} A^{-1}$.

21. Under what condition is $I + A + A^2$ the inverse of $I - A$?

22. Matrices $A, B \in \mathbb{R}_n$ are similar if there exists an invertible matrix $P$ such that $P^{-1} BP = A$. Show that if $A$ and $B$ are similar, $\det(A) = \det(B)$.

23. Let $A \in \mathbb{F}_n$ be upper triangular. Show that $A$ has an inverse iff there are no zeros on the diagonal.

24. Show if $A$ is invertible and $AB = 0$, then $B = 0$.

25. Show that if $A$ is invertible and $AC = I$, then $C = A^{-1}$.

26. Let $A, B \in \mathbb{F}_n$ s.t. $AB = I$. Show that $A = B^{-1}$ and $B = A^{-1}$ (That is, for a square matrix, a one sided inverse is a two sided inverse.)

2 **Vector Spaces**

1. Show that any line (plane) in $\mathbb{R}^2$ containing the origin is a subspace of $\mathbb{R}^3$.

2. Prove that in any vector space $V$, $a \in \mathbb{F}$, $x \in V$:

   (a) If $a \neq 0$ and $ax = 0$, then $x = \theta$.
   
   (b) If $x \neq \theta$ and $ax = 0$, then $a = 0$.

3. A square matrix is **idempotent** if $A^2 = A$. Is the set of all idempotent $n \times n$ matrices a subspace of $\mathbb{R}_n^n$?
3 Linear Independence, Spanning Sets, and Bases

1. Prove that $x_1, x_2, x_3 \in \mathbb{R}^3$ form a basis for $\mathbb{R}^3$ iff $\det(x_1, x_2, x_3) \neq 0$.

2. Let $V = \mathbb{C}_2^n = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} : a_{ij} \in \mathbb{R} \right\}$.

   (a) Find a basis for $V$.
   (b) What is $\dim(V)$?
   (c) Show that $W = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} : a_{11} + a_{22} = 0 \right\}$ is a subspace of $V$.
   (d) Find a basis for $W$.

3. For each of the following, give an example if it is possible, or give an argument showing that it is not possible.
   (a) $A, B \in \mathbb{R}_2^n$ such that $\text{rank}(A) = \text{rank}(B) = 3$ and $\text{rank}(AB) = 2$.
   (b) $A, B \in \mathbb{R}_2^n$ such that $\text{rank}(A) = \text{rank}(B) = 1$ and $\text{rank}(AB) = 0$.

4 Linear Transformations

1. Given a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ with $T(1, 2) = (1, 1, 1)$ and $T(2, 3) = (3, 0, 1)$, find $T(3, 2)$.

2. Given a linear transformation $T : V \to W$ and $S \subset V$ a linearly independent subset of $V$, prove or disprove:
   (a) If $T$ is injective, then $T(S) \subset W$ is linearly independent.
   (b) If $T$ is surjective, then $T(S) \subset W$ is linearly independent.

3. Show that the trace is a linear transformation from $\mathbb{R}^n \to \mathbb{R}$.

4. Given $T : \mathcal{P}^2 \to \mathcal{P}^2$ defined by $T(f) = f'' + f'$
   (a) Determine if $T$ is invertible
   (b) Compute $T^{-1}$ if it exists

5 Matrix Representations of Linear Transformations

1. Given the linear transformation $T : \mathcal{P}^2 \to \mathcal{P}^3$ defined by $T(f)(x) = \int_0^x f(t)dt + f'(x) + x^2f''(x)$ with $\beta = \{1, x, x^2\}$ and $\gamma = \{1, x, x^2, x^3\}$,
   (a) Find $[T]_{\beta}$
   (b) Use $[T]_{\beta}$ to find $T(3x^2 + 4x + 6)$.
2. Given the linear transformation \( T : \mathcal{P}^2 \to \mathcal{P}^2 \) defined by \( T(f)(x) = (x + 2)f'(x) + 3f \) and \( U : \mathcal{P}^2 \to \mathbb{R}^3 \) defined by \( U(ax^2 + bx + c) = (a + c, b, c) \) with \( \beta = \{1, x, x^2\} \) and \( \gamma = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} \).

(a) Find \([U]_{\beta}^\gamma\), \([T]_{\beta}^\gamma\), and \([UT]_{\beta}^\gamma\).

(b) Use \([UT]_{\beta}^\gamma\) to find \( UT(x^2 + 2x + 1) \).

6 Eigenvectors

Find the eigenvalues and eigenvectors of the following matrices:

1. \[
\begin{pmatrix}
3 & -2 \\
2 & -2
\end{pmatrix}
\]

5. \[
\begin{pmatrix}
4 & -3 \\
8 & -6
\end{pmatrix}
\]

9. \[
\begin{pmatrix}
1 & 1 & 2 \\
1 & 2 & 1 \\
2 & 1 & 1
\end{pmatrix}
\]

2. \[
\begin{pmatrix}
2 & -1 \\
3 & 2
\end{pmatrix}
\]

6. \[
\begin{pmatrix}
1 & 1 \\
1 & 2
\end{pmatrix}
\]

10. \[
\begin{pmatrix}
1 & 1 & 1 \\
2 & 1 & -1 \\
-8 & -5 & -3
\end{pmatrix}
\]

3. \[
\begin{pmatrix}
-2 & 1 \\
1 & -2
\end{pmatrix}
\]

7. \[
\begin{pmatrix}
1 & i \\
-i & 1
\end{pmatrix}
\]

11. \[
\begin{pmatrix}
1 & 1 & 4 \\
3 & 2 & -1 \\
2 & 1 & -1
\end{pmatrix}
\]

4. \[
\begin{pmatrix}
5 & -1 \\
3 & 1
\end{pmatrix}
\]

8. \[
\begin{pmatrix}
-2 & 1 \\
5 & 4
\end{pmatrix}
\]

12. \[
\begin{pmatrix}
1 & 1 & 2 \\
0 & 2 & 2 \\
-1 & 1 & 3
\end{pmatrix}
\]

13. Show that if \( \lambda \) is an eigenvalue of \( T : V \to V \), then \( \lambda^n \) is an eigenvalue of \( T^n : V \to V \).

14. Let \( T : V \to V \ (\ < \infty) \) be a linear transformation. Show that \( T \) is invertible iff zero is not an eigenvalue of \( T \).

15. Let \( T : \mathbb{R}_2^2 \to \mathbb{R}_2^2 \) be defined by \( T(A) = A^t \). Find the eigenvalues of \( T \).

16. Let \( A \in \mathbb{R}_n^2 \). Prove that if \( n \) is odd, \( A \) has at least one real eigenvalue.

17. Let \( \lambda_1 \) and \( \lambda_2 \) be eigenvalues of \( T : \nu \to V \ (\ dim \nu \ < \infty) \). If \( \dim(E_{\lambda_1}) = k \) and \( \dim(E_{\lambda_2}) = m \), what is \( \dim(E_{\lambda_1} \cap E_{\lambda_2}) \)?

18. Let \( T : \mathbb{R}^2 \to \mathbb{R}^2 \) be the rotation of each point in \( \mathbb{R}^2 \) counterclockwise by 90°.

(a) Find the characteristic polynomial for \( T \).

(b) Determine whether \( T \) is diagonalizable.

19. Diagonalize the matrices from problems 1, 5, and 11 above.

7 Inner Product Spaces

1. Let \( V \) be an IPS. Show that if \( x \in V \), \( x^\perp = \{y : \langle x, y \rangle = 0\} \) is a subspace of \( V \).

2. Find two orthogonal vectors in \( C[0, 1] \) with the inner product \( \int_0^1 f(x)g(x)dx \).
3. Find the norm of $x^2$ in the norm induced by the inner product given in exercise 1.

4. Show that a real normed linear space is an inner product space if and only if its norm satisfies the parallelogram law

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

(a) Show that the parallelogram law holds in an inner product space (use the fact that $\|x\| = \sqrt{\langle x, x \rangle}$ and simplify).

(b) Show that $4\langle x, y \rangle = \|x + y\|^2 - \|x - y\|^2$.

(c) Show that $\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2$ is an inner product if the parallelogram law holds. Hence we can define an inner product in terms of a norm which satisfies the parallelogram law.

5. Show that $\|f\|_1 = \int_0^1 |f(x)|dx$ is a norm on $C[0,1]$ and use the parallelogram law to show that it does not come from an inner product.

6. Prove the Pythagorean theorem for inner product spaces: If $x \perp y$, then $\|x\|^2 + \|y\|^2 = \|x + y\|^2$.

7. Show that if $\|x + y\| = \|x - y\|$, then $x \perp y$.

8. Show that $\langle f, g \rangle = \int_0^\infty e^{-x}f(x)g(x)dx$ is an inner product on $P^2$. Use the Gram-Schmidt procedure on the standard basis for $P^2$ under this inner product to get the first 3 Laguerre polynomials.
Hints and Solutions to Additional Exercises

Matrices and Systems of Equations

1. \[
\begin{pmatrix}
1 & 5 & 4 \\
13 & 19 & 19 \\
42 & 24 & 12
\end{pmatrix}
\]
2. \[
\begin{pmatrix}
13 & 11 & 11 \\
-4 & -7 & 13 \\
23 & 24 & 90
\end{pmatrix}
\]
3. \[
\left\{ \begin{pmatrix}
-1 \\
2 \\
1
\end{pmatrix} t + \begin{pmatrix}
1 \\
-9 \\
0
\end{pmatrix} : t \in \mathbb{R}
\right\}
\]
4. \[
\left\{ \begin{pmatrix}
-2 \\
-6 \\
3
\end{pmatrix} t : t \in \mathbb{R}
\right\}
\]
5. \[
\left\{ \begin{pmatrix}
2 \\
-1 \\
1
\end{pmatrix} s + \begin{pmatrix}
-5 \\
1 \\
0
\end{pmatrix} t + \begin{pmatrix}
8 \\
3 \\
0
\end{pmatrix} : s, t \in \mathbb{R}
\right\}
\]

6. a. \(x_1 = -2/5, x_2 = 1\). b. \(x_1 = 0, x_2 = 11/5\). c. \(x_1 = 2/5, x_2 = 17/5\). d. \(x_1 = 4/5, x_2 = 23/5\).

7. a. \(x = 2, y = 3\). b. \(x = 2, y = -1\). c. \(x = 5, y = -2, z = -1\).

8. \(\det = 5\). Inverse: \[
\begin{pmatrix}
2/5 & 3/5 \\
-9/5 & 11/5
\end{pmatrix}
\]

9. \(\det = 2\). Inverse: \[
\begin{pmatrix}
2 & 1/2 \\
-1 & 0
\end{pmatrix}
\]

10. \(\det = -3\). Inverse: \[
\begin{pmatrix}
4 & -5/3 \\
-1 & 1/3
\end{pmatrix}
\]

11. \(\det(\cdot) = -1\). Inverse: \[
\begin{pmatrix}
2 & 2 & 5 \\
0 & 1 & 1 \\
-1 & -6 & -12
\end{pmatrix}
\]

12. \(\det(\cdot) = 0\). No inverse.

13. \(\det(\cdot) = 1\). Inverse: \[
\begin{pmatrix}
-1 & -13 & 8 \\
1 & 8 & -1
\end{pmatrix}
\]

14. \(\det(\cdot) = 2\). Inverse: \[
\begin{pmatrix}
1 & 1 & 0 & 3/2 \\
0 & 1 & -1 & 1/2 \\
0 & -3 & 4 & 3/2 \\
-1/2 & 1/2 & 1/2 & 1/2
\end{pmatrix}
\]

15. \(\det(\cdot) = -8\). Inverse: \[
\begin{pmatrix}
4 & 1 & 0 & 3 \\
0 & 1 & 0 & 0 \\
-1 & -1/4 & 1/4 & 3/4 \\
-1 & 0 & 0 & -1
\end{pmatrix}
\]

16. \(\det(\cdot) = 0\). No inverse.

17. b. No. 18. \textit{Hint}: Use c to get d. 19. \textit{Hint}: Use problem 18 to get a. Use a to get b. 21. \textit{Hint}: Multiply the two quantities out. 22. \textit{Hint}: \(\det(AB) = \det(A) \det(B)\). 23. \textit{Hint}: How do you calculate the determinant of an upper triangular matrix? 26. \textit{Hint}: Establish that \(A\) and \(B\) are invertible and use 25. The fact that \(\det(AB) = \det(A) \det(B)\) might help.
Vector Spaces

1. **Hint**: Recall that a line through the origin has the form \( \begin{pmatrix} x \\ y \\ z \end{pmatrix} \).

2. b. **Hint**: Consider \( ax + x \). 3. No.

Linear Independence, Spanning Sets, and Bases

1. **Hint**: If \( \det(x_1, x_2, x_3) \neq 0 \), rank\( (x_1, x_2, x_3) = n \)
2. a. \( \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \) b. 4
3. a. possible b. possible.

Linear Transformations

1. \((7, -5, -1)\). 2. a. true b. false. 4. a. \( T \) is not invertible (**Hint**: what is \( \ker T \)).

Matrix Representations of Linear Transformations

1. a. \( [T] = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 1/2 & 2 \\ 0 & 0 & 1/3 \end{pmatrix} \) b. \( 4 + 12x + 8x^2 + x^3 \).

2. a. \( [T] = \begin{pmatrix} 3 & 2 & 0 \\ 0 & 4 & 4 \\ 0 & 0 & 5 \end{pmatrix} \), \( [U] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \), \( [UT] = \begin{pmatrix} 3 & 2 & 5 \\ 0 & 4 & 4 \\ 1 & 0 & 1 \end{pmatrix} \), b. \((4, 4.6)\).

Eigenvectors

1. \( \left( 1, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right), \left( 2, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right) \)
2. \( \left( 1, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right), \left( 1, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \)
3. \( \left( 3, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right), \left( 4, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \)
4. \( \left( 2, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right), \left( 4, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \)
5. \( \left( -3, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right), \left( -3, \begin{pmatrix} 1 \\ -4 \end{pmatrix} \right) \)
6. \( \left( 0, \begin{pmatrix} 1 \\ 4 \end{pmatrix} \right), \left( 2, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \)
7. \( \left( 0, \begin{pmatrix} 1 \\ i \end{pmatrix} \right), \left( 2, \begin{pmatrix} 1 \\ -i \end{pmatrix} \right) \)
8. \( \left( 1, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right), \left( 3, \begin{pmatrix} 1 \\ 5 \end{pmatrix} \right) \)
9. \( \left( 1, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right), \left( 1, \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right), \left( 4, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \)
10. \( \left( -1, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right), \left( -2, \begin{pmatrix} 4 \\ -5 \end{pmatrix} \right), \left( 2, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \)
11. \[ \begin{pmatrix} -2, & \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ 1 \end{pmatrix}, \begin{pmatrix} 1, & \begin{pmatrix} 1 \\ -4 \end{pmatrix} \\ 1 \end{pmatrix}, \begin{pmatrix} 3, & \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ 1 \end{pmatrix} \].
12. \[ \begin{pmatrix} 1, & \begin{pmatrix} 0 \\ -2 \end{pmatrix} \\ 1 \end{pmatrix}, \begin{pmatrix} 2, & \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ 0 \end{pmatrix} \].

13. *Hint:* Use induction. 14. If $T$ is invertible, its determinant is nonzero; consider what this implies about the characteristic polynomial. 15. *Hint:* Recall that if $a + bi$ is an eigenvalue of $T$, so is $a - bi$. 16. 0. 17. a. $\lambda^2 + 1$ b. yes. 18. *Hint:* Recall that eigenvectors corresponding to distinct eigenvalues are linearly independent.

**Inner Product Spaces**

2. e.g. $f(x) = 1$ and $g(x) = 2x - 1$. 3. $1/\sqrt{5}$. 5. To show $\| \cdot \|_1$ does not come from an inner product, consider the functions (two “tents” of area 1):

\[
f(x) = \begin{cases} 16x & x \in [0, 1/4] \\ 8 - 16x & x \in [1/4, 1/2] \\ 0 & x \in [1/2, 1] \end{cases}
\]

\[
g(x) = \begin{cases} 0 & x \in [0, 1/2] \\ 16x - 8 & x \in [1/2, 3/4] \\ 16 - 16x & x \in [3/4, 1] \end{cases}
\]

6. *Hint:* Expand the norms and simplify. 8. $\{1, 1 - x, x^2 - 4x + 2\}$.