These notes are adapted from lecture notes taught by Dr. Alan Thompson and from “Elementary Linear Algebra: 10th Edition” by Howard Anton. Picture above sourced from (http://i.imgur.com/RgmnA.gif)
# Table of Contents

Chapter 3 – Euclidean Vector Spaces

3.1 – Vectors in 2-space, 3-space, and n-space

   - Theorem 3.1.1 – Algebraic Vector Operations without components
   - Theorem 3.1.2

3.2 – Norm, Dot Product, and Distance

   - Definition 1 – Norm of a Vector
   - Definition 2 – Distance in \( \mathbb{R}^n \)

   - Dot Product

   - Definition 3 – Dot Product

   - Definition 4 – Dot Product, Component by component

   - Theorems 3.2.2 and 3.2.3

Geometry in \( \mathbb{R}^n \)

   - Theorem 3.2.5 – Triangle Inequality
   - Theorem 3.2.6
   - Theorem 3.2.7

   - Dot Products as Matrix Multiplication

3.3 – Orthogonality

   - Point-Normal forms of lines and planes

   - Theorem 3.3.1 – Point-Normal Forms of Lines and Planes

Orthogonal Projections

   - Theorem 3.3.3 – Pythagorean Theorem in \( \mathbb{R}^n \)

   - Magnitudes of projections

Distance Problems

   - Theorem 3.3.4

3.4 – Geometry of Linear Systems

   - Vector & Parametric Equations of Lines and Planes

   - Theorem 3.4.1

   - Theorem 3.4.2

   - Dot Product form of a Linear System

3.5 – Cross Product

   - Theorem 3.5.1 – Relationships between Cross Product and Dot Product

   - Theorem 3.5.2 – Properties of Cross Product

   - Magnitude of a Cross Product

   - Geometric Interpretation of Determinants

Chapter 4 – General Vector Spaces

4.1 – Real Vector Spaces

   - Vector Space Axioms

   - Definition 1 – Vector Space Axioms

   - Theorem 4.1.1

4.2 – Subspaces

   - Definition 1

   - Theorem 4.2.1
Building Subspaces

Theorem 4.2.2
Definition 2
Theorem 4.2.3
Definition 3
Theorem 4.2.4
Theorem 4.2.5

4.3 – Linear Independence

Definition 1
Theorem 4.3.1
Theorem 4.3.2

Geometric Interpretation of Linear Independence

Theorem 4.3.3

4.4 – Co-ordinates and Basis

Definition 1
Theorem 4.4.1 – Uniqueness of Basis Representation
Definition 2

4.5 – Dimension

Number of Vectors in a Basis
Theorem 4.5.1
Theorem 4.5.2
Definition 1
Example 1
Example 2 – Dimension of a solution space

Plus / Minus Theorem

Theorem 4.5.3 – The Plus/Minus Theorem
Theorem 4.5.4
Theorem 4.5.5
Theorem 4.5.6

4.6 – Change of Basis

Change of Basis Problem (The Rubbish Way of Solving for the Transition Matrix)
Theorem 4.6.1
Solving for the transition matrix quickly
Theorem 4.6.2

4.7 – Row Space, Column Space, and Null Space

Definition 2
Theorem 4.7.1
Theorem 4.7.2

Bases for Row, Column, and Null Spaces
Theorem 4.7.3
Theorem 4.7.4
Theorem 4.7.5
Example

Bases Formed from Row and Column Vectors of a Matrix
Problem

4.8 – Rank, Nullity, and the Fundamental Matrix Spaces

Theorem 4.8.1
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank and Nullity</td>
<td>27</td>
</tr>
<tr>
<td><strong>Definition 1</strong></td>
<td>27</td>
</tr>
<tr>
<td>Maximum Value for Rank</td>
<td>27</td>
</tr>
<tr>
<td><strong>Theorem 4.8.2 – Dimension theorem for Matrices</strong></td>
<td>28</td>
</tr>
<tr>
<td><strong>Theorem 4.8.3</strong></td>
<td>28</td>
</tr>
<tr>
<td><strong>Theorem 4.8.4 – Invertible Matrix Theorem</strong></td>
<td>28</td>
</tr>
<tr>
<td>Overdetermined and Underdetermined Systems</td>
<td>28</td>
</tr>
<tr>
<td><strong>Theorem 4.8.5</strong></td>
<td>28</td>
</tr>
<tr>
<td><strong>Theorem 4.8.6</strong></td>
<td>28</td>
</tr>
<tr>
<td><strong>The Fundamental Spaces of a Matrix</strong></td>
<td>29</td>
</tr>
<tr>
<td><strong>Theorem 4.8.7</strong></td>
<td>29</td>
</tr>
<tr>
<td><strong>Definition 2</strong></td>
<td>29</td>
</tr>
<tr>
<td><strong>Theorem 4.8.8</strong></td>
<td>29</td>
</tr>
<tr>
<td><strong>A Geometric Link Between the Fundamental Spaces</strong></td>
<td>30</td>
</tr>
<tr>
<td><strong>Theorem 4.8.9</strong></td>
<td>30</td>
</tr>
<tr>
<td><strong>4.9 – Matrix Transformations From Rn to Rm</strong></td>
<td>31</td>
</tr>
<tr>
<td>Functions and Transformations</td>
<td>31</td>
</tr>
<tr>
<td><strong>Definition 1</strong></td>
<td>31</td>
</tr>
<tr>
<td>Notational Matters</td>
<td>32</td>
</tr>
<tr>
<td><strong>Properties of Matrix Transformations</strong></td>
<td>32</td>
</tr>
<tr>
<td><strong>Theorem 4.9.1</strong></td>
<td>32</td>
</tr>
<tr>
<td><strong>Theorem 4.9.2</strong></td>
<td>32</td>
</tr>
<tr>
<td><strong>A Procedure for Finding Standard Matrices</strong></td>
<td>33</td>
</tr>
<tr>
<td>Finding the Standard Matrix for a Matrix Transformation</td>
<td>33</td>
</tr>
<tr>
<td><strong>Reflection Operators</strong></td>
<td>33</td>
</tr>
<tr>
<td><strong>Projection Operators</strong></td>
<td>33</td>
</tr>
<tr>
<td>Rotation Operators</td>
<td>34</td>
</tr>
<tr>
<td><strong>4.10 – Properties of Matrix Transformations</strong></td>
<td>34</td>
</tr>
<tr>
<td>Compositions of Matrix Transformations</td>
<td>34</td>
</tr>
<tr>
<td><strong>One–to–One Matrix Transformations</strong></td>
<td>35</td>
</tr>
<tr>
<td><strong>Definition 1</strong></td>
<td>35</td>
</tr>
<tr>
<td><strong>Theorem 4.10.1</strong></td>
<td>35</td>
</tr>
<tr>
<td><strong>Inverse of a one-to-one matrix operator</strong></td>
<td>35</td>
</tr>
<tr>
<td><strong>Linearity Properties</strong></td>
<td>36</td>
</tr>
<tr>
<td>Problem</td>
<td>36</td>
</tr>
<tr>
<td><strong>Theorem 4.10.2</strong></td>
<td>36</td>
</tr>
<tr>
<td><strong>Theorem 4.10.3</strong></td>
<td>36</td>
</tr>
<tr>
<td><strong>Chapter 5 – Eigenvalues and Eigenvectors</strong></td>
<td>37</td>
</tr>
<tr>
<td><strong>5.1 – Eigenvalues and Eigenvectors</strong></td>
<td>37</td>
</tr>
<tr>
<td><strong>Definition 1</strong></td>
<td>37</td>
</tr>
<tr>
<td><strong>Theorem 5.1.1</strong></td>
<td>37</td>
</tr>
<tr>
<td><strong>Theorem 5.1.2</strong></td>
<td>37</td>
</tr>
<tr>
<td><strong>Theorem 5.1.3</strong></td>
<td>37</td>
</tr>
<tr>
<td><strong>Bases for Eigenspaces</strong></td>
<td>37</td>
</tr>
<tr>
<td><strong>Powers of a Matrix</strong></td>
<td>38</td>
</tr>
<tr>
<td><strong>Theorem 5.1.4</strong></td>
<td>38</td>
</tr>
<tr>
<td><strong>Eigenvalues and Invertibility</strong></td>
<td>38</td>
</tr>
</tbody>
</table>

4/52
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle between Vectors</td>
<td>47</td>
</tr>
<tr>
<td>Properties of Length and Distance in General Inner Product</td>
<td>47</td>
</tr>
<tr>
<td>Spaces</td>
<td></td>
</tr>
<tr>
<td>Theorem 6.2.2 – Triangle Inequality</td>
<td>47</td>
</tr>
<tr>
<td>Orthogonality</td>
<td>48</td>
</tr>
<tr>
<td>Theorem 6.2.3 – Generalized Theorem of Pythagoras</td>
<td>48</td>
</tr>
<tr>
<td>Orthogonal Complements</td>
<td>48</td>
</tr>
<tr>
<td>Theorem 6.2.4</td>
<td></td>
</tr>
<tr>
<td>Theorem 6.2.5</td>
<td></td>
</tr>
<tr>
<td>6.3 – Gram-Schmidt Process</td>
<td>49</td>
</tr>
<tr>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td>Theorem 6.3.1</td>
<td>49</td>
</tr>
<tr>
<td>Theorem 6.3.2</td>
<td></td>
</tr>
<tr>
<td>Orthogonal Projections</td>
<td>50</td>
</tr>
<tr>
<td>Theorem 6.3.3</td>
<td></td>
</tr>
<tr>
<td>Theorem 6.3.4</td>
<td></td>
</tr>
<tr>
<td>Theorem 6.3.5</td>
<td></td>
</tr>
<tr>
<td>The Gram-Schmidt Process</td>
<td>50</td>
</tr>
<tr>
<td>Chapter 7 – Diagonalization</td>
<td>51</td>
</tr>
<tr>
<td>7.1 – Orthogonal Matrices</td>
<td>51</td>
</tr>
<tr>
<td>Definition 1</td>
<td>51</td>
</tr>
<tr>
<td>Theorem 7.1.1</td>
<td>51</td>
</tr>
<tr>
<td>Theorem 7.1.2</td>
<td>51</td>
</tr>
<tr>
<td>Orthogonal Matricies as Linear Operators</td>
<td>51</td>
</tr>
<tr>
<td>Theorem 7.1.3</td>
<td>51</td>
</tr>
<tr>
<td>Change of Orthonormal Basis</td>
<td>52</td>
</tr>
<tr>
<td>Theorem 7.1.4</td>
<td>52</td>
</tr>
<tr>
<td>Theorem 7.1.5</td>
<td>52</td>
</tr>
<tr>
<td>7.2 – Orthogonal Diagonalization</td>
<td>52</td>
</tr>
<tr>
<td>Definition 1</td>
<td>52</td>
</tr>
<tr>
<td>Conditions for Orthogonal Diagonalizability (Theorem 7.2.1)</td>
<td>52</td>
</tr>
<tr>
<td>Orthogonally Diagonalizing an nxn Symmetric Matrix</td>
<td>52</td>
</tr>
</tbody>
</table>
Chapter 3 – Euclidean Vector Spaces

3.1 – Vectors in 2-space, 3-space, and n-space

Theorem 3.1.1 – Algebraic Vector Operations without components

- \( U + v = v + u \)
- \( (u+v)+w = u + (v+w) \)
- \( u + 0 = 0 + u = u \)
- \( u + (-u) = 0 \)
- \( k(u + v) = ku + kv \)
- \( (k + m)u = ku + mu \)
- \( k(mu) = (km)u \)
- \( 1u = u \)

Theorem 3.1.2

- \( 0v = 0 \)
- \( k0 = 0 \)
- \( (-1)v = -v \)

3.2 – Norm, Dot Product, and Distance

Definition 1 – Norm of a Vector

\[
\|v\| = \sqrt{v_1^2 + v_2^2 + v_3^2 + \ldots + v_n^2}
\]

Definition 2 – Distance in \( \mathbb{R}^n \)

Given vectors \( u, v \), the distance \( d(u,v) \) is:

\[
d(u,v) = \|u - v\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}
\]
Dot Product

If \( \mathbf{u}, \mathbf{v} \) start at the same point and we want to find the angle between the two vectors such that \( 0 \leq \theta \leq \pi \), we use the dot product:

**Definition 3 – Dot Product**

\[
\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta
\]

Manipulating above for the cosine term, we see that

\[
\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}
\]

And since \( 0 \leq \theta \leq \pi \),

- \( \theta \) is acute if \( \mathbf{u} \cdot \mathbf{v} > 0 \)
- \( \theta \) is obtuse if \( \mathbf{u} \cdot \mathbf{v} < 0 \)
- \( \theta = \frac{\pi}{2} \) if \( \mathbf{u} \cdot \mathbf{v} = 0 \)

If we look at the dot product, *component-by-component*, we arrive at

**Definition 4 – Dot Product, Component by component**

\[
\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n
\]

Relating the dot product to the norm, we arrive at

\[
\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}
\]

*Theorem's 3.2.2 and 3.2.3* deal with the properties of the dot product.

**Theorems 3.2.2 and 3.2.3**

- \( \mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u} \) (Symmetry)
- \( \mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} \) (Distributive)
- \( k (\mathbf{u} \cdot \mathbf{v}) = k \mathbf{u} \cdot \mathbf{v} \) (Homogeneity)
- \( \mathbf{v} \cdot \mathbf{v} \geq 0 \) and \( \mathbf{v} \cdot \mathbf{v} = 0 \iff \mathbf{v} = 0 \) (Positivity)
- See theorem 3.2.3 on page 136 for more of the same.

If we solve the formula for the dot product for theta, shown below, we notice that the argument for arccos must be between \([-1, 1]\). This is called the **Cauchy – Schwarz Inequality**.

\[
\theta = \arccos \left( \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \right)
\]
Geometry in $\mathbb{R}^n$

Knowing that the shortest distance between two points is a straight line, we find the **triangle inequality**.

**Theorem 3.2.5 – Triangle Inequality**

- $\|u + v\| \leq \|u\| + \|v\|$
- $d(u, v) \leq d(u, w) + d(w, v)$

On page 138, **Theorem 3.2.6** is pretty funky.

*It is proved in plane geometry that for any parallelogram, the sum of the squares of the diagonals is equal to the sum of the squares of the four sides.*

**Theorem 3.2.6**

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$$

**Theorem 3.2.7**

$$u \cdot v = \frac{1}{4} ||u + v||^2 - \frac{1}{4} ||u - v||^2$$

**Dot Products as Matrix Multiplication**

There is a massive table on page 139. I’m not sure if I need to know these properties.
3.3 – Orthogonality

If \( u \cdot v = 0 \), \( u \) and \( v \) are orthogonal.

**Point-Normal forms of lines and planes.**

To find the equation of a line or plane, we take an arbitrary point \( P_0 = (X_o, Y_o, Z_o) \), and another point, \( P(x,y,z) \). We form a vector \( \vec{P_0P} = (x-x_0, y-y_0, z-z_0) \).

Then we know that the normal must be orthogonal to this vector (and the plane/line), so that \( n \cdot \vec{P_0P} = 0 \). If this normal \( n \), is defined as \( n = (a, b, c) \), then the above equation becomes (by the component dot product):

\[
a(x-x_0) + b(y-y_0) + c(z-z_0) = 0
\]

[Point-Normal Form of a Plane]

The above equation, \( a(x-x_0) + b(y-y_0) + c(z-z_0) = 0 \) [Point-Normal Form of a Plane], can be simplified. If we multiply the terms out and simplify we get the following theorem.

**Theorem 3.3.1 – Point-Normal Forms of Lines and Planes**

- \( Ax + By + C = 0 \), is a line in \( \mathbb{R}^2 \) with normal \( n=(a,b) \).
- \( Ax + By + Cz + D = 0 \), is a plane in \( \mathbb{R}^3 \) with normal \( n=(a, b, c) \).
Orthogonal Projections.

If we have a vector \( \mathbf{v} \) in \( \mathbb{R}^2 \) expressed as a point \((x,y)\) that is \textit{not} a standard vector, we can break it up into components. Typically these are \(x\)-and-\(y\) components.

- Construct the vector \( \mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1 \).

\[
\begin{align*}
\text{(a)} & \quad \mathbf{u} \\
\text{(b)} & \quad \mathbf{w}_2 \\
\text{(c)} & \quad \mathbf{w}_1 \\
\text{(d)} & \quad \mathbf{w}_2
\end{align*}
\]

Figure 3.3.2 In parts (b) through (d), \( \mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 \), where \( \mathbf{w}_1 \) is parallel to \( \mathbf{a} \) and \( \mathbf{w}_2 \) is orthogonal to \( \mathbf{a} \).

Since

\[
\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w}_1 + (\mathbf{u} - \mathbf{w}_1) = \mathbf{u}
\]

we have decomposed \( \mathbf{u} \) into a sum of two orthogonal vectors, the first term being a scalar multiple of \( \mathbf{a} \) and the second being orthogonal to \( \mathbf{a} \).

Let’s call the \(x\)-axis \( \mathbf{a} \). Another way to think of the above is that \( \mathbf{v} \) can be broken up into \( \mathbf{w}_1 \) and \( \mathbf{w}_2 \), both of which are perpendicular to each other. If \( \mathbf{w}_1 \) is along \( \mathbf{a} \), we know that \( \mathbf{w}_1 = k \mathbf{a} \), and that

\[
\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{v}.
\]

Follow the Proof on Page 146 to yield:

1. \( \text{proj}_a \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} \) \text{ [Component of } \mathbf{u} \text{ along } \mathbf{a}] \]
2. \( \mathbf{u} - \text{proj}_a \mathbf{u} = \mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} \) \text{ [Component of } \mathbf{u} \text{ orthogonal to } \mathbf{a}] \]

3. To remember the above, think of \( \mathbf{1} \) as \( k \mathbf{a} \), as described in the proof.
4. To remember \( \mathbf{2} \), remember in the proof that \( \mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w}_1 + (\mathbf{u} - \mathbf{w}_1) = \mathbf{u} \). If \( \mathbf{w}_1 = \text{proj}_a \mathbf{U} \), then the remaining bit must be the \textbf{orthogonal} component.

Theorem 3.3.3 – Pythagorean Theorem in \( \mathbb{R}^n \)

\[
\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2
\]

Magnitudes of projections

\[
\|\text{proj}_a \mathbf{u}\| = \frac{\|\mathbf{u} \cdot \mathbf{a}\|}{\|\mathbf{a}\|} \quad \text{And} \quad \|\text{proj}_a \mathbf{u}\| = \|\mathbf{u}\| |\cos \theta|\]
Distance Problems

We want to find the shortest distance between an arbitrary point \( P_0 \), and a plane \( Q \). Since we know the plane has a normal \( n \), we want to take the vector \( \overrightarrow{QP_0} \) and project it onto the normal \( n \).

![Diagram of Distance from \( P_0 \) to Plane]

Notice how the line D, the projection, is transmissible?

**Proof of Theorem 3.3.4**

\[
D = \| \text{proj}_n \overrightarrow{QP_0} \| = \frac{|\overrightarrow{QP_0} \cdot n|}{\| n \|}
\]

Recall that the numerator is the slope-point equation of a plane, and the bottom is simply the magnitude of a 3-dimensional vector.

\[
D = \frac{|a(x_0 - x) + b(y_0 - y) + c(z_0 - z)|}{\| n \|} = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}
\]

**Theorem 3.3.4**

The Distance \( D \) between a plane and a point \( P_0 = (x_0, y_0, z_0) \) is

\[
D = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}
\]
3.4 – Geometry of Linear Systems

Vector & Parametric Equations of Lines and Planes

Theorem 3.4.1
The equation of a line through \( \mathbf{x}_0 \) that is parallel to \( \mathbf{v} \) is
\[
\mathbf{x} = \mathbf{x}_0 + t \mathbf{v}
\]

Theorem 3.4.2
We require two non-collinear vectors to express a plane.

The equation of a plane through \( \mathbf{x}_0 \) that is parallel to \( \mathbf{v}_1 \) and \( \mathbf{v}_2 \) is
\[
\mathbf{x} = \mathbf{x}_0 + t_1 \mathbf{v}_1 + t_2 \mathbf{v}_2
\]
This plane is parallel to both \( \mathbf{v}_1 \) and \( \mathbf{v}_2 \).

We can find the equation of a line through two points by taking \( \mathbf{v} = \mathbf{x}_1 - \mathbf{x}_0 \) above. This is equivalent to
\[
\mathbf{x} = (1-t) \mathbf{x}_0 + t \mathbf{x}_1,
\] page 155.

Dot Product form of a Linear System

If we solve a homogenous system \( \mathbf{A} \mathbf{x} = \mathbf{0} \), then the solution set of that system is orthogonal to every row vector of \( \mathbf{A} \), **Theorem 3.4.3**.
3.5 – Cross Product.

The cross product is defined only in 3-space, for this course, as

\[ \mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \]

**Theorem 3.5.1 -- Relationships between Cross Product and Dot Product.**

- \( \mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = 0 \) (\( \mathbf{u} \times \mathbf{v} \) is orthogonal to \( \mathbf{u} \))
- \( \mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0 \) (\( \mathbf{u} \times \mathbf{v} \) is orthogonal to \( \mathbf{v} \))
- \( \|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2 \) (Lagrange's Identity)
- \( \mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w}) \mathbf{v} - (\mathbf{u} \cdot \mathbf{v}) \mathbf{w} \)
- \( (\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \cdot \mathbf{w}) \mathbf{v} - (\mathbf{v} \cdot \mathbf{w}) \mathbf{u} \)

**Theorem 3.5.2 – Properties of Cross Product**

- \( \mathbf{u} \times \mathbf{v} = - (\mathbf{v} \times \mathbf{u}) \)
- \( \mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w}) \)
- \( (\mathbf{u} + \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \times \mathbf{w}) + (\mathbf{v} \times \mathbf{w}) \)
- \( k(\mathbf{u} \times \mathbf{v}) = (k \mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (k \mathbf{v}) \)
- \( \mathbf{u} \times 0 = 0 \times \mathbf{u} = 0 \)
- \( \mathbf{u} \times \mathbf{u} = 0 \)

**Magnitude of a Cross Product**

\( \|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin \theta \)

If \( \mathbf{u} \) and \( \mathbf{v} \) are vectors in 3-space, then the above is equal to the **area of the parallelogram** formed by \( \mathbf{u} \) and \( \mathbf{v} \)

**Scalar Triple Product**

\[ \mathbf{a} \cdot (\mathbf{u} \times \mathbf{v}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \]

**Geometric Interpretation of Determinants.**

**Theorem 3.5.4**

- The value of \( \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \) is the **area of the parallelogram** in 2-space.

- The value of \( \mathbf{a} \cdot (\mathbf{u} \times \mathbf{v}) \) is the **volume of the parallelepiped** in 3-space formed by \( \mathbf{a}, \mathbf{u}, \mathbf{v} \).
Chapter 4 – General Vector Spaces

4.1 – Real Vector Spaces

Vector Space Axioms

**Definition 1 – Vector Space Axioms**

V is an arbitrary set of objects on which addition and scalar multiplication are defined. If all of the following axioms are satisfied by \( u, v, w \) in V, and all scalars \( k \) and \( m \), then V is a **vector space**.

1. If \( u \) and \( v \) are objects in V, then \( u + v \) is in V
2. \( u + v = v + u \)
3. \( u + (v + w) = (u + v) + w \)
4. There is an object \( 0 \) in V, called the **zero vector** for V, such that \( 0 + u = u + 0 = u \) for all \( u \).
5. For each \( u \) in V, the **negative** of \( u \), \(-u\) exists such that \( u + (-u) = (-u) + u = 0 \).
6. If \( k \) is any scalar and \( u \) is any object, \( ku \) is in V
7. \( k(u + v) = ku + kv \)
8. \( (k+m)u = ku + mu \)
9. \( k(mu) = (km)u \)
10. \( 1u = u \)

To verify a set with two operations is a vector space, first check **axioms 1 and 6**. Then check the rest.

**Theorem 4.1.1**

- \( 0u = 0 \)
- \( k0 = 0 \)
- \( (-1)u = -u \)
- If \( ku = 0 \), then \( k = 0 \) or \( u = 0 \).
4.2 – Subspaces

Definition 1
A subset \( W \) of a vector space \( V \) is a **subspace** of \( V \) if \( W \) is itself a vector space using the addition and scalar multiplication defined for \( V \).

Most axioms are inherited by \( W \), except for the following:

- Axiom 1 – Closure of \( W \) under addition
- Axiom 4 – Existence of a zero vector in \( W \)
- Axiom 5 – Existence of a negative in \( W \) for every vector in \( W \)
- Axiom 6 – Closure of \( W \) under scalar multiplication.

Theorem 4.2.1
If axioms 1 and 6 hold, axioms 4 and 5 hold, so you only need to **prove axioms 1, and 6, to verify if \( W \) is a valid subspace**.

1) If \( u \) and \( v \) are objects in \( V \), then \( u + v \) is in \( V \)

6) If \( k \) is any scalar and \( u \) is any object, \( ku \) is in \( V \)

Building Subspaces

Theorem 4.2.2
If \( W_1, W_2, ..., W_r \) are subspaces of \( V \), then the intersection of these subspaces is also a subspace of \( V \).

Definition 2
If \( w \) is a vector in a vector space \( V \), then \( w \) is a **linear combination** of the vectors \( v_1, v_2, ..., v_r \) in \( V \) if it can be expressed in the form

\[ w = k_1 v_1 + k_2 v_2 + ... + k_r v_r \]

where the \( k \)'s are scalars, called the **coefficients** of the linear combination.

Theorem 4.2.3
If \( S = \{ w_1, w_2, ..., w_r \} \) is a non-empty set of vectors in a vector space \( V \), then

- The set \( W \) of all possible linear combinations of the vectors in \( S \) is a subspace of \( V \)
- The set \( W \) in part (a) is the “smallest” subspace of \( W \) that contains all of the vectors in \( S \) in the sense that any other subspace that contains those vectors contains \( W \).
Definition 3
The subspace of a vector space V that is formed from all possible linear combinations of the vectors of a nonempty set S is called the span of S.

See page 186, Example 15 to test for spanning.

Theorem 4.2.4
The solution set of a homogeneous linear system \(Ax = 0\) in n-unknowns is a subspace of \(R^n\).

Theorem 4.2.5
If \(S = \{v_1, v_2, \ldots, v_r\}\), and \(S' = \{w_1, w_2, \ldots, w_r\}\), then
\[
\text{span}\{v_1, v_2, \ldots, v_r\} = \text{span}\{w_1, w_2, \ldots, w_r\}
\]
iff each vector in S is a linear combination of those in S', and each vector in S' is a linear combination of those in S.

4.3 – Linear Independence

Definition 1
Suppose \(S = \{v_1, v_2, \ldots, v_r\}\) is a nonempty set of vectors in a vector space V. If the equation
\[
k_1v_1 + k_2v_2 + \ldots + k_rv_r = 0
\]
has
- only the trivial solution, then S is a linearly independent set
- additional solutions, then S is a linearly dependent set.

Protip from Josh: Take a look at the lab on linear independence. In the RREF, only the vectors without free variables form the span of the subspace. The columns that are free-variables are linearly dependent. Check the lab manual to make sure this statement is correct.

Theorem 4.3.1
A set S with two or more vectors is
- Linearly Dependent iff at least one of the vectors in S is expressible as a linear combination of the other vectors in S.
- Linearly Independent iff no vector in S is expressible as a linear combination of a the other vectors in S.
Example 7 in the text is useless. See the lab on Linear Independence to see how to do these questions.

**Theorem 4.3.2**

- A finite set that contains 0 is linearly independent
- A set with exactly one vector is linearly independent iff that vector is not 0
- A set with exactly two vectors is linearly independent iff neither vector is a scalar multiple of the other.

**Geometric Interpretation of Linear Independence.**

In $\mathbb{R}^2$ and $\mathbb{R}^3$:

- Two vectors are linearly independent iff they do not lie on the same line when they have their initial points at the origin. Otherwise one would be a scalar multiple of the other.
  - In other words, if vectors $\mathbf{a}$ and $\mathbf{b}$ are defined such that
    $$\mathbf{a} = (0,0,0) + s(2,3,4) \quad \mathbf{b} = (0,0,0) + t(2,3,4)$$
    Then they are linearly dependent as they lie on the same line, $(2,3,4)$, but are scalar multiples of each other.
- Three vectors in 3-space are linearly independent iff they do not lie in the same plane when they have their initial points at the origin.

**Theorem 4.3.3**

Let $S = \{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \}$ be a set of vectors in $\mathbb{R}^n$. If $r > n$, then $S$ is linearly dependent.

**4.4 – Co-ordinates and Basis**

Typically we use a rectangular co-ordinate system in 2-space and 3-space. This is not necessary, as the co-ordinate axes can be skewed. We define the co-ordinate axes as a **basis**.

**Definition 1**

If $V$ is any vector space and $S = \{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \}$ is a finite set of vectors in $V$, then $S$ is a **basis** for $V$ if the following conditions hold

- $S$ is linearly independent
- $S$ spans $V$

**Theorem 4.4.1 – Uniqueness of Basis Representation**

If $S = \{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \}$ is a basis for a vector space $V$, then every vector $\mathbf{v}$ in $V$ can be expressed in the form $\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \ldots + c_n \mathbf{v}_n$ in exactly one way.
Definition 2

If \( S = \{v_1, v_2, \ldots, v_r\} \) is a basis for the vector space \( V \), and
\[
v = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n
\]
is the expression for a vector \( v \) in terms of a basis \( S \), then the scalars \( c_i \) are the co-
ordinates of \( v \) relative to the basis \( S \). The vector \( (c_1, c_2, \ldots, c_n) \) in \( \mathbb{R}^n \) constructed from
these co-ordinates is called the co-ordinate vector of \( v \) relative to \( S \), denoted by
\[
(v)_s = (c_1, c_2, \ldots, c_n)
\]

In example 9 (page 206), a method for finding co-ordinate vectors relative to another basis is
shown. It involves solving the system of equations
\[
v = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n
\]
for \( (c_1, c_2, \ldots, c_n) \), where \( v \) is given & forming an augmented matrix.

4.5 – Dimension.

Number of Vectors in a Basis

Theorem 4.5.1

All bases for a finite-dimensional vector space have the same number of vectors.

Theorem 4.5.2

Let \( V \) be a finite-dimensional vector space, and let \( S = \{v_1, v_2, \ldots, v_n\} \) be any basis.

- If a set has more than \( n \) vectors, then it is linearly dependent
- If a set has fewer than \( n \) vectors, it does not span \( V \).

Definition 1

The dimension of a finite-dimensional vector space \( V \) is denoted by \( \dim(V) \) and is defined to be
the number of vectors in a basis for \( V \). In addition, the zero vector space is defined to have a
dimension zero.

Example 1

\[
\dim(\mathbb{R}^n) = n \quad \dim(P_n) = n + 1 \quad \dim(M_{mn}) = mn
\]

If we have a set \( S = \{v_1, v_2, \ldots, v_n\} \), which is linearly independent, then \( S \) is automatically a
basis for \( \text{span}(S) \). \( \dim(\text{span}(s)) = n \)
Example 2 – Dimension of a solution space.

On page 210, we solve the given system and find that:

\[(x_1, x_2, x_3, x_4, x_5) = (-s-t, s, -t, 0, t) = s(-1, 1, 0, 0, 0) + t(-1, 0, -1, 0, 1)\]

Since there are two parameters, the vectors \(s\) and \(t\) span the solution space. Since they are linearly independent and span the solution space, they form a basis with a dimension of 2.

**Plus / Minus Theorem.**

If we have a set of vectors \(S\), we can add and remove vectors from this set. If we begin with \(S\) being linearly independent, and add another vector that is not a linear combination of the ones in \(S\), \(S'\) remains linearly independent.

If we start with a set \(S\) that has one vector that is linearly dependent on another one, we can remove the linearly dependent vector, and leave only the independent vectors. This set, \(S'\), containing only the linearly independent vectors has the same span as \(S\), \(\text{span}(S') = \text{span}(S)\).

**Theorem 4.5.3 – The Plus/Minus Theorem.**

The Plus/Minus Theorem states the above results very formally.

- If \(S\) is a linearly independent set, and if \(v\) is a vector in \(V\) that is outside of \(\text{span}(S)\), then the set \(S \cup \{v\}\) that results by inserting \(V\) into \(S\) is still linearly independent.
- If \(v\) is a vector in \(S\) that is expressible as a linear combination of the other vectors in \(S\), and if 

  \[S - \{v\}\] 
  denotes the set obtained by removing \(v\) from \(S\), then \(S - \{v\}\) span the same space, 

  \[\text{span}(S) = \text{span}(S - \{v\})\]

If we want to show that a set of vectors \(\{v_1, v_2, ..., v_n\}\) is a basis for \(V\), we must show that the vectors are **linearly independent** and **span** \(V\). If we know that \(V\) has a dimension of \(n\), so that \(\{v_1, v_2, ..., v_n\}\) contains \(n\) entities, then:

- Check for either **linear independence** or **spanning**
**Theorem 4.5.4**

Let \( V \) be an \( n \)-dimensional vector space, and let \( S \) be a set in \( V \) with exactly \( n \) vectors. Then \( S \) is a basis for \( V \) if and only if \( S \) spans \( V \) or \( S \) is linearly independent.

The next theorem shows two important facts about vectors in a finite-dimensional space \( V \):

1. Every spanning set for a subspace is either a basis for that subspace, or has a basis as a subset.
2. Every linearly independent set in a subspace is either a basis for that subspace or can be extended to a basis for it.

Diagrammatically showing number 1:

\[
\begin{array}{c}
\text{..................................................} \\
\text{Basis} \\
\text{..................................................} \\
\end{array} = \text{Spanning Set for a subspace}
\]

Part two is somewhat of a tautology. Basically you have a linearly independent subset that is a basis, just like 1. Or, you can extend that “incomplete” basis to form a proper basis.

**Theorem 4.5.5**

- If \( S \) spans \( V \) but is not a basis for \( V \), then \( S \) can be reduced to a basis for \( V \) by removing appropriate vectors from \( S \).
- If \( S \) is a linearly independent set that is not already a basis for \( V \), \( S \) can be enlarged to a basis for \( V \) by adding appropriate vectors into \( S \).

**Theorem 4.5.6**

If \( W \) is a subspace of a finite-dimensional vector space \( V \),

- \( W \) is finite-dimensional
- \( \dim(W) \leq \dim(V) \)
- \( W = V \) iff \( \dim(W) = \dim(V) \)
4.6 – Change of Basis

We need a way to map co-ordinates between a vector space $V$ to $\mathbb{R}^n$. We do this by defining notation called a co-ordinate map. (Page 217 for notation)

Change of Basis Problem (The Rubbish Way of Solving for the Transition Matrix)

If $v$ is in $V$, and we change the basis for $V$ from $B$ to $B'$, how is $[v]_B$ related to $[v]_{B'}$?

**Solution:** Solve this problem in two dimensions first.

Let the following be the old and new bases.

$B = \{ u_1, u_2 \}$  $B' = \{ u'_1, u'_2 \}$

We are interested in forming a Transition Matrix from $B$ to $B'$.

Take $B' = \{ u'_1, u'_2 \}$, and find how each cell is related to the vectors in $B = \{ u_1, u_2 \}$. That is, identify $a, b, c, d$ in:

$u'_1 = au_1 + bu_2$

$u'_2 = cu_1 + du_2$

Take the cells $a, b$, and form it into a column $\begin{bmatrix} a \\ b \end{bmatrix}$ of the transition matrix. Take the cells $c, d$ and form it into a column $\begin{bmatrix} c \\ d \end{bmatrix}$ of the transition matrix.

Therefore, the Transition Matrix is defined as:

$P_{B' \rightarrow B} = \begin{bmatrix} [u'_1]_B & [u'_2]_B & \ldots & [u'_n]_B \\ [u'_1]_{B'} & [u'_2]_{B'} & \ldots & [u'_n]_{B'} \end{bmatrix}$

The column vectors of the transition matrix from old -> new are the co-ordinate vectors of old relative to new.

However, the above method is complete rubbish, you’ll be pretty scuppered on an exam trying to find $a, b, c, d$ in $u'_1 = au_1 + bu_2$

$u'_2 = cu_1 + du_2$.

Theorem 4.6.1 is related to the next procedure.

**Theorem 4.6.1**

If $P$ is the transition matrix from a basis $B'$ to a basis $B$ for a finite-dimensional vector space $V$, then $P$ is invertible and $P^{-1}$ is the transition matrix from $B$ to $B'$. 

22/52
Solving for the transition matrix quickly.

- Form the matrix \([B’ \mid B]\)
- RREF
- The result is \([I \mid P_{B \rightarrow B’}]\)

\[
\begin{bmatrix}
\text{New Basis} & \text{Old Basis}
\end{bmatrix} \rightarrow \text{RREF} \rightarrow \begin{bmatrix} I & P_{\text{old} \rightarrow \text{New}} \end{bmatrix}
\]

In Example 3, Page 221, there was a change of basis from \(B = \{e_1, e_2\}\) to \(B’ = \{(1,1), (2,1)\}\). After solving for the transition using the method above, it was found that

\[
P_{B’ \rightarrow B} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}
\]

When we convert from the standard basis to another basis, the transition matrix is the basis \(B’\), as the left-side is already the identity matrix.

**Theorem 4.6.2**

In less formal words than the textbook (page 222),

When we convert from the standard basis to another basis, the transition matrix is the basis \(B’\), as the left-side is already the identity matrix.

### 4.7 – Row Space, Column Space, and Null Space

If we have a matrix, we can find its row and column vectors.

**Definition 2**

If \(a\) is an \(m \times n\) matrix, then:

- The subspace of \(\mathbb{R}^n\) spanned by the row vectors of \(A\) is the row space of \(A\).
- Subspace of \(\mathbb{R}^m\) spanned by the column vectors of \(A\) is the column space of \(A\).
- The solution space of the system \(A\mathbf{x} = \mathbf{0}\), which is a subspace of \(\mathbb{R}^n\) is the null space of \(A\).
We will answer two questions in this section

1. What relationships exist among the solutions of \(Ax = b\), and the spaces above for the coefficient matrix \(A\)?

2. What relationships exist between the row, column, and null space of a matrix?

Examining the first question with a system \(Ax = 0\), we can expand \(A\) so that:

\[
A \begin{bmatrix} a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n \\ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n \\ \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}
\]

Therefore we know that:

\[
A x = x_1 c_2 + x_2 c_2 + \cdots + x_n c_n = b
\]

\(Ax = b\) is consistent iff \(b\) is expressible as a linear combination of the column vectors of \(A\). Therefore:

**Theorem 4.7.1**

*A system of linear equations \(Ax = b\) is consistent iff \(b\) is in the column space of \(A\).*

Recall Theorem 3.4.4, which states that the solution to a consistent linear system \(Ax = b\) can be obtained by adding any specific solution of this system to the solution of \(Ax = 0\). Remembering that the null space is the same solution space of \(Ax = 0\),

**Theorem 4.7.2**

*If \(x_0\) is any solution for the consistent linear system \(Ax = b\), and if \(S = \{v_1, v_2, \ldots, v_k\}\) is a basis for the null space of \(A\), then every solution of \(Ax = b\) can be expressed in the form

\[
x = x_0 + c_1 v_1 + c_2 v_2 + \cdots + c_k v_k
\]

Conversely, for all choices of scalars, the vector \(x\) in this formula is a solution of \(Ax = b\).*

The equation above gives a formula for the *general solution* of \(Ax = b\). The vector \(x_0\) in that formula is the *particular solution of \(Ax = b\)*, and the rest is the *general solution of \(Ax = 0\).*

*Therefore, general solution = particular solution + general solution of homogenous system.*
Bases for Row, Column, and Null Spaces

<table>
<thead>
<tr>
<th>Theorem 4.7.3</th>
<th>Theorem 4.7.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary row operations do not change the null space of a matrix.</td>
<td>Elementary row operations do not change the row space of a matrix.</td>
</tr>
</tbody>
</table>

However, these row operations do change the column space. We'll get back to this. We can easily find the row and column spaces for a matrix already in REF.

Theorem 4.7.5

If a matrix R is in REF, then

- Row vectors with leading 1's form a basis for the rowspace.
- Column vectors with the leading 1's of the row vectors form a basis for the column space.

Example

Given a matrix

\[
R = \begin{bmatrix}
1 & -2 & 5 & 0 & 3 \\
0 & 1 & 3 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

- Basis for Row Space:
  - \(r_1 = \begin{bmatrix} 1 & -2 & 5 & 0 & 3 \end{bmatrix}\)
  - \(r_2 = \begin{bmatrix} 0 & 1 & 3 & 0 & 0 \end{bmatrix}\)
  - \(r_3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix}\)

- Basis for Column Space
  - \(c_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}\)
  - \(c_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\)
  - \(c_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}\)

This is great if your matrix is already in REF.

If we are not given a matrix already in REF:

- RREF
- Read off the basis for the row space.
- Note the position of the leading one's. Don't read off those columns as the column space, instead read off the respective columns in the original matrix.

The above is outlined in Theorem 4.7.6.
We can use the knowledge in this section to find a basis for the space spanned by a set of vectors in \( \mathbb{R}^n \).

See example 8, page 232.

In this example, you are given 4 vectors that span a subspace in \( \mathbb{R}^5 \). These vectors are put into a matrix, and we RREF to find the row space.

These vectors then form a basis for the row space, and consequently form a basis for the subspace of \( \mathbb{R}^5 \) spanned by the 4 vectors given.

**Bases Formed from Row and Column Vectors of a Matrix**

In the previous examples, we used row reduction to find the bases for the row and column spaces. We did not pose any restrictions on which vectors must make up these bases.

If we were to place restrictions such that

- The basis for the row space must consist of some of the row vectors
- The basis for the column space must consist of some of the column vectors

Then we must change our procedure.

This change in procedure is outlined in *Example 9, Page 232*.

Essentially, we are given a matrix \( A \).

We transpose \( A \), converting the row space of \( A \) into the column space of \( A^T \). Once we find the basis of the column space of \( A^T \), we will convert those vectors from column vectors into row vectors.

**Problem**

Given a set of vectors \( S = \{ v_1, v_2, \ldots, v_k \} \), in \( \mathbb{R}^n \), find a subset of these vectors that forms a basis for \( \text{span}(S) \), and express those vectors that are not in that basis as a linear combination of the basis vectors.

See *Example 10, page 233*. Also, this was covered on the labs.

Essentially,

- Construct a matrix that has \( S = \{ v_1, v_2, \ldots, v_k \} \) as column vectors. RREF this matrix.
- Then express using parameters, like we did in Chapter 1 (with S’s and T’s and free variables). These are called *dependency equations*.

Check the labs for this, they serve as excellent examples. The book’s method sucks because it pulls the “by inspection, \( v_1 = 2v_2 - v_5 + 4v_4 \)” crap.
4.8 – Rank, Nullity, and the Fundamental Matrix Spaces.

We found that the row and column spaces of the matrix \( A \) (used in examples 6 and 7 of Section 4.7) both have 3 basis vectors and therefore, are both 3 dimensional.

Theorem 4.8.1

The row and column space of a matrix have the same dimension.

Rank and Nullity

**Definition 1**

- **Rank(A):** The common dimension of the row space and column space.
- **Nullity(A):** The dimension of the null space.

**Example 1, Page 238** illustrates the general procedure for finding rank and nullity.

Essentially, we RREF a matrix \( A \), and find that

\[
\begin{bmatrix}
   a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\
   a_{21} & a_{22} & a_{23} & a_{24} & a_{15} \\
   \vdots & \vdots & \vdots & \vdots & \vdots \\
   a_{n1} & a_{n2} & a_{n3} & a_{n4} & a_{n5}
\end{bmatrix}
\begin{bmatrix}
   x_1 \\
   x_2 \\
   \vdots \\
   x_n
\end{bmatrix}
= \begin{bmatrix}
   P_1 \\
   P_2 \\
   \vdots \\
   P_n
\end{bmatrix}
\begin{bmatrix}
   R_{11} \\
   R_{21} \\
   \vdots \\
   R_{m1}
\end{bmatrix}
\begin{bmatrix}
   R_{11} \\
   R_{21} \\
   \vdots \\
   R_{m1}
\end{bmatrix}
\]

where \( P_n \) denotes a parameter (such as r, s, t u) and \( R_{mn} \) denotes the appropriate columns from the RREF.

In the above condensing of the example, we note that nullity(A)=\( n \) as there are \( n \) parameters.

**Maximum Value for Rank**

The maximum possible rank for an \( m \times n \) matrix is

\[
\text{rank}(A) \leq \text{min}(m, n)
\]

- The row vectors are in \( \mathbb{R}^n \) and are \textit{n-dimensional}
- The column vectors are in \( \mathbb{R}^m \) and are \textit{m-dimensional}.
- **Rank(A)** is the highest common dimension of the row and column space, \( \text{rank}(A) \leq \min(m, n) \), where min() is the minimum.
**Theorem 4.8.2 – Dimension theorem for Matrices**

If \( A \) is a matrix with \( n \) columns,

\[
\text{rank}(A) + \text{nullity}(A) = n
\]

As well,

\[
\left| \text{number of leading variables} \right| + \left| \text{number of free variables} \right| = n
\]

**Theorem 4.8.3**

If \( A \) is an \( m \times n \) matrix,

- \( \text{Rank}(A) = \) number of leading variables in the general solution of \( Ax=0 \)
- \( \text{Nullity}(A) = \) number of parameters in the general solution of \( Ax=0 \).

**Theorem 4.8.4 – Invertible Matrix Theorem.**

This theorem is listed on page 240. It is not listed here because it will be expanded in Chapter 5 anyways. It will also be listed in the back pages of this document.

**Overdetermined and Underdetermined Systems**

- **Overdetermined** systems have constraints > unknowns
- **Underdetermined** systems have constraints < unknowns, equivalent to unknowns > constraints.

**Theorem 4.8.5**

If \( Ax=b \) is a consistent linear system of \( m \) equations in \( n \) unknowns, and if \( A \) has rank \( r \), then the system has \( n - r \) parameters.

**Theorem 4.8.6**

Let \( A \) be an \( mxn \) matrix.

- **(Overdetermined Case)** If \( m > n \), the linear system is inconsistent for at least one vector \( b \) in \( \mathbb{R}^n \).
- **(Underdetermined Case)**. If \( m < n \), then for each vector \( b \) in \( \mathbb{R}^m \), the system \( Ax=b \) is either inconsistent or has infinitely many solutions.

See **Example 6, Page 242** for the Meta-RREF.

Essentially,

A linear system is given, and it is RREF’d into a mess with 5 parameters. Those parameters are then put into an augmented matrix and RREF’d, leaving only two parameters that decide the other 5 parameters.
The Fundamental Spaces of a Matrix

There are six important vector spaces associated with a matrix.

- **Row space of** \( A \)
- **Column Space of** \( A \)
- **Null Space of** \( A \)

**Equivalent**
- **Row space of** \( A^T \)
- **Column Space of** \( A^T \)
- **Null Space of** \( A^T \)

**Theorem 4.8.7**

If \( A \) is any matrix, then \( \text{rank}(A) = \text{rank}(A^T) \)

If \( A \) is \( m \times n \), then

\[
\text{Rank}(A) + \text{nullity}(A^T) = m
\]

Using **Theorem 4.8.2**, \( (\text{rank}(A) + \text{nullity}(A) = n) \) we can write

\[
\begin{array}{ccc}
\text{dim}[\text{row}(A)] = r & \text{dim}[\text{col}(A)] = r \\
\text{dim}[\text{null}(A)] = n - r & \text{dim}[\text{null}(A^T)] = m - r
\end{array}
\]

The above formulas provide an algebraic relation between the size of a matrix and the dimension of its fundamental spaces.

Recall from **Theorem 3.4.3**: If \( A \) is \( m \times n \), the null space of \( A \) consists of vectors that are orthogonal to each of the row vectors of \( A \).

**Definition 2**

If \( W \) is a subspace of \( \mathbb{R}^n \), then the set of all vectors in \( \mathbb{R}^n \) that are orthogonal to every vector in \( W \) is called the **Orthogonal Complement** of \( W \) and is denoted \( W^\perp \).

**Theorem 4.8.8**

If \( W \) is a subspace of \( \mathbb{R}^n \), then:

- \( W^\perp \) is a subspace of \( \mathbb{R}^n \)
- The only vector common to \( W \) and \( W^\perp \) is \( 0 \).
- The orthogonal complement of \( W^\perp \) is \( W \).
In each of the following spaces, the orthogonal complement of \( W \) through the origin is

- In \( \mathbb{R}^2 \), the orthogonal complement of a line is the \textit{line that is perpendicular} to \( W \).
- In \( \mathbb{R}^3 \), the orthogonal complement of a plane is the \textit{line through the origin that is perpendicular to the plane}.

A Geometric Link Between the Fundamental Spaces

The following Theorem provides a link between the fundamental spaces of a matrix.

Part A is a restatement of \textit{Theorem 3.4.3},

If we solve a homogenous system \( A x = 0 \), then the solution set of that system is orthogonal to every row vector of \( A \), \textit{Theorem 3.4.3}.

\textbf{Theorem 4.8.9}

If \( A \) is \( m \times n \):

- \( \text{null}(A) \) and \( \text{row}(A) \) are orthogonal complements in \( \mathbb{R}^n \)
- \( \text{null}(A^T) \) and \( \text{col}(A) \) are orthogonal complements in \( \mathbb{R}^m \)
4.9 – Matrix Transformations From \( \mathbb{R}^n \) to \( \mathbb{R}^m \)

**Functions and Transformations**

Let’s say we have two sets, \( A \) and \( B \).

A **function** associates each element of set \( A \) with one and only one element of set \( B \).

Therefore, \( b \) is the **image** of \( a \) under \( f \).

\( f(a) \) is the **value** of \( f \) at \( a \).

The set \( A \) is the **domain** and set \( B \) is the **codomain** of \( f \).

The subset of the codomain that consists of all images of points in the domain is the **range** of \( f \).

**Definition 1**

If \( V \) and \( W \) are vector spaces, and if \( f \) is a function with domain \( V \) and codomain \( W \), then:

- \( f \) is a **transformation** from \( V \) to \( W \), or that \( f \) **maps** \( V \) to \( W \), denoted \( f : V \rightarrow W \)
- In the special case where \( V=W \), the transformation is called an **operator** on \( V \).

Using our common knowledge of function notation from High School, let’s define a function as

\[
\begin{align*}
    w_1 &= f_1(x_1, x_2, \cdots, x_n) \\
    w_2 &= f_2(x_1, x_2, \cdots, x_n) \\
    w_3 &= f_3(x_1, x_2, \cdots, x_n) \\
    &\vdots \\
    w_m &= f_m(x_1, x_2, \cdots, x_n)
\end{align*}
\]

There are \( m \) equations here that take an \( n \)-tuple point \( (x_1, x_2, \cdots, x_n) \) \( (\mathbb{R}^n) \) and define a corresponding \( m \)-tuple point in \( \mathbb{R}^m \). \( (w_1, w_2, \cdots, w_m) \)

If these transformations are linear, we can write the transformations as a system of equations, which we could then write as a matrix as

\[
    \begin{bmatrix}
        w_1 \\
        w_2 \\
        \vdots \\
        w_m
    \end{bmatrix} =
    \begin{bmatrix}
        a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\
        a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\
        \vdots & \vdots & \vdots & \ddots & \vdots \\
        a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}
    \end{bmatrix}
    \begin{bmatrix}
        x_1 \\
        x_2 \\
        \vdots \\
        x_n
    \end{bmatrix}
\]

Rather than viewing this as a system, the matrix \( A \) will be viewed as an operator, denoted

\[ w = T_A(x) \text{ or } x \xrightarrow{T_A} w \]

\( T_A \) is called **multiplication by** \( A \), and \( A \) is called the **standard matrix** for this transformation.
Notational Matters

If we want to denote a matrix transformation \( w = T_A(x) \) without giving the standard matrix a name, we can write \( T(x) = [T]x \).

Properties of Matrix Transformations

**Theorem 4.9.1**

For every matrix \( A \), the matrix transformation \( T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m \) has the following properties for vectors \( u \) and \( v \) in \( \mathbb{R}^n \) and for every scalar \( k \):

- \( T_A(0) = 0 \)
- \( T_A(ku) = kT_A(u) \) \[Homogeneity Property\]
- \( T_A(u+v) = T_A(u) + T_A(v) \) \[Additivity Property\]
- \( T_A(u-v) = T_A(u) - T_A(v) \)

If two matrix transformations from \( \mathbb{R}^n \) to \( \mathbb{R}^m \) have the same image of every point in \( \mathbb{R}^n \), then the matrices must be the same.

**Theorem 4.9.2**

If \( T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m \) and \( T_B : \mathbb{R}^n \rightarrow \mathbb{R}^m \) are matrix transformations, and if \( T_A(x) = T_B(x) \) for every \( x \) in \( \mathbb{R}^n \), then \( A = B \).

The **zero transformation** has a standard matrix of \( 0 \).
A Procedure for Finding Standard Matrices

To find the standard matrix of a transform in \( \mathbb{R}^n \), suppose it's standard matrix is unknown. Using the standard basis vectors for \( \mathbb{R}^n \), \( \{ e_1, e_2, \cdots, e_n \} \), we then know that:

\[
T_A(e_1) = A e_1 \quad T_A(e_2) = A e_2 \quad \cdots \quad T_A(e_n) = A e_n
\]

Next, we need to use the results from Theorem 1.3.1.

**Theorem 1.3.1**

Given a matrix multiplication

\[
A x = \begin{bmatrix}
  a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n \\
  a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n \\
  \vdots \\
  a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n
\end{bmatrix} = x_1 \begin{bmatrix}
  a_{11} \\
  a_{21} \\
  \vdots \\
  a_{m1}
\end{bmatrix} + x_2 \begin{bmatrix}
  a_{12} \\
  a_{22} \\
  \vdots \\
  a_{m2}
\end{bmatrix} + \cdots + x_n \begin{bmatrix}
  a_{1n} \\
  a_{2n} \\
  \vdots \\
  a_{mn}
\end{bmatrix}
\]

if \( x \) is an \( nx1 \) column vector, then \( Ax \) is a linear combination of the column vectors of \( A \) in which the coefficients are the entries of \( x \). Shown below,

\[
A x = \begin{bmatrix}
  a_{11} x_1 \\
  a_{21} x_1 \\
  \vdots \\
  a_{m1} x_1
\end{bmatrix} = x_1 \begin{bmatrix}
  a_{11} \\
  a_{21} \\
  \vdots \\
  a_{m1}
\end{bmatrix}
\]

Going back to the procedure, we know that \( Ae_j \) is a linear combination of \( A \), where \( j \) is an iteration from 1 to \( n \). From the textbook:

*But all of the entries \( e_j \) are zero except the \( j \)th, so the product \( Ae_j \) is just the \( j \)th column of the matrix \( A \).*

**Finding the Standard Matrix for a Matrix Transformation**

\[
A = [ T_A(e_1) | T_A(e_2) | \cdots | T_A(e_n) ]
\]

**Reflection Operators**

A *reflection operator* maps each point to it's symmetric image about a fixed line or plane.

You can derive reflection operators by using the method above to find a standard matrix.

Just write out \( T(e_1) = T(1,0) = (-1,0) \)

\( T(e_2) = T(0,1) = (0,1) \) , for example, and put the results in columns.

**Projection Operators**

*Projection Operators* take a vector and map each point along a fixed line or plane. If we're mapping on co-ordinate planes, then it'll make that component zero. (For instance, if we have a 3-space vector and we want it's projection on the xy-plane, we'll make it's z component zero). Use the same method as above.
Rotation Operators

Our standard matrix in $\mathbb{R}^2$ makes use of the properties of the unit circle.

When calculating the standard matrix, first take the vector $e_1$, and rotate it just like a pole in a circle. It's co-ordinates are the familiar $[\cos \theta, \sin \theta]$.

When we rotate the vector $e_2$, however, notice in the diagram that theta is not in standard position. For this reason, the co-ordinates are different. $e_2$'s co-ordinates are $[-\sin \theta, \cos \theta]$.

Therefore, the standard matrix is: 

$$
\begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}
$$

also called the rotation matrix.

Using the properties of matrix multiplication, we can find the rotation equations.

$$
\begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}
\begin{bmatrix}
w_1 \\
w_2
\end{bmatrix}
= 
\begin{bmatrix}
w_1 = x_1 \cos \theta - x_2 \sin \theta \\
w_2 = x_1 \sin \theta + x_2 \cos \theta
\end{bmatrix}
$$

4.10 – Properties of Matrix Transformations

Compositions of Matrix Transformations

We can compose matrix operators just like we can compose functions. For instance,

$$
T_B(T_A(x)) = (T_B \circ T_A)(x) = T_B \circ T_A = T_{BA}
$$

Careful to keep your compositions in the correct order! Remember that $AB \neq BA$ and $(T_B \circ T_A)(x) = B(T_A(x)) = B(Ax) = (BA)x$.

therefore $(T_B \circ T_A)(x) \neq (T_A \circ T_B)(x)$  

[Keep your compositions in order!]

The above figure states that $T_C \circ T_B \circ T_A = T_{CBA}$
The book now runs through several examples of compositions, but these are just plain old matrix multiplication.

One–to–One Matrix Transformations

**Definition 1**

A matrix transformation \( T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m \) is **one-to-one** if \( T_A \) maps distinct vectors in \( \mathbb{R}^n \) to distinct vectors in \( \mathbb{R}^m \).

Rotation operators are one-to-one, since they take points and keep distinct images.

Projection operators are not one-to-one as they destroy data to project the vector.

The above sentences relate the invertibility of a matrix to whether or not it’s one to one. (For instance, why would rotation operators be one-to-one but not projection operators?)

**Theorem 4.10.1**

If \( A \) is an \( n \times n \) matrix and \( T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is the matrix operator, then the following are equivalent:

- \( A \) is invertible.
- The range of \( T_A \) is in \( \mathbb{R}^n \)
- \( T_A \) is one-to-one.

**Inverse of a one-to-one matrix operator**

If \( T_A \) is one-to-one, then it is invertible.

- \( T_A \) co-responds to a standard matrix \( A \), so therefore
- \( T_A^{-1} \) co-responds to a standard matrix \( A^{-1} \).

\[
T_A^{-1} = T_A^{-1}
\]
Linearity Properties

So far we've looked at linear transformations only. What properties do non-linear transformations have, and how can we check to see if a non-linear transformation actually is a transformation?

Problem

Is there a way to check whether a mapping $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a matrix transformation without computing the standard matrix?

The following theorem outlines two Linearity Conditions.

Theorem 4.10.2

$T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a matrix transformation iff the following relationships hold for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ and every scalar $k$:

- $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ [Additivity Property]
- $T(k \mathbf{u}) = kT(\mathbf{u})$ [Homogeneity Property]

A transformation that satisfies the above linearity conditions is a Linear Transformation.

Theorem 4.10.3

Every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a matrix transformation. Every matrix transform from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a linear transformation.
Chapter 5 – Eigenvalues and Eigenvectors

5.1 – Eigenvalues and Eigenvectors

Definition 1
If A is an $n \times n$ matrix, then a non-zero vector $x$ in $\mathbb{R}^n$ is called an eigenvector of A if $Ax$ is a scalar multiple of $x$:

$$Ax = \lambda x$$

for some scalar lambda. The scalar lambda is an eigenvalue of A.

Theorem 5.1.1
If A is an $n \times n$ matrix, then lambda is an eigenvalue of A iff it satisfies

$$\det(\lambda I - A) = 0$$

This is the characteristic equation of A.
To find the eigenvalues of A, solve the characteristic equation.

Theorem 5.1.2
If A is an $n \times n$ triangular matrix, then the eigenvalues of A are on the main diagonal of A.

Theorem 5.1.3
If A is an $n \times n$ matrix, the following are equivalent

- $\lambda$ is an eigenvalue of A
- The system of equations $(\lambda I - A)x = 0$ has non-trivial solutions.
- There is a non-zero vector such that $Ax = \lambda x$
- $\lambda$ is a solution of the characteristic equation $\det(\lambda I - A) = 0$

Bases for Eigenspaces.
Once you have found the eigenvalues, take those eigenvalues and sub them into $(\lambda I - A)x = 0$.
From there, you can find the bases just like in Chapter 4.
Powers of a Matrix

We can use eigenvalues to find powers of a matrix.

Given that \( Ax = \lambda x \), we can see that \( A^2 x = A(Ax) = A(\lambda x) = \lambda (Ax) = \lambda^2 x \)

**Theorem 5.1.4**

If \( k \) is a positive integer, and \( \lambda \) is an eigenvalue of a matrix \( A \), and \( x \) is a corresponding eigenvector, then \( \lambda^k \) is an eigenvalue of \( A^k \) and \( x \) is a corresponding eigenvector.

Eigenvalues and Invertibility

**Theorem 5.1.5**

A square matrix \( A \) is invertible iff \( \lambda = 0 \) is not an eigenvalue of \( A \).

**Theorem 5.1.6**

The Matrix Invertibility Theorem in its glory.

5.2 – Diagonalization

Recall that a **diagonal matrix** is a matrix which has entries only on the leading diagonal, all other entries are zero.

The Matrix Diagonalization Problem

Given an \( n \times n \) matrix \( A \), is there a matrix \( P \) such that \( P^T AP \) is diagonal?

Given an \( n \times n \) matrix \( A \), does \( A \) have \( n \) linearly independent eigenvectors?

Similarity

**B is similar to A if** there is an invertible matrix \( P \) such that \( B = P^{-1} AP \)
**Definition – Diagonalizability**

A square matrix $A$ is **diagonalizable** if it is similar to some diagonal matrix, if there is an invertible matrix $P$ such that $P^{-1}AP$ is diagonal. In this case matrix $P$ diagonalizes $A$.

**Similarity Invariants**

The following conditions are equivalent (so both of $A$ and $P^{-1}AP$ have the following)

- Same Determinants
- Both are equivalently invertible
- Rank
- Nullity
- Trace
- Characteristic Polynomial
- Eigenvalues
- Eigenspace dimension.

**Theorem 5.2.1**

If $A$ is an $n \times n$ matrix, then

- $A$ is diagonalizable
- $A$ has $n$ linearly independent eigenvectors.

**Procedure for Diagonalizing a Matrix**

1. Confirm the matrix is actually diagonalizable by finding $n$ linearly independent eigenvectors.

   1. Do this by finding a basis for each eigenspace. Take these vectors and form them into a set $S$.
   2. Form the matrix $P = [p_1 \ p_2 \ \cdots \ p_3]$ that has the vectors in $S$ as its column vectors.
   3. The matrix $P^{-1}AP$ will be diagonal and have the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ corresponding to the eigenvectors $P = [p_1 \ p_2 \ \cdots \ p_3]$ as its successive diagonal entries.

**Theorem 5.2.2**

If $v_1, v_2, \cdots, v_k$ are eigenvectors of a matrix $A$ corresponding to distinct eigenvalues, then $\{v_1, v_2, \cdots, v_k\}$ is a linearly independent set.
**Theorem 5.2.3**

If an n×n matrix has n distinct eigenvalues, then A is diagonalizable.

**Powers of a Matrix**

To calculate powers of a matrix, let us manipulate \( D = P^{-1} A P \).

\[
P^{-1} A P = \begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n
\end{bmatrix} = D
\]

Square both sides

\[
(P^{-1} A P)^2 = \begin{bmatrix}
\lambda_1^2 & 0 & \cdots & 0 \\
0 & \lambda_2^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n^2
\end{bmatrix} = D^2
\]

If we expand the left side, \((P^{-1} A P)^2 = P^{-1} A (P P^{-1}) A P = P^{-1} A I A P = P^{-1} A^2 P\). In general, we can write \( P^{-1} A^k P = D^k = \begin{bmatrix}
\lambda_1^k & 0 & \cdots & 0 \\
0 & \lambda_2^k & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n^k
\end{bmatrix} \).

*If we manipulate, we get* \( A^k = PD^k P^{-1} \)

**Eigenvalues of Powers of a matrix.** \( A x = \lambda x \)

Recall from our earlier theorem that \( A^2 x = A( A x) = A(\lambda x) = \lambda ( A x) = \lambda (\lambda x) = \lambda^2 x \).

In general, we can expand this \( A^k x = \lambda^k x \)

**Theorem 5.2.4**

If we have a square matrix A with eigenvalues, and k is a positive integer, \( A^k x = \lambda^k x \)


**Multiplicities**

In Theorem 5.2.3, we say that \( n \) distinct eigenvalues \( \rightarrow \) diagonalizability. However, we can also have diagonalizable matrices with fewer than \( n \) distinct eigenvalues.

For example, if a matrix yields the characteristic polynomial \( (\lambda - 1)(\lambda - 2)^2 \), then

- The eigenspace co-responding to \( \lambda = 1 \) will be at most one-dimensional (1 parameter)
- The eigenspace co-responding to \( \lambda = 2 \) will be at most two-dimensional (2 parameters)

**Geometric Multiplicity:** Dimension of the Eigenspace (Number of Parameters)

**Algebraic Multiplicity:** Number of times the \( (\lambda - \text{whatever}) \) term appears in the characteristic polynomial.

**Theorem 5.2.5 – Geometric and Algebraic Multiplicity**

If \( A \) is a square matrix, then:

- For every eigenvalue of \( A \), the geometric multiplicity \( \leq \) algebraic multiplicity
- \( A \) is diagonalizable iff geometric multiplicity = algebraic multiplicity.
5.i – Complex Numbers

Complex numbers arise when solving the quadratic equation when the discriminant under the root sign is negative.

\[ x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \]

Use the basic definition that \( i^2 = -1 \).

Complex numbers are written in the form \( a + bi \). Complex numbers are treated as binomials.

\[(a + bi) + (c + di) = (a + c) + (b + d)i \quad \text{[Collect Like Terms]}\]

\[(a + bi)(c + di) = (ac - bd) + (ad + bc)i \quad \text{[Binomial Expansion]}\]

Complex Conjugates

If you have a number \( z = a + bi \) then the complex conjugate is \( z = a - bi \).

Polar Form

We can find the magnitude of \( |z| \).

\[ z \bar{z} = (a + bi)(a - bi) = a^2 + b^2 \]

\[ |z| = \sqrt{z \bar{z}} = \sqrt{a^2 + b^2} \]

Reciprocals and Division

The reciprocal is defined as \( \left( \frac{1}{z} \right ) z = 1 \). To solve for \( 1/z \), multiply both sides by \( z \) and use the fact that \( z \bar{z} = |z|^2 \).

\[ \frac{1}{z} = \frac{z}{|z|^2} \]

\( z_1/z_2 \) is defined as the product of \( z_1 \) and \( 1/z_2 \).

\[ \frac{z_1}{z_2} = \frac{\overline{z}_2}{|z_2|^2} z_1 \]
**Theorem B.0.1**

- $z_1 + z_2 = z_1 + z_2$
- $z_1 - z_2 = z_1 - z_2$
- $z_1 z_2 = z_1 z_2$
- $\overline{z}_1 z_2 = \overline{z}_1 \overline{z}_2$
- $\overline{z}_1 / z_2 = \overline{z}_1 / z_2$
- $\overline{z} = z$

**Theorem B.0.2**

- $|z| = |z|
- $|z_1 z_2| = |z_1||z_2|
- $|z_1 / z_2| = |z_1|/|z_2|
- $|z_1 + z_2| \leq |z_1| + |z_2|$

**Polar Form of a Complex Number**

Seeing that $a = |z| \cos \theta$ and $b = |z| \sin \theta$

$z = |z| (\cos \theta + i \sin \theta)$

$z_1 z_2 = |z_1||z_2| [\cos (\theta_1 + \theta_2) + i \sin (\theta_1 + \theta_2)]$

*For Multiplication, we are multiplying the moduli and adding the arguments.*

For *division*, we are *dividing moduli, subtracting arguments.*

**Demoivre's Formula**

$z^n = |z|^n (\cos n \theta + i \sin n \theta)$

$(\cos \theta + i \sin \theta)^n = \cos n \theta + i \sin n \theta$

**Euler's Formula**

$e^{i \theta} = \cos \theta + i \sin \theta$
5.3 – Complex Vector Spaces

The Complex Euclidean Inner Product

**Definition 2**

If \( u = (u_1, u_2, \ldots, u_n) \) and \( v = (v_1, v_2, \ldots, v_n) \) are vectors in \( \mathbb{C}^n \), then the **complex Euclidean inner product** of \( u \) and \( v \), \( u \cdot v \) is

\[
u \cdot v = u_1 \overline{v}_1 + u_2 \overline{v}_2 + \cdots + u_n \overline{v}_n\]

The **Euclidean norm** is

\[
\|v\| = \sqrt{\overline{v} \cdot v} = \sqrt{|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2}
\]

Recall that if \( u \) and \( v \) are column vectors in \( \mathbb{R}^n \), then \( u \cdot v = u^T v = v^T u \).

Analogously in \( \mathbb{C}^n \), (if \( u \) and \( v \) are column vectors), \( u \cdot v = u^T \overline{v} = \overline{v}^T u \).

**Theorem 5.3.3**

If \( u, v \) and \( w \) are vectors in \( \mathbb{C}^n \), and if \( k \) is a scalar, then the complex dot product has the following properties.

- \( u \cdot v = \overline{u} \cdot \overline{v} \) \hspace{1cm} [Anti-symmetry property]
- \( u \cdot (v + w) = u \cdot v + u \cdot w \) \hspace{1cm} [Distributive Property]
- \( k (u \cdot v) = (k u) \cdot v \) \hspace{1cm} [Homogeneity Property]
- \( u \cdot k v = \overline{k} (u \cdot v) \) \hspace{1cm} [Antihomogeneity Property]
- \( v \cdot v \geq 0 \) and \( v \cdot v = 0 \) iff \( v = 0 \) \hspace{1cm} [Positivity Property]

**Vector Concepts in \( \mathbb{C}^n \)**

If the equation \( \det (\square I - A) = 0 \) has complex roots, we have **complex eigenvalues**.

**Theorem 5.3.4**

FINISH SECTION PAGE 319
Chapter 6 – Inner Product Spaces

6.1 – Inner Products

**Definition 1**
An *inner product* on a real vector space $V$ is a function that associates a real number $\langle u, v \rangle$ with each pair of vectors in $V$, such that the following axioms are satisfied.

1. $\langle u, v \rangle = \langle v, u \rangle$ [Symmetry Axiom]
2. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ [Additivity Axiom]
3. $\langle ku, v \rangle = k \langle u, v \rangle$ [Homogeneity Axiom]
4. $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0$ iff $v = 0$ [Positivity Axiom]

Inner products are flexible, we can define weighted inner products.

The above definition can be used to define the *Euclidean Inner Product* as the standard dot product.

**Definition 2**
If $V$ is a real inner product space, then the *norm* of $v$ in $V$ is

$$\|v\| = \sqrt{\langle v, v \rangle}$$

The *distance* between two vectors $d(u, v)$ is

$$d(u, v) = \|u - v\| = \sqrt{\langle u - v, u - v \rangle}$$

**Theorem 6.1.1**
If $u$ and $v$ are vectors in a real inner product space $V$, and if $k$ is a scalar:

- $\|v\| \geq 0$ iff $v = 0$
- $d(u, v) = d(v, u)$
- $d(u, v) \leq 0$ with equality iff $u = v$

**Weighted Inner Products**
We can define a row vector $w$ for weighting.

The *weighted Euclidean Inner Product with weights* $w$ is

$$\langle u, v \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \cdots + w_n u_n v_n$$
Inner Products Generated by Matrices

We can do a really fancy weighting system by doing an *inner product on* $\mathbb{R}^n$ *generated by* $A$

$$ \langle u, v \rangle = A u \cdot v $$

Recall that if $u$ and $v$ are column vectors in $\mathbb{R}^n$, then $u \cdot v = u^T v = v^T u$.

Use this form to write

$$ \langle u, v \rangle = (Av)^T u = v^T A^T A u $$

Other examples of Inner Products

An Inner Product on $M_{n \times n}$

Recall that a *trace* is the sum of the entries on the main diagonal of a matrix. 

$$ \langle U, V \rangle = tr(U^T V) $$

Inner Product on $P_n$

If we take two polynomials in $P_n$ where

- $p$ has coefficients $a_n$
- $q$ has coefficients $b_n$

then

$$ \langle p, q \rangle = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n $$

$$ \| p \| = \sqrt{\langle p, p \rangle} = \sqrt{a_0^2 + a_1^2 + \cdots + a_n^2} $$

We can *evaluate an inner product at a point*

$$ \langle p, q \rangle = p(x_0) q(x_0) + p(x_1) q(x_1) + \cdots + p(x_n) q(x_n) $$

Inner Products of Functions

We can take two functions on $C[a,b]$. $P_n$ is a subspace of $C[a,b]$ because polynomials are continuous functions.

Let $f = f(x)$ and $g = g(x)$ be two functions.

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

$$ \| f \| = \langle f, f \rangle^{1/2} = \sqrt{\int_a^b f^2(x) \, dx} $$
Algebraic Properties of Inner Products

**Theorem 6.1.2**

If \( u, v, \) and \( w \) are vectors in a real inner product space \( V \), and if \( k \) is a scalar, then:

- \( \langle 0, v \rangle = \langle v, 0 \rangle = 0 \)
- \( \langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle \)
- \( \langle u, v - w \rangle = \langle u, v \rangle - \langle u, w \rangle \)
- \( \langle u - v, w \rangle = \langle u, w \rangle - \langle v, w \rangle \)
- \( k \langle u, v \rangle = \langle u, kv \rangle \)

6.2 – Angle and Orthogonality in Inner Product Spaces

If we solve the formula for the *Euclidean dot product* for theta, shown below, we notice that the argument for \( \arccos \) must be between [-1, 1]. This is called the **Cauchy – Schwarz Inequality**.

\[
\theta = \arccos \left( \frac{u \cdot v}{\|u\| \|v\|} \right)
\]

The above leads to **Theorem 3.2.4** which states that \( |u \cdot v| \leq \|u\| \|v\| \)

**Theorem 6.2.1**

If \( u \) and \( v \) are vectors in a real inner product space \( V \), then

\[
|\langle u, v \rangle| \leq \|u\| \|v\|
\]

Other forms include:

\[
\langle u, v \rangle^2 \leq \langle u, u \rangle \langle v, v \rangle
\]

\[
\langle u, v \rangle \leq \|u\|^2 \|v\|^2
\]

**Angle between Vectors**

Extend the formula above to get

\[
\theta = \arccos \left( \frac{\langle u, v \rangle}{\|u\| \|v\|} \right)
\]

**Properties of Length and Distance in General Inner Product Spaces**

**Theorem 6.2.2 – Triangle Inequality**

- \( \|u + v\| \leq \|u\| + \|v\| \)
- \( d(u, v) \leq d(u, w) + d(w, v) \)
Orthogonality

Two vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space are **orthogonal** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$

**Theorem 6.2.3 – Generalized Theorem of Pythagoras**

If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors in an inner product space

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Orthogonal Complements

If $W$ is a subspace of an inner product space $V$, then the set of all vectors in $V$ that are orthogonal to every vector in $W$ is called the **orthogonal complement** of $W$ and is denoted by $W^\perp$.

**Theorem 6.2.4**

If $W$ is a subspace of an inner product space $V$, then

- $W^\perp$ is a subspace of $V$
- $W \cap W^\perp = \{ 0 \}$

**Theorem 6.2.5**

If $W$ is a subspace of a finite-dimensional inner product space $V$, then the orthogonal complement of $W^\perp$ is in $W$

$$W^\perp \cap W^\perp = W$$

*Figure 6.2.2*  Each vector in $W$ is orthogonal to each vector in $W^\perp$ and conversely
6.3 – Gram-Schmidt Process

**Definition**

If we take a set of vectors, the set is an **orthogonal set** iff all the vectors in the set are orthogonal to each other. To test for an orthogonal set, you must check that \( \langle u, v \rangle = 0 \) for all combinations of vectors in the set.

An orthogonal set where all vectors have a **norm 1** is an **orthonormal set**.

In \( \mathbb{R}^n \), any mutually perpendicular vectors are linearly independent. This theorem generalizes to an inner product space.

**Theorem 6.3.1**

If \( S = \{ v_1, v_2, \cdots, v_n \} \) is an orthogonal set of non-zero vectors in an inner product space, then \( S \) is linearly independent.

We can have **orthonormal** and **orthogonal bases** for an inner product space. We can take any vector and make co-ordinates relative to these bases.

**Theorem 6.3.2**

Given that \( S = \{ v_1, v_2, \cdots, v_n \} \) is the **orthogonal basis** for an inner product space \( V \), and that \( u \) is any vector:

\[
\begin{align*}
  u &= \frac{\langle u, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle u, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle u, v_n \rangle}{\|v_n\|^2} v_n \\
  (u)_s &= \frac{\langle u, v_1 \rangle}{\|v_1\|^2}, \frac{\langle u, v_2 \rangle}{\|v_2\|^2}, \cdots, \frac{\langle u, v_n \rangle}{\|v_n\|^2}
\end{align*}
\]

If \( S = \{ v_1, v_2, \cdots, v_n \} \) is just an **orthonormal basis**, we can omit the dividing by the magnitude of \( v \)

\[
\begin{align*}
  u &= \langle u , v_1 \rangle v_1 + \langle u , v_2 \rangle v_2 + \cdots + \langle u , v_n \rangle v_n \\
  (u)_s &= \langle u , v_1 \rangle, \langle u , v_2 \rangle, \cdots, \langle u , v_n \rangle
\end{align*}
\]
Orthogonal Projections

**Theorem 6.3.3**

Let $W$ be a finite-dimensional subspace of an inner product space. Every vector $u$ in $V$ can be expressed as

$$u = w_1 + w_2$$

where $w_1$ is in $W$ and $w_2$ is in $W^\perp$.

**Theorem 6.3.4**

If we have $S = \{v_1, v_2, \ldots, v_r\}$ be an orthogonal basis for $W$, then

$$\text{proj}_u u = \frac{\langle u, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle u, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle u, v_r \rangle}{\|v_r\|^2} v_r$$

If the basis $S$ is orthonormal, not that the norm of every vector in $S = 1$.

**Theorem 6.3.5**

Every non-zero finite-dimensional inner product space has an orthonormal basis.

**The Gram-Schmidt Process**

To convert a basis $\{u_1, u_2, \ldots, u_r\}$ into an orthogonal basis $\{v_1, v_2, \ldots, v_r\}$:

1. $v_1 = u_1$
2. $v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\|v_1\|^2} v_1$
3. $v_3 = u_3 - \frac{\langle u_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_3, v_2 \rangle}{\|v_2\|^2} v_2$
4. $v_4 = u_4 - \frac{\langle u_4, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_4, v_2 \rangle}{\|v_2\|^2} v_2 - \frac{\langle u_4, v_3 \rangle}{\|v_3\|^2} v_3$
5. ...

**Optional Step:** You can normalize each of these basis vectors to get an orthonormal basis.
Chapter 7 – Diagonalization

7.1 – Orthogonal Matricies

Definition 1
A square matrix $A$ is **orthogonal** if $A^{-1} = A^T$, equivalently $AA^T = A^TA = I$

Theorem 7.1.1
The following are equivalent for an $n \times n$ matrix $A$:

- $A$ is orthogonal
- The row vectors of $A$ form an orthonormal set in $\mathbb{R}^n$ with the Euclidean inner product.
- The column vectors of $A$ form an orthonormal set in $\mathbb{R}^n$ with the Euclidean inner product.

Theorem 7.1.2
- The inverse of an orthogonal matrix is orthogonal.
- A product of orthogonal matrices is orthogonal
- If $A$ is orthogonal, $\det(A) = 1$ or $\det(A) = -1$

Orthogonal Matricies as Linear Operators
When we did reflection and rotation operators, those matrices were actually orthogonal because of these useful properties

Theorem 7.1.3
If $A$ is an $n \times n$ matrix then the following are equivalent.

- $A$ is orthogonal
- $\|Ax\| = \|x\|$ for all $x$ in $\mathbb{R}^n$
- $Ax \cdot Ay = x \cdot y$ for all $x$ and $y$ in $\mathbb{R}^n$. 
Change of Orthonormal Basis

Orthonormal bases for inner product spaces are convenient because many familiar formulas hold for such bases.

Theorem 7.1.4
If $S$ is an orthonormal basis for an $n$-dimensional inner product space $V$, and if

$$\begin{align*}
(u)_S &= (u_1, u_2, \ldots, u_n) \\
(v)_S &= (v_1, v_2, \ldots, v_n)
\end{align*}$$

then

- $\|u\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$ [Norm of a vector still holds]
- $d(u, v) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$ [Distance Formula still holds]
- $\langle u, v \rangle = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$

Theorem 7.1.5
Let $V$ be a finite-dimensional inner product space. If $P$ is the transition matrix from one orthonormal basis for $V$ to another orthonormal basis for $V$, then $P$ is an orthogonal matrix.

7.2 – Orthogonal Diagonalization

In section 5.2, we talked about diagonalizability with two matrices being similar if $B = P^{-1} A P$. Now that we know that for orthogonal matrices that $A^T = A^{-1}$, we can say that

Definition 1
If $A$ and $B$ are square matrices, they are orthogonally similar if there is an orthogonal matrix $P$ such that $B = P^T A P$

Conditions for Orthogonal Diagonalizability (Theorem 7.2.1)
- $A$ has an orthonormal set of $n$ eigenvectors.
- $A$ is symmetric.

Orthogonally Diagonalizing an $n \times n$ Symmetric Matrix

1. Find a basis for each eigenspace of $A$
2. Apply the Gram-Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.
3. Form the matrix $P$ whose columns are the vectors constructed in Step 2. This matrix will orthogonally diagonalize $A$, and the eigenvalues on the diagonal of $D = P^T A P$ will be in the same order as their co-responding eigenvectors in $P$.  

52/52