Math 235 Notes

Syed Mustafa Raza Rizvi

June 17, 2020

These are my 2nd year Linear Algebra 2 notes at the University of Waterloo (MATH 235). They are pretty similar to the content you may see in the course notes by D. Wolczuk.
You will find that these aren't very useful as notes, in the sense that they are not significantly shorter than the content in the course notes, they're really just a way for me to type down the content I am learning and absorb it. Also, I won't be including the proofs, it's best to read the course notes for that.
Thanks to Prof. Dan Wolczuk for providing me with the macros to typeset this LaTeX document.
If the university or staff feel that I should take down this document, please feel free to contact me on github

7. Fundamental Subspaces

Definition
Fudamental Subspaces

Let A be an $m \times n$ matrix. The **four fundamental subspaces** of A are

1. $\text{Col}(A) = \{A\vec{x} \in \mathbb{R}^m | \vec{x} \in \mathbb{R}^n\}$, called the **column space**

2. $\text{Row}(A)\{A^T\vec{x} \in \mathbb{R}^n | \vec{x} \in \mathbb{R}^m\}$, called the **row space**

3. $\text{Null}(A) = \{\vec{x} \in \mathbb{R}^n | A\vec{x} = \vec{0}\}$, called the **null space**

4. $\text{Null}(A^T) = \{\vec{x} \in \mathbb{R}^m | A^T\vec{x} = \vec{0}\}$, called the **left nullspace**

Theorem 7.1.1

If A is an $m \times n$ matrix, then $\text{Col}(A)$ and $\text{Null}(A^T)$ are subspaces of $\mathbb{R}^m$, and $\text{Row}(A)$ and $\text{Null}(A)$ are subspaces of $\mathbb{R}^n$

Theorem 7.1.2

If A is an $m \times n$ matrix, then the colums of A which correspond to leading ones in the RREF of A form a basis for $\text{Col}(A)$. Moreover,

$$\dim \text{Col}(A) = \text{rank } A$$

Theorem 7.1.3

If R is an $m \times n$ matrix and E is an $n \times n$ invertible matrix, then

$$\{RE\vec{x} | \vec{x} \in \mathbb{R}^n\} = \{R\vec{y} | \vec{y} \in \mathbb{R}^n\}$$

Theorem 7.1.4

If A is an $m \times n$ matrix, then the non-zero rows in the reduced row echelon form of A form a basis for $\text{Row}(A)$. Hence,

$$\dim \text{Row}(A) = \text{rank } A$$

Corollary 7.1.5

For any $m \times n$ matrix A we have $\text{rank } A = \text{rank } A^T$

Dimension Theorem

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

$$\text{rank } A + \dim \text{Null}(A) = n$$

8. Linear Mappings

Some review

Definition Linear Mapping

A mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ is called **linear** if

$$L(s\vec{x} + t\vec{y}) = sL(\vec{x}) + tL(\vec{y})$$

for all $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $s, t \in \mathbb{R}$
Two linear mappings $L : \mathbb{R}^n \to \mathbb{R}^m$ and $M : \mathbb{R}^n \to \mathbb{R}^m$ are said to be **equal** if $L(\vec{v}) = M(\vec{v}), \forall \vec{v} \in \mathbb{R}^n$

Definition Range

The **range** of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ is defined by

$$\text{Range}(L) = \{L(\vec{x})|\vec{x} \in \mathbb{R}^n\}$$

Definition Kernel

The **kernel** of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ is defined by

$$\text{Ker}(L) = \{\vec{x} \in \mathbb{R}^n | L(\vec{x}) = \vec{0}\}$$

Definition Standard Matrix

The **standard matrix** of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ is defined by

$$[L] = [L(\vec{e}_1) \ \ldots \ L(\vec{e}_n)]$$

It satisfies

$$L(\vec{x}) = [L]\vec{x}$$

for all $\vec{x} \in \mathbb{R}^n$

Theorem 7.2.1

If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear mapping, then $\text{Range}(L) = \text{Col}([L])$

Theorem 7.2.2

If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear mapping, then $\text{Ker}(L) = \text{Null}([L])$

Theorem 7.2.3

Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. Then,

$$\dim(\text{Range}(L)) + \dim(\text{Ker}(L)) = \dim(\mathbb{R}^n)$$

8.1 General Linear Mappings

Linear Mappings $L : \mathbb{V} \to \mathbb{W}$
We will extend our definition of a linear mapping to the case where the domain and codomain are general vector spaces instead of just $\mathbb{R}^n$

Definition
Linear Mapping

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces. A mapping $L : \mathbb{V} \to \mathbb{W}$ is called **linear** if

$$L(s\vec{x} + t\vec{y}) = sL(\vec{x}) + tL(\vec{y})$$

for all $\vec{x}, \vec{y} \in \mathbb{V}$ and $s, t \in \mathbb{R}$
Two linear mappings $L : \mathbb{V} \to \mathbb{W}$ and $M : \mathbb{V} \to \mathbb{W}$ are said to be **equal** if $L(\vec{v}) = M(\vec{v}), \forall \vec{v} \in \mathbb{V}$

Note: A linear mapping $L : \mathbb{V} \to \mathbb{V}$ is called a **linear operator**

Definition
Addition Scalar Multiplication

Let $L : \mathbb{V} \to \mathbb{W}$ and $M : \mathbb{V} \to \mathbb{W}$ be linear mappings. We define $L + M : \mathbb{V} \to \mathbb{W}$ by

$$(L + M)(\vec{v}) = L(\vec{v}) + M(\vec{v})$$

and for any $t \in \mathbb{R}$ we define $tL : \mathbb{V} \to \mathbb{W}$ by

$$(tL)(\vec{v}) = tL(\vec{v})$$

Theorem 8.1.1

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces. The set $\mathbb{L}$ of all linear mappings $L : \mathbb{V} \to \mathbb{W}$ with standard addition and scalar multiplication of mappings is a vector space

Definition
Composition

Let $L : \mathbb{V} \to \mathbb{W}$ and $M : \mathbb{W} \to \mathbb{U}$ be linear mappings. We define $M \circ L : \mathbb{V} \to \mathbb{U}$ by

$$(M \circ L)(\vec{v}) = M(L(\vec{v})),\ \forall \vec{v} \in \mathbb{V}$$

Theorem 8.1.2

If $L : \mathbb{V} \to \mathbb{W}$ and $M : \mathbb{W} \to \mathbb{U}$ are linear mappings, then $M \circ L : \mathbb{V} \to \mathbb{U}$ is a linear mapping

Definition
Invertible Mapping

Let $L : \mathbb{V} \to \mathbb{W}$ and $M : \mathbb{V} \to \mathbb{W}$ be linear mappings. If $(M \circ L)(\vec{v}) = \vec{v}, \forall \vec{v} \in \mathbb{V}$ and $(L \circ M)(\vec{w}) = \vec{w}, \forall \vec{w} \in \mathbb{W}$, then L and M are said to be **invertible**.
We write $M = L^{-1}$ and $L = M^{-1}$

8.2 Rank-Nullity Theorem

Let's extend the definitions of the range and kernel to general linear mappings.

Definition Range Kernel

For a linear mapping $L : \mathbb{V} \to \mathbb{W}$ the **kernel** of L is defined to be

$$\mathrm{Ker}(L) = \{\vec{v} \in \mathbb{V} | L(\vec{v}) = \vec{0}_{\mathbb{W}}\}$$

and the **range** of L is defined to be

$$\mathrm{Range}(L) = \{L(\vec{v}) \in \mathbb{W} | \vec{v} \in \mathbb{V}\}$$

Theorem 8.2.1

If $\mathbb{V}$ and $\mathbb{W}$ are vector spaces and $L : \mathbb{V} \to \mathbb{W}$ is a linear mapping, then

$$L(\vec{0}_{\mathbb{V}}) = \vec{0}_{\mathbb{W}}$$

Theorem 8.2.2

If $L : \mathbb{V} \to \mathbb{W}$ is a linear mapping, then $\mathrm{Ker}(L)$ is a subspace of $\mathbb{V}$ and $\mathrm{Range}(L)$ is a subspace of $\mathbb{W}$

Definition Rank Nullity

Let $L : \mathbb{V} \to \mathbb{W}$ be a linear mapping. We define the **rank** of L to be

$$\mathrm{rank}(L) = \dim(\mathrm{Range}(L))$$

We define the **nullity** of L to be

$$\mathrm{nullity}(L) = \dim(\mathrm{Ker}(L))$$

Rank-Nullity Theorem

Let $\mathbb{V}$ be an n-dimensional vector space and let $\mathbb{W}$ be a vector space. If $L : \mathbb{V} \to \mathbb{W}$ is linear, then

$$\mathrm{rank}(L) + \mathrm{nullity}(L) = n$$

Remark: The proof for the Rank-Nullity Theorem is identical to the proof for the dimension theorem, only that this time it is a generalisation for general vector spaces. You could see the Rank-Nullity theorem as an analog of the Dimension theorem for general linear mappings.

8.3 Matrix of a Linear Mapping

We will now show that every linear mapping $l : \mathbb{V} \to \mathbb{W}$ can also be represented as a matrix mapping. However, we must be careful when dealing with general vector spaces as our domain and codomain. For example, it is certainly impossible to represent a linear mapping that maps polynomials to matrices, since we cannot multiply a matrix by a polynomial.
Thus, if we are going to convert vectors in $\mathbb{V}$ to vectors in $\mathbb{R}^n$ in order to define matrix representations of general linear mappings. Recall the coordinate vector of $\vec{x} \in \mathbb{V}$ with respect to a basis $\mathcal{B}$ is a vector in $\mathbb{R}^n$. In particular, if $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ and $\vec{x} = b_1\vec{v}_1 + \cdots + b_n\vec{v}_n$ then the coordinate vector of $\vec{x}$ with respect to $\mathcal{B}$ is defined to be

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Using coordinates, we can write a matrix mapping representaion of a linear mapping $L : \mathbb{V} \to \mathbb{W}$. We want to find a matrix A such that

$$[L(\vec{x})]_{\mathcal{C}} = A[\vec{x}]_{\mathcal{B}}$$

for every $\vec{x} \in \mathbb{V}$, where $\mathcal{B}$ is a basis for $\mathbb{V}$ and $\mathcal{C}$ is a basis for $\mathbb{W}$
Considering $[L(\vec{x})]_{\mathcal{C}}$ and using the properties of linear mappings and coordinates, we get

$$\begin{aligned} [L(\vec{x})]_{\mathcal{C}} = [L(b_1\vec{v}_1 + \cdots + b_n\vec{v}_n)]_{\mathcal{C}} &= b_1[L(\vec{v}_1)]_{\mathcal{C}} + \cdots + b_n[L(\vec{v}_n)]_{\mathcal{C}} \\ &= \begin{bmatrix} [L(\vec{v}_1)]_{\mathcal{C}} & \ldots & [L(\vec{v}_n)]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \end{aligned}$$

This, we have the matrix $\begin{bmatrix} [L(\vec{v}_1)]_{\mathcal{C}} & \ldots & [L(\vec{v}_n)]_{\mathcal{C}} \end{bmatrix}$ being matrix-vector multiplied by the vector $[\vec{x}]_{\mathcal{B}}$ as desired.

Definition Matrix of a Linear Mapping

Suppose $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for a vector space $\mathbb{V}$ and $\mathcal{C}$ is any basis for a finite dimensional vector space $\mathbb{W}$. For a linear mapping $L : \mathbb{V} \to \mathbb{W}$, the **matirx of L with respect to basis $\mathcal{B}$ and $\mathcal{C}$** is defined by

$${}_{\mathcal{C}}[L]_{\mathcal{B}} = \begin{bmatrix} [L(\vec{v}_1)]_{\mathcal{C}} & \ldots & [L(\vec{v}_n)]_{\mathcal{C}} \end{bmatrix}$$

and satisfies

$$[L(\vec{x})]_{\mathcal{C}} = {}_{\mathcal{C}}[L]_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$$

for all $\vec{x} \in \mathbb{V}$

In the special case of a linear operator L acting on a finite dimensional vector space $\mathbb{V}$ with basis $\mathcal{B}$, we often wish to find the matrix ${}_{\mathcal{B}}[L]_{\mathcal{B}}$

Definition Matrix of a Linear Operator

Suppose $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is any basis for an n-dimensional vector space $\mathbb{V}$ and let $L : \mathbb{V} \to \mathbb{V}$ be a linear operator. The **$\mathcal{B}$-matrix of** L (or the matrix of L with respect to the basis $\mathcal{B}$) is defined by

$$[L]_{\mathcal{B}} = \begin{bmatrix} [L(\vec{v}_1)]_{\mathcal{B}} & \ldots & [L(\vec{v}_n)]_{\mathcal{B}} \end{bmatrix}$$

and satisfies

$$[L(\vec{x})]_{\mathcal{B}} = [L]_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}$$

for all $\vec{x} \in \mathbb{V}$

(Makes more sense if you refer to the course notes) In example 4 we found that the matrix of L with respect to basis $\mathcal{B}$ is diagonal since the vectors in $\mathcal{B}$ are eigenvectors of L. Such a basis is called a **geometrically natural basis**

8.4 Isomorphisms

The ten vector space axioms define a "structure' for the set based on the operations of addition and scalar multiplication. Since all vector spaces satisfy the same ten properties, we expect that all n-dimensional vector spaces should have the same structure, and they do as seen by the work done with coordinate vectors. Whatever basis we use for an n-dimensional vector space $\mathbb{V}$, we have a nice way of relating vectors in $\mathbb{V}$ to vectors in $\mathbb{R}^n$. Moreover, we see that the operations of addition and scalar multiplication are preserved with respect to this basis.
No matter which vector space we are using and which basis for the vector space, any linear linear combination of vectors is really just performed on the coordinates of the vectors with respect to the defined basis.

Let's now look at how to use general linear mappings to mathematically prove these observations.

Let $L : \mathbb{V} \to \mathbb{W}$ be a linear mapping. L is called **one-to-one** (**injective**) if for every $\vec{u}, \vec{v} \in \mathbb{V}$ such that $L(\vec{u}) = L(\vec{v})$, we must have $\vec{u} = \vec{v}$

Definition
One-To-One
Onto

L is called **onto** (**surjective**) if for every $\vec{w} \in \mathbb{W}$, there exists $\vec{v} \in \mathbb{V}$ such that $L(\vec{v}) = \vec{w}$

We observe that for a linear mapping $L : \mathbb{V} \to \mathbb{W}$ being onto means that Range$(L) = \mathbb{W}$
And L being one-to-one means that for each $\vec{w} \in$ Range(L) there exists exactly one $\vec{v} \in \mathbb{V}$ such that $L(\vec{v}) = \vec{w}$. We now establish a relationship between a mapping being one-to-one and its kernel

Let $L : \mathbb{V} \to \mathbb{W}$ be a linear mapping. L is one-to-one (injective) if and only if Ker$(L)= \{\vec{0}_{\mathbb{V}}\}$

Lemma 8.4.1

For two vector spaces $\mathbb{V}$ and $\mathbb{W}$ to have the same structure we need each vector $\vec{v} \in \mathbb{V}$ to be identified with a unique vector $\vec{w} \in \mathbb{W}$ such that linear combinations are preserved. So to do that we need to find a mapping L from $\mathbb{V}$ to $\mathbb{W}$ which is both one-to-one and onto, and for linear combinations to be preserved we need the linear mapping to be linear.

A vector space $\mathbb{V}$ is said to be **isomorphic** to a vector space $\mathbb{W}$ if there exists a linear mapping $L : \mathbb{V} \to \mathbb{W}$ which is one-to-one and onto. L is called an **isomorphism** from $\mathbb{V}$ to $\mathbb{W}$

Definition
Isomorphism
Isomorphic

Looking at examples of isomorphisms will show that it is mapping basis vectors to basis vectors. And since it is that basis vectors are mapping to basis vectors, the dimensions of the two vector spaces should be the same

Let $\mathbb{V}$ and $\mathbb{W}$ be finite dimensional vector spaces. $\mathbb{V}$ is isomorphic to $\mathbb{W}$ if and only if $\dim \mathbb{V} = \dim \mathbb{W}$

Theorem 8.4.2

If $\mathbb{V}$ and $\mathbb{W}$ are both n-dimensional vector spaces and $L : \mathbb{V} \to \mathbb{W}$ is linear, then L is one-to-one if and only if L is onto — Theorem 8.4.3

Let $\mathbb{V}$ and $\mathbb{W}$ be isomorphic vector spaces and let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be a basis for $\mathbb{V}$. A linear mapping $L : \mathbb{V} \to \mathbb{W}$ is an isomorphism if and only if $\{L(\vec{v}_1), \ldots, L(\vec{v}_n)\}$ is a basis for $\mathbb{W}$ — Theorem 8.4.4

Linear Extensions

Suppose you know that the linear transformation $T : P_2(\mathbb{R}) \to P_1(\mathbb{R})$
Where, $T(1) = -3 + 2x, T(x) = 2 + 3x, T(x^2) = x^2$
In order to compute $ax^2 + bx + c$ you would rewrite $T(ax^2 + bx + c) = aT(x^2) + bT(x) + cT(1)$
And so, by assuming linearity, you **determined** the image of $ax^2 + bx + c$ **by the action of T on a basis** $\{1, x, x^2\}$

Let $S, T : \mathbb{V} \to \mathbb{W}$ be linear transformations and let $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be a basis of $\mathbb{V}$ — Lemma 8.4.X
If $S(\vec{b}_i) = T(\vec{b}_i)$ for each i, then $S = T$
This can be proven by showing that $S(\vec{v}) = T(\vec{v})$ for every $\vec{v} \in \mathbb{V}$

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces, and let $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be a basis of $\mathbb{V}$ — Theorem 8.4.Y
Furthermore, let $f : \mathcal{B} \to \mathbb{W}$ be a function
Then there exists a **unique** linear transformation $T : \mathbb{V} \to \mathbb{W}$ such that

$$T(\vec{b}_i) = f(\vec{b}_i)$$

for each i. Moreover, the general formula for an image of T is

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

After proving that a T exists, uniqueness follows immediately from Lemma 8.4.X

Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces, and let $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be a basis of $\mathbb{V}$ — Definition Linear Extension
Furthermore, let $f : \mathcal{B} \to \mathbb{W}$ be a function.
Then the **linear extension** of f is the linear transformation $T : \mathbb{V} \to \mathbb{W}$ where

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

Every linear transformation is the linear extension of its action on a basis

9. Inner Products

9.1 Inner Product Spaces

Definition — Inner Product, Inner Product Space

Let $\mathbb{V}$ be a vector space. An **inner product** on $\mathbb{V}$ is a function $\langle , \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ that has the following properties: for every $\vec{v}, \vec{u}, \vec{w} \in \mathbb{V}$ and $s, t \in \mathbb{R}$ we have

- I1 $\langle \vec{v}, \vec{v} \rangle \geq 0$, and $\langle \vec{v}, \vec{v} \rangle = 0$ iff $\vec{v} = \vec{0}$ (**Positive Definite**)
- I2 $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$ (**Symmetric**)
- I3 $\langle s\vec{v} + t\vec{u}, \vec{w} \rangle = s\langle \vec{v}, \vec{w} \rangle + t\langle \vec{u}, \vec{w} \rangle$ (**Left linear**)

A vector space $\mathbb{V}$ with an inner product $\langle , \rangle$ on $\mathbb{V}$ is called an **inner product space**

Remark: Since an inner product is left linear and symmetric, then it is also **right linear**:

$$\langle \vec{w}, s\vec{v} + t\vec{u} \rangle = s\langle \vec{w}, \vec{v} \rangle + t\langle \vec{w}, \vec{u} \rangle$$

Thus, we say that an inner product is **bilinear**

The dot product is an inner product on $\mathbb{R}^n$, called the **standard inner product** on $\mathbb{R}^n$
$\langle A, B \rangle = tr(B^T A)$ on $M_{m \times n}(\mathbb{R})$ is called the **standard inner product** on $M_{m \times n}(\mathbb{R})$

Observe that calculating $\langle A, B \rangle$ corresponds exactly to finding the dot product on vectors in $\mathbb{R}^{mn}$ under the obvious isomorphism (one that matches the trace with the dot product expression). Hence, when using this inner product you do not actually have to compute $B^T A$ (and can get away with using the isomorphism and calculating the dot product)

The inner product $\langle f(x), g(x) \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$ on the vector space $C[-\pi, \pi]$ of continuous functions defined on the closed interval from $-\pi$ to π. It is extremely important in applied math, physics and engineering. This is the foundation for Fourier Series.

Theorem 9.1.1

If $\mathbb{V}$ is an inner product space with inner product $\langle , \rangle$, then for any $v \in \mathbb{V}$ we have

$$\langle \vec{v}, \vec{0} \rangle = 0$$

Remark: Whenever one considers an inner product space, they must define which inner product space they are using. For $\mathbb{R}^n$ or $M_{m \times n}(\mathbb{R})$ it's generally the standard inner product, this is the one we assume is being used if no other inner product is defined.

9.2 Orthogonality and Length

Length

Let $\mathbb{V}$ be an inner product space with inner product $\langle , \rangle$. For any $\vec{v} \in \mathbb{V}$ we define the **length** (or **norm**) of $\vec{v}$ by

Definition
Length

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

The length of a vector is well defined as the inner product must be positive definite.

Let $\mathbb{V}$ be an inner product space with inner product $\langle , \rangle$. For any $\vec{x}, \vec{y} \in \mathbb{V}$ and $t \in \mathbb{R}$ we have

Theorem
9.2.1

(1) $||\vec{x}|| \geq 0$, and $||\vec{x}|| = 0$ iff $\vec{x} = \vec{0}$

(2) $||t\vec{v}|| = |t|||\vec{v}||$

(3) $\langle \vec{x}, \vec{y} \rangle \leq ||\vec{x}||||\vec{y}||$ Cauchy-Schwarz-Bunyakovski Inequality

(4) $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$ Triangle Inequality

Let $\mathbb{V}$ be an inner product space with inner product $\langle , \rangle$. If $\vec{v} \in \mathbb{V}$ is a vector such that $||\vec{v}|| = 1$, then $\vec{v}$ is called a **unit vector**

Definition
Unit vector

We often find a unit vector, $\hat{v}$, in the direction of a certain vector $\vec{v} \in \mathbb{V}$, this is called **normalizing** the vector. By theorem 9.2.1 we see

$$\hat{v} = \frac{1}{||\vec{v}||}\vec{v}$$

Orthogonality

Let $\mathbb{V}$ be an inner product space with inner product $\langle , \rangle$. If $\vec{x}, \vec{y} \in \mathbb{V}$ such that

Definition
Orthogonal
vectors

$$\langle \vec{x}, \vec{y} \rangle = 0$$

then $\vec{x}$ and $\vec{y}$ are said to be **orthogonal**

Just like length, whether two vectors are orthogonal or not is dependent on its definition of the inner product.

If $\mathcal{S} = \{\vec{v}_1, \dots, \vec{v}_k\}$ is a set of vectors in an inner product space $\mathbb{V}$ with inner product $\langle , \rangle$ such that $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ for all $i \neq j$, then $\mathcal{S}$ is called an **orthogonal set**

Definition
Orthogonal
Set

If $\{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthogonal set in an inner product space $\mathbb{V}$, then

Theorem
9.2.2

$$||\vec{v}_1 + \cdots + \vec{v}_k||^2 = ||\vec{v}_1||^2 + \cdots + ||\vec{v}_k||^2$$

If $S = \{\vec{v}_1, \dots, \vec{v}_k\}$ is an orthogonal set in an inner product space $\mathbb{V}$ with inner product $\langle , \rangle$ such that $\vec{v}_i \neq 0$ for all $1 \leq i \leq k$, then S is linearly independent

Theorem
9.2.3

If $\mathcal{B}$ is an orthogonal set in an inner product space $\mathbb{V}$ that is a basis for $\mathbb{V}$, then $\mathcal{B}$ is called an **orthogonal basis** for $\mathbb{V}$ — Definition: Orthogonal Basis

If $\mathcal{S} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthogonal basis for an inner product space $\mathbb{V}$ with linear product $\langle , \rangle$ and $\vec{v} \in \mathbb{V}$, then the coefficient of $\vec{v}_i$ when $\vec{v}$ is written as a linear combination of the vectors in $\mathcal{S}$ is $\frac{\langle \vec{v}, \vec{v}_i \rangle}{||\vec{v}_i||^2}$. In particular, — Theorem 9.2.4

$$\vec{v} = \frac{\langle \vec{v}, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 + \cdots + \frac{\langle \vec{v}, \vec{v}_n \rangle}{||\vec{v}_n||^2}\vec{v}_n$$

Orthonormal Bases

Since the formula for the coordinates would be simpler if the vectors in the orthogonal basis were unit vectors. So we will **normalize** vectors so that they are of unit length. $\hat{v} = \frac{\vec{v}}{||\vec{v}||}$

If $\mathcal{S} = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal set in the inner product space $\mathbb{V}$ such that $||\vec{v}_i|| = 1$ for $1 \leq i \leq k$, then $\mathcal{S}$ is called an **orthonormal set** — Definition: Orthonormal Set

A basis for an inner product space $\mathbb{V}$ which is an orthonormal set is called an **orthonormal basis** of $\mathbb{V}$ — Definition: Orthonormal Basis

If $\vec{v}$ is any vector in an inner product space $\mathbb{V}$ with inner product $\langle , \rangle$ and $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{V}$, then — Corollary 9.2.5

$$\vec{v} = \langle \vec{v}, \vec{v}_1 \rangle \vec{v}_1 + \cdots + \langle \vec{v}, \vec{v}_n \rangle \vec{v}_n$$

Orthogonal Matrices

For an $n \times n$ matrix P, the following are equivalent — Theorem 9.2.6

1. The columns of P form an orthonormal basis for $\mathbb{R}^n$
2. $P^T = P^{-1}$
3. The rows of P form an orthonormal basis for $\mathbb{R}^n$

If the columns of an $n \times n$ matrix P form an orthonormal basis for $\mathbb{R}^n$, then P is called an **orthogonal matrix** — Definition: Orthogonal Matrix

If P and Q are $n \times n$ orthogonal matrices and $\vec{x}, \vec{y} \in \mathbb{R}^n$, then — Theorem 9.2.7

1. $(P\vec{x}) \cdot (P\vec{y}) = \vec{x} \cdot \vec{y}$
2. $||P\vec{x}|| = ||\vec{x}||$
3. $\det P = \pm 1$
4. All real eigenvalues of P are 1 or -1
5. PQ is an orthogonal matrix

9.3 The Gram-Schmidt Procedure

(**Gram-Schmidt Orthogonalization Theorem**) Theorem 9.3.1

Let $\{\vec{w}_1, \ldots, \vec{w}_n\}$ be a basis for an inner product space $\mathbb{W}$. If we define $\vec{v}_1, \ldots, \vec{v}_n$ successively as follows:

$$\vec{v}_1 = \vec{w}_1$$
$$\vec{v}_2 = \vec{w}_2 - \frac{\langle \vec{w}_2, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1$$
$$\vec{v}_i = \vec{w}_i - \frac{\langle \vec{w}_i, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 - \frac{\langle \vec{w}_i, \vec{v}_2 \rangle}{||\vec{v}_2||^2}\vec{v}_2 - \cdots - \frac{\langle \vec{w}_i, \vec{v}_{i-1} \rangle}{||\vec{v}_{i-1}||^2}\vec{v}_{i-1}$$

for $3 \leq k \leq n$, then $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal basis for $\text{Span}\{\vec{w}_1, \ldots, \vec{w}_k\}$ for $1 \leq k \leq n$

Note: the theorem implies that every finite dimensional inner product space has an orthogonal basis.
The process for finding an orthogonal basis for an inner product space in the theorem is called the **Gram-Schmidt procedure**
The order in which one processes the vectors whilst applying the procedure changes what the resultant basis looks like.

(QR-Decomposition) Theorem 9.3.2

Let $A = [\vec{a}_1 \ \ldots \ \vec{a}_n] \in M_{m\times n}(\mathbb{R})$ where dim $\text{Col}(A) = n$. Let $\vec{q}_1, \ldots, \vec{q}_n$ denote the vectors that result from applying the Gram-Schmidt procedure to the columns of A (in order) and then normalizing. If we define

$$Q = [\vec{q}_1 \ \cdots \ \vec{q}_n]$$

and

$$R = \begin{bmatrix} \vec{a}_1 \cdot \vec{q}_1 & \vec{a}_2 \cdot \vec{q}_1 & \ldots & \vec{a}_n \cdot \vec{q}_1 \\ 0 & \vec{a}_2 \cdot \vec{q}_2 & \ldots & \vec{a}_n \cdot \vec{q}_2 \\ 0 & 0 & \ddots & \vec{a}_n \cdot \vec{q}_{n-1} \\ 0 & \ldots & 0 & \vec{a}_n \cdot \vec{q}_n \end{bmatrix}$$

then Q is orthogonal, R is invertible, and

$$A = QR$$

9.4 General Projections

As the title suggests, we will abstract the ideas of projections from our works that used the standard inner product in $\mathbb{R}^n$. Given a vector $\vec{x} \in \mathbb{R}^3$ and a plane $P \in \mathbb{R}^3$ which passes through the origin (a subspace), we wanted to write $\vec{x}$ as the sum of a vector in P and a vector orthogonal to every vector in P. Therefore, for a given subspace $\mathbb{W}$ of an inner product space $\mathbb{V}$ and any vector $\vec{v} \in \mathbb{V}$

$$\vec{v} = \text{proj}_{\mathbb{W}}(\vec{v}) + \text{perp}_{\mathbb{W}}(\vec{v})$$

Definition
Orthogonal Complement

Let $\mathbb{W}$ be a subspace of an inner product space $\mathbb{V}$. The **orthogonal complement** $\mathbb{W}^{\perp}$ of $\mathbb{W}$ in $\mathbb{V}$ is defined by

$$\mathbb{W}^{\perp} = \{\vec{v} \in \mathbb{V} | \langle \vec{w}, \vec{v} \rangle = 0 \text{ for all } \vec{w} \in \mathbb{W}\}$$

Theorem 9.4.1

Let $\{\vec{v}_1, \ldots, \vec{v}_k\}$ be a spanning set for a subspace $\mathbb{W}$ of an inner product space $\mathbb{V}$, and let $\vec{x} \in \mathbb{V}$. We have that $\vec{x} \in \mathbb{W}^{\perp}$ if and only if $\langle \vec{x}, \vec{v}_i \rangle = 0$ for $1 \leq i \leq k$

Theorem 9.4.2

If $\mathbb{W}$ is a subspace of an inner product space $\mathbb{V}$, then

1. $\mathbb{W}^{\perp}$ is a subspace of $\mathbb{V}$
2. If dim $\mathbb{V} = n$, then dim $\mathbb{W}^{\perp} = n -$dim $\mathbb{W}$
3. If dim $\mathbb{V} = n$, then $(\mathbb{W}^{\perp})^{\perp} = \mathbb{W}$
4. $\mathbb{W} \cap \mathbb{W}_{\perp} = \{\vec{0}\}$
5. If dim $\mathbb{V} = n$, $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal basis for $\mathbb{W}$, and $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is an orthogonal basis for $\mathbb{W}^{\perp}$, then $\{\vec{v}_1, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is an orthogonal basis for $\mathbb{V}$

Definition
Projection, Perpendicular

Suppose $\mathbb{W}$ is a k-dimensional subspace of an inner product space $\mathbb{V}$ and $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal basis for $\mathbb{W}$. For any $\vec{v} \in \mathbb{V}$ we define the **projection** of $\vec{v}$ onto $\mathbb{W}$ by

$$\text{proj}_{\mathbb{W}}(\vec{v}) = \frac{\langle \vec{v}, \vec{v}_1 \rangle}{||\vec{v}_1||^2}\vec{v}_1 + \cdots + \frac{\langle \vec{v}, \vec{v}_k \rangle}{||\vec{v}_k||^2}\vec{v}_k$$

and the **perpendicular** of the projection by

$$\text{perp}_{\mathbb{W}}(\vec{v}) = \vec{v} - \text{proj}_{\mathbb{W}}(\vec{v})$$

Notice that we need an *orthogonal basis* for $\mathbb{W}$ to calculate the projection. Hence, these are often called **orthogonal projections**.
To save ourselves some work, we defined the perpendicular of the projection in such a way that we do not need an orthogonal basis (or any basis) for $\mathbb{W}^{\perp}$.

Theorem 9.4.3

If $\mathbb{W}$ is a k-dimensional subspace of an inner product space $\mathbb{V}$, then for any $\vec{v} \in \mathbb{V}$ we have

$$\text{perp}_{\mathbb{W}}(\vec{v}) = \vec{v} - \text{proj}_{\mathbb{W}}(\vec{v}) \in \mathbb{W}^{\perp}$$

If $\mathbb{W}$ is a k-dimensional subspace of an inner product space $\mathbb{V}$, then $\mathrm{proj}_{\mathbb{W}}$ is a linear operator on $\mathbb{V}$ with kernel $\mathbb{W}^{\perp}$ Theorem 9.4.4

If $\mathbb{W}$ is a subspace of a finite dimensional inner product space $\mathbb{V}$, then for any $\vec{v} \in \mathbb{V}$ we have Theorem 9.4.5

$$\mathrm{proj}_{\mathbb{W}^{\perp}}(\vec{v}) = \mathrm{perp}_{\mathbb{W}}(\vec{v})$$