PMath 225 - Applied Linear Algebra 2

Course Notes, University of Waterloo

Tyrone Ghaswala

Fall 2016

Contents

These notes are for the Fall 2016 offering of PMath 225 - Applied Linear Algebra 2. They will be updated as I go, and are definitely not free of typos and mistakes. If you find any, please let me know about it! These notes are by no means comprehensive, and are a bare bones version of the content in the course. Your notes (or your friend's notes) from class are to serve as the official text for the course.

For the sake of having a cross reference, I have included in parentheses the part(s) of the textbook that closest resemble what I covered in class for that section.

For this course, $\mathbb{F}$ will denote the complex numbers $\mathbb{C}$ or the real numbers $\mathbb{R}$.

1 Vector Spaces (§4.1,4.2, 9.3)

Linear algebra is the study of vector spaces. Here are some to get your toes wet.

The vector space $\mathbb{R}^n$ is given by

$$\mathbb{R}^n := \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} : a_i \in \mathbb{R} \text{ for all } i \right\}.$$

Addition and scalar multiplication are given by

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix} \quad \text{and} \quad \alpha \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \alpha a_1 \\ \vdots \\ \alpha a_n \end{pmatrix}.$$

The vector space $\mathcal{P}_n(\mathbb{R})$ is the set of polynomials of degree at most n with real coefficients. That is

$$\mathcal{P}_n(\mathbb{R}) := \{a_n x^n + \cdots + a_1 x + a_0 : a_i \in \mathbb{R} \text{ for all } i\}$$

with addition and scalar multiplication defined by

$$(a_n x^n + \cdots + a_0) + (b_n x^n + \cdots + b_0) = (a_n + b_n)x^n + \cdots + (a_0 + b_0) \quad \text{and} \quad \alpha(a_n x^n + \cdots + a_0) = (\alpha a_n x^n + \cdots + \alpha a_0).$$

The vector space of n by m matrices with real coefficients is given by

$$M_{n\times m}(\mathbb{R}) := \left\{ \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} : a_{ij} \in \mathbb{R} \text{ for all } i, j \right\}.$$

Addition and scalar multiplication are given by matrix addition and scalar multiplication of matrices as usual.

The vector space of continuous functions is denoted by

$$\mathcal{C}([0,1]) := \{f : [0,1] \to \mathbb{R} : f \text{ is continuous}\}.$$

Addition and scalar multiplication are defined by

$$(f+g)(x) := f(x) + g(x) \quad \text{and} \quad (\alpha f)(x) := \alpha(f(x)).$$

Here's a slightly more interesting one. Let $\mathbb{V}$ be the set of all lines in $\mathbb{R}^2$ with slope 1. Each line has equation $y = x + d$. Addition and scalar multiplication in $\mathbb{V}$ is defined by

$$(y = x + d_1) + (y = x + d_2) := (y = x + (d_1 + d_2)) \quad \text{and} \quad \alpha(y = x + d) := (y = x + \alpha d).$$

Vector spaces can also have the complex numbers as the scalars, instead of the real numbers. The complex 2-dimension vector space is given by

$$\mathbb{C}^2 := \left\{ \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} : a_1, a_2 \in \mathbb{C} \right\}$$

with addition and scalar multiplication exactly the same as in $\mathbb{R}^2$, except all the numbers in sight are in $\mathbb{C}$.

Definition. A **vector space** over a field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$ for this course) is a set $\mathbb{V}$ together with operations $+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ (addition) and $\cdot : \mathbb{F} \times \mathbb{V} \to \mathbb{V}$ (scalar multiplication) such that for all $\vec{x}, \vec{y}, \vec{z} \in \mathbb{V}$ and $s, t \in \mathbb{F}$ the following hold.

1. $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$,
2. There exists a vector $\vec{0} \in \mathbb{V}$ such that $\vec{0} + \vec{x} = \vec{x} + \vec{0} = \vec{x}$,
3. For every $\vec{x} \in \mathbb{V}$, there exists a $\vec{-x} \in \mathbb{V}$ such that $\vec{x} + \vec{-x} = \vec{-x} + \vec{x} = \vec{0}$,
4. $\vec{x} + \vec{y} = \vec{y} + \vec{x}$,
5. $s(t\vec{x}) = (st)\vec{x}$,
6. $(s + t)\vec{x} = s\vec{x} + t\vec{x}$,
7. $s(\vec{x} + \vec{y}) = s\vec{x} + s\vec{y}$, and
8. $1\vec{x} = \vec{x}$.

Fact 1. *Let $\mathbb{V}$ be a vector space over a field $\mathbb{F}$. Then*

- $0\vec{x} = \vec{0}$ *for all* $\vec{x} \in \mathbb{V}$,
- $(-1)\vec{x} = \vec{-x}$ *for all* $\vec{x} \in \mathbb{V}$, *and*
- $t\vec{0} = \vec{0}$ *for all* $t \in \mathbb{F}$.

2 Subspaces (§4.2, 9.3)

Definition. Let $\mathbb{V}$ be a vector space and $\mathbb{U} \subset \mathbb{V}$ a subset. We call $\mathbb{U}$ a **subspace** of $\mathbb{V}$ if $\mathbb{U}$, endowed with the addition and scalar multiplication from $\mathbb{V}$, is a vector space.

Theorem 2 (The subspace test)**.** *Suppose $\mathbb{U}$ is a subset of a vector space $\mathbb{V}$ over $\mathbb{F}$. If*

- $\mathbb{U}$ *is non-empty,*
- *For all* $\vec{u_1}, \vec{u_2} \in \mathbb{U}$, $\vec{u_1} + \vec{u_2} \in \mathbb{U}$, *and*
- *For all* $\alpha \in \mathbb{F}$ *and for all* $\vec{u} \in \mathbb{U}$, $\alpha\vec{u} \in \mathbb{U}$,

then $\mathbb{U}$ is a subspace of $\mathbb{V}$.

Definition. Let $\mathcal{B} = \{\vec{v_1}, \ldots, \vec{v_k}\}$ be a subset of a vector space $\mathbb{V}$. Define the **span** of $\mathcal{B}$ by

$$\mathrm{Span}(\mathcal{B}) := \{t_1\vec{v_1} + \cdots + t_k\vec{v_k} : t_1, \ldots, t_k \in \mathbb{F}\}.$$

Proposition 3. *Let $\mathcal{B} = \{\vec{v_1}, \ldots, \vec{v_k}\}$ be a subset of a vector space $\mathbb{V}$. Then* $\mathrm{Span}(\mathcal{B})$ *is a subspace of* $\mathbb{V}$.

3 Bases and Dimension (§4.2, 4.3, 4.4, 9.3)

Definition. A set of vectors $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_k\}$ in a vector space $\mathbb{V}$ is a **spanning set** for $\mathbb{V}$, and we say $\mathcal{B}$ **spans** $\mathbb{V}$, if $\text{Span}(\mathcal{B}) = \mathbb{V}$.

Definition. A set of vectors $\{\vec{v}_1, \dots, \vec{v}_k\}$ in a vector space $\mathbb{V}$ is **linearly independent** if the only solution to the equation

$$t_1\vec{v}_1 + \cdots + t_k\vec{v}_k = \vec{0}$$

is $t_1 = \cdots = t_k = 0$. The set is **linearly dependent** otherwise.

Definition. A **basis** for a vector space $\mathbb{V}$ is a linearly independent subset that spans $\mathbb{V}$.

Fact 4. *Every vector space has a basis.*

Lemma 5. *Let $\mathbb{V}$ be a vector space and $\text{Span}(\{\vec{v}_1, \dots, \vec{v}_n\}) = \mathbb{V}$. If $\{\vec{u}_1, \dots, \vec{u}_k\}$ is a linearly independent set in $\mathbb{V}$, then $k \leq n$.*

Theorem 6. *Suppose $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ and $\mathcal{C} = \{\vec{u}_1, \dots, \vec{u}_k\}$ are both bases of a vector space $\mathbb{V}$. Then $k = n$.*

Definition. The **dimension** of a vector space $\mathbb{V}$, denoted $\dim(\mathbb{V})$, is the size of any basis.

Note that theorem 6 shows that this definition makes sense.

Theorem 7. *Let $\mathbb{V}$ be an n-dimensional vector space. Then*

1. *A set of more than n vectors must be linearly dependent.*
2. *A set of fewer than n vectors cannot span $\mathbb{V}$.*
3. *A set with n elements in $\mathbb{V}$ is a spanning set for $\mathbb{V}$ if and only if it is linearly independent.*

3.1 Obtaining Bases

There are many ways you could find a basis for a finite-dimensional vector space, here are a couple of important ways.

1. **Extending a linearly independent subset**. Suppose you have a linearly independent subset $\{\vec{v}_1, \dots, \vec{v}_k\}$ in a finite dimensional vector space $\mathbb{V}$. If it is a spanning set, then you have a basis. If not, choose a vector $\vec{v}_{k+1}$ not in the span of $\{\vec{v}_1, \dots, \vec{v}_k\}$. Then $\{\vec{v}_1, \dots, \vec{v}_{k+1}\}$ is must be linearly independent. If this new set spans, then it's a basis. If not, then repeat. This process must eventually stop since our vector space is finite dimensional, and you will be left with a basis containing $\{\vec{v}_1, \dots, \vec{v}_k\}$.

2. **Reducing an arbitrary finite spanning set**. Suppose you have a finite spanning set $\{\vec{v}_1, \dots, \vec{v}_k\}$ for your vector space, and let's assume that it doesn't contain $\vec{0}$. If it is linearly independent, it is a basis! If not, you can write one of them, say v_i, as a linear combination of the others. Now $\text{Span}(\{\vec{v}_1, \dots, \vec{v}_k\}) = \text{Span}(\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_k\})$, so $\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_k\}$ spans our vector space. If this new set is linearly independent, then it is a basis! If not, repeat to remove another vector. This process must eventually stop since we started with finitely many vectors in our spanning set. The final product will be a basis made up entirely out of vectors from our original spanning set.

Coordinates with Respect to a Basis

Theorem 8. *Let $\mathbb{V}$ be a vector space, let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ be a subset of $\mathbb{V}$, and let $\mathbb{U} = \mathrm{Span}(\mathcal{B})$. Then every vector in $\mathbb{U}$ can be expressed in a unique way as a linear combination of the vectors in $\mathcal{B}$ if and only if $\mathcal{B}$ is linearly independent.*

Definition. Let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ be a basis for a vector space $\mathbb{V}$. If $\vec{x} \in \mathbb{V}$ with $\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n$, then the **coordinate vector** of $\vec{x}$ with respect to $\mathcal{B}$ is

$$[\vec{x}]_{\mathcal{B}} := \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Note that we are assuming the order of our vectors in our basis is fixed.

4 Linear Maps (§4.5, 4.6, 9.3)

Definition. If $\mathbb{V}$ and $\mathbb{W}$ are vector spaces over $\mathbb{F}$, a function $L : \mathbb{V} \to \mathbb{W}$ is a **linear map** if it satisfies the linearity properties:

1. $L(\vec{x} + \vec{y}) = L(\vec{x}) + L(\vec{y})$, and
2. $L(t\vec{x}) = tL(\vec{x})$

for all $\vec{x}, \vec{y} \in \mathbb{V}$, $t \in \mathbb{F}$.

Definition. Let $L : \mathbb{V} \to \mathbb{W}$ be a linear map. The **range** of L is

$$\mathrm{Range}(L) := \{L(\vec{x}) \in \mathbb{W} : \vec{x} \in \mathbb{V}\}.$$

The **kernel** or **nullspace** of L is

$$\mathrm{Null}(L) := \{\vec{x} \in \mathbb{V} : L(\vec{x}) = \vec{0}\}.$$

Theorem 9. *Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$, and let $L : \mathbb{V} \to \mathbb{W}$ be a linear map. Then*

1. $L(\vec{0}) = \vec{0}$,
2. $\mathrm{Null}(L)$ *is a subspace of* $\mathbb{V}$, *and*
3. $\mathrm{Range}(L)$ *is a subspace of* $\mathbb{W}$.

Definition. Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$. The **rank** of a linear map $L : \mathbb{V} \to \mathbb{W}$ is the dimension of the range of L. The **nullity** of L is the dimension of the nullspace of L. That is,

$$\mathrm{rank}(L) = \dim(\mathrm{Range}(L)) \quad \text{and} \quad \mathrm{nullity}(L) = \dim(\mathrm{Null}(L)).$$

Theorem 10 (Rank-Nullity Theorem)**.** *Let $\mathbb{V}$ and $\mathbb{W}$ be vector spaces over $\mathbb{F}$ with $\dim(\mathbb{V}) = n$. Let $L : \mathbb{V} \to \mathbb{W}$ be a linear map. Then $\mathrm{rank}(L) + \mathrm{nullity}(L) = n$.*

4.1 Linear Maps as Matrices

Theorem 11. *Let $\mathbb{V}$ be an n-dimensional vector space with basis $\mathcal{B}$. Let $\mathbb{W}$ be an m-dimensional vector space with basis $\mathcal{C}$. Then for every linear map $L : \mathbb{V} \to \mathbb{W}$, there exists an $m \times n$ matrix A such that $[L(\vec{v})]_{\mathcal{C}} = A[\vec{v}]_{\mathcal{B}}$ for all $\vec{v} \in \mathbb{V}$. Conversely, every $m \times n$ matrix A defines a linear map $L : \mathbb{V} \to \mathbb{W}$ by $[L(\vec{v})]_{\mathcal{C}} := A[\vec{v}]_{\mathcal{B}}$.*

Corollary 12. *Let $\mathbb{V}$ be a vector space with basis $\mathcal{B} = \{\vec{\beta}_1, \ldots, \vec{\beta}_n\}$. Let $\mathbb{W}$ be a vector space with basis $\mathcal{C} = \{\vec{\gamma}_1, \ldots, \vec{\gamma}_m\}$. Let $L : \mathbb{V} \to \mathbb{W}$ be a linear map. Then the $m \times n$ matrix A such that $[L(\vec{v})]_{\mathcal{C}} = A[\vec{v}]_{\mathcal{B}}$ for all $\vec{v} \in \mathbb{V}$, which we denote ${}_{\mathcal{C}}[L]_{\mathcal{B}}$, is given by*

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

Definition. We call the matrix ${}_{\mathcal{C}}[L]_{\mathcal{B}}$ the **matrix of the linear mapping** L with respect to the bases $\mathcal{B}$ and $\mathcal{C}$. If $L : \mathbb{V} \to \mathbb{V}$ and we are choosing the same basis $\mathcal{B}$ for both the domain and codomain of L, then we may write $[L]_{\mathcal{B}} = {}_{\mathcal{B}}[L]_{\mathcal{B}}$.

Fact 13. *Let $\mathbb{V}$, $\mathbb{U}$, and $\mathbb{W}$ be vector spaces with bases $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ respectively. Let $L : \mathbb{V} \to \mathbb{U}$ and $M : \mathbb{U} \to \mathbb{W}$ be linear maps. Then ${}_{\mathcal{D}}[M \circ L]_{\mathcal{B}} = {}_{\mathcal{D}}[M]_{\mathcal{C}}\,{}_{\mathcal{C}}[L]_{\mathcal{B}}$.*

Definition. Let $\mathbb{V}$ be a finite dimensional vector space, and let $\mathcal{B}$ and $\mathcal{C}$ be two bases for $\mathbb{V}$. The **change of coordinate matrix** ${}_{\mathcal{C}}P_{\mathcal{B}}$ is the matrix of the linear mapping $\mathrm{id} : \mathbb{V} \to \mathbb{V}$ where $\mathrm{id}(\vec{v}) = \vec{v}$ for all $v \in \mathbb{V}$. This name makes sense since $[\vec{v}]_{\mathcal{C}} = {}_{\mathcal{C}}P_{\mathcal{B}}[\vec{v}]_{\mathcal{B}}$ for all $\vec{v} \in \mathbb{V}$.

Fact 14. *Let $\mathbb{V}$ be a finite dimensional vector space with bases $\mathcal{B}$ and $\mathcal{C}$. Then ${}_{\mathcal{C}}P_{\mathcal{B}} = ({}_{\mathcal{B}}P_{\mathcal{C}})^{-1}$.*

5 Isomorphisms of Vector Spaces (§4.7)

Definition. Let $L : \mathbb{V} \to \mathbb{W}$ be a linear map between vector spaces. We say L is **injective** (or one-to-one) if $L(\vec{v}_1) = L(\vec{v}_2)$ implies $\vec{v}_1 = \vec{v}_2$. We say L is **surjective** (or onto) if $\mathrm{Range}(L) = \mathbb{W}$.

Lemma 15. *A linear map L is injective if and only if* $\mathrm{Null}(L) = \{\vec{0}\}$.

Definition. Let $L : \mathbb{V} \to \mathbb{W}$ be a linear map. If L is injective and surjective, we say L is an **isomorphism**. If L is an isomorphism, we say the vector spaces $\mathbb{V}$ and $\mathbb{W}$ are **isomorphic** and write $\mathbb{V} \cong \mathbb{W}$.

Lemma 16. *A linear map $L : \mathbb{V} \to \mathbb{W}$ is an isomorphism if and only if there exists a linear map $L^{-1} : \mathbb{W} \to \mathbb{V}$ such that $L \circ L^{-1}(\vec{w}) = \vec{w}$ for all $\vec{w} \in \mathbb{W}$ and $L^{-1} \circ L(\vec{v}) = \vec{v}$ for all $\vec{v} \in \mathbb{V}$. In this case we call L^{-1} the* **inverse linear map** *to L.*

Theorem 17. *Suppose $\mathbb{U}$ and $\mathbb{V}$ are finite dimensional vector spaces. Then $\mathbb{U}$ and $\mathbb{V}$ are isomorphic if and only if* $\dim(\mathbb{U}) = \dim(\mathbb{V})$.

6 Determinants (§5.1, 5.2, 5.4)

It only makes sense to talk about determinants and inverses of matrices when the matrices in question are square. Because of this, we will assume all matrices are square in this section.

6.1 Cofactor Expansion

Definition. Let A be a $n \times n$ matrix. Let $A_{i,j}$ denote the $(n-1)\times(n-1)$ submatrix obtained from A by deleting the i-th row and j-th column. The **cofactor** of a_{ij} is defined to be $C_{ij} = (-1)^{i+j} \det A_{i,j}$.

Definition. The **determinant** of an $n \times n$ matrix A is defined to be

$$\det A := a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$$

and we define $\det[a] = a$. We sometimes write $\det A = |A|$.

Computing the determinant this way is called **cofactor expansion** along the first row.

Fact 18. *Suppose A is an $n \times n$ matrix. Then the determinant is given by cofactor expansion along any row or column. That is*

$$\begin{aligned} \det A &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}, \quad \text{and} \\ \det A &= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}, \end{aligned}$$

for all $1 \le i, j \le n$.

Using this last fact to compute the determinant would be called **cofactor expansion** along the ith row or jth column.

Fact 19. *Let A and B be $n \times n$ matrices with entries in $\mathbb{F}$.*

1. *If one of A's rows or columns is 0, then $|A| = 0$.*
2. *If A is an upper or lower triangular matrix, then the determinant of A is the product of the diagonal entries.*
3. $|A| = \left|A^T\right|$.
4. $|AB| = |A|\,|B|$.
5. $\left|A^{-1}\right| = |A|^{-1}$.

6.2 Elementary Row Operations and the Determinant

Every row operation has a matrix associated to it, called the elementary matrix. If E is the elementary matrix for a certain row operation, then this matrix has the property that if B is obtained from A by that particular row operation, then $EA = B$. Using this and the fact that $|AB| = |A|\,|B|$, we can figure out what happens to the determinant of a matrix when row operations are performed.

Proposition 20. *The elementary matrix for a row operation is obtained by applying the row operation to the identity matrix.*

Proposition 21. *Suppose A is an $n \times n$ matrix and B is obtained from A by an elementary row operation. Then*

- *If the row operation takes the form $R_k \to R_k + tR_j$ for any $t \in \mathbb{F}$ and for rows $k \neq j$, then* $\det A = \det B$.
- *If the row operation is multiplying a row by t, then $t\,|A| = |B|$.*
- *If the row operation is given by switching any two rows, then $|A| = -\,|B|$.*

6.3 Invertibility

Definition. A matrix A is **invertible** if there exists a matrix A^{-1} such that $A^{-1}A = AA^{-1} = I$, where I is the $n \times n$ identity matrix.

Definition. For any matrix A, define the **rank** of A, denoted $\text{rank}(A)$, as the number of leading 1s in its row reduced form.

Fact 22. *Let A be an $n \times n$ matrix. The following are equivalent.*

1. $\det A \neq 0$.
2. $\text{rank}(A) = n$.
3. *A is invertible.*

6.4 Determinants and Volume

So far, we only seem to care whether or not a determinant is 0. Here is a geometric interpretation of the actual value of the determinant.

Definition. Given n vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ in $\mathbb{R}^n$, we can talk about the volume of the **parallelotope** defined by these n vectors. This volume is the n-dimensional volume of the n-dimensional shape with vertices given by the sum of every subset of $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Denote this volume by $\text{vol}\{\vec{v}_1, \ldots, \vec{v}_n\}$.

For example, in $\mathbb{R}^2$, the $\text{vol}(\{\vec{v}_1, \vec{v}_2\})$ is the volume (in this case area) of the parallelogram with vertices $\vec{0}$, $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_1 + \vec{v}_2$.

In $\mathbb{R}^3$, the $\text{vol}(\{\vec{v}_1, \vec{v}_2, \vec{v}_3\})$ is the volume of the parallelopiped with vertices $\{\vec{0}, \vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_1 + \vec{v}_2, \vec{v}_1 + \vec{v}_3, \vec{v}_2 + \vec{v}_3, \vec{v}_1 + \vec{v}_2 + \vec{v}_3\}$.

Fact 23. *Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be n vectors in $\mathbb{R}^n$. Then* $\text{vol}(\{\vec{v}_1, \ldots, \vec{v}_n\}) = \left|\det \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}\right|$.

Fact 24. *Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map, and let A be the matrix of L with respect to the standard bases. Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be n vectors in $\mathbb{R}^n$. Then* $\text{vol}(\{L(\vec{v}_1), \ldots, L(\vec{v}_n)\}) = |\det A| \, \text{vol}(\{\vec{v}_1, \ldots, \vec{v}_n\})$.

7 Eigenvectors and Diagonalization (§6.1, 6.2, 6.3, 6.4, 9.4)

Definition. Let $L : \mathbb{V} \to \mathbb{V}$ be a linear map. A non-zero vector $\vec{v} \in \mathbb{V}$ such that $L(\vec{v}) = \lambda\vec{v}$ for some $\lambda \in \mathbb{F}$ is called an **eigenvector** of L. The number λ is called an **eigenvalue** of L.

Definition. Let $L : \mathbb{V} \to \mathbb{V}$ be a linear map, and let λ be an eigenvalue of L. Define the **eigenspace** of L corresponding to λ to be $\{\vec{v} \in \mathbb{V} : L(\vec{v}) = \lambda\vec{v}\}$.

Proposition 25. *Let $L : \mathbb{V} \to \mathbb{V}$ be a linear map, and let λ be an eigenvalue of L. The eigenspace corresponding to λ is a subspace of $\mathbb{V}$.*

Once we've picked a basis, we can think of these definitions purely as definitions for matrices. In this case, we can think of a square matrix as a linear map from $\mathbb{F}^n$ to itself, and column matrices as vectors in $\mathbb{F}^n$.

Definition. Let A be an $n \times n$ matrix. A non-zero vector $\vec{v} \in \mathbb{F}^n$ is called an **eigenvector** of A if $A\vec{v} = \lambda\vec{v}$ for some $\lambda \in \mathbb{F}$. The scalar λ is called an **eigenvalue**.

Definition. Let A be an $n \times n$ matrix and let λ be an eigenvalue of A. Define the **eigenspace** of A corresponding to λ to be $\{\vec{v} \in \mathbb{F}^n : A\vec{v} = \lambda\vec{v}\}$.

The eigenspace corresponding to an eigenvalue λ is also the set of all eigenvectors corresponding to λ along with $\vec{0}$.

7.1 Finding Eigenvectors and Eigenvalues

To find eigenvectors and eigenvalues for a linear map, first pick a basis so you have an $n \times n$ matrix. Now the problem becomes finding eigenvalues and eigenvectors for a square matrix.

Definition. Let A be an $n \times n$ matrix. The **characteristic polynomial** of A is the polynomial in λ given by $\det(A - \lambda I)$.

Proposition 26. *Let A be an $n \times n$ matrix. The eigenvalues of A are the values of λ that are solutions to the equation $\det(A-\lambda I) = 0$. That is, they are the roots of the characteristic polynomial of A.*

Fact 27. *The determinant of a matrix is the product of its eigenvalues.*

Proposition 28. *Let A be an $n \times n$ matrix, and let λ be an eigenvalue of A. The eigenspace corresponding to λ is equal to $\text{Null}(A - \lambda I)$.*

Fact 29. *If A is an $n \times n$ matrix, then the characteristic polynomial is a polynomial of degree n.*

Fact 30 (Fundamental Theorem of Algebra)**.** *Every degree n polynomial factors as $a(x-c_1)^{k_1}(x-c_2)^{k_2}\cdots(x-c_m)^{k_m}$ where $c_1, \ldots, c_m \in \mathbb{C}$ are distinct, $k_1 + \cdots + k_m = n$ and a is some non-zero complex number. The c_i are the* **roots** *of the polynomial and k_i is the* **multiplicity** *of the root c_i.*

Another way to phrase this fact is to say that every degree n polynomial has n roots over the complex numbers, when they are counted with multiplicity.

7.2 Diagonalisation

This section will take place with $\mathbb{F} = \mathbb{C}$, and we will make a comment at the end about diagonalisation over $\mathbb{R}$.

Definition. A square matrix A is **diagonalisable** if there exists an invertible matrix P such that $P^{-1}AP = D$ where D is a diagonal matrix.

Definition. If A and B are $n \times n$ matrices such that $P^{-1}AP = B$ for some invertible $n \times n$ matrix P, then we say A and B are **similar** matrices.

This definition is motivated by the fact that A and B are two matrix representatives of the same linear map, and P is the change of basis matrix.

Theorem 31. *If A and B are similar matrices, then they have the same determinant, same eigenvalues, same rank, and same trace.*

Theorem 32. *An $n \times n$ matrix A is diagonalisable if and only if there exists a basis for $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $\mathbb{C}^n$ such that each $\vec{v}_i$ is an eigenvector for A. If such a basis exists, then $P = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ and*
$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$
where $\vec{v}_i$ is an eigenvector with eigenvalue λ_i.

Definition. Let A be an $n \times n$ matrix and let λ be an eigenvalue of A. The **algebraic multiplicity** of λ is the multiplicity of λ as a root of the characteristic polynomial of A. The **geometric multiplicity** of λ is defined to be the dimension of $\text{Null}(A - \lambda I)$.

Fact 33. *Let λ be an eigenvalue of an $n \times n$ matrix A. Then*

$$1 \leq \textit{geometric multiplicity of } \lambda \leq \textit{algebraic multiplicity of } \lambda.$$

Theorem 34. *Suppose $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of a square matrix A with corresponding eigenvectors $\vec{v}_1, \ldots, \vec{v}_k$. Then $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent.*

Fact 35. *A matrix A is diagonalisable if and only if every eigenvalue of A has its geometric multiplicity equal to its algebraic multiplicity.*

Corollary 36. *If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalisable.*

Remark. If you have a matrix A with real entries, you can ask whether or not there are matrices P and D with real entries such that $P^{-1}AP = D$. The complication in this case is that you may not even get n real eigenvectors. However, if you do, then all the results hold. If you don't, then there is no hope for diagonalisation!

7.3 Applications of Diagonalisation

Taking Powers of Matrices

Suppose we have a diagonalisable matrix A, and we wish to take arbitrarily large powers of A. Then one way we can do this is by observing $A = PDP^{-1}$ for some diagonal matrix D so

$$A^n = PDP^{-1}PDP^{-1} \cdots PDP^{-1} = PD^nP^{-1}.$$

We have reduced the problem to computing the nth power of a diagonal matrix, but this is much simpler than for an arbitrary matrix since if

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k \end{bmatrix}$$

then

$$D^n = \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k^n \end{bmatrix}.$$

Taking powers of matrices like this arises when studying dynamical systems, and in an area of probability called Markov chains or the Markov process to name a couple. In particular the Google page rank system relies on taking large powers of matrices!

Decoupling Differential Equations

Suppose you have a system of differential equations of the form

$$\begin{aligned}\frac{dx}{dt} &= ax + by \\ \frac{dy}{dt} &= cx + dy.\end{aligned}$$

Here the variables x and y depend on each other, but it would be great if they didn't, since then you could solve two separate differential equations. If the matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is diagonalisable, then by a simple change of coordinates, we can decouple the differential equation.

Let $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\frac{d\vec{x}}{dt} = \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix}$. Then the set of equations above can be written as $\frac{d\vec{x}}{dt} = A\vec{x}$.

Suppose $P^{-1}AP = D$ for some diagonal matrix $D = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. Let $\vec{x} = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ and let $\vec{u} = \left[\begin{smallmatrix} u \\ w \end{smallmatrix}\right] = P^{-1}\vec{x}$. Then

$$\frac{d\vec{x}}{dt} = A\vec{x} \iff P\frac{d\vec{u}}{dt} = PDP^{-1}\vec{x} \iff \frac{d\vec{u}}{dt} = D\vec{u}.$$

Rewriting this last equation as a pair of differential equations we get

$$\begin{aligned}\frac{du}{dt} &= \lambda_1 u \\ \frac{dw}{dt} &= \lambda_2 w.\end{aligned}$$

No we have two decoupled equations, each which can be solved independently of the other. This process of decoupling is easily generalised to n variables and n equations.

8 Inner Product Spaces (§7.1, 7.2, 7.3, 7.4, 9.5)

Definition. Let $\mathbb{V}$ be a vector space over $\mathbb{F}$. An **inner product** on $\mathbb{V}$ is a function

$$\langle\ ,\ \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{F}$$

such that for all $u, v, w \in \mathbb{V}$ and $\alpha \in \mathbb{F}$,

1. $\langle \vec{v}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{v} \rangle}$.
2. $\langle \alpha\vec{v}, \vec{w} \rangle = \alpha \langle \vec{v}, \vec{w} \rangle$.
3. $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$.
4. (a) $\langle \vec{v}, \vec{v} \rangle \geq 0$.
   (b) If $\langle \vec{v}, \vec{v} \rangle = 0$ then $\vec{v} = \vec{0}$.

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

Definition. For any vector $\vec{v}$ in an inner product space, define the **norm** as $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$.

If $\mathbb{F} = \mathbb{R}$, then all the scalars are equal to their own conjugate, that is $\alpha = \overline{\alpha}$ so the first property becomes $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$.

Proposition 37. *Every vector space admits an inner product.*

Fact 38. *Let $\mathbb{V}$ be an inner product space. For all $\vec{v}, \vec{u}, \vec{w} \in \mathbb{V}$ and $\alpha \in \mathbb{F}$ the following is true.*

- $\langle \vec{0}, \vec{v} \rangle = 0$.
- $\langle \vec{v}, \alpha\vec{w} \rangle = \overline{\alpha} \langle \vec{v}, \vec{w} \rangle$.
- $\langle \vec{v}, \vec{u} + \vec{w} \rangle = \langle \vec{v}, \vec{u} \rangle + \langle \vec{v}, \vec{w} \rangle$.

8.1 Orthogonality

Definition. Let $\mathbb{V}$ be an inner product space. We say $\vec{v}$ and $\vec{w}$ are **orthogonal**, and write $\vec{v} \perp \vec{w}$, if $\langle \vec{v}, \vec{w} \rangle = 0$.

Proposition 39. *Let $\mathbb{V}$ be an inner product space. Suppose $\vec{v} \perp \vec{w}$. Then $\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2$.*

Lemma 40. *Let $\mathbb{V}$ be an inner product space and let $\vec{v}, \vec{w} \in \mathbb{V}$ such that $\vec{w} \neq \vec{0}$. Then $\vec{w}$ is perpendicular to $\vec{v} - \frac{\langle \vec{v}, \vec{w} \rangle}{\langle \vec{w}, \vec{w} \rangle} \vec{w}$.*

Proposition 41 (Cauchy-Schwartz Inequality)**.** *Let $\mathbb{V}$ be an inner product space. Then $|\langle \vec{v}, \vec{w} \rangle| \leq \|\vec{v}\| \, \|\vec{w}\|$.*

Definition. Let $\mathbb{V}$ be a real vector space. Define the angle θ between two vectors $\vec{v}$ and $\vec{w}$ by $\cos(\theta) = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{v}\| \|\vec{w}\|}$.

There are various ways to define the angle between vectors in a complex vector space, each serving a different purpose. We won't be talking about the angle between vectors in a complex vector space in this course, except for the case when vectors are orthogonal which appears a little later in the course.

Proposition 42. *Let $\mathbb{V}$ be an inner product space. Then the norm satisfies the following conditions.*

1. $\|\alpha\vec{v}\| = |\alpha| \, \|\vec{v}\|$,
2. $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$, *(this is called the* **triangle inequality***)*
3. *If* $\|\vec{v}\| = 0$*, then* $\vec{v} = 0$.

8.2 Projections and Orthonormal Bases

Definition. Let $\mathbb{V}$ be an inner product space. A set $\{\vec{v}_1, \ldots, \vec{v}_k\} \subset \mathbb{V}$ is called **orthogonal** if $\langle \vec{v}_i, \vec{v}_j \rangle = 0$ whenever $i \neq j$.

Definition. A vector $\vec{v}$ in an inner product space is a **unit vector** if $\|\vec{v}\| = 1$.

Definition. A set $\{\vec{v}_1, \ldots, \vec{v}_k\}$ in an inner product space is an **orthonormal set** if it is an orthogonal set and each vector is a unit vector.

Proposition 43. *Suppose $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is orthogonal and $\vec{v}_i \neq \vec{0}$ for all i. Then $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent.*

Definition. A set $\{\vec{v}_1, \ldots, \vec{v}_n\} \subset \mathbb{V}$ is an **orthonormal basis** if it is an orthonormal set and it is a basis.

Definition. Let $\mathbb{V}$ be an inner product space, let $\vec{w}, \vec{v} \in \mathbb{V}$ with $\vec{w} \neq \vec{0}$. Define the **projection** of $\vec{v}$ onto $\vec{w}$ as

$$\text{proj}_{\vec{w}}(\vec{v}) := \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w}.$$

Define the **perpendicular** vector of $\vec{v}$ with respect to $\vec{w}$ by

$$\text{perp}_{\vec{w}}(\vec{v}) := \vec{v} - \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w}.$$

Definition. Let $\mathbb{V}$ be an inner product space and let $\mathbb{W} \subset \mathbb{V}$ be a subspace. Let $\{\vec{w}_1, \ldots, \vec{w}_k\}$ be an orthogonal basis for $\mathbb{W}$. Let $\vec{v} \in \mathbb{V}$. Define the **projection** of $\vec{v}$ onto $\mathbb{W}$, and the **perpendicular** vector of $\vec{v}$ with respect to $\mathbb{W}$ by

$$\text{proj}_{\mathbb{W}}(\vec{v}) := \text{proj}_{\vec{w}_1}(\vec{v}) + \cdots + \text{proj}_{\vec{w}_k}(\vec{v}) \quad \text{and} \quad \text{perp}_{\mathbb{W}}(\vec{v}) := \vec{v} - \text{proj}_{\mathbb{W}}(\vec{v}).$$

respectively.

Note that this projection only works for an orthogonal basis.

Definition. Let $\mathbb{V}$ be an inner product space and let $\mathbb{W} \subset \mathbb{V}$ be a subspace. Define the **orthogonal complement** of $\mathbb{W}$ by

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

Fact 44. *Let $\mathbb{V}$ be an inner product space and $\mathbb{W} \subset \mathbb{V}$ a subspace. The following statements are true.*

- $\mathbb{W}^{\perp}$ *is a subspace of* $\mathbb{V}$.
- $\mathbb{W} \cap \mathbb{W}^{\perp} = \{\vec{0}\}$.
- *If* $\{\vec{v}_1, \ldots, \vec{v}_k\}$ *is a basis for* $\mathbb{W}$ *and* $\{\vec{w}_1, \ldots, \vec{w}_m\}$ *is a basis for* $\mathbb{W}^{\perp}$, *then* $\{\vec{v}_1, \ldots, \vec{v}_k, \vec{w}_1, \ldots, \vec{w}_m\}$ *is a basis for* $\mathbb{V}$.
- $\dim(\mathbb{W}) + \dim(\mathbb{W}^{\perp}) = \dim(\mathbb{V})$.

8.3 The Gram-Schmidt Orthogonalisation Procedure

We now describe a method to start with a basis in an inner product space and obtain an orthonormal basis.

Let $\mathbb{V}$ be an inner product space with basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$. To obtain an orthogonal basis $\mathcal{C} = \{\vec{w}_1, \ldots, \vec{w}_n\}$ proceed as follows.

$$\begin{aligned}
\vec{w}_1 &= \vec{v}_1 \\
\vec{w}_2 &= \vec{v}_2 - \frac{\langle \vec{v}_2, \vec{w}_1 \rangle}{\|\vec{w}_1\|^2} \vec{w}_1 \\
\vec{w}_3 &= \vec{v}_3 - \frac{\langle \vec{v}_3, \vec{w}_1 \rangle}{\|\vec{w}_1\|^2} \vec{w}_1 - \frac{\langle \vec{v}_3, \vec{w}_2 \rangle}{\|\vec{w}_2\|^2} \vec{w}_2 \\
&\vdots \\
\vec{w}_n &= \vec{v}_n - \frac{\langle \vec{v}_n, \vec{w}_1 \rangle}{\|\vec{w}_1\|^2} \vec{w}_1 - \cdots - \frac{\langle \vec{v}_n, \vec{w}_{n-1} \rangle}{\|\vec{w}_{n-1}\|^2} \vec{w}_{n-1}.
\end{aligned}$$

Then $\mathcal{C} = \{\vec{w}_1, \ldots, \vec{w}_n\}$ is an orthogonal basis for $\mathbb{V}$. To obtain an orthonormal basis $\mathcal{D} = \{\vec{u}_1, \ldots, \vec{u}_n\}$ let $\vec{u}_i = \frac{1}{\|\vec{w}_i\|} \vec{w}_i$.

Corollary 45. *Every inner product space admits an orthonormal basis.*

8.4 Method of Least Squares

Suppose we have some data points

x	x_1	$\cdots$	x_n
y	y_1	$\cdots$	y_n

and we want to find the equation $y = a_0 + a_1 x + \cdots + a_k x^k$ of best fit to this data. Let

$$\vec{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \vec{x^2} = \begin{bmatrix} x_1^2 \\ \vdots \\ x_n^2 \end{bmatrix}, \dots, \vec{x^k} = \begin{bmatrix} x_1^k \\ \vdots \\ x_n^k \end{bmatrix}, \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \text{and} \quad \vec{a} = \begin{bmatrix} a_0 \\ \vdots \\ a_k \end{bmatrix}.$$

Let $X = \begin{bmatrix} \vec{1} & \vec{x} & \cdots & \vec{x^k} \end{bmatrix}$, then $\vec{a} = (X^T X)^{-1} X^T \vec{y}$ gives the equation of best fit.

9 Hermitian, Unitary and Normal Matrices (§7.1, 8.1, 9.6)

Definition. Let A be a matrix. Define the **adjoint** of A, denoted A^* by the conjugate transpose. That is, $A^* = \overline{A}^T$.

Fact 46. *Let A and B be $n \times n$ matrices. Then*

- $(AB)^* = B^* A^*$,
- $(A^*)^* = A$,
- *If A is a real matrix, then $A^* = A^T$,*
- $(\alpha A)^* = \overline{\alpha} A^*$.

Definition. A matrix A is **Hermitian** if $A = A^*$. If A is real, this is the same as $A = A^T$ and we say A Is **symmetric**.

Definition. A square matrix U is called **unitary** if $U^* = U^{-1}$. If the matrix O is real, then $O^T = O^{-1}$ and we say O is **orthogonal**.

Proposition 47. *Let U be an $n \times n$ matrix. The following are equivalent.*

- *U is unitary.*
- *The rows form an orthonormal basis of $\mathbb{C}^n$.*
- *The columns form an orthonormal basis of $\mathbb{C}^n$.*

Fact 48. *Let U be a unitary matrix and consider $\mathbb{C}^n$ with the standard Hermitian inner product. Then*

$$\langle U\vec{v}, U\vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle$$

for all $\vec{v}, \vec{w} \in \mathbb{C}^n$.

9.1 Spectral Theorem for Hermitian Matrices

Theorem 49 (Schur's Triangulation Theorem). *Let A be a square matrix. There is a unitary matrix U and an upper triangular matrix T such that $U^*AU = T$.*

Corollary 50. *For any matrix A, the determinant is the product of its eigenvalues and the trace is the sum of its eigenvalues.*

Corollary 51. *Let A be an $n \times n$ matrix and let the characteristic polynomial be*

$$\det(A - \lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0.$$

Then $a_n = (-1)^n$, $a_{n-1} = (-1)^{n-1}\operatorname{tr}(A)$, and $a_0 = |A|$.

Proposition 52. *Let A be a Hermitian matrix. Then A is unitarily diagonalisable (that is there is a unitary U and a diagonal D such that $U^*AU = D$) and all its eigenvalues are real.*

Fact 53. *Let A be a matrix such that there exists a unitary U and a diagonal matrix D with real entries such that $U^*AU = D$. Then A is Hermitian.*

Theorem 54 (Spectral Theorem for Hermitian Matrices). *A square matrix is Hermitian if and only if it is unitarily diagonalisable with real eigenvalues.*

Theorem 55 (Spectral Theorem for Orthogonal Matrices). *A real square matrix is symmetric if and only if it is orthogonally diagonalisable over $\mathbb{R}$.*

Proposition 56. *Let A be an $n \times n$ Hermitian matrix, and suppose $A\vec{v} = \lambda\vec{v}$ and $A\vec{w} = \mu\vec{w}$ with $\lambda \neq \mu$. Then $\langle \vec{v}, \vec{w} \rangle = 0$ with respect to the standard Hermitian inner product on $\mathbb{C}^n$.*

With this proposition, we now have an algorithm to unitarily diagonalise a Hermitian matrix. Suppose we want to find a unitary U and a diagonal D such that $U^*AU = D$ for a unitarily diagonalisable matrix A.

1. Diagonalise as usual, obtaining $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ and a basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of eigenvectors for $\mathbb{C}^n$.
2. Perform Gram-Schmidt on each eigenspace to obtain $\{\vec{w}_1, \ldots, \vec{w}_n\}$, which is an orthonormal basis of eigenvectors for $\mathbb{C}^n$.
3. Let $U = \begin{bmatrix} \vec{w}_1 & \cdots & \vec{w}_n \end{bmatrix}$.

9.2 Inner Products and Hermitian Matrices

Proposition 57. *Let $\mathbb{V}$ be a vector space with basis $\mathcal{B} = \{\vec{\beta}_1, \ldots, \vec{\beta}_n\}$. Then $\langle\ ,\ \rangle$ is an inner product if and only if there is an $n \times n$ Hermitian matrix A with positive eigenvalues such that*

$$\langle \vec{v}, \vec{w} \rangle = [\vec{w}]_{\mathcal{B}}^* A [\vec{v}]_{\mathcal{B}}.$$

If there is such a matrix, it must be given by

$$A = \begin{bmatrix} \left\langle \vec{\beta}_1, \vec{\beta}_1 \right\rangle & \cdots & \left\langle \vec{\beta}_n, \vec{\beta}_1 \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle \vec{\beta}_1, \vec{\beta}_n \right\rangle & \cdots & \left\langle \vec{\beta}_n, \vec{\beta}_n \right\rangle \end{bmatrix}.$$

9.3 Normal Matrices

Definition. A square matrix A is **normal** if $AA^* = A^*A$.

Fact 58. *Hermitian, unitary, and diagonal matrices are all normal.*

Theorem 59 (Spectral Theorem for Normal Matrices)**.** *A square matrix is normal if and only if it is unitarily diagonalisable.*