Algebra I

Lecture Notes

Syllabus

Definition of an abstract vector space over an arbitrary field. Examples. Linear maps. Division Algorithm in $F[x]$. Characteristic polynomials and minimal polynomials. Coincidence of roots. [2]

Quotient vector spaces. The first isomorphism theorem for vector spaces and rank-nullity. Induced linear maps. Applications: Triangular form for matrices over $\mathbb{C}$. Cayley-Hamilton Theorem. [2.5]

Bezout's Lemma in $F[x]$. Primary Decomposition Theorem. Diagonalizability and Triangularizability in terms of minimal polynomials. Proof of existence of Jordan canonical form over $\mathbb{C}$ (using primary decomposition and inductive proof of form for nilpotent linear maps). [3.5]

Dual spaces of finite-dimensional vector spaces. Dual bases. Dual of a linear map and description of matrix with respect to dual basis. Natural isomorphism between a finite-dimensional vector space and its second dual. Annihilators of subspaces, dimension formula. Isomorphism between U^* and V^*/U°. [3]

Recap on real inner product spaces. Definition of non-degenerate symmetric bilinear forms and description as isomorphism between V and V^*. Hermitian forms on complex vector spaces. Review of Gram-Schmidt. Orthogonal Complements. [2]

Adjoints for linear maps of inner product spaces. Uniqueness. Concrete construction via matrices [1]

Definition of orthogonal/unitary maps. Definition of the groups O_n, SO_n, U_n, SU_n. Diagonalizability of self-adjoint and unitary maps. [2]

Stephen Thatcher

Ed. Sam Adam-Day

Lecturer: Prof Ulrike Tillmann

Contents

Chapter 1

Vector Spaces

Let $\mathbb{F}$ be a field, then both $(\mathbb{F}, +, 0)$ and $(\mathbb{F}\backslash\{0\}, \times, 1)$ are abelian groups and the distribution law holds:

$$\forall a, b, c \in \mathbb{F} : a(b + c) = ab + ac$$

The smallest integer p such that

$$\underbrace{1 + 1 + \cdots + 1}_{p \text{ times}} = 0$$

is called the **characteristic** of $\mathbb{F}$.

If no such p exists, then $\mathbb{F}$ is said to have **characteristic zero**.

Example 1.1. Characteristic zero: $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Q} \cup \{i\}$

Example 1.2. Characteristic p: $\{0, 1, \ldots, p\}$

A **vector space** V over a field $\mathbb{F}$ is an abelian group $(V, +, \underline{0})$ together with scalar multiplication $\mathbb{F} \times V \to V$ such that for all $a, b \in \mathbb{F}$, $v, w \in V$:

(1) $a(v + w) = av + aw$

(2) $(a + b)v = av + bv$

(3) $(ab)v = a(bv)$

(4) $1v = v$

Definition 1.1. A set $S \subset V$ is **Linearly Independent** if for all $a_i \in \mathbb{F}, s_i \in S$:

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = 0 \Rightarrow a_i = 0 \forall i$$

Definition 1.2. A set $S \subset V$ is **Spanning** if for all $v \in V$ there exists $a_i \in \mathbb{F}$ and $S_i \in S$ such that

$$v = a_1s_1 + a_2s_2 + \cdot + a_ns_n$$

Definition 1.3. S is a basis of V if S is **spanning and linearly independent**

Definition 1.4. The span of S is the smallest vector space containing S

Example 1.3.

$V = \mathbb{F}^n$ with **standard basis** $\{(1,0,\ldots,0),\ldots,(0,0,\ldots,1)\}$

$V = \mathbb{F}[x]$ with **standard basis** $\{1,x,x^2,\ldots,x^n,\ldots\}$

$V = \mathbb{N}^{\mathbb{R}} = \{(a_0,a_1,\ldots) : a_i \in \mathbb{R}\}$

$V \supset S = \{(1,0,\ldots),(0,1,0,\ldots),\ldots,(0,\ldots,0,1,0,\ldots),\ldots)$
S is an infinite **linearly independent** subset
Note that $\mathrm{Span}(S) \neq V$ as $(1,1,\ldots,1) \notin \mathrm{Span}(S)$ - no finite sum

Suppose V and W are vector spaces over $\mathbb{F}$

Definition 1.5. A **map** $T : V \to W$ is a **linear transformation** if for all $a \in \mathbb{F}, v, v' \in V$

$$T(av + v') = aTv + Tv'$$

Definition 1.6. A **bijective** linear transformation is a linear **isomorphism** of vector spaces

Example 1.4. $T : \mathbb{R}[x] \to \mathbb{R}[x]$ given by $f(x) \mapsto xf(x)$

$$\begin{aligned} T(af(x) + g(x)) &= x(af(x) + g(x)) \\ &= axf(x) + xg(x) \qquad\qquad = aT(f(x)) + T(g(x)) \end{aligned}$$

T is **injective** and defines a **linear isomorphism** from $\mathbb{R}[x]$ to $x\mathbb{R}[x]$ the subspace of polynomials with zero constant term: $x\mathbb{R}[x] \lneq \mathbb{R}[x]$
$W \leq \mathbb{N}^{\mathbb{R}} \to \mathbb{R}[x]$ given by $e_i = \underbrace{(0,0,\ldots,1,0,\ldots)}_{\text{1 in the i}^{\text{th}}\text{ place}} \mapsto x^i$ defines a linear isomorphism

Challenge 1. Prove that there is no isomorphism $T : W \to V = \mathbb{N}^{\mathbb{R}}$. Hence, V has no countable basis

Remark. Every linear map $T : V \to W$ is determined by its values on a basis $\mathcal{B}$ of V (since $\mathcal{B}$ is a spanning set of V). Indeed, can be determined by any spanning set. Given any map $T : \mathcal{B} \to W$ we can extend to a linear transformation $T : V \to W$.

Let $\mathrm{Hom}(V, W)$ be the set of linear transformations from V to W.

For $a \in \mathbb{F}^n, v \in V, T, S \in \mathrm{Hom}(V, W)$ define

$$(aT)(v) = a(Tv)$$
$$(S + T)(v) = Sv + Tv$$

Lemma 1.1. With these operations, $\mathrm{Hom}(V, W)$ is a **vector space** over $\mathbb{F}$

Proof. Assume V and W are finite dimensional and let $\mathcal{B} = \{e_1, \dots, e_m\}$ and $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ be bases for V and W respectively

Denote by ${}_{\mathcal{B}'}[T]_{\mathcal{B}}$ the matrix (a_{ij}) such that

$$Te_i = a_{i1}e'_1 + \dots + a_{in}e'_n$$

Note:

$${}_{\mathcal{B}}[aT]_{\mathcal{B}} = a_{\mathcal{B}'}[T]_{\mathcal{B}'}$$
$${}_{\mathcal{B}'}[T + S]_{\mathcal{B}} = {}_{\mathcal{B}'}[T]_{\mathcal{B}} + {}_{\mathcal{B}'}[S]_{\mathcal{B}}$$

Theorem 1.2.

*The map that takes T to ${}_{\mathcal{B}'}[T]_{\mathcal{B}}$ is an **isomorphism** of **vector spaces** from Hom(V, W) to the $n \times m$ matrices over $\mathbb{F}$. Furthermore, this correspondence is compatible with composition, taking composition to multiplication of matrices:*

Proof. If $T : V \to W$, $S : W \to U$ with $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ bases for V, W, U respectively then

$${}_{\mathcal{B}''}[S \circ T]_{\mathcal{B}} =_{\mathcal{B}''} [S]_{\mathcal{B}'\mathcal{B}'}[T]_{\mathcal{B}}$$

Chapter 2

Polynomials

Definition 2.1. $\mathbb{F}[x]$ is the space of polynomials over a field $\mathbb{F}$

Example 2.1.

$$\begin{array}{l} \big(\quad 2x^3 + 4x^2 \quad + 9x + 7\big) \div \big(x^2 - 2x\big) = 2x + 8 + \dfrac{25x + 7}{x^2 - 2x} \\ \underline{\; - 2x^3 + 4x^2 \;} \\ \qquad\qquad 8x^2 \quad + 9x \\ \qquad\quad \underline{\; - 8x^2 + 16x \;} \\ \qquad\qquad\qquad 25x \end{array}$$

Proposition 2.1. (Division Algorithm for Polynomials)
Let $f(x), g(x)$ be polynomials over a field $\mathbb{F}$ such that $g(x) \neq 0$
Then, $\exists q(x), r(x) \in \mathbb{F}[x]$ with

$$f(x) = q(x)g(x) + r(x) \qquad \text{(with deg } r(x) < \text{deg } g(x))$$

Proof. If $\deg(f) < \deg(g)$ then we take $q(x) = 0, r(x) = f(x)$ and we are done.
Hence we can now assume that $\deg(g) \leq \deg(f)$, then

$$f(x) = a_n x^n + \cdots + a_1 x + a_0$$
$$g(x) = b_m x^m + \cdots + b_1 x + b_0$$

with $n \leq m$. Then $\deg\left(f(x) - \frac{a_n}{b_m} x^{n-m} g(x)\right) < \deg f(x)$.
Then by induction on $\deg f$ we have that

$$\exists s, r \in \mathbb{F}[x] : s(x)g(x) + r(x) = f(x) - \frac{a_n}{b_m} x^{n-m} g(x) \text{ and } \deg r < \deg g$$

Now setting $q(x) = \frac{a_n}{b_m} x^{n-m} + s(x)$ the result follows. $\square$

Corollary 2.2. For $f(x) \in \mathbb{F}[x], a \in \mathbb{F}$, if $f(a) = 0$ then $(x - a)|f(x)$

Proof. By the division algorithm there exist $q, r \in \mathbb{F}[x]$ with $\deg r < \deg(x - a) = 1$ with $f(x) = q(x)(x - a) + r(x)$. Since $\deg r < 1$, we must have that $r \in \mathbb{F}$, that is r is constant, then:

$$f(x) = q(x)(x - a) + r$$
$$f(a) = q(a)(a - a) + r$$

By assumption $f(a) = 0$, hence

$$0 = r$$

Thus $f(x) = q(x)(x - a)$ and thus $(x - a)|f(x)$ as required

□

Corollary 2.3. If $\deg f(x) \leq n$ then f has, at most n roots

Proof. From the above by induction

Definition 2.2. A field $\mathbb{F}$ is algebraically closed if every polynomial in $\mathbb{F}[x]$ has a root in $\mathbb{F}$

Example 2.2. By the fundamental theorem of algebra, $\mathbb{C}$ is an algebraically closed field

Theorem 2.4. *Any field $\mathbb{F}$ has an algebraic closure $\overline{\mathbb{F}}$, which, by definition, is the smallest algebraically closed field containing $\mathbb{F}$.*

Example 2.3.

$\mathbb{R}$ - not algebraically closed since $x^2 + 1$ has no real solutions - $\overline{\mathbb{R}} = \mathbb{R} \cup \{i\}$

$\overline{\mathbb{Q}} \lneq \mathbb{C}$ - does not require anything from $\mathbb{R} \backslash \mathbb{Q}$, e.g. $\pi \notin \overline{\mathbb{Q}}$

Challenge 2. Prove that no finite field is algebraically closed

Let $A \in M_{n \times n}(\mathbb{F})$ - the set of $n \times n$ matrices over $\mathbb{F}$
Let $f(x) = a_m x^m + \ldots + a_0 \in \mathbb{F}[x]$
Define $f(A) = a_m A^m + \ldots + a_1 A + a_0 I \in M_{n \times n}(\mathbb{F})$

Remark. For $f(x), g(x) \in \mathbb{F}[x]$ we have $f(A)g(A) = g(A)f(A)$ [1]

Remark. If for $v \in \mathbb{F}^n$ we have $Av = \lambda v$ for some $\lambda \in \mathbb{F}$ then $f(A)v = f(\lambda)v$ [2]

Lemma 2.5. For all $A \in M_{n^2}(\mathbb{F})$ there exists a polynomial $f(x) \in \mathbb{F}[x]$ such that $f(A) = 0 \in M_{n^2}(\mathbb{F})$

Proof. $\dim(M_{n^2}(\mathbb{F})) = n^2 < \infty$, hence $I, A, A^2, \ldots, A^k$ must be linearly dependent for $k > n^2$. So there exists $a_i \in \mathbb{F}$ such that

$$a_0 I + a_1 A + \cdots + a_k A^k = 0$$

Hence we can set $f(x) = \sum_{i=0}^{k} a_i x^i$ and we are done

□

[1] $A^k A^l = A^l A^k = A^{k+l}$ and $A(aI) = (aI)A$
[2] $a_k A^k(v) = a_k(\lambda^k v) = (a_k \lambda^k)v$

Definition 2.3. A **minimal polynomial** is a *monic* polynomial of least degree with $m_A(A) = 0$

Theorem 2.6. *If $f(A) = 0$ for $f(x) \in \mathbb{F}[x]$, then $m_A(x)|f(x)$. Furthermore, $m_A(x)$ is unique.*

Proof. Suppose $f(A) = 0$ for some $f(x) \in \mathbb{F}[x]$.
Applying polynomial long division to $f(x) \div m_A(x)$ we have that there exists $q(x), r(x) \in \mathbb{F}[x]$ such that $f(x) = q(x)m_A(x) + r(x)$ with $\deg r(x) < \deg m_A(x)$. Now, evaluating at A we obtain $r(A) = 0$. Since r has degree less than m_A, it must be identically zero, else it would contradict the choice of m_A as a minimal polynomial. Thus $f(x) = q(x)m_A(x)$ and so $m_A(x)|f(x)$ as required.
It follows that if there were two monic polynomials m_A, m'_A such that $m_A(A) = m'_A(A) = 0$ then as they must both divide each other they must be equal, hence m_A is unique. □

Definition 2.4. The **characteristic polynomial** of $A \in M_{n^2}(\mathbb{F})$ is given by:

$$\mathcal{X}_A(x) = \det(A - xI)$$

Lemma 2.7. $\mathcal{X}_A(x) = (-1)^n x^n + \mathrm{tr}A(-1)^{n-1}x^{n-1} + \cdots + \det A$

Proof. We prove this result by showing that

$$\mathcal{X}_A(x) = \det \begin{pmatrix} a_{11}-x & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn}-x \end{pmatrix} = (a_{11} - x) \cdots (a_{nn} - x) + f(x)$$

For some $f(x)$, a polynomial of degree at most $n - 2$. We proceed by induction.
Base case: $n = 2$
When $n = 2$ we have $\mathcal{X}_A(x) = (a_{11} - x)(a_{22} - x) - a_{12}a_{21}$, with $f(x) = a_{12}a_{21}$ the result follows.
Inductive case: $n = k$
We suppose that the result holds for all $n < k$ then we can calculate $\mathcal{X}_A(x)$ by expanding by minors along the first row:

$$\mathcal{X}_A(x) = (a_{11} - x)\det(A - x)_{11} - a_{12}\det(A - x)_{12} + \cdots + (-1)^{n-1}a_{1n}\det(A - x)_{1n}$$

The first term $(a_{11} - x)\det(A - x)_{11}$ is just the characteristic polynomial of some $n - 1 \times n - 1$ matrix and hence by the induction hypothesis we have

$$(a_{11} - x)\det(A - x)_{11} = (a_{11} - x)(a_{22} - x)\cdots(a_{nn} - x) + g(x)$$

where $g(x)$ is some polynomial of degree at most $n - 2$.
For $i \neq 1$ we see that $\det(A - x)_{1i}$ is a polynomial of degree at most $n - 2$ and hence setting $f(x) = g(x) - a_{12}\det(A - x)_{12} + \cdots + (-1)^{n-1}a_{1n}\det(A - x)_{1n}$ we have

$$\mathcal{X}_A(x) = (a_{11} - x)\cdots(a_{nn} - x) + f(x)$$

with $\deg f(x) \leq n - 2$ as required.
Now, returning to our general case, we note that expanding out we obtain

$$\mathcal{X}_A(x) = (-1)^n x^n + (-1)^{n-1}x^{n-1}(a_{11} + \cdots + a_{nn}) + a_{n-2}x^{n-2} + \cdots + a_0$$

Hence $a_n = (-1)^n$, $a_{n-1} = (-1)^{n-1}(a_{11} + \cdots + a_{nn}) = (-1)^{n-1}\mathrm{tr}A$ and $a_0 = \mathcal{X}_A(0) = \det A$ □

Theorem 2.8. *The following are equivalent:*

$$\begin{aligned}&(a)\quad \lambda \textit{ is an eigenvalue of } A\\ &(b)\quad \lambda \textit{ is a root of } \mathcal{X}_A(x)\\ &(c)\quad \lambda \textit{ is a root of } m_A(x)\end{aligned}$$

Proof. a $\iff$ b

$$\begin{aligned}\mathcal{X}_A(\lambda) = 0 &\iff \det(A - \lambda I) = 0\\ &\iff (A - \lambda I) \text{ is singular}\\ &\iff \exists v \in \mathbb{F}^n : (A - \lambda I)v = 0, v \neq 0\\ &\iff \exists v \in \mathbb{F}^n : Av = \lambda v, v \neq 0\end{aligned}$$

c $\Rightarrow$ a

First we note that by a corollary to the division algorithm $m_A(\lambda) = 0 \Rightarrow m_A(x) = (x - \lambda)g(x)$ for some $g(x) \in \mathbb{F}[x]$ with $\deg g < \deg m_A$. Then by the minimality of m_A it must hold that $g(A) \neq 0$. Hence there exists some $w \in \mathbb{F}^n$ such that $g(A)w \neq 0$, putting $v = g(A)w$ we have

$$\begin{aligned}(A - \lambda I)v &= (A - \lambda I)(g(A)w)\\ &= m_A(A)w\\ &= 0\end{aligned}$$

Hence $Av = \lambda v$, i.e. λ is an eigenvalue.

a $\Rightarrow$ c

Assume λ is an eigenvalue. Then there exists a non-zero vector $v \in \mathbb{F}^n$ such that $Av = \lambda v$. Then

$$m_A(\lambda)v = m_A(A)v = 0 \cdot v = 0$$

and since $v \neq 0$ we have that $m_A(\lambda) = 0$ and hence λ is a root of $m_A(x)$

□

Chapter 3

Quotient Spaces

Let V be a vector space over a field $\mathbb{F}$

Let U be a subspace of V

Definition 3.1. The set of cosets $V/U = \{v + U : v \in V\}$ is a vector space, the **quotient space**, with operations:

$$(v + U) + (w + U) = (v + w) + U \qquad (\forall v, w \in V)$$
$$a(v + U) = av + U \qquad (\forall a \in \mathbb{F}, v \in V)$$

Proof. $(V/U, +)$ is just the quotient group associated to V and U, hence we need only check well-definedness of scalar multiplication: first we note that $v+U = v'+U$ if and only if $v = v'+u$ for some element $u \in U$. Then,

$$\begin{aligned} a(v + U) &= av + U \\ &= a(v' + u) + U \\ &= av' + au + U \\ &= av' + U \qquad \text{(since } U \text{ is a vector space it is closed under linear multipication)} \\ &= a(v' + U) \end{aligned}$$

Thus all the vector space axioms are satisfied as they hold for V and U

$\square$

Let $\mathcal{E}$ be a basis for U and let $\mathcal{B}$ be a basis for V containing $\mathcal{E}$

Define $\overline{\mathcal{B}} = \{e + U : e \in \mathcal{B}\backslash\mathcal{E}\}$

Proposition 3.1. $\overline{\mathcal{B}}$ is a basis for V/U

Proof. Take $v + U \in V/U$
As $v \in V$ there exist $a_i \in \mathbb{F}$, $e_1, \ldots, e_k \in \mathcal{E}$ and $e_{k+1}, \ldots, e_n \in \mathcal{B}\backslash\mathcal{E}$ such that

$$v = a_1e_1 + \cdots + a_ke_k + a_{k+1}e_{k+1} + \cdots + a_ne_n$$

Then

$$\begin{aligned} v + U &= a_1e_1 + \cdots + a_ne_n + U \\ &= a_{k+1}e_{k+1} + \cdots + a_ne_n + U && (a_1e_1 + \cdots + a_ke_k \in U) \\ &= a_{k+1}(e_{k+1} + U) + \cdots + a_n(e_n + U) \in \mathrm{Sp}\left(\overline{\mathcal{B}}\right) \end{aligned}$$

Hence $\overline{\mathcal{B}}$ spans V/U, it remains to show that $\overline{\mathcal{B}}$ is linearly independent
Suppose that we have $a_1(e_1 + U) + \cdots + a_n(e_n + U) = 0$ with $e_1 + U, \ldots, e_n + U \in \overline{\mathcal{B}}$ and $e_1, \ldots, e_n \in \mathcal{B}\backslash\mathcal{E}$. Then

$$\begin{aligned} & a_1e_1 + \ldots + a_ne_n + U = U \\ \Rightarrow & a_1e_1 + \ldots + a_ne_n \in U \\ \Rightarrow & a_1e_1 + \ldots + a_ne_n = b_1e'_1 + \ldots + b_ke'_k && (b_i \in \mathbb{F}, e'_i \in \mathcal{E}) \\ \Rightarrow & a_1 = \cdots = a_n = b_1 = \cdots = b_k = 0 && (\mathcal{E} \text{ is linearly independent}) \end{aligned}$$

Hence $\overline{\mathcal{B}}$ is linearly independent

Example 3.1.

$V = \mathbb{F}[x] : \mathcal{B} = \{1, x, x^2, \ldots\}$

$U =$ even polynomials : $\mathcal{E} = \{1, x^2, x^4, \ldots\}$

$V/U =$ odd polynomials : $\overline{\mathcal{B}} = \{x + U, x^3 + U, \ldots\}$

Corollary 3.2. If V is finite dimensional, then

$$\dim V = \dim U + \dim V/U$$

Theorem 3.3. First Isomorphism Theorem (For Vector Spaces)
Let $T : V \to W$ be a linear transformation of vector spaces. Then,

$$\begin{aligned}\overline{T} : V/\ker T &\to \textit{Im } T \\ v + \ker T &\mapsto T(v)\end{aligned}$$

is a linear isomorphism

Proof. We first show that $\overline{T}$ is well defined.
Suppose $v + \ker T = v' + \ker T$, then $v = v' + k$ for some $k \in \ker T$ and hence:

$$\begin{aligned}\overline{T}(v + \ker T) &= T(v) \\ &= T(v' + k) \\ &= T(v') + T(k) \\ &= T(v') \\ &= \overline{T}(v' + \ker T)\end{aligned}$$

Moreover, $\overline{T}$ is a homomorphism since

$$\begin{aligned}\overline{T}\big(a(v + \ker T) + (v' + \ker T)\big) &= \overline{T}(av + v' + \ker T) \\ &= T(av + v') \\ &= aT(v) + T(v') \\ &= a\overline{T}(v + \ker T) + \overline{T}(v' + \ker T)\end{aligned}$$

Now, $\overline{T}$ is injective as it has a trivial kernel:

$$\begin{aligned}\overline{T}(v + \ker T) = 0 &\iff T(v) = 0 \\ &\iff v \in \ker T \\ &\iff v + \ker T = \ker T \qquad \text{(Since } \ker T \text{ is a vector subspace)}\end{aligned}$$

Finally, $\overline{T}$ is surjective as its image is Im T. □

Corollary 3.4. Rank-Nullity Theorem
If $T : V \to W$ is a linear map and V is a finite dimensional vector space then

$$\dim V = \dim \ker T + \dim \mathrm{Im}\ T$$

Proof. We take $U = \ker T$ and apply Corollary 3.2.

$$\begin{aligned}\dim V &= \dim U + \dim V/U \\ &= \dim \ker T + \dim V/\ker T \\ &= \dim \ker T + \dim \mathrm{Im}\ T \qquad \text{(By First Isomorphism Theorem)}\end{aligned}$$

□

Example 3.2. Let $V = \mathbb{R}^3$ and $U = \left\langle \left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right) \right\rangle$, then

$$\dim V/U = \dim V - \dim U = 3 - 1 = 2$$

A basis for V/U is given by $\mathcal{B} = \left\{ \left(\begin{smallmatrix}0\\0\\1\end{smallmatrix}\right) + U, \left(\begin{smallmatrix}0\\1\\1\end{smallmatrix}\right) + U \right\}$
We can visualise V/U as the space of lines parallel to U.

Let $T : V \to W$ be a linear map and let $A \subset V$ and $B \subset W$ be subspaces.

Lemma 3.5. The formula $\overline{T}(v + A) := T(v) + B$ defines a linear map $\overline{T} : V/A \to W/B$ if and only if $T(A) \subseteq B$

Proof. Assume $T(A) \subset B$. Then $\overline{T}$ will be linear if well defined.
Let $v + A = v' + A$
Then $v = v' + a$ for some $a \in A$

$$\begin{aligned}
\overline{T}(v + a) &= T(v) + B \\
&= T(v' + a) + B \\
&= T(v') + T(a) + B \\
&= T(v') + B && (T(a) \in B) \\
&= \overline{T}(v' + A)
\end{aligned}$$

Conversely, assume $\exists a \in A : T(a) \notin B$, then:

$$\begin{aligned}
B = \overline{T}(A) &= \overline{T}(a + A) \\
&= T(a) + B \\
&\Rightarrow T(a) \in B && \text{(CONTRADICTION)}
\end{aligned}$$

□

Let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be a basis for V containing $\mathcal{E} = \{e_1, \ldots, e_k\}$, a basis for A.
Let $\mathcal{B}' = \{e'_1, \ldots, e'_m\}$ be a basis for W containing $\mathcal{E}' = \{e'_1, \ldots, e'_l\}$ a basis for B.

Proposition 3.6. Assume $T : V \to W$ satisfies $T(A) \subset B$, then T can be restricted to a linear map $T|_A : A \to B$ by $a \mapsto T(a)$.
Then we have the following block matrix composition of T:

$$_{\mathcal{B}'}[T]_{\mathcal{B}} = \begin{pmatrix} _{\mathcal{E}'}[T|_A]_{\mathcal{E}} & \star \\ 0 & _{\overline{\mathcal{B}}'}[\overline{T}]_{\overline{\mathcal{B}}} \end{pmatrix}$$

Remark.

$$\begin{aligned}
\overline{T}(e_j + A) &= T(e_j) + B \\
&= a_{1j}e'_1 + \cdots + a_{mj}e'_m + B \\
&= a_{l+1,j}e'_{l+1} + \cdots + a_{mj}e'_m + B && (a_{1j}e'_j + \cdots + a_{lj}e_l \in B) \\
&= (a_{l+1,j}e_{l+1,j} + B) + \cdots + (a_{mj}e'_m + B)
\end{aligned}$$

Chapter 4

Triangular Form and Cayley-Hamilton Theorem

Let $T : V \to V$ be a linear transformation. A subspace $U \subseteq V$ is called **T-invariant** if $T(U) \subseteq U$. Let $S : V \to V$ be another transformation.

Lemma 4.1. If U is $T-$ and $S-$ invariant, then it is also invariant in the following:

(1) **zero map**, since $U \leq V$ we have $0 \in U$

(2) **identity map**, clearly $U \subseteq U$

(3) $\boldsymbol{aT}$ **for any** $\boldsymbol{a} \in \mathbb{F}$, U subspace $\to$ closed under scalar multiplication

(4) $\boldsymbol{S+T}$, $S(U), T(U) \in U$, U closed under addition

(5) $\boldsymbol{T \circ S}$, $S(U) \in U \Rightarrow T(S(U)) \subseteq T(U) \subseteq U$

In particular, U in invariant for any $\rho(T)$ where $\rho(x) \in \mathbb{F}[x]$. Moreover, $\rho(T)$ restricts to a map $U \to U$ and also induces a linear map of quotient spaces:

$$\overline{\rho(T)} : V/U \to V/U$$

Example 4.1. If λ is a root of characteristic polynomial, $\mathcal{X}_T(x)$, then $\exists v \in V$ with $v \neq 0$ such that $Tv = \lambda v$. Then $\langle v \rangle$ is $T-$invariant.

Remark. More generally, $V_\lambda := \ker(T - \lambda I)$, the $\boldsymbol{\lambda}$**-eigenspace** of T is $T-$invariant

Recall: $\mathcal{E}$ basis for U, $\mathcal{B}$ basis for V, with $\mathcal{E} \subseteq \mathcal{B}$. The $\overline{\mathcal{B}} = \{v + U : v \in \mathcal{B}/\mathcal{E}\}$ is a basis for V/U with

$${}_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} {}_{\mathcal{E}}[T|_U]_{\mathcal{E}} & \star \\ 0 & {}_{\overline{\mathcal{B}}}[\overline{T}]_{\overline{\mathcal{B}}} \end{pmatrix}$$

Remark. Determinant is independent of basis:
Say $P^{-1}AP = B$, then

$$\begin{aligned}\det(B - xI) &= \det(P^{-1}AP - xI)\\ &= \det(P^{-1}(A - xI)P)\\ &= \det(P^{-1})\det(A - xI)\det(P)\\ &= \frac{\det(A - xI)\det(P)}{\det(P)}\\ &= \det(A - xI)\end{aligned}$$

Proposition 4.2.

$$\begin{aligned}\mathcal{X}_T(x) &= \det\left({}_{\mathcal{B}}[T]_{\mathcal{B}} - xI\right)\\ &= \det\left({}_{\mathcal{E}}[T|_U]_{\mathcal{E}} - xI\right) \cdot \det\left({}_{\overline{\mathcal{B}}}[\overline{T}]_{\overline{\mathcal{B}}}\right)\\ &= \mathcal{X}_{T|_U}(x) \cdot \mathcal{X}_{\overline{T}}(x)\end{aligned}$$

Remark. The relation between the minimal polynomials is not so straight forward!

Definition 4.1. $A = (a_{ij}) \in M_{n\times n}(\mathbb{F})$ is upper triangular if $a_{ij} = 0$ for all $i > j$

Theorem 4.3. *Let V be a finite vector space and $T : V \to V$ a linear transformation. Assume that $\mathcal{X}_T(x)$ is a product of linear factors. Then there exists a basis $\mathcal{B}$ for V such that ${}_{\mathcal{B}}[T]_{\mathcal{B}}$ is upper triangular*

Remark. If our field $\mathbb{F}$ is algebraically closed, such as $\mathbb{C}$, then the characteristic polynomial is always a product of linear factors

Proof. We proceed by induction on $\dim V = n$
If $n = 1$, then clearly ${}_{\mathcal{B}}[T]_{\mathcal{B}}$ is upper triangular for any basis $\mathcal{B}$
In general, $\mathcal{X}_T$ has a root λ and hence $\exists v_1 \in V$ such that $Tv_1 = \lambda v_1$.
Now, let $U = \langle v_1 \rangle$, then U is $T-$invariant. Consider $\overline{T} : V/U \to V/U$; by proposition, $\mathcal{X}_{\overline{T}}(x)$ is also a product of linear factors. By the induction hypothesis, $\exists\overline{\mathcal{B}} = \{v_2 + U, \ldots, v_n + U\}$ such that ${}_{\overline{\mathcal{B}}}[\overline{T}]_{\overline{\mathcal{B}}}$ is upper-triangular. We can now put $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$, so $\mathcal{B}$ is a basis for V and

$${}_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} \lambda & \star \\ 0 & {}_{\overline{\mathcal{B}}}[\overline{T}]_{\overline{\mathcal{B}}} \end{pmatrix}$$

Since ${}_{\overline{\mathcal{B}}}[\overline{T}]_{\overline{\mathcal{B}}}$ is upper triangular then so is ${}_{\mathcal{B}}[T]_{\mathcal{B}}$.

□

Corollary 4.4. If $A \in M_{n\times n}(\mathbb{F})$ with characteristic polynomial a product of linear factors, then there exists P such that $P^{-1}AP$ is upper-triangular

Proposition 4.5. Let A be an upper-triangular matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. Then $(A - \lambda_1 I) \cdots (A - \lambda_n I) = 0$

Proof. Let $e_1, \dots, e_n$ be the standard basis for $\mathbb{F}^n$.

$(A - \lambda_n I)v \in \langle e_1, \dots, e_{n-1} \rangle$ for all $v \in V$

More generally,

$(A - \lambda_i I)w \in \langle e_1, \dots, e_{i-1} \rangle$ for all $w \in \langle e_1, \dots, e_i \rangle$

Why... more explanation required here

Hence,

$\underbrace{(A - \lambda_1 I)}_{\in \langle e_1 \rangle} \underbrace{\cdots\cdots\cdots}_{\in \langle e_1, \dots, e_{n-2} \rangle} \underbrace{(A - \lambda_{n-1} I)}_{\in \langle e_1, \dots, e_{n-1} \rangle}(A - \lambda_n I)v$

$\square$

Theorem 4.6. Cayley-Hamilton Theorem

If $T : V \to V$ is a linear transformation and V finite dimensional, then $\mathcal{X}_T(T) = 0$ and hence $m_T(x) \div \mathcal{X}_T(x)$

Proof. We work over the algebraic closure $\overline{\mathbb{F}} \supseteq \mathbb{F}$

Now, $\mathcal{X}_T(x) = (x - \lambda_1) \cdots (x - \lambda_n)$ for some $\lambda_i \in \overline{\mathbb{F}}$.

By the above theorem, for some basis $\mathcal{B}$, $A = {}_{\mathcal{B}}[T]_{\mathcal{B}}$ is upper-triangular.

Hence $\mathcal{X}_T(T) = \mathcal{X}_T(A) = (A - \lambda_1 I) \cdots (A - \lambda_n I) = 0$

As the minimal polynomial divides any annihilating polynomial it must divide $\mathcal{X}_T(x)$.

$\square$

Example 4.2.

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \qquad \Longrightarrow \qquad \mathcal{X}_A(x) = (1-x)^2(2-x)^2$$

Possible minimal polynomials:

$$\begin{aligned} &(x-1)(x-2) && (A-I)(A-2I) \neq 0 \\ &(x-1)(x-2)^2 && (A-I)(A-2I)^2 \neq 0 \\ &(x-1)^2(x-2) && (A-I)^2(A-2I) = 0 \end{aligned}$$

Hence $m_A(x) = (x-1)^2(x-2)$

Chapter 5

The Primary Decomposition Theorem

Proposition 5.1. Let $a, b \in \mathbb{F}[x]$ be non-zero polynomials. Assume that $\gcd(a, b) = c \in \mathbb{F}[x]$. Then $\exists s, t \in \mathbb{F}[x]$ such that

$$a(x)s(x) + b(x)t(x) = c(x)$$

Proof. Without loss of generality, assume that $\deg a \geq \deg b$ and $\gcd(a, b) = 1$

Proceed by induction on $\deg a + \deg b$

By the division algorithm for polynomials we have that there exist $q, r \in \mathbb{F}[x]$ such that

$$a(x) = q(x)b(x) + r(x) \quad \text{and} \quad \deg r < \deg b$$

Now, if $r(x) = 0$, then we have that $b(x) = \lambda \in \mathbb{F}$, some constant (since $\gcd(a, b) = 1$) and hence

$$a(x) + b(x)\left(\frac{1}{\lambda}\right)(1 - a(x)) = 1$$

and then we are done. So, assume now that $r \neq 0$, note:

- $\deg r + \deg b < \deg a + \deg b$ (since $\deg r < \deg b \leq \deg a$)
- $\gcd(a, b) = 1 \Longrightarrow \gcd(r, b) = 1$

Hence, by the induction hypothesis, there exists $s', t' \in \mathbb{F}[x]$ such that

$$b(x)s'(x) + r(x)t'(x) = 1$$

Then, by $(*)$ we have

$$b(x)s'(x) + \big(a(x) - q(x)b(x)\big)t'(x) = 1$$
$$a(x)t'(x) + b(x)\left(s'(x) - q(x)t'(x)\right) = 1$$

Hence setting $s = t'$ and $t = s' - qt$ we are done.

□

Remark. Direct Sum Decompositions:

- $V = W_1 \oplus \cdots \oplus W_r$ is the direct sum of subspaces W_i if every $v \in V$ can be written as $v = w_1 + \cdots + w_r$ with $w_i \in W_i$ in a unique way
- Let $\mathcal{B}_i$ be a basis for W_i for $i = 1, \ldots, r$
  Then $\bigcup_i \mathcal{B}_i = \mathcal{B}$ is a basis for $V = \bigoplus_i W_i$

From now on we assume that $\dim V < \infty$

Let $T : V \to V$ be a linear transformation such that W_i is $T-$invariant: $T(W_i) \subseteq W_i$ for all i.
Then,

$$_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{pmatrix} \text{ where } A_i = {}_{\mathcal{B}_i}[T|_{W_i}]_{\mathcal{B}_i}$$

Also note that:

$$\mathcal{X}_T(x) = \mathcal{X}_{T|_{W_1}}(x) \times \cdots \times \mathcal{X}_{T|_{W_r}}(x)$$

Proposition 5.2. Assume $f(x) = a(x)b(x)$ with $\gcd(a, b) = 1$ and $f(T) = 0$.
Then $V = \ker a(T) \oplus \ker b(T)$ is a $T-$invariant direct sum decomposition

Proof. Suppose that $v \in \ker a(T)$
Then

$$\begin{aligned} a(T)(Tv) &= (a(T)T)(v) \\ &= (Ta(T))(v) \\ &= T(a(T)(v)) \\ &= T(\mathbf{0}) = \mathbf{0} \end{aligned}$$

Hence $Tv \in \ker a(T)$ and so $\ker a(T)$ (and similarly $\ker b(T)$) is $T-$invariant.
By Proposition 5.1. there exist $s, t \in \mathbb{F}[x]$ with $as + bt = 1$

$$\Rightarrow a(T)s(T) + b(T)t(T) = \mathrm{Id_v}$$

$$\Rightarrow v = \mathrm{Id_v}(v) = a(T)s(T)v + b(T)t(T)v \qquad (*)$$

Moreover, for all $v \in V$ we have

$$\begin{aligned} b(T)[(a(T)s(T))(v)] &= s(T)[a(T)b(T)(v)] \\ &= s(T)f(T)(v) && (ab = f) \\ &= s(T) \cdot \mathbf{0} && (f(T) = 0) \\ &= \mathbf{0} \end{aligned}$$

Hence $a(T)s(T)v \in \ker b(T)$, and similarly $b(T)t(T)v \in \ker a(T)$ so $V = \ker a(T) + \ker b(T)$.

It remains to show that $\ker a(T) \cap \ker b(T) = \{0\}$

Suppose that $v \in \ker a(T) \cap \ker b(T)$, then

$$\begin{aligned}
v &= a(T)s(T)v + b(T)t(T)v && (\text{by } *)\\
&= \quad \mathbf{0} \quad + b(T)t(T)v && (v \in \ker a(T))\\
&= \quad \mathbf{0} \quad + \quad \mathbf{0} && (v \in \ker b(T))\\
&= \quad \mathbf{0}
\end{aligned}$$

Thus $V = \ker a(T) \oplus \ker b(T)$ as required

□

Remark. If $f(x) = m_T(x)$ is the minimal polynomial of T in the above proposition, then we obtain:

$$m_{T|_{\ker a(T)}}(x) = a(x) \qquad\qquad m_{T|_{\ker b(T)}}(x) = b(x)$$

and

$$m_T(x) = m_{T|_{\ker a(T)}}(x) \cdot m_{T|_{\ker b(T)}}(x) = a(x)b(x)$$

Proof. Call $m_1(x) = m_{T|_{\ker a(T)}}(x)$ and $m_2(x) = m_{T|_{\ker b(T)}}(x)$.

By definition a is annihilating for $\ker a(T)$ so $m_1 | a$, similarly $m_2 | b$
Further, for any $v \in V$ there exists $w_1 \in \ker a(T), w_2 \in \ker b(T)$ with $v = w_1 + w_2$, thus:

$$\begin{aligned}
m_1(T)m_2(T)v &= m_1(T)m_2(T)w_1 + m_1(T)m_2(T)w_2 \\
&= \qquad 0 \qquad + m_1(T)m_2(T)w_2 && (m_1(T) \text{ annihilates } \ker a(T))\\
&= \qquad 0 \qquad + \qquad 0 && (m_2(T) \text{ annihilates } \ker b(T))
\end{aligned}$$

Hence $m_1(T)m_2(T) = 0$ and so $m | m_1 m_2$

By degree and minimality we have $m = m_1 \cdot m_2 = ab$ with $m_1 = a$ and $m_2 = b$.

□

Theorem 5.3. Primary Decomposition Theorem

Assume that the minimal polynomial has the form

$$m_T(x) = f_1(x)^{m_1} \cdots f_r(x)^{m_r}$$

Where the f_i are distinct irreducible monic polynomials
Put $W_i = \ker f_i(T)^{m_i}$, then

- *W_i is $T-$invariant*
- $V = W_1 \oplus \cdots \oplus W_r$
- $m_{T|_{W_i}} = f_i(x)^{m_i}$

Proof. Put $a = f_1 \cdots f_{r-1}$ and $b = f_r$ and proceed by induction using Proposition 5.2. □

Remark.

- Given $m_T(x) = f_1(x)^{m_1} \cdots f_r(x)^{m_r}$ as in the theorem,

$$\mathcal{X}(x) = f_1(x)^{n_1} \cdots f_r(x)^{n_r} \text{ with } n_i \geq m_i$$

Proof.

- T is triagonalizable
  $\Leftrightarrow$ $\mathcal{X}_T$ factors as a product of linear polynomials
  $\Leftrightarrow$ each f_i is linear
  $\Leftrightarrow$ m_T factors as a product of linear polynomials

Let $T : V \to V$ be a linear map on a finite dimensional vector space

Theorem 5.4. *T is diagonalizable $\iff$ m_T factors as a product of distinct linear polynomials*

Proof.

$\Leftarrow$ Assume $m_T(x) = (x - \lambda_1)\cdots(x - \lambda_n)$ for some $\lambda_i \neq \lambda_j$

By Primary Decomposition Theorem we have

$$\begin{aligned} V &= \ker(T - \lambda_1 I) \oplus \cdots \oplus \ker(T - \lambda_n I) \\ &= E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_n} \end{aligned}$$

is the direct sum of the eigenspaces

Let $\mathcal{B}_i$ be a basis for E_{λ_i}: $\mathcal{B} = \cup_i \mathcal{B}_i$ is a basis for V and ${}_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ is diagonal

Note that the λ_i may be repeated many times

$\Rightarrow$ T is diagonal $\to$ $\exists \mathcal{B}$ a basis of eigenvectors and every $v \in V$ is $v = \sum_i a_i v_i$ for these eigenvectors v_i.

Define $f(x) = (x-\lambda_1)\cdots(x-\lambda_n)$, with λ_i the distinct eigenvalues of the v_i. Then $f(T) = 0$ since $f(T)$ annihilates every element of $\mathcal{B}$. Now, since we must have $m_T | f$ then $m_T = f$, a product of distinct linear factors.

$\square$

Example 5.1. P is a projection $\iff$ $P^2 = P$ $\iff$ $P^2 - P = P(P - I) = 0$

$$\implies m_P(x) = \begin{cases} x(x-1) & V = E_0 + E_1, \exists \mathcal{B} : {}_{\mathcal{B}}[P]_{\mathcal{B}} = \left(\begin{smallmatrix} 0 & 0 \\ 0 & I \end{smallmatrix}\right) \\ x & P = 0 \\ (x-1) & P = I \end{cases}$$

Example 5.2. Suppose $A = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ and $\mathcal{X}_A = (1-x)^2 + 1 = x^2 - 2x + 2$, then:

$$\begin{aligned}
(1)\ \ \mathbb{F} = \mathbb{R} &\implies \text{no roots} \\
&\implies \mathcal{X}_A(x) = m_A(x) \\
&\implies \text{not triangulizable}
\end{aligned}$$

$$\begin{aligned}
(2)\ \ \mathbb{F} = \mathbb{C} &\implies \mathcal{X}_A(x) = (x-(1+i))(x-(1-i)) \\
&\implies \mathcal{X}_A(x) = m_A(x) && \text{(distinct roots)} \\
&\implies \text{triangulizable and diagonalizable}
\end{aligned}$$

$$\begin{aligned}
(3)\ \ \mathbb{F} = \mathbb{F}_5 &\implies \mathcal{X}_A(x) = (x-3)(x-4) \\
&\implies \mathcal{X}_A(x) = m_A(x) && \text{(distinct roots)} \\
&\implies \text{triangulizable and diagonalizable}
\end{aligned}$$

Challenge 3. Find a basis of eigenvectors in $(\mathbb{F}_5)^2$

Chapter 6

Jordan Canonical Form

Let V be finite dimensional and $T : V \to V$ a linear map

Definition 6.1. If $T^m = 0$ for some $m > 0$ then T is **nilpotent**

Theorem 6.1. *If T is nilpotent and $m_T(x) = x^m$ for some $m > 0$, then there exists a basis $\mathcal{B}$ of V such that*

$$_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} 0 & * & & \\ & \ddots & \ddots & \\ & & 0 & * \\ & & & 0 \end{pmatrix} \text{ where } * = 0, 1$$

Proof.
Note that $0 \subset \ker T \subset \ker T^2 \subset \cdots \subset \ker T^{m-1} \subset \ker T^m = V$
Let $\mathcal{B}_i$ be such that $\overline{\mathcal{B}_i} = \{w + \ker T^{i-1} : w \in \mathcal{B}_i\}$ is a basis for $\ker T^i / \ker T^{i-1}$

$\boxed{\text{Claim}_1\text{: } \mathcal{B} = \bigcup_i \mathcal{B}_i \text{ is a basis for } V}$

Proof. Since V is finite dimensional we have $\dim V = \dim U + \dim V/U$

$$\begin{aligned} \dim V = \dim T^m &= \dim(\ker T^m / \ker T^{m-1}) + \dim(\ker T^{m-1}) \\ &= \dim(\ker T^m / \ker T^{m-1}) + \dim(\ker T^{m-1} / \ker T^{m-2}) + \dim(\ker T^{m-2}) \\ &\dots \\ &= \dim(\ker T^m / \ker T^{m-1}) + \ldots + \dim(\ker T^2 / \ker T) + \dim(\ker T / \{0\}) \\ &= \left|\overline{\mathcal{B}_m}\right| + \cdots + \left|\overline{\mathcal{B}_2}\right| + \left|\overline{\mathcal{B}_1}\right| \\ &= |\mathcal{B}_m| + \cdots + |\mathcal{B}_2| + |\mathcal{B}_1| \end{aligned}$$

$$\underbrace{\underbrace{\underbrace{\mathcal{B}_1}_{\ker T} \cup \mathcal{B}_2}_{\ker T^2} \cup \cdots \cup \mathcal{B}_m}_{\ker T^m = V}$$

$\boxed{\text{CLAIM}_2\text{: } \{Tw + \ker T^{i-1} : w \in \mathcal{B}_{i+1}\} \subset \ker T^i / \ker T^{i-1} \text{ is linearly independent}}$

Proof.

Assume $\sum_s a_s(Tw_s + \ker T^{i-1}) = \ker T^{i-1}$

$\implies \sum_s a_s T w_s \in \ker T^{i-1}$

$\implies T \sum_s a_s w_s \in \ker T^{i-1}$

$\implies \sum_s a_s w_s \in \ker T^i$

$\implies \sum_s a_s(w_s + \ker T^i) = \ker T^i$

$\implies a_s = 0$ for all s as $\overline{\mathcal{B}_{i+1}}$ is a basis for $\ker T^{i+1}/\ker T^i$

Now, we can inductively find $\mathcal{E}_i = \{w_1^i, \ldots, w_k^i\}$ such that $\mathcal{B}_i = \mathcal{E}_i \sqcup T(\mathcal{B}_{i+1})$ with $\overline{\mathcal{B}_i}$ a basis for $\ker T^i / \ker T^{i+1}$ as above. Such $\mathcal{E}_i$ exist as, by CLAIM$_2$, $T(\mathcal{B}_{i+1})$ is linearly independent.

Then, by CLAIM$_1$, we have that $\mathcal{B} = \bigcup_i \mathcal{B}_i$ is a basis for V and furthermore:

$$_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_s \end{pmatrix}$$

is a block diagonal matrix with $|\mathcal{E}_i|$ Jordan blocks of size i with the form:

$$\mathcal{J}_i = \left.\underbrace{\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}}_{i}\right\} i$$

□

Theorem 6.2. *If T is nilpotent and $m_T(x) = x^m$ for some m then there exists a basis $\mathcal{B}$ such that $_{\mathcal{B}}[T]_{\mathcal{B}}$ is block diagonal with blocks equal to*

$$\mathcal{J}_i = \left.\underbrace{\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}}_{i}\right\} i$$

of size $i \leq m$, with at least one block of size m.

Proof. Follows from above that we can write the matrix in block diagonal form, since the minimal polynomial is x^m it is clear that $\ker T^m / \ker T^{m-1}$ has dimension at least one, which is to say $|\mathcal{E}_m| \geq 1$ and so there must be at least Jordan block of size m.

□

Example 6.1. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by:

$$A = \begin{pmatrix} -2 & -1 & 1 \\ 14 & 7 & -7 \\ 10 & 5 & -5 \end{pmatrix}$$

Note that $A^2 = 0$ and hence $\mathcal{X}_A(x) = x^3$ and $m_A(x) = x^2$.

We also have: $0 \subsetneq \ker T \subsetneq \ker T^2 = \mathbb{R}^3$

We can observe that $\ker T = \left\langle \left(\begin{smallmatrix} 1 \\ 0 \\ 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right) \right\rangle$ and so $\dim \ker T = 2$

Further, $\dim \ker T^2 / \ker T = 3 - 2 = 1$ and thus, since $w = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) \notin \ker T$, we have

$$\ker T^2 / \ker T = \left\langle \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) + \ker T \right\rangle$$

So we have $\mathcal{B}_2 = \mathcal{E}_2 = \{w\} = \left\{ \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) \right\}$

Then with $\mathcal{B}_1 = \mathcal{E}_1 \cup T(\mathcal{B}_2)$ we have $T(\mathcal{B}_2) = \{Tw\} = \left\{ \left(\begin{smallmatrix} -2 \\ 14 \\ 10 \end{smallmatrix}\right) \right\}$ and letting $\mathcal{E}_1 = \{u\} = \left\{ \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right) \right\}$ we see that $\mathcal{B}_1 = \overline{\mathcal{B}_1}$ is a basis for $\ker T / \{0\} = \ker T$.

Hence $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 = \{Tw, w, u\}$ is a basis for $\mathbb{R}^3$ and

$${}_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Corollary 6.3. If $m_T(x) = (x - \lambda)^m$ for some m then there is a basis $\mathcal{B}$ for V such that ${}_{\mathcal{B}}[T]_{\mathcal{B}}$ is block diagonal with blocks:

$$\mathcal{J}_i(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}$$

of size $i \leq m$ with at least one block of size m.

Proof. $T - \lambda I$ is nilpotent with $m_{T-\lambda I}(x) = x^m$

Apply theorem to get a basis $\mathcal{B}$ for V such that

$${}_{\mathcal{B}}[T - \lambda I]_{\mathcal{B}} = \begin{pmatrix} \mathcal{J}_m & & \\ & \ddots & \\ & & \mathcal{J}_i \end{pmatrix}$$

Then

$$\begin{aligned} {}_{\mathcal{B}}[T]_{\mathcal{B}} &= {}_{\mathcal{B}}[T - \lambda I + \lambda I]_{\mathcal{B}} \\ &= {}_{\mathcal{B}}[T - \lambda I]_{\mathcal{B}} + {}_{\mathcal{B}}[\lambda I]_{\mathcal{B}} \\ &= {}_{\mathcal{B}}[T - \lambda I]_{\mathcal{B}} + \lambda I \end{aligned}$$

Which is the form required $\square$

Definition 6.2. The $\mathcal{J}_i$ are called **Jordan Blocks**

Example 6.2. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by $A = \left(\begin{smallmatrix} 3 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & 2 \end{smallmatrix}\right)$

$$\mathcal{X}_A(x) = \begin{vmatrix} 3-x & 0 & 1 \\ -1 & 1-x & -1 \\ 0 & 1 & 2-x \end{vmatrix} = (2-x)^3$$

We consider $A - 2I$:

$$\begin{aligned} A - 2I &= \left(\begin{smallmatrix} 1 & 0 & 1 \\ -1 & -1 & -1 \\ 0 & 1 & 0 \end{smallmatrix}\right) \\ (A-2I)^2 &= \left(\begin{smallmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{smallmatrix}\right) \\ (A-2I)^3 &= 0 \end{aligned}$$

So $m_A = (x-2)^3$ and we can read off the Jordan Form: $\left(\begin{smallmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{smallmatrix}\right)$

To construct the basis $\mathcal{B}$ we first note that for $S = A - 2I$

$$0 \subsetneq \ker S \subsetneq \ker S^2 \subsetneq \ker S^3 = \mathbb{R}^3$$

Thus $\dim \ker S^3 / \ker S^2 = 1$ and as $\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) \notin \ker T^2$ we can set $\mathcal{B}_3 = \mathcal{E}_3 = \left\{\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)\right\}$

Further, $\dim \ker T^2 / \ker T = 1$ and hence $B_2 = S(\mathcal{B}_3) = \left\{\left(\begin{smallmatrix} 1 \\ -1 \\ 0 \end{smallmatrix}\right)\right\}$

Finally we have $\dim \ker T / \{0\} = 1$ and hence $B_1 = S(\mathcal{B}_2) = \left\{\left(\begin{smallmatrix} 1 \\ 0 \\ -1 \end{smallmatrix}\right)\right\}$

Hence for $\mathcal{B} = \left\{\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ -1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 0 \\ -1 \end{smallmatrix}\right)\right\}$ we have

$${}_{\mathcal{B}}[T - 2I]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow {}_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Lemma 6.4. Consider $\mathcal{J}_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}$ and put $V_n = (v_n^1, \dots, v_n^k)$

Suppose $V_n = \mathcal{J}_k(\lambda) V_{n-1} = (\mathcal{J}_k(\lambda))^n V_0$

Then $v_n^{k-i} = \lambda^n v_0^{k-i} + \binom{n}{1} \lambda^{n-1} v_0^{k-i+1} + \dots + \binom{n}{i} \lambda^{n-i} v_0^k$

Proof. By induction on n

Base case: $n = 0$

$\binom{0}{n} = 0$ thus, for any i, we have: $v_0^{k-i} = \lambda^0 v_0^{k-i} = v_0^{k-i}$ which is clearly true

Case n:

$$v_n^{k-i} = \lambda v_{n-1}^{k-i} + v_{n-1}^{k-i+1}$$

By induction hypothesis the lemma holds for $n-1$

$$\begin{aligned}
&= \lambda \left[\lambda^{n-1} v_0^{k-i} + \binom{n-1}{1} \lambda^{n-2} v_0^{k-i+1} + \dots + \binom{n-1}{i} \lambda^{n-1-i} v_0^k \right] \\
&\quad + \left[\lambda^{n-1} v_0^{k-i+1} + \binom{n-1}{1} \lambda^{n-2} v_0^{k-i+2} + \dots + \binom{n-1}{i-1} \lambda^{n-i} v_0^k \right] \\
&= \lambda^n v_0^{k-i} + \left[\binom{n-1}{0} + \binom{n-1}{1} \right] \lambda^{n-1} v_0^{k-i+1} + \dots + \left[\binom{n-1}{i-1} + \binom{n-1}{i} \right] \lambda^{n-i} v_0^k
\end{aligned}$$

We now use the identity $\binom{n-1}{j-1} + \binom{n-1}{j} = \binom{n}{j}$

$$= \lambda^n v_0^{k-i} + \binom{n}{1} \lambda^{n-1} v_0^{k-i+1} + \dots + \binom{n}{i} \lambda^{n-i} v_0^k$$

Which is the desired identity

□

Chapter 7

Dual Spaces

Definition 7.1. Let V be a vector space over $\mathbb{F}$, then its **dual**, V', is the vector space of maps from V to $\mathbb{F}$. i.e. $V' = \hom(V, \mathbb{F})$

Definition 7.2. The elements of V' are called **linear functionals**

Example 7.1. Let $V = \mathcal{C}([0,1])$ be the vector space of continuous functions on $[0,1]$

Then $\int : V \to \mathbb{R}$ given by $f \mapsto \int_0^1 f(t)dt$ is a linear functional

Proof.

$$\begin{aligned}
\int\big(f + \lambda g\big) &= \int_0^1 \big(f + \lambda g\big)(t)dt \\
&= \int_0^1 \big(f(t) + \lambda g(t)\big)dt \\
&= \int_0^1 f(t)dt + \lambda \int_0^1 g(t)dt \\
&= \int(f) + \lambda \int(g)
\end{aligned}$$

Example 7.2. Let V be the vector space of finite sequences: $V = \{(a_0, a_1, \ldots) : \text{finitely many } a_i \neq 0\}$

Let $\bar{b} = (b_0, b_1, \ldots)$ be any sequence, then $\bar{b}\big((a_0, a_1, \ldots)\big) = \sum_1^\infty a_i b_i$ defines a linear functional

Proof.

$$\begin{aligned}
\bar{b}\big((a_0, a_1, \ldots) + \lambda(a_0', a_1', \ldots)\big) &= \sum_i \big(a_i + \lambda a_i'\big)b_i \\
&= \sum_i a_i b_i + \lambda \sum_i a_i' b_i \\
&= \bar{b}\big((a_0, a_1, \ldots)\big) + \lambda \bar{b}\big((a_0', a_1', \ldots)\big)
\end{aligned}$$

Theorem 7.1. *Let V be a finite dimensional vector space and let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be a basis for V For each i define the dual of e_i (with respect to $\mathcal{B}$) to be the linear functional*

$$e_i'(e_j) = \delta_{ij} = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}$$

*Then $\mathcal{B}' = \{e_1', \ldots, e_n'\}$ is a basis for V' called the **dual basis** of $\mathcal{B}$*

Remark. In particular $e_i \mapsto e_i'$ defines an isomorphism from V to V'

Proof. We first show linear independence. Assume that $\sum a_i e_i' = 0$, then

$$\begin{aligned} \sum a_i e_i' = 0 &\implies \forall j : \left(\sum a_i e_i'\right)(e_j) = 0 \\ &\iff \forall j : \sum a_i e_i'(e_j) = 0 \\ &\iff \forall j : a_j = 0 \end{aligned}$$

Next we show that $\mathcal{B}'$ spans V'. Suppose $f \in V'$
We put $a_i = f(e_i)$ for each i.
Then $f = \sum_i a_i e_i'$ as both evaluate to the same on the basis elements:

$$f(e_j) = a_j; \quad \left(\sum a_i e_i'\right)(e_j) = \sum a_i e_i' e_j = a_j$$

□

Example 7.3. Let $V = \mathbb{R}^n$ with basis $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\}$
Then the dual basis is given by:

$$\mathcal{B}' = \{(1, 0, \ldots, 0), \ldots, (0, \ldots, 1)\} \in V' = M_{1\times n}(\mathbb{R})$$

Remark. If V is the vector space of finite sequences then V' is the vector space of infinite sequences.

Since any linear functional is uniquely determined by its values on the basis elements $e_i = (0, \ldots, 0, 1, 0, \ldots)$

i.e. f is determined by $\bar{b} = (b_0, b_1, \ldots)$ where $b_i = f(e_i)$

Remark. In this case V is **not isomorphic** to V'
Though the dual basis elements $\{e_0', e_1', \ldots\}$ are linearly independent, they do not span: $(1, 1, \ldots) \notin \langle e_i \rangle$

Definition 7.3. A **natural** linear map is independent of choice of basis
(in contrast to the dual map: $v \mapsto v'$)

Theorem 7.2. *Let V be a finite dimensional vector space, then $V \to (V')' = V''$ defined by $v \mapsto E_v$ where $E_v : V' \to \mathbb{F}$ is defined by $f \mapsto f(v)$, taking a vector v to it's evaluation map E_v is a natural linear isomorphism.*

Proof.

- E_v is a linear map:

$$\begin{aligned} E_v(f + \lambda g) &= (f + \lambda g)(v) \\ &= f(v) + \lambda g(v) \\ &= E_v(f) + \lambda E_v(g) \end{aligned}$$

- $v \mapsto E_v$ is injective:
  Assume $E_v = 0$, then $\forall f \in V'$ we have $E_v(f) = f(v) = 0$. We want to show that this obtains iff $v = 0$
  Assume that $v \neq 0$, then we can extend to a basis $\mathcal{B} = \{v, e_2, \ldots, e_n\}$ for V. Then for $f = v'$ with respect to $\mathcal{B}$ we get $E_v(f) = E_v(v') = v'(v) = 1 \neq 0$ which is a contradiction. Hence $v = 0$ and thus the map is injective.
- $v \mapsto E_v$ is surjective:
  Observe that $\dim V = \dim V' = \dim V''$. Then by injectivity and the rank-nullity theorem the map must be surjective.

□

Definition 7.4. Let $U \leq V$. The we define the **annihilator** of U to be

$$U^0 = \{f \in V' : f|_U \equiv 0\}$$

Proposition 7.3. Let $U \leq V$. Then the annihilator of U is a subspace of V'

Proof. First note that $f \equiv 0 \in U^0$, so that $U^0 \neq \varnothing$
Now, suppose $f, g \in U^0$ and $\lambda \in \mathbb{F}$, then

$$\begin{aligned} (f + \lambda g)(U) &= f(U) + \lambda g(U) \\ &= 0 + \lambda 0 && (f, g \in U^0) \\ &= 0 \end{aligned}$$

Thus $f + \lambda g \in U^0$ and hence $U^0 \leq V'$

□

Theorem 7.4. *If V is finite dimensional and $U \leq V$ then* $\dim U^0 = \dim V - \dim U$

Proof. Let $\mathcal{B}_U = \{e_1, \ldots, e_m\}$ be a basis for U and extend to a basis $\mathcal{B}_V = \{e_1, \ldots, e_m, \ldots, e_n\}$ for V. If we consider the dual basis $B'_V = \{e'_1, \ldots, e'_n\}$ then the theorem follows from the claim that $\{e'_{m+1}, \ldots, e'_n\}$ is a basis for U^0.

Proof. $\mathcal{B}'_U = \{e'_{m+1}, \ldots, e'_n\} \subset \mathcal{B}'_V$ hence $\mathcal{B}'_U$ is linearly independent

For $j = m+1, \ldots, n$ and $i = 1, \ldots, m$ we have $e'_j(e_i) = 0$, thus $\langle \mathcal{B}'_U \rangle \subset U^0$

Now, let $f \in U^0 \leq V'$, then there exist $a_i \in \mathbb{F}$ such that $f = \sum_i a_i e'_i$ and, since $\mathcal{B}_U$ is a basis for U, we have that for $i = 1, \ldots, m$:

$$f(e_i) = 0 = \sum_{j=1}^{n} a_j e'_j e_i = a_i$$

Hence we must have $a_i = 0$ for $i = 1, \ldots, m$ and hence $\mathcal{B}'_U$ is also spanning

□

Theorem 7.5. *If $U, W \leq V$ then:*

(1) $U \leq W \Longrightarrow W^0 \leq U^0$

(2) $(U + W)^0 = U^0 \cap W^0$

(3) $(U \cap W)^0 = U^0 + W^0$ *if* $\dim V < \infty$

Proof.

(1)
$$\begin{aligned} f \in W^0 &\iff \forall w \in W : f(w) = 0 \\ &\Longrightarrow \forall u \in U \leq W : f(u) = 0 \\ &\iff f \in U^0 \end{aligned}$$

□

(2)
$$\begin{aligned} f \in (U+W)^0 &\iff \forall u \in U : f(u) = 0, \forall w \in W : f(w) = 0 \\ &\iff f \in U^0, f \in W^0 \\ &\iff f \in U^0 \cap W^0 \end{aligned}$$

□

(3) $\underline{U^0 + W^0 \leq (U \cap W)^0}$

$$\begin{aligned} f \in U^0 + W^0 &\iff \exists g \in U^0, \exists h \in W^0 : f = g + h \\ &\implies \forall x \in U \cap W : f(x) = g(x) + h(x) = 0 \\ &\iff f \in (U \cap W)^0 \end{aligned}$$

$\underline{(U \cap W)^0 = U^0 + W^0}$

$$\begin{aligned} \dim(U^0 + W^0) &= \dim U^0 + \dim W^0 - \dim(U^0 \cap W^0) \\ &= \dim V - \dim U + \dim V - \dim W - \dim(U + W)^0 \\ &= \dim V - \dim U + \dim V - \dim W - \dim V + \dim(U + W) \\ &= \dim V - \dim U - \dim W + \dim U + \dim W - \dim(U \cap W) \\ &= \dim V - \dim(U \cap W) \\ &= \dim(U \cap W)^0 \end{aligned}$$

$\square$

Theorem 7.6. *Suppose V is finite dimensional and $U \leq V$.*
Then, under the isomorphism $\tau : V \xrightarrow{\sim} V''$ given by $v \mapsto E_v$ we have that $U \cong U^{00}$

Proof.

$$\begin{aligned} E_x \in U^{00} &\iff \forall f \in U^0 : E_x(f) = f(x) = 0 \\ &\implies x \in U \to E_x \in U^{00} \\ &\implies \tau(U) \subseteq U^{00} \end{aligned}$$

Further we have that

$$\begin{aligned} \dim U^{00} &= \dim V - \dim U^0 \\ &= \dim V - (\dim V - \dim U) \\ &= \dim U \end{aligned}$$

Thus $U \cong U^{00}$ as required

$\square$

Theorem 7.7. *Let $U \leq V$ with V finite dimensional. Then there exists an isomorphism such that*

$$U' \cong V'/U^0$$

Proof. Consider $\Phi : V' \to U'$ given by $f \mapsto f|_U$

Then Φ is linear as for all $f, g \in V'$, $\lambda \in \mathbb{F}$ we have

$$\Phi(f + \lambda g) = (f + \lambda g)|_U = f|_U + \lambda g|_U = \Phi(f) + \lambda \Phi(g)$$

Furthermore, we have

$$\begin{aligned} f \in \ker \Phi &\iff f|_U = 0 \\ &\iff f \in U^0 \end{aligned}$$

Hence $\ker \Phi = U^0$ and so we can apply the first isomorphism theorem to get

$$\tilde{\Phi} : V'/U^0 \xrightarrow{\sim} \text{Im } \Phi \subseteq U'$$

Now, since V is finite dimensional, any basis $\mathcal{B}_U = \{e_1, \dots e_k\}$ of U can be extended to a basis $\mathcal{B}_V = \{e_1, \dots, e_n\}$. Then any $g \in U'$ is the image under Φ of $\tilde{g} \in V'$, defined by

$$\tilde{g}(e_i) = \begin{cases} g(e_i) & i = 1, \dots, m \\ 0 & i = m+1, \dots, n \end{cases}$$

Thus Im $\Phi = U'$ and so we are done.

Moreover, $\{e'_{m+1}, \dots, e'_n\}$ is a basis for U^0 and $\{e'_1 + U^0, \dots, e'_m + U^0\}$ is a basis for V'/U^0 such that $\tilde{\Phi} : U' \to V'/U^0$ defined by $e'_i \mapsto e'_i + U^0$ is an isomorphism as required.

$\square$

Remark. This result is also true in infinite dimensional case

Definition 7.5. Let $T : V \to W$ be a linear transformation. We define the **dual map** by:

$$\begin{aligned} T' : W' &\to V' \\ f &\mapsto f \circ T \end{aligned}$$

Remark. Since $f \circ T$ is linear T' is well defined

Proposition 7.8. T' is linear

Proof. Let $f, g \in W'$, $\lambda \in \mathbb{F}$ and $v \in V$

$$\begin{aligned} T'(f + \lambda g)(v) &= ((f + \lambda g) \circ T)(v) \\ &= (f + \lambda g)(Tv) \\ &= f(Tv) + \lambda g(Tv) \\ &= (f \circ T)(v) + \lambda (g \circ T)(v) \\ &= T'(f)(v) + \lambda T'(g)(v) \end{aligned}$$

□

Proposition 7.9. The map $\hom(V, W) \to \hom(W', V')$ given by $T \mapsto T'$ is linear

Proof. Let $S, T \in \hom(V, W), \lambda \in \mathbb{F}, f \in W'$ and $v \in V$, then

$$\begin{aligned} (T + \lambda S)'(f)(v) &= (f \circ (T + \lambda S))(v) \\ &= f((T + \lambda S)(v)) \\ &= f(T(v)) + \lambda f(S(v)) \\ &= T'(f)(v) + \lambda S'(f)(v) \\ &= (T' + \lambda S')(f)(v) \end{aligned}$$

□

Theorem 7.10. *Suppose V and W are finite dimensional, then $T \mapsto T'$ defines a natural isomorphism between* $\hom(V, W)$ *and* $\hom(W', V')$

Proof. Assume $T' = 0$

But now, $T'(f)(v) = 0$ for all $f \in W', v \in V$ if and only if $f(T(v)) = 0$ for all f and v.
Suppose $T(v) \neq 0$, then we can extend $T(v)$ to a basis $\mathcal{B}_W$ of W.
Then the corresponding element of the dual basis $\mathcal{B}'_W$ satisfies $(T(v))'(T(v)) = 1$ contradicting that $f(T(v)) = 0$ for all f. Thus $T(v) = 0$ for all $v \in V$, or $T \equiv 0$, and hence $T \mapsto T'$ is injective.

As

$$\begin{aligned} \dim \hom(V, W) &= \dim V \cdot \dim W \\ &= \dim W' \cdot \dim V' \\ &= \dim \hom(W', V') \end{aligned}$$

We have that $T \mapsto T'$ is also surjective and hence is the isomorphism required. □

Theorem 7.11. *Let V, W be finite dimensional vector spaces*
Let $\mathcal{B}_V, \mathcal{B}_W$ be bases of V and W respectively
Let $\mathcal{B}'_V, B'_W$ be the corresponding dual bases of V' and W'
Then, for any linear map $T : V \to W$

$$\left({}_{\mathcal{B}_W}[T]_{\mathcal{B}_V}\right)^t = {}_{\mathcal{B}'_V}[T']_{\mathcal{B}'_W}$$

where A^t denotes the transpose of A.

Proof. Let $\mathcal{B}_V = \{e_1, \dots, e_n\}$ and $\mathcal{B}_W = \{x_1, \dots, x_m\}$

Put ${}_{\mathcal{B}_W}[T]_{\mathcal{B}_V} = A = (a_{ij})_{m\times n}$

Then $T(e_j) = \sum_{i=1}^{m} a_{ij}x_i$ and $x'_i(T(e_j)) = a_{ij}$

Put ${}_{\mathcal{B}'_v}[T']_{\mathcal{B}'_W} = B = (b_{ij})_{n\times m}$

Then $T'(x'_i) = \sum_{j=1}^{n} b_{ji}e'_j$ and $T'(x'_i)(e_j) = b_{ji}$

Hence $b_{ji} = T'(x'_i)(e_j) = x'_i(T(e_j)) = a_{ij}$ thus $B = A^t$

□

Remark. The above theorem is the isomorphism from $M_{n\times m}(\mathbb{F}) \to M_{m\times n}(\mathbb{F})$ given by $A \mapsto A^t$

Chapter 8

Bilinear Forms and Inner Products

Definition 8.1. Let V be a vector space over $\mathbb{F}$

A **bilinear form** on V is a map $\mathcal{F} : V \times V \to \mathbb{F}$ such that for all $u, v, w \in V, \lambda \in \mathbb{F}$

(1) $\mathcal{F}(u + v, w) = \mathcal{F}(u, w) + \mathcal{F}(v, w)$

(2) $\mathcal{F}(u, v + w) = \mathcal{F}(u, v) + \mathcal{F}(u, w)$

(3) $\mathcal{F}(\lambda v, w) = \lambda\mathcal{F}(v, w) = \mathcal{F}(v, \lambda w)$

$\mathcal{F}$ is called **symmetrical** if $F(v, w) = F(w, v)$ for all $v, w \in V$

$\mathcal{F}$ is called **non-degenerate** if $(\forall w \in V : F(v, w) = 0) \Rightarrow v = 0$

$\mathcal{F}$ is called **positive definite** if for all $v \in V : v \neq 0 \Rightarrow F(v, v) > 0$

Remark. Positive Definite $\Rightarrow$ Non-Degenerate: $\mathcal{F}(v, v) = 0 \Rightarrow v = 0$

Example 8.1. Minkowski Space: $V = \mathbb{R}^3 \times \mathbb{R}$

$\mathcal{F}\left[((x, y, z), t), ((x', y', z'), t')\right] = xx' + yy' + zz' - c^2tt'$

$\mathcal{F}$ is bilinear, symmetric, non-degenerate, NOT positive definite

Example 8.2. $V = \mathbb{R}^3$

$\mathcal{F}((x, y, z), (x', y', z')) = xx' + yy' + zz'$

$\mathcal{F}$ is bilinear, symmetric and positive definite

Example 8.3. $V = \mathcal{C}([0, 1])$

$\mathcal{F}(f, g) = \int_0^1 f(x)g(x)dx$

$\mathcal{F}$ is bilinear, symmetric and positive definite

Definition 8.2. Let V be a vector space over $\mathbb{C}$

A **sesquilinear form** on V is a map $\mathcal{F} : V \times V \to \mathbb{C}$ such that for all $u, v, w \in V$, $\lambda \in \mathbb{C}$

(1) $\mathcal{F}(u + v, w) = \mathcal{F}(u, w) + \mathcal{F}(v, w)$

(2) $\mathcal{F}(u, v + w) = \mathcal{F}(u, v) + \mathcal{F}(u, w)$

(3) $\mathcal{F}(\bar{\lambda} v, w) = \lambda \mathcal{F}(v, w) = \mathcal{F}(v, \lambda w)$

$\mathcal{F}$ is **conjugate symmetric** if $\mathcal{F}(v, w) = \mathcal{F}(\bar{w}, v)$ for all $v, w \in V$

$\mathcal{F}$ is **non-degenerate** if $(\forall w \in V : \mathcal{F}(v, w) = 0) \Rightarrow v = 0$

$\mathcal{F}$ is **positive definite** if $\mathcal{F}(v, v) \in \mathbb{R}, \mathcal{F}(v, v) > 0$ for all $v \in V$

Example 8.4. $V = \mathbb{C}^n$

$\mathcal{F}(v, w) = \bar{v}^t A w$ for some $A \in M_{n \times n}(\mathbb{C})$

- $\mathcal{F}$ is sesquilinear, conjugate symmetric iff $A = \bar{A}^t$

  Observe that $\mathcal{F}(e_i, e_j) = \bar{e_i}^t A e_j = a_{ij}$ and $\overline{\mathcal{F}(e_j, e_i)} = \overline{\bar{e_j}^t A e_i} = \overline{a_{ji}}$.
  Thus $\mathcal{F}$ is conjugate symmetric iff $a_{ij} = \overline{a_{ji}}$ for all i, j, that is $A = \bar{A}^t$

- $\mathcal{F}$ is non-degenerate $\iff$ A is non-singular

$$\begin{aligned} A \text{ singular } &\iff \exists w \neq 0 \in V : Aw = 0 \\ &\iff \exists w \neq 0 \in V : \forall v \in V : \bar{v}^t A w = 0 \\ &\iff \mathcal{F} \text{ degenerate} \end{aligned}$$

Definition 8.3. A real (complex) vector space V with a positive definite, symmetric (conjugate symmetric), bilinear (sesquilinear) form $\mathcal{F} = \langle \, , \, \rangle$ is called an **inner product space**

Definition 8.4. $\{w_1, \dots, w_n\}$ are mutually **orthogonal** if $\langle w_i, w_j \rangle = 0$ for all $i \neq j$

Definition 8.5. $\{w_1, \dots, w_n\}$ are mutually **orthonormal** if $\langle w_i, w_j \rangle = \delta_{ij}$ for all i, j

Proposition 8.1. Suppose V is an inner product space over $\mathbb{R}$ or $\mathbb{C}$ and $\{w_1, \dots, w_n\}$ are orthogonal with $w_i \neq 0$ for all i. Then $\{w_1, \dots, w_n\}$ is linearly independent.

Proof. Assume $\sum_i \lambda_i w_i = 0$ for some $\lambda_i \in \mathbb{F}$

$$\begin{aligned} &\Rightarrow \left\langle w_j, \sum_i \lambda_i w_i \right\rangle = 0 && \forall j \\ &\Rightarrow \sum_i \lambda_i \langle w_j, w_i \rangle = 0 && \forall j \\ &\Rightarrow \lambda_j \langle w_j, w_j \rangle = 0 && \forall j \\ &\Rightarrow \lambda_j = 0 && \forall j \end{aligned}$$

Theorem 8.2. Gram-Schmidt Process

Let $\{v_1, \dots, v_n\}$ be a basis of the inner product space V

Put $w_1 = v_1$

$$w_2 = v_2 - \frac{\langle w_1, v_2 \rangle}{\langle w_1, w_1 \rangle} w_1$$

$$\vdots$$

$$w_k = v_k - \sum_{i=1}^{k-1} \frac{\langle w_i, v_k \rangle}{\langle w_i, w_i \rangle} w_i \qquad (*)$$

Clearly $\langle w_i \rangle = \langle v_i \rangle$

Editor's note: need to show that $w_k \neq 0$

Assuming $\langle w_1, \dots, w_{k-1} \rangle = \langle v_1, \dots, v_{k-1} \rangle$ *we have by* $(*)$

$$\langle w_1, \dots, w_k \rangle = \langle w_1, \dots, w_{k-1}, v_k \rangle$$

Then by the inductive hypothesis we have

$$\langle w_1, \dots, w_k \rangle = \langle w_1, \dots, w_{k-1}, v_k \rangle = \langle v_1, \dots, v_k \rangle$$

Now, if we assume that $\{w_1, \dots, w_{k-1}\}$ *is orthogonal then for* $j < k$ *we have from* $(*)$ *that*

$$\begin{aligned} \langle w_k, w_j \rangle &= \langle v_k, w_j \rangle - \sum_{i=1}^{k-1} \frac{\langle w_i, v_k \rangle}{\langle w_i, w_i \rangle} \langle w_i, w_j \rangle \\ &= \langle v_k, w_j \rangle - \frac{\langle w_j, v_k \rangle}{\langle w_j, w_j \rangle} \langle w_j, w_j \rangle \\ &= 0 \end{aligned}$$

Hence by induction we have that $\{w_1, \dots, w_n\}$ *is an orthogonal basis for* V

Now we put $u_i = \dfrac{w_i}{||w_i||}$ *where* $||w_i|| = \sqrt{\langle w_i, w_i \rangle} \in \mathbb{R}$ *for each* i

Then $\{u_1, \dots, u_n\}$ *is an* ***orthonormal basis*** *of* V

Remark. The change of basis matrix from $\{v_i\}$ to $\{u_i\}$ is upper-triangular with positive entries on the diagonal.

Theorem 8.3. Bessel's Inequality

Let $\dim V < \infty$ *and* $\{u_1, \dots, u_n\}$ *be an orthonormal basis*

Then $\forall v \in V$:

$$||v||^2 \geq \sum_{i=1}^{k} \left| \langle v, u_i \rangle \right|^2$$

With equality holding iff $k = \dim V$

8.1 Duals of Inner Product Spaces

Let V be an inner product space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$)

Then for all $v \in V$

$$\begin{aligned}\langle v, \_ \rangle : V &\to \mathbb{F} \\ w &\mapsto \langle v, w \rangle\end{aligned}$$

Is a linear functional on V as $\langle\ ,\ \rangle$ is linear in the second co-ordinate

Theorem 8.4. *For $\mathbb{F} = \mathbb{R}$, the map $v \mapsto \langle v, \_ \rangle$ is a **natural** injective linear map $\Phi : V \to V'$ which is an isomorphism when* $\dim V < \infty$

Proof. Φ is linear as for all $v, w \in V, \lambda \in \mathbb{R}$

$$\langle v + \lambda w, \_ \rangle = \langle v, \_ \rangle + \lambda \langle w, \_ \rangle$$

Since $\langle\ ,\ \rangle$ is non-degenerate we have that $\langle v, \_ \rangle = \langle \_, v \rangle$ is the zero function iff $v = 0$
Hence Φ is injective.

If $\dim V < \infty$ then we have $\dim V = \dim V'$, therefore $\mathrm{Im}\ \Phi = V'$ and so Φ is an isomorphism in the finite dimensional case

□

Remark. For $\mathbb{F} = \mathbb{C}$, Φ defines a conjugate linear map: $\Phi(\lambda v) = \overline{\lambda}\Phi(v)$

Definition 8.6. Let $U \leq V$ be a finite dimensional subspace of V

The **orthogonal complement** of U is defined as

$$U^{\perp} := \{v \in V : \langle u, v \rangle = 0 \text{ for all } u \in U\}$$

Proposition 8.5. $U^{\perp}$ is a linear subspace

Proof. For all $v, w \in U^{\perp}, \lambda \in \mathbb{F}$ and for all $u \in U$:

$$\langle u, v + \lambda w \rangle = \langle u, v \rangle + \lambda \langle u, w \rangle = 0$$

Thus $v + \lambda w \in U^{\perp}$

□

Proposition 8.6. Let U, W be finite dimensional subspaces of an inner product space V

(1) $U \cap U^\perp = \{0\}$

Proof. If $u \in U \cap U^\perp$ then $\langle u, u \rangle = 0$. By positive definiteness of $\langle\ ,\ \rangle$ we have $u = 0$

□

(2) $\dim V < \infty \Rightarrow U \oplus U^\perp = V$

Proof. Take $\{e_1, \dots, e_k\}$ an orthonormal basis of U and let $\{e_1, \dots, e_k, \dots, e_n\}$ be an orthonormal basis for V.

Now, assume $v = \sum_i a_i e_i \in U^\perp$. Then,

$$\langle e_j, v \rangle = \langle e_j, \textstyle\sum_i a_i e_i \rangle = \langle e_j, a_j e_j \rangle = a_j$$

By definition of $U^\perp$ we have $\langle e_j, v \rangle = a_j = 0$ for $j = 1, \dots, k$, thus $v \in \langle e_{k+1}, \dots, e_n \rangle$

Vice versa, if $v \in \langle e_{k+1}, \dots, e_n \rangle$ then for all $u \in U$ clearly $\langle u, v \rangle = 0$ that is $v \in U^\perp$

Thus $U^\perp = \langle e_{k+1}, \dots, e_n \rangle$ and hence $V = U \oplus U^\perp$

□

(3) $(U + W)^\perp = U^\perp \cap W^\perp$

Proof. Take $v \in (U+W)^\perp$. Since U and W are subspaces of V they both contain 0. Then for all $u \in U$ and all $w \in W$ we have $u + 0 = u, 0 + w = w \in U + W$ and hence $\langle v, u \rangle = 0 = \langle v, w \rangle$ and so $v \in U^\perp \cap W^\perp$ as required.

Conversely, take $v \in U^\perp \cap W^\perp$. Then for all $\omega \in U + W$ we have $\omega = u + w$ for some $u \in U, w \in W$ and hence $\langle v, \omega \rangle = \langle v, u + w \rangle = \langle v, u \rangle + \langle v, w \rangle = 0 + 0 = 0$ and so $v \in (U + W)^\perp$.

□

(4) $(U \cap W)^\perp \geq U^\perp + W^\perp$ with equality if $\dim V < \infty$

Proof. Take $v \in U^\perp + W^\perp$, then there exist $u \in U^\perp, w \in W^\perp$ such that $v = u + w$. Now, for $\omega \in U \cap W$ we have $\omega \in U$ and $\omega \in W$, therefore

$$\langle v, \omega \rangle = \langle u + w, \omega \rangle = \langle u, \omega \rangle + \langle w, \omega \rangle = 0 + 0 = 0$$

Hence $v \in (U \cap W)^\perp$ as required.

Further, if $\dim V < \infty$ then we can apply the dimension formula to obtain

$$\begin{aligned}
\dim (U \cap W)^\perp &= \dim V - \dim (U \cap W) \\
&= \dim V - \dim U + \dim V - \dim W \\
&= \dim U^\perp + \dim W^\perp \\
&= \dim (U^\perp + W^\perp)
\end{aligned}$$

Hence, by dimensionality, we have $(U \cap W)^\perp = U^\perp + W^\perp$ when $\dim V < \infty$

□

(5) $U \leq (U^\perp)^\perp$ with equality if $\dim V < \infty$

Proof. Let $u \in U$

Then for all $w \in U^\perp : \langle u, w \rangle = \overline{\langle w, u \rangle} = 0$ and hence $\langle w, u \rangle = 0$ and thus $u \in (U^\perp)^\perp$

If $\dim V < \infty$ then

$$\dim(U^\perp)^\perp = \dim V - \dim U^\perp = \dim V - \dim V + \dim U = \dim U$$

Thus, by dimensionality, equality holds when $\dim V < \infty$

$\square$

Example 8.5. Let $U, W \leq \mathbb{R}^3$ be defined

$$U := \left\{ \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} \right\}_{x \in \mathbb{R}} \qquad W := \left\{ \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \right\}_{y \in \mathbb{R}}$$

$U^\perp = yz\text{-plane} = \left\{ \begin{pmatrix} 0 \\ y \\ z \end{pmatrix} \right\}_{y,z \in \mathbb{R}}$

$W^\perp = xz\text{-plane} = \left\{ \begin{pmatrix} x \\ 0 \\ z \end{pmatrix} \right\}_{x,z \in \mathbb{R}}$

$U^\perp \cap W^\perp = z\text{-axis} = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \right\}_{z \in \mathbb{R}}$

$(U + W)^\perp = (xy\text{-plane})^\perp = z\text{-axis} = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \right\}_{z \in \mathbb{R}}$

$(U \cap W)^\perp = \{0\}^\perp = \mathbb{R}^3$

$U^\perp + W^\perp = \{yz\text{-plane} + xz\text{-plane}\} = \mathbb{R}^3$

Proposition 8.7. Let $\dim V < \infty$ and $\mathbb{F} = \mathbb{R}$

Then, under the isomorphism $\Phi : V \to V'$ given by $v \mapsto \langle v, \_ \rangle$

$$U^\perp \cong U^0$$

Proof. Let $v \in U^\perp$, then for all $u \in U$ we have

$$\langle v, u \rangle = 0 = \langle u, v \rangle$$

Thus $\Phi(v) = \langle v, \_ \rangle \in U^0$

Moreover, $\dim U^\perp = \dim V - \dim U = \dim U^0$, giving $U^\perp \cong U^0$ as required.

$\square$

Example 8.6. Let V be the vector space of real polynomials with degree ≤ 2, so $V = \langle 1, t, t^2 \rangle$

Define $\langle f, g\rangle = f(1)g(1) + f(2)g(2) + f(3)g(3)$

Then $\langle\ ,\ \rangle$ is bilinear, symmetric and positive definite:

$$\begin{aligned} \langle f, f\rangle = f(1)^2 + f(2)^2 + f(3)^2 = 0 &\iff f(1) = f(2) = f(3) = 0 \\ &\Rightarrow f \text{ has 3 roots} \end{aligned}$$

Since f has degree ≤ 2 it cannot have 3 roots, thus $f \equiv 0$

Let $U = \langle 1, t\rangle$ and take $f \in U, g \in U^\perp$ such that $f + g = t^2$, then for orthonormal basis $\{u_1, u_2\}$ of U

$$g = t^2 - \Big(\langle t^2, u_1\rangle u_1 + \langle t^2, u_2\rangle u_2 \Big)$$

Put

Let $u_1 = \dfrac{1}{\sqrt{\langle 1, 1\rangle}} = \dfrac{1}{\sqrt{3}}$

$$u_2 = \frac{t - \langle t, u_1\rangle u_1}{||\uparrow||} = \frac{t - \frac{1}{\sqrt{3}}\left(\frac{1}{\sqrt{3}} + \frac{2}{\sqrt{3}} + \frac{3}{\sqrt{3}}\right)}{||\uparrow||} = \frac{t-2}{||t-2||} = \frac{t-2}{\sqrt{2}}$$

Then,

$$f = \left\langle t^2, \frac{1}{\sqrt{3}} \right\rangle \frac{1}{\sqrt{3}} + \left\langle t^2, \frac{t-2}{\sqrt{2}} \right\rangle \frac{t-2}{\sqrt{2}} = \frac{14}{3} + 4(t-2) = 4t - \frac{10}{3}$$

8.2 Adjoint Maps

Let V be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$

Definition 8.7. A linear map $T : V \to V$ has an **adjoint map** $T^* : V \to V$ if for all $v, w \in V$

$$\langle v, Tw\rangle = \langle T^* v, w\rangle$$

Lemma 8.8. If T^* exists, then it is unique

Proof. Suppose T' is another adjoint map, then for all $v, w \in V$

$$\begin{aligned} \langle T^* v - T' v, w\rangle &= \langle T^* v, w\rangle - \langle T' v, w\rangle \\ &= \langle v, Tw\rangle - \langle v, Tw\rangle \\ &= 0 \end{aligned}$$

Thus $T^* v - T' v = 0$ for all $v \in V$, that is $T^* \equiv T'$

□

Theorem 8.9. *Let $T : V \to V$ be linear and $\dim V < \infty$, then T^* exists and is also linear*

Proof. Fix $v \in V$ and consider the map $\phi : V \to \mathbb{F}$ by $\phi(w) = \langle v, Tw \rangle$

ϕ is a linear functional as T is linear and $\langle\ ,\ \rangle$ is linear in its second coordinate

As $\dim V < \infty$, $\Phi : V \to V'$ given by $v \mapsto \langle v, \_ \rangle$ is a linear isomorphism when $\mathbb{F} = \mathbb{R}$ and is a conjugate linear bijection when $\mathbb{F} = \mathbb{C}$

Then, $\exists u \in V$ such that $\phi = \langle u, \_ \rangle$ - we define $\langle T^*v, \_ \rangle = \langle u, \_ \rangle$

For all $v_1, v_2, w \in V, \lambda \in \mathbb{F}$:

$$\begin{aligned}
\langle T^*(v_1 + \lambda v_2), w \rangle &= \langle v_1 + \lambda v_2, Tw \rangle \\
&= \langle v_1, Tw \rangle + \bar{\lambda} \langle v_2, Tw \rangle \\
&= \langle T^*v_1, w \rangle + \bar{\lambda} \langle T^*v_2, w \rangle \\
&= \langle T^*v_1 + \lambda T^*v_2, w \rangle
\end{aligned}$$

□

Proposition 8.10. Let $T : V \to V$ be linear and $\mathcal{B} = \{e_1, \ldots, e_n\}$ be an orthonormal basis for V:

$$_{\mathcal{B}}[T^*]_{\mathcal{B}} = {}_{\mathcal{B}}\overline{[T]}_{\mathcal{B}}^{\,t}$$

Proof. Let $A = {}_{\mathcal{B}}[T]_{\mathcal{B}}$ and $B = {}_{\mathcal{B}}[T^*]_{\mathcal{B}}$
Then,

$$\begin{aligned}
b_{ij} &= \langle e_i, T^*e_j \rangle \\
&= \overline{\langle T^*e_j, e_i \rangle} \\
&= \overline{\langle e_j, Te_i \rangle} \\
&= \overline{a_{ji}}
\end{aligned}$$

And hence $B = \overline{A}^t$ as required

□

Remark.

(1) The theorem is false if V is not finite dimensional

(2) The proposition is false if $\mathcal{B}$ is not orthonormal

(3) For $\mathbb{F} = \mathbb{R}$, under the linear isomorphism $\phi : V \to V'$ by $\phi(v) = \langle v, \_ \rangle$, T^* is identified with T':

An orthonormal basis $\mathcal{B}$ is taken to its dual basis and hence

$$_{\mathcal{B}}[T']_{\mathcal{B}} = ({}_{\mathcal{B}}[T]_{\mathcal{B}})^t = {}_{\mathcal{B}}[T^*]_{\mathcal{B}}$$

Proposition 8.11. Let $S, T : V \to V$ be linear, $\lambda \in \mathbb{F}$ and $\dim V < \infty$, then:

(1) $(S+T)^* = S^* + T^*$

(2) $(\lambda T)^* = \bar{\lambda} T^*$

(3) $(ST)^* = T^* S^*$

(4) $(T^*)^* = T$

(5) If m_T is the minimal polynomial of T then $m_{T^*} = \overline{m_T}$

Proof. Follow straightforwardly from Proposition 8.10 and standard properties of matrices

Definition 8.8. A linear map $T : V \to V$ is **self-adjoint** if $T^* = T$

Lemma 8.12. If λ is an eigenvalue of a self-adjoint linear transformation, then $\lambda \in \mathbb{R}$

Proof. Assume $w \neq 0, Tw = \lambda w$, then

$$\begin{aligned}
\lambda \langle w, w \rangle &= \langle w, \lambda w \rangle && \text{(linearity in second coordinate)} \\
&= \langle w, Tw \rangle && (Tw = \lambda w) \\
&= \langle T^* w, w \rangle && \text{(definition of adjoint)} \\
&= \langle Tw, w \rangle && (T \text{ self-adjoint}) \\
&= \langle \lambda w, w \rangle && (Tw = \lambda w) \\
&= \bar{\lambda} \langle w, w \rangle && \text{(conjugate linearity in first coordinate)}
\end{aligned}$$

Since $\langle w, w \rangle \neq 0$ we must have $\lambda = \bar{\lambda}$ and hence $\lambda \in \mathbb{R}$

□

Lemma 8.13. If $T : V \to V$ is self-adjoint and $U \leq V$ is T-invariant, then $U^\perp$ is also T-invariant

Proof. Let $w \in U^\perp, u \in U$, then

$$\begin{aligned}
\langle u, Tw \rangle &= \langle T^* u, w \rangle && \text{(definition of adjoint)} \\
&= \langle Tu, w \rangle && (T \text{ self-adjoint}) \\
&= 0 && (Tu \in U, w \in U^\perp)
\end{aligned}$$

Hence $Tw \in U^\perp$, and thus $U^\perp$ is T-invariant

□

Theorem 8.14. *Suppose $T : V \to V$ is self-adjoint over a complex vector space with $\dim V < \infty$, then there exists an orthonormal basis of eigenvectors*

Proof. By Lemma 8.12 there exists $\lambda \in \mathbb{R}$ and $w \neq 0 \in V$ with $Tw = \lambda w$.

Clearly $\langle w \rangle$ is T-invariant, and hence, by Lemma 8.13., $\langle w \rangle^\perp$ is also T-invariant

Let $e_1 = \frac{w}{||w||}$, then $\{e_1\}$ is an orthonormal basis for $\langle w \rangle$.
This is the base case for an induction on the dimension of the subspace.

By inductive hypothesis, there is an orthonormal basis of eigenvectors, $\{e_2, \dots, e_n\}$, for $T|_{\langle w \rangle^\perp}$

Now, since $V = U \oplus U^\perp$, $\{e_1, \dots, e_n\}$ is an orthonormal basis for V of eigenvectors of T.

□

Corollary 8.15. Any $n \times n$ matric A satisfying $A = \overline{A}^t$ is diagonalizable by an orthonormal change of basis

Proof. By theorem, there exists $P = \begin{pmatrix} | & & | \\ e_1 & \cdots & e_n \\ | & & | \end{pmatrix}$ with $\{e_1, \dots, e_n\}$ an orthonormal basis for $\mathbb{F}^n$ such that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

8.3 Orthogonal and Unitary Transformations

Definition 8.9. A is called **orthogonal** when $A^{-1} = \bar{A}^t$ and $\mathbb{F} = \mathbb{R}$

Definition 8.10. A is called **unitary** when $A^{-1} = \bar{A}^t$ and $\mathbb{F} = \mathbb{C}$

Example 8.7. Let $\{e_1, \dots, e_n\}$ be an orthonormal basis for $\mathbb{F}^n$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) under the normal dot product

Put $A = \begin{pmatrix} | & & | \\ e_1 & \cdots & e_n \\ | & & | \end{pmatrix}$, then $\bar{A}^t A = I$ as $\bar{A}^t = \begin{pmatrix} — & \bar{e_1} & — \\ & \vdots & \\ — & \bar{e_2} & — \end{pmatrix}$

Hence $A^{-1} = \bar{A}^t$

Definition 8.11. Let V be a finite dimensional vector space with inner product and $T : V \to V$ be a linear transformation satisfying $T^*T = I = TT^*$

Then T is **orthogonal** if $\mathbb{F} = \mathbb{R}$ or **unitary** if $\mathbb{F} = \mathbb{C}$

Theorem 8.16. *The following are equivalent:*

(1) $T^* = T^{-1}$

(2) *T preserves inner products:* $\langle v, w\rangle = \langle Tv, Tw\rangle \quad \forall v, w \in V$

(3) *T preserves lengths:* $||v|| = ||Tv|| \quad \forall v \in V$

Proof.

(1) $\Rightarrow$ (2) $\langle v, w\rangle = \langle v, T^*Tw\rangle = \langle Tv, Tw\rangle$

(2) $\Rightarrow$ (3) $||v||^2 = \langle v, v\rangle = \langle Tv, Tv\rangle = ||Tv||^2$

(2) $\Rightarrow$ (1) $\langle T^*Tv - v, w\rangle = \langle T^*Tv, w\rangle - \langle v, w\rangle = \langle Tv, Tw\rangle - \langle v, w\rangle = 0$
By non-degeneracy of $\langle\ ,\ \rangle$: $T^*T = I$

(3) $\Rightarrow$ (2) See Proposition 8.17.

□

Remark. Orthogonal/Unitary linear transformations are isometries:

$$d(v, w) = ||v - w|| = ||T(v - w)|| = ||Tv - Tw|| = d(Tv, Tw)$$

Remark. Let $\mathcal{B}$ be an orthonormal basis for V and T be an orthogonal/unitary linear transformation. Then ${}_{\mathcal{B}}[T]_{\mathcal{B}}$ is an orthogonal/unitary matrix - the columns (and rows) form an orthonormal basis.

Proposition 8.17. The length function uniquely determines the inner product:

$$\langle v, v\rangle_1 = \langle v, v\rangle_2 \ \ \forall v \in V \iff \langle v, w\rangle_1 = \langle v, w\rangle_2 \ \ \forall v, w \in V$$

Proof. ($\Longleftarrow$) is clear, remains to show ($\Longrightarrow$), we have

$$\langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \overline{\langle v, w\rangle} + \langle w, w\rangle$$

Hence, when $\mathbb{F} = \mathbb{R}$

$$\langle v, w\rangle = \frac{1}{2}\Big(||v + w||^2 - ||v||^2 - ||w||^2\Big)$$

Alternatively, when $\mathbb{F} = \mathbb{C}$ we also consider

$$\langle v + iw, v + iw\rangle = \langle v, v\rangle + i\,\langle v, w\rangle - i\overline{\langle v, w\rangle} + \langle w, w\rangle$$

To obtain:

$$\Re\,\langle v, w\rangle = \frac{1}{2}\Big(||v + w||^2 - ||v||^2 - ||w||^2\Big)$$
$$\Im\,\langle v, w\rangle = \frac{1}{2}\Big(||v + iw||^2 - ||v||^2 - ||w||^2\Big)$$

□

Definition 8.12. The following are groups:

$$O_n = \{A \in M_{n\times n}(\mathbb{R}) : A^{-1} = A^t\} \quad \text{(Orthogonal)}$$
$$SO_n = \{A \in M_{n\times n}(\mathbb{R}) : A^{-1} = A^t, \det A = 1\} \quad \text{(Special Orthogonal)}$$
$$U_n = \{A \in M_{n\times n}(\mathbb{C}) : A^{-1} = \bar{A}^t\} \quad \text{(Unitary)}$$
$$SU_n = \{A \in M_{n\times n}(\mathbb{C}) : A^{-1} = \bar{A}^t, \det A = 1\} \quad \text{(Special Unitary)}$$

Lemma 8.18. Let $T : V \to V$ be an orthogonal/unitary linear map on a finite dimensional inner product space V. If λ is an eigenvalue of T, then $|\lambda| = 1$.

Proof. Take $v \neq 0$, an eigenvector for λ, then

$$||v|| = ||Tv|| = ||\lambda v|| = |\lambda| \cdot ||v|| \Rightarrow |\lambda| = 1$$

□

Corollary 8.19. If A is an orthogonal/unitary matrix then:

$\det A = \pm 1$ for $\mathbb{F} = \mathbb{R}$

$\det A \in S^1$ for $\mathbb{F} = \mathbb{C}$

Proof. Working over $\mathbb{C}$, A can be upper-triangulized with eigenvalues on the diagonal (with repetitions), that is there exists P such that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * & * \\ & \ddots & * \\ 0 & & \lambda_n \end{pmatrix}$$

Then, $\det A = \det P^{-1}AP = \lambda_1 \cdots \lambda_n$

Now, by Lemma 8.18, $|\lambda_i| = 1$ for all i, hence $|\det A| = 1$

So $\det A = \pm 1$ for $\mathbb{F} = \mathbb{R}$ or $\det A \in S^1$ for $\mathbb{F} = \mathbb{C}$

□

Remark.

$$\det : O_n \to \{\pm 1\} \cong \mathbb{Z}_2 \quad (\ker \det = SO_n)$$
$$\det : U_n \to S^1 \quad (\ker \det = SU_n)$$

Lemma 8.20. Let $T : V \to V$ be a linear map on a finite dimensional inner product space V and assume $T^* = T^{-1}$. Then if $U \leq V$ is T-invariant, then $U^\perp$ is also.

Proof. Let $u \in U, w \in U^\perp$ and let $Tu = u' \in U$. Then,

$$0 = \langle u, w \rangle = \langle Tu, Tw \rangle = \left\langle u', Tw \right\rangle$$

Now, as T is invertible it must be a bijection and thus $T(U) \subseteq U \implies T(U) = U$.

Hence, $Tw \in U^\perp$ as required.

□

Theorem 8.21. *Let $T : V \to V$ be a unitary linear transformation on a finite dimensional inner product space. Then there exists an orthonormal basis of eigenvectors.*

Proof. There exists $v \neq 0$ such that $Tv = \lambda v$ for some eigenvalue λ

Then $\langle v \rangle$ is T-invariant and hence, by Lemma 8.20, so is $\langle v \rangle^\perp$

$\dim \langle v \rangle^\perp < \dim V$ thus by induction $\langle v \rangle^\perp$ has an orthonormal basis of eigenvectors, $\{e_2, \ldots, e_n\}$

Setting $e_1 = \dfrac{v}{||v||}$ we obtain $\{e_1, \ldots, e_n\}$, an orthonormal basis of eigenvectors.

□

Corollary 8.22. If $A \in U_n$ then there exists $P \in U_n$ such that $P^{-1}AP$ is diagonal

Remark. If $A \in O_n$, then $A \in U_n$ **but** A may not be diagonalizable over $\mathbb{R}$

Example 8.8. Let $A \in O_2$ and $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$.
Then $A^t = A^{-1}$ and hence:

$$a^2 + c^2 = 1 = b^2 + c^2 \qquad ab + cd = 0 \qquad ad - bc = \pm 1$$

Solving these gives:

$$\underset{\text{ROTATION}}{A = R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}}_{\det = 1} \text{ or } \underset{\text{REFLECTION}}{A = S_\theta = \begin{pmatrix} \sin\theta & \cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix}}_{\det = -1}$$

Further,

$$\mathcal{X}_{S_\theta} = x^2 - \sin^2\theta - \cos^2\theta = x^2 - 1 = (x+1)(x-1)$$

$\Rightarrow S_\theta$ is diagonalizable (the eigenvector for 1 gives the line of reflection)

$$\mathcal{X}_{R_\theta} = x^2 - 2x\cos\theta + \cos^2\theta + \sin^2\theta = x^2 - 2x\cos\theta + 1 = (x - \lambda)(x - \bar{\lambda}) \qquad (\lambda = e^{2\pi i\theta})$$

$\Rightarrow R_\theta$ has real eigenvalues $\iff \theta = 0, \pi$

R_θ is **not** diagonalizable over $\mathbb{R}$ for $\theta \neq 0, \pi$

Theorem 8.23. *Let $T : V \to V$ be an orthogonal map over a finite dimensional, real inner product space V. Then there exists an orthonormal basis $\mathcal{B}$ such that:*

$$
{}_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} I & & & & \\ & -I & & & \\ & & R_{\theta_i} & & \\ & & & \ddots & \\ & & & & R_{\theta_k} \end{pmatrix}
$$

Proof. Let $S = T + T^*$. Then $S^* = (T + T^*)^* = T^* + T = S$, thus S is self-adjoint, and, by Theorem 8.21, there exists an orthonormal basis of eigenvectors. Accordingly we can write $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_n}$ with $\lambda_i \neq \lambda_j$ for $i \neq j$ and where V_{λ_i} is the λ_i-eigenspace of V.

Now suppose $v \in V_\lambda$, then:

$$
S(Tv) = (T + T^*)(Tv) = T(T + T^*)(v) = T(Sv) = T(\lambda v) = \lambda Tv
$$

Thus Tv is a λ eigenvector of S, that is $Tv \in V_\lambda$, and hence V_λ is T-invariant for each λ. Accordingly, we may restrict the problem to $T|_{V_\lambda}$

For $v \in V_\lambda$:

$$
\begin{aligned}
(T + T^{-1})v &= \lambda v \\
\Rightarrow T(T + T^{-1})v &= \lambda Tv \\
\Rightarrow (T^2 - \lambda T + I)v &= 0
\end{aligned}
$$

If $\lambda = \pm 2$, then $(T - \mu I)^2 = 0$ or $(T + \mu I)^2 = 0$ with $\mu = \pm 1$, and thus $T|_{V_\lambda} = \pm I$

If $\lambda \neq \pm 2$, then $T|_{V_\lambda}$ has no real eigenvalues - note: real eigenvalues $= \pm 1$
So $\{v, Tv\}$ are linearly independent for $V \neq 0$.
Consider $W = \langle v, Tv \rangle$. W is T-invariant:

$$
\begin{aligned}
v &\mapsto Tv \in W \\
Tv &\mapsto T^2 v = \lambda Tv - v \in W
\end{aligned}
$$

Hence, $W^\perp$ is also T invariant.

By induction, V_λ splits into two-dimensional T-invariant subspaces.
Moreover, $\mathcal{X}_{T|_W}(x) = x^2 - \lambda_i x + 1$ and hence $\det T|_W = 1$.
Thus, by Example 8.8, each $T|_W$, with respect to some orthonormal basis of W, is of the form $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ for some $\theta \neq 0, \pi$

$\square$

8.4 Normal Transformations

Definition 8.13. Let $T : V \to V$ be a linear transformation and V be a finite dimensional complex inner product space.
T is **normal** if it commutes with its adjoint:

$$T^*T = TT^*$$

Example 8.9.

$$\begin{aligned} T \text{ unitary } &\Rightarrow T^* = T^{-1} \Rightarrow T \text{ is } \textbf{normal} \\ T \text{ self adjoint } &\Rightarrow T^* = T \quad\;\; \Rightarrow T \text{ is } \textbf{normal} \end{aligned}$$

Lemma 8.24. Let T be normal, then:

(1) $Tv = 0 \iff T^*v = 0$

Proof.

$$\begin{aligned} Tv = 0 &\iff \langle Tv, Tv \rangle = 0 \\ &\iff \langle T^*Tv, v \rangle = 0 \\ &\iff \langle TT^*v, v \rangle = 0 \\ &\iff \langle T^*v, T^*v \rangle = 0 \\ &\iff T^*v = 0 \end{aligned}$$

(2) $T - \lambda I$ is normal for all $\lambda \in \mathbb{C}$

Proof. $(T - \lambda I)^* = T^* - \bar{\lambda} I$, this commutes with $T - \lambda I$ as T commutes with T^* and both the identity matrix and scalar multiplication commute with anything

(3) $Tv = \lambda v \Rightarrow T^*v = \bar{\lambda} v$

Proof.

$$\begin{aligned} Tv = \lambda v &\iff (T - \lambda I)v = 0 \\ &\iff (T - \lambda I)^*v = 0 && \text{(by (1))} \\ &\iff T^*v = \bar{\lambda} v && \text{(by (2))} \end{aligned}$$

(4) $Tv = \lambda_1 v$, $Tw = \lambda_2 v$, $\lambda_1 \neq \lambda_2 \Rightarrow \langle v, w \rangle = 0$

Proof.

$$\begin{aligned} \lambda_1 \langle v, w \rangle &= \left\langle \bar{\lambda_1} v, w \right\rangle \\ &= \langle T^*v, w \rangle && \text{(by (3))} \\ &= \langle v, Tw \rangle \\ &= \langle v, \lambda_2 w \rangle \\ &= \lambda_2 \langle v, w \rangle \end{aligned}$$

Since $\lambda_1 \neq \lambda_2$ we must have $\langle v, w \rangle = 0$

$\square$

Theorem 8.25. *Let $T : V \to V$ be a normal linear transformation over a finite dimensional complex inner product space V. Then there is an orthonormal basis of eigenvectors for V*

Proof. As V is complex there is an eigenvalue λ and corresponding normed eigenvector $v \in V$ with $||v|| = 1$, such that $Tv = \lambda v$.

Consider $U = \langle v \rangle$. By Lemma 8.24(3) we have that U is both T- and T^*- invariant.

Consider $U^\perp$. $U^\perp$ is also T- and T^*-invariant since for all $u \in U, w \in U^\perp$:

$$\begin{aligned}\langle u, Tw \rangle &= \langle T^* u, w \rangle \\ &= \langle u', w \rangle && \text{(for some } u' \in U \text{ since } U \text{ is } T^* \text{ invariant)} \\ &= 0\end{aligned}$$

$$\begin{aligned}\langle u, T^* w \rangle &= \langle Tu, w \rangle \\ &= \langle u', w \rangle \\ &= 0\end{aligned}$$

Now we proceed by induction on the dimension of V.

We have $\dim U^\perp = \dim V - 1 < \dim V$ and we know $T|_{U^\perp}$ is normal, thus, by induction hypothesis, there exists an orthonormal basis of eigenvectors of $T|_{U^\perp}$ for $U^\perp$, $\mathcal{B}' = \{e_2, \ldots, e_n\}$.

Then, putting $e_1 = v$ we obtain $\mathcal{B} = \{e_1, \ldots, e_n\}$ is an orthonormal basis of eigenvectors of T

□

Theorem 8.26. Spectral Theorem

Let $T : V \to V$ be a normal (symmetric) linear transformation on a finite dimensional complex (real) inner product space.

Then there exist orthogonal projections $E_1, \ldots, E_r$ on V and $\lambda_1, \ldots, \lambda_r \in \mathbb{C}(\mathbb{R})$ such that:

(1) $T = \lambda_1 E_1 + \ldots + \lambda_r E_r$

(2) $E_1 + \ldots + E_r = I$

(3) $E_i E_j = 0$ *for all* $i \neq j$

Remark. This is just a reformulation of Theorem 8.23

Proof. By Theorem 8.23, $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$

Then, each E_i is a projection from V to V_{λ_i}, that is $E_i = \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & I & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}$

8.5 Simultaneous Diagonalization

Remark. If $\mathcal{B}$ is a basis wrt which both S and T are diagonal, for $S, T : V \to V$, then $ST = TS$

$$ {}_{\mathcal{B}}[ST]_{\mathcal{B}} = \underbrace{{}_{\mathcal{B}}[S]_{\mathcal{B}\mathcal{B}}[T]_{\mathcal{B}} = {}_{\mathcal{B}}[T]_{\mathcal{B}\mathcal{B}}[S]_{\mathcal{B}}}_{\text{(diagonal matrices commute)}} = {}_{\mathcal{B}}[TS]_{\mathcal{B}} $$

Theorem 8.27. *If $S, T : V \to V$ are normal (symmetric) linear transformations on a finite dimensional complex (real) inner product space with $ST = TS$, then there exists and orthonormal basis of eigenvectors for S and T simultaneously*

Proof. V decomposes to λ-eigenspaces for S: $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$

Let $v \in V_\lambda$, then

$$ S(Tv) = T(Sv) = T(\lambda v) = \lambda T(v) $$

So Tv is an eigenvector of S and hence $Tv \in V_\lambda$

Now, there exists an orthonormal basis of eigenvectors of V_λ for $T|_{V_\lambda}$, $\mathcal{B}_\lambda$

Then $\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \cdots \cup \mathcal{B}_{\lambda_s}$ is an orthonormal basis of eigenvectors for T and S simultaneously.

□

Challenge 4.
If $S_1, \ldots, S_r : V \to V$ are normal (symmetric) for $\dim V < \infty$ over $\mathbb{C}(\mathbb{R})$ with $S_i S_j = S_j S_i$ for all i, j. Then there exists an orthonormal basis of eigenvectors for all S_k simultaneously.

Challenge 5.
If $A_1, \ldots, A_r \in O_n$ then there exists $P \in O_n$ such that

$$ P^{-1} A_i P = \begin{pmatrix} I & & & & \\ & -I & & & \\ & & R_{\theta_1} & & \\ & & & \ddots & \\ & & & & R_{\theta_s} \end{pmatrix} $$