MT 2022

Ao: Linear Algebra

Jiaming (George) Yu
jiaming.yu@jesus.ox.ac.uk

June 17, 2023

Contents

0 Vector Spaces and Linear Maps

We quickly recap some basic definitions.

Definition 0.1 (Field)

A set $\mathbb{F}$ with two binary operations $+$ and $\times$ is a *field* if both $(\mathbb{F}, +, 0)$ and $(\mathbb{F} \setminus \{0\}, \times, 1)$ are abelian groups, and $\times$ distributes over $+$:

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

Definition 0.2 (Characteristic of field)

Let V be a field. If there exists an integer[1] p such that $\overbrace{1 + 1 + \cdots + 1}^{p \text{ times}} = 0$, we call the smallest such p the *characteristic* of $\mathbb{F}$. If no such p exist, we define the characteristic to be 0.

[1] In fact, if such an integer exists, it is necessarily prime

Definition 0.3 (Vector space)

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

(i) $1_{\mathbb{F}} \cdot v = v$

(ii) $\lambda(v + w) = \lambda v + \lambda w$

(iii) $(\lambda + \mu)v = \lambda v + \mu v$

(iv) $(\lambda\mu)v = \lambda(\mu v)$

Proposition 0.1 (Subspace test)

Let V be a vector space over $\mathbb{F}$ and $U \subseteq V$. Then $U \leqslant V$ if and only if $0_V \in U$ and $\lambda v + w \in U$ for any $v, w \in V, \lambda \in \mathbb{F}$.

Definition 0.4 (Basis)

Let V be a vector space over $\mathbb{F}$ and $S \subseteq V$.

- S is *linearly independent* if for any $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and $s_1, \ldots, s_n \in S$,

$$\lambda_1 s_1 + \cdots + \lambda_n s_n = 0 \implies \lambda_1 = \cdots = \lambda_n = 0$$

- S is *spanning* if for any $v \in V$, there exists some $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and $s_1, \ldots, s_n \in S$ such that

$$v = \lambda_1 s_1 + \cdots + \lambda_n s_n$$

- S is a *basis* of V (usually denoted by $\mathcal{B}$ rather than S) if S is spanning and linearly independent. We call the size of any such S the *dimension* of V, denoted by $\dim V$.[2]

[2] It can be shown that $\dim V$ is well-defined

Definition 0.5 (Linear map)

Let V, W be vector spaces over $\mathbb{F}$. A map $T : V \to W$ is a *linear map* if for all $\lambda \in \mathbb{F}, v, w \in V$,

$$T(\lambda v + w) = \lambda T(v) + T(w)$$

If T is also bijective, then it is a *isomorphism* of V and W.

Any linear map $T : V \to W$ is determined by its image of a basis $\mathcal{B}$ of V. On the other hand, any map $T : \mathcal{B} \to W$ can be extended to a well-defined linear map $T' : V \to W$.

Definition 0.6

Let V, W be vector spaces with bases $\mathcal{B} = \{e_1, \ldots, e_n\}$ and $\mathcal{B}' = \{f_1, \ldots, f_m\}$ respectively, and let $T \in \mathrm{Hom}(V, W)$. Define ${}_{\mathcal{B}'}[T]_{\mathcal{B}}$ to be the *change of basis* matrix

$$T(e_j) = \sum_{i=1}^{m} ???$$

1 Rings

Rings are a natural extension of fields.

Definition 1.1 (Ring)

A *ring* is a non-empty set R equipped with binary operations $+, \times$ such that the following holds:

(i) $(R, +, 0)$ is an abelian group

(ii) $(R, \times, 1)$ is a monoid[3]

(iii) $\times$ distributes over $+$ (in both directions)

If $\forall a, b \in R : ab = ba$ then we also say R is *commutative*.

[3] Apart from requiring $1 \in R$, we also require $\times$ to be associative. Alternatively, a monoid is a group without invertibility

Definition 1.2 (Ring homomorphism)

For rings R, S, a map $\phi : R \to S$ is a *ring homomorphism* if for any $r, r' \in R$,

$$\phi(r + r') = \phi(r) + \phi(r'), \quad \phi(rr') = \phi(r)\phi(r')$$

If ϕ is also bijective, then it is a *ring isomorphism*.

Definition 1.3 (Ideal)

For a ring R, a non-empty subset $I \subseteq R$ is an *ideal* if $(I, +)$ is a subgroup of $(R, +)$ and that for any $s \in I$, $r \in R$, we have $sr, rs \in I$.

We remark on the relationship between ideals and subrings. Any ideal which contains 1 is the whole ring. Therefore, if we require $1 \in R$, then

Ideals are to rings what normal subgroups are to subgroups in that, I is an ideal exactly when then R/I inherits a ring structure from R.[4]

[4] This is just like $H \trianglelefteq G$ exactly when G/H inherits a group structure from G

Theorem 1.1 (First Isomorphism Theorem)

Let $\phi : R \to S$ be a ring homomorphism. Then $\ker \phi := \phi^{-1}(0)$ is an ideal, and $\operatorname{Im} \phi$ is a subring of S, and ϕ induces the following isomorphism:

$$R/\ker \phi \cong \operatorname{Im} \phi$$

Proof. [Show that the underlying isomorphisms of abelian groups is compatible with the multiplication, i.e. is a ring homomorphism.] □

1.1 Polynomial Rings

Theorem 1.2 (Division algorithm for polynomials)

Let $f, g \in \mathbb{F}[x]$ be polynomials with $g \neq 0$. Then there exists $q, r \in \mathbb{F}[x]$ with $\deg r < \deg g$ such that

$$f(x) = q(x)g(x) + r(x)$$

We now formalize evaluation of polynomials on matrices. Given $A \in M_n(\mathbb{F})$ and $f(x) = a_n x^n + \cdots + a_0 \in \mathbb{F}[x]$, we define

$$f(A) = a_n A^n + \cdots + a_0 I \in M_n(\mathbb{F})$$

For any polynomials $f, g \in \mathbb{F}[x]$, we have $f(A)g(A) = g(A)f(A)$. For any $v \in \mathbb{F}^n$ and $\lambda \in \mathbb{F}$, we have $Av = \lambda v \implies f(A)v = f(\lambda)v$.

Lemma 1.3

For any $A \in M_n(\mathbb{F})$, there exists a non-zero polynomial $f \in \mathbb{F}[x]$ such that $f(A) = 0$.
In other words, the group homomorphism $E_A : \mathbb{F}[x] \to M_n(\mathbb{F}), f(x) \mapsto f(A)$ has a non-zero kernel.

Proof. Since $\dim M_n(\mathbb{F}) = n^2$ is finite, we have that for any $k \geqslant n^2$, the set $\left\{I, A, \ldots, A^k\right\} \subseteq M_n(\mathbb{F})$ must be linearly dependent. Therefore there are scalars $a_0, \ldots, a_k$ which are not all zero such that

$$a_k A^k + \cdots + a_0 I = 0$$ □

Since all kernels are ideals, and that all ideals in $\mathbb{F}[x]$ can be generated by a single element, we are now interested in the unique polynomial that generates $\ker E_A$.

Definition 1.4 (Algebraically closed field)

A field $\mathbb{F}$ is *algebraically closed* if every non-constant polynomial in $\mathbb{F}[x]$ has a root in $\mathbb{F}$.

Definition 1.5 (Algebraic closure)

An algebraically closed field $\overline{\mathbb{F}}$ is an *algebraic closure* of a field $\mathbb{F}$ if $\overline{\mathbb{F}}$ is the smallest algebraically closed field containing $\mathbb{F}$ — in other words, there doesn't exist a smaller algebraically closed field L with $\mathbb{F} \subseteq L \subset \overline{\mathbb{F}}$.

Indeed, $\mathbb{C}$ is algebraically closed but $\mathbb{R}$ is not. Also, every field has an algebraic closure $\overline{\mathbb{F}}$; and no finite field is algebraically closed.

1.2 Minimal and Characteristic Polynomials

Definition 1.6 (Minimal polynomial)

Let $A \in M_n(\mathbb{F})$. The *minimal polynomial* $m_A(x)$ of A is the monic polynomial p of least degree such that $p(A) = 0$.

Theorem 1.4

Let $A \in M_n(\mathbb{F})$ and $f \in \mathbb{F}[x]$. If $f(A) = 0$ (that is, $f \in \ker E_A$), then $m_A \mid f$. Furthermore, m_A is unique.

Proof. By Theorem 1.2 (division algorithm), there exists $q, r \in \mathbb{F}[x]$ with $\deg r < \deg m_A$ such that $f = qm_A + r$. Since $f(A) = m_A(A) = 0$, we have $r(A) = 0$. If $r \neq 0$ it would contradict the minimality of m_A (since $\deg r < \deg m_A$), so $r = 0$, and $m_A \mid f$ follows.
To show uniqueness, let n_A be another minimal polynomial. Then $\deg m_A = \deg n_A$ by definition, and $m_A \mid n_A$ by the above, so $n_A = \lambda m_A$ for some $\lambda \in \mathbb{F}$, but both polynomials are monic by definition, so $\lambda = 1$ and $n_A = m_A$. □

Definition 1.7 (Characteristic polynomial)

Let $A \in M_n(\mathbb{F})$. The *characteristic polynomial* $\chi_A(x)$ of A is defined as

$$\chi_A(x) = \det(A - xI)$$

Lemma 1.5

Let $A \in M_n(\mathbb{F})$. Then

$$\chi_A(x) = (-1)^n x^n + (-1)^{n-1} \operatorname{tr}(A) x^{n-1} + \cdots + \det(A)$$

Theorem 1.6

Let $A \in M_n(\mathbb{F})$ and $\lambda \in \mathbb{F}$. Then the following claims are equivalent.

(i) λ is an eigenvalue of A

(ii) λ is a root of $\chi_A(x)$

(iii) λ is a root of $m_A(x)$

Proof. First we show (i) $\iff$ (ii).

$$\begin{aligned}\chi_A(\lambda) = 0 &\iff \det(A - \lambda I) = 0 \\ &\iff A - \lambda I \text{ is singular} \\ &\iff \exists v \neq 0 : (A - \lambda I)v = 0 \\ &\iff \exists v \neq 0 : Av = \lambda v\end{aligned}$$

The last statement implies $\exists v \neq 0 : m_A(A)v = m_A(\lambda)v = 0$, so we can conclude $m_A(\lambda) = 0$, which is (iii). Conversely, suppose λ is a root of m_A, then $(x - \lambda) \mid m_A$, so there exists a polynomial g such that $m_A(x) = g(x)(x - \lambda)$. Furthermore, we know $g(A) \neq 0$ since $\deg g < \deg m_A$. So we can find $v = g(A)w \neq 0$ for some $w \in \mathbb{F}^n$, and we will have $(A - \lambda I)v = m_A(A)w = 0$, so v is an eigenvalue of A. □

2 Quotient Spaces

Definition 2.1

Let V be a vector space over $\mathbb{F}$ and U be a subspace of V. The set of cosets $\{v + U \mid v \in V\}$, denoted by V/U, with operations

(i) $(v + U) + (w + U) = (v + w) + U$

(ii) $\lambda(v + U) = (\lambda v) + U$

is called a *quotient space*, and is a vector space.

Proposition 2.1

Let V be a vector space and $U \leqslant V$. Let $\mathcal{E}$ be a basis of U, assuming we can extend this to a basis $\mathcal{B}$ of V, we can then define

$$\overline{\mathcal{B}} = \{e + U : e \in \mathcal{B} \setminus \mathcal{E}\} \subseteq V/U$$

Then $\overline{\mathcal{B}}$ is a basis of V/U.

Proof. Suppose $\mathcal{E} = \{e_1, \ldots, e_n\}$ and $\mathcal{B} = \{e_1, \ldots, e_m\}$ with $m \geqslant n$. Therefore $\overline{\mathcal{B}} = \{e_{n+1}, \ldots, e_m\}$.
(Spanning) Let $v + U \in V/U$ with $v = \sum_{i=1}^m a_i e_i$ for $a_i \in \mathbb{F}$. Then $v + U = \sum_{i=1}^m a_i(e_i + U)$. But for any $1 \leqslant i \leqslant n$, $e_i \in \mathcal{E}$ and hence $e_i + U = U$. So we have $v + U = U + \sum_{i=n+1}^m a_i(e_i + U)$.
(Linear independence) Suppose for some $a_{n+1}, \ldots, a_m \in \mathbb{F}$ we have that $\sum_{i=n+1}^m a_i(e_i + U) = 0 + U = U$. Then we must have $\sum_{i=n+1}^m a_i e_i \in U$, so there are some $b_1, \ldots, b_n$ such that $\sum_{i=n+1}^m a_i e_i = \sum_{j=1}^n b_i e_i$ which implies $\sum_{i=n+1}^m a_i e_i + \sum_{j=1}^n (-b_i)e_i = 0$. Since $\mathcal{B}$ is linearly independent and each $e_i \in \mathcal{B}$, we conclude that $b_1 = \cdots = b_n = a_{n+1} = \cdots = a_m = 0$ and hence $\overline{\mathcal{B}}$ is linearly independent. □

Conversely, we also have the following result.

Proposition 2.2

Let V be a vector space and $U \leqslant V$. Let $\mathcal{E}$ be a basis of U and $\mathcal{F} \subseteq V$ be such that $\{v + U : v \in \mathcal{F}\} \subseteq V/U$ forms a basis of V/U. Then $\mathcal{E} \cup \mathcal{F}$ forms a basis of V.

From these, we can easily arrive

Corollary 2.3

Let V be a finite-dimensional vector space and $U \leqslant V$. Then

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

We now provide the First Isomorphism Theorem for vector spaces.

Theorem 2.4 (First Isomorphism Theorem for vector spaces)

Let V, W be vector spaces over $\mathbb{F}$ and $T : V \to W$ be a linear map. Then T induces the following isomorphism:

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

Recall also the rank-nullity theorem.

Theorem 2.5 (First Isomorphism Theorem for vector spaces)

Let V, W be vector spaces over $\mathbb{F}$ where V is finite-dimensional and let $T : V \to W$ be a linear map. Then

$$\dim V = \dim(\ker T) + \dim(\operatorname{Im} T)$$

3 The Cayley–Hamilton Theorem and Triangularization

Previously we have shown that the minimal polynomial is annihilating and divides any annihilating polynomial. We now want to show that the characteristic polynomial is also annihilating.

Definition 3.1

Let V be a vector space and $T : V \to V$ a linear map. A subspace $U \leqslant V$ is called *T-invariant* if $T(U) \subseteq U$.

Proposition 3.1

Let V be a vector space under $\mathbb{F}$ and $U \leqslant V$, and let $T, S : V \to V$ be distinct linear maps. If U is both T- and S-invariant, then U is also invariant under the following maps:

- the zero map
- the identity map
- aT for any $a \in \mathbb{F}$
- $S + T$
- $S \circ T$
- (from the above) any polynomial $p(x)$ evaluated at T

Definition 3.2

A matrix $A = (a_{i,j})$ is *upper triangular* if $a_{i,j} = 0$ for all $i > j$.

Theorem 3.2

Let V be a finite-dimensional vector space and $T : V \to V$ be a linear map. If χ_T is a product of linear factors[5], then there exists a basis $\mathcal{B}$ of V such that ${}_{\mathcal{B}}[V]_{\mathcal{B}}$ is upper-triangular.

[5] Note that in an algebraically closed field, this condition is always satisfied.

Proof. □

Lemma 3.3

Let A be an upper triangular $n \times n$ matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$, then

$$\prod_{i=1}^{n} (A - \lambda_i I) = 0$$

Proof. Let $e_1, \ldots, e_n$ be the canonical basis for $\mathbb{F}^n$. Then by definition, for all $v \in \mathbb{F}^n$, we have

$$(A - \lambda_n I)v \in \langle e_1, \ldots, e_{n-1} \rangle$$

□

Theorem 3.4 (Cayley–Hamilton)

Let V be a finite-dimensional vector space over $\mathbb{F}$ and $T : V \to V$ be a linear map. Then $\chi_T(T) = 0$. As a consequence, $m_T(x) \mid \chi_T(x)$.

Proof. Let A be the matrix of T w.r.t. some basis of V. We assume that $\mathbb{F}$ can be embedded in an algebraically closed field $\overline{\mathbb{F}}$. In $\overline{\mathbb{F}}[x]$, every polynomial factors into linear terms. Thus, (by a previous result), there exists a matrix $P \in M_n(\overline{\mathbb{F}})$ such that $P^{-1}AP$ is upper-triangular with diagonal

entries $\lambda_1, \ldots, \lambda_n$. Therefore, we now have

$$\chi_{P^{-1}AP}(x) = (-1)^{\dim V} \prod_{k=1}^{n} (x - \lambda_k)$$

Now we want to show the product above evaluates to 0. Let $e_1, \ldots, e_n$ be the standard basis vectors for $\overline{\mathbb{F}}^n$. Then,

$$(P^{-1}AP - \lambda_i I)v \in \langle e_1, \ldots, e_{i-1} \rangle \text{ for any } v \in \overline{\mathbb{F}}^n$$

So, this means that

$$\mathrm{Im}(P^{-1}AP - \lambda_n I) \subseteq \langle e_1, \ldots, e_{n-1} \rangle$$

Then we have

$$\chi_{P^{-1}AP}(P^{-1}AP) = 0$$

□

4 The Primary Decomposition Theorem

With the knowledge now equipped, we are able to decompose a vector space into T-invariant subspaces. We will focus our discussion on finite-dimensional vector spaces.

4.1 Direct Sums

First recall that for a vector space V and subspaces $W_1, \ldots, W_r$, if any $v \in V$ can be **uniquely** written as a sum $v = w_1 + \cdots + w_r$ where $w_i \in W_i$, then V is the direct sum $W_1 \oplus \cdots \oplus W_r$.

Proposition 4.1

Suppose $V = W_1 \oplus \cdots \oplus W_r$, then

(i) let $\mathcal{B}_i$ be a basis of W_i, then $\mathcal{B} = \cup_i \mathcal{B}_i$ is a basis of V.

(ii) let $T : V \to V$ be a linear map such that each W_i is T-invariant, then

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

where each $A_i := {}_{\mathcal{B}_i}[T|_{W_i}]_{\mathcal{B}_i}$ is a block matrix.

(iii) $\chi_T(x) = \chi_{T|_{W_1}}(x) \cdots \chi_{T|_{W_r}}(x)$

Proposition 4.2

Let $T : V \to V$ be a linear map. If f is a polynomial such that $f(T) = 0$ and $f(x) = a(x)b(x)$ for coprime polynomials a, b, then

$$V = \ker a(T) \oplus \ker b(T)$$

gives a T-invariant direct sum (that is, both $\ker a(T)$ and $\ker b(T)$ are T-invariant). Furthermore, if $f = m_T$ and a, b are monic, then $a = m_{T|_{\ker a(T)}}$ and $b = m_{T|_{\ker b(T)}}$.

4.2 Primary decomposition theorem

Primary decomposition theorem provides a generalization of the above proposition.

Theorem 4.3 (Primary decomposition theorem)

Let m_T be the minimal polynomial in the form

$$m_T(x) = f_1^{q_1}(x) \cdots f_r^{q_r}(x)$$

where each f_i is a **distinct monic irreducible** polynomial. Write $W_i := \ker\left(f_i^{q_i}(T)\right)$, then

(i) $V = W_1 \oplus \cdots \oplus W_r$

(ii) each W_i is T-invariant

(iii) $m_{T|_{W_i}} = f_i^{q_i}$

Proof. We will prove for the case $m_T(x) = f^n(x)g^m(x)$ □

Corollary 4.4

Let $T : V \to V$ be a linear map, then:

$$\begin{aligned} & T \text{ is triangularizable (over } \mathbb{F}) \\ \iff & \chi_T \text{ factors into a product of linear polynomials (over } \mathbb{F}) \\ \iff & \text{each } f_i \text{ is linear} \\ \iff & m_T \text{ factors into a product of linear polynomials} \end{aligned}$$

Theorem 4.5

T is diagonalizable if and only if m_T factors into a product of **distinct** linear polynomials.

Proof. ($\Rightarrow$) Suppose T is diagonalizable. Then there exists a basis $\mathcal{B}$ of eigenvectors of V such that ${}_{\mathcal{B}}[T]_{\mathcal{B}}$ is diagonal with eigenvalues $\lambda_1, \ldots, \lambda_n$

as entries. So,

$$m(x) = (x - \lambda_1) \dots (x - \lambda_r)$$

is annilhating as $m(T)v = 0$ for any $v \in \mathcal{B}$ and hence for any $v \in V$. Furthermore, m is minimal as every eigenvalue must be a root of the minimal polynomial. Therefore $m = m_T$ factors into a product of distinct linear polynomials.
($\Leftarrow$) Suppose $m_T(x) = (x - \lambda_1) \dots (x - \lambda_r)$. By Theorem 4.3 (Primary decomposition theorem), we can decompose V into eigenspaces:

$$V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_r}$$

where each $E_{\lambda_i} := \ker(T - \lambda_i I)$. Let $\mathcal{B}_i$ be a basis of E_{λ_i}, then $\mathcal{B} := \bigcup_{i=1}^{r} \mathcal{B}_i$ gives a basis of eigenvectors such that ${}_{\mathcal{B}}[T]_{\mathcal{B}}$ is diagonal. □

5 Jordan Normal Form

Definition 5.1 (Nilpotent)

Let V be a finite-dimensional vector space and $T : V \to V$ a linear map. We say T is *nilpotent* if $T^n = 0$ for some $n > 0$.

Theorem 5.1

Let $T : V \to V$ be a nilpotent linear map. Then the minimal polynomial of T has the form $m_T(x) = x^m$ for some m, and there exists a basis $\mathcal{B}$ of V such that

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

Corollary 5.2

Let V be a finite-dimensional vector space and $T : V \to V$ be a linear map. If $m_T(x) = (x - \lambda)^m$ for some m, then there exists a basis $\mathcal{B}$ of V such that ${}_{\mathcal{B}}[T]_{\mathcal{B}}$ is block diagonal and each block J_i for $i = 1, \dots, m$ is of the form

$$J_i = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

Theorem 5.3

6 Dual Spaces

Definition 6.1 (Dual space)

Let V be a vector space over $\mathbb{F}$. The *dual space* of V is the vector space V' of linear maps from V to $\mathbb{F}$, these maps are called *linear functionals*. In other words, $V' = \mathrm{Hom}(V, \mathbb{F})$.

Definition 6.2 (Dual basis)

Let V be a vector space with a basis $\mathcal{B} = \{e_1, \ldots, e_n\}$. The *dual basis* of $\mathcal{B}$ is a basis $\mathcal{B}' = \{f_1, \ldots, f_n\}$ of V', where each f_i is a linear functional such that

$$f_i(e_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Definition 6.3 (Dual map)

Let V, W be vector spaces and $T : V \to W$ a linear map. The *dual map* of T is the linear map $T' : W' \to V'$ defined by $f \mapsto f \circ T$.

We will now justify a few assumptions made in the definitions above.

Proposition 6.1

A dual basis $\mathcal{B}' = \{f_1, \ldots, f_n\}$ is a basis of V'.

Proof. (Linear independence) Suppose for some $a_1, \ldots, a_n \in \mathbb{F}$ we have $a_1 f_1 + \ldots a_n f_n = 0$, where 0 is the zero map. For each $i = 1, \ldots, n$, we have that $(a_1 f_1 + \cdots + a_n f_n)(e_i) = a_i = 0$ since $f_j(e_i) = 0$ for any $j \neq i$. Therefore $a_1 = \cdots = a_n = 0$.
(Spanning) Let $f \in V'$. For each $i = 1, \ldots, n$, write $a_i := f(e_i)$. Then, since any linear map is determined by its values on any basis, we have that $f = \sum_i a_i f_i$. □

Proposition 6.2

A dual map $T' : W' \to V'$ is a linear map.

Proof. Let $f, g \in W'$ and $\lambda \in \mathbb{F}$. Since T is linear, then,

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

□

Proposition 6.3

Let V be a finite-dimensional vector space. Then $E : V \to V''$, $v \mapsto E_v$ defines a natural linear isomorphism, where $E_v : V' \to \mathbb{F}$, $f \mapsto f(v)$ is the evaluation map at v. This isomorphism is natural in that it is independent of a choice of basis.

6.1 Annihilators

Definition 6.4 (Annihilator)

Let V be a vector space and $U \leqslant V$. The *annihilator* U^0 of U is

$$U^0 = \{f \in V' : f(u) = 0 \text{ for all } u \in U\}$$

Proposition 6.4

Let V be a finite-dimensional vector space and $U \leqslant V$. Then,

$$\dim(U^0) = \dim V - \dim U$$

Proof. Let $\{e_1, \dots, e_m\}$ be a basis of U, which can be extended to a basis $\{e_1, \dots, e_n\}$ of V. Let $\{f_1, \dots, f_n\}$ be a corresponding dual basis. For any $i = 1, \dots, m$ and $j = m+1, \dots, n$, we have $j > i$ and hence $f_j(e_i) = 0$. Therefore $f_j \in U^0$. So we have $\langle f_{m+1}, \dots, f_n \rangle \subseteq U^0$.
Conversely, let $f \in U^0$.
We conclude that $U^0 = \langle f_{m+1}, \dots, f_n \rangle$, thus $\dim(U^0) = n - m = \dim V - \dim U$. □

Proposition 6.5

Let V be a vector space and $U, W \leqslant V$. Then,

(i) $U \subseteq W \implies W^0 \subseteq U^0$

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

(iii) $U^0 + W^0 \subseteq (U \cap W)^0$ where equality holds if $\dim V$ is finite.

Proof. □

Proposition 6.6

Let V be a finite-dimensional vector space and $U \leqslant V$. Then the natural map $E : V \to V''$, $v \mapsto E_v$ maps U isomorphically to U^{00}.

Proof. Let $u \in U$, then $E(u) = E_u$. The elements of U^{00} maps all linear functionals in $U^0 \leqslant V'$ to 0.[6] Let $f \in U^0$, then $E_u(f) = f(u) = 0$ for all $u \in U$, hence $E_u \in U^{00}$. □

[6] Note that dual spaces always map to $\mathbb{F}$, e.g. $V'' : V' \to \mathbb{F}$, so 0 here is in $\mathbb{F}$ and not the zero function.

Proposition 6.7

Let V be a vector space and $U \leqslant U$. Then there exists a natural isomorphism $U^0 \cong (V/U)'$ given by $f \mapsto \bar{f}$ where $\bar{f} : V/U \to V,\ v + U \mapsto v$.

Proof. □

6.2 Dual Maps

Definition 6.5 (Dual map)

Let $T : V \to W$ be a linear map. The *dual map* of T is the linear map[7] $T' : W' \to V'$ defined by $T'(f) = f \circ T$.

[7] This requires proof! Also linearity implies that $T'(f) \in V'$.

Proposition 6.8

Let V, W be finite-dimensional vector spaces. Then $T \mapsto T'$ defines a natural isomorphism from $\mathrm{Hom}(V, W)$ to $\mathrm{Hom}(W', V')$.

Proof. □

Proposition 6.9

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

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

Proof. □

In summary, we have the following isomorphisms:

- $V \cong V''$ by $v \mapsto E_v$
- $U \cong U^{00}$ by $u \mapsto E_u$
- $U^0 \cong (V/U)'$ by $f \mapsto \bar{f}$ where $\bar{f}(v + U) = v$ (not necessarily finite-dimensional)
- $\mathrm{Hom}(V, W) \cong \mathrm{Hom}(W', V')$ by $T \mapsto T'$
- $(U^0)' \cong V/U$ by

7 Inner Product Spaces

Recall that on $\mathbb{R}^n$, we define the inner product as $\langle v, w \rangle := v^T w$, and on $\mathbb{C}^n$, we define it as $\langle v, w \rangle := \overline{v}^T w$. These are the usual inner products on their respective spaces, and they endow the spaces with notions of length and distance (and for $\mathbb{R}^n$, also angles), we shall study other inner products which behave similarly.

Definition 7.1 (Bilinear form)

Let V be a vector space over $\mathbb{F}$. A *bilinear form* on V is a map $F : V \times V \to \mathbb{F}$ such that for all $u, v, w \in V$ and $\lambda \in \mathbb{F}$,

(i) $F(u+v, w) = F(u, w) + F(v, w)$

(ii) $F(u, v+w) = F(u, v) + F(u, w)$

(iii) $F(\lambda v, w) = F(v, \lambda w) = \lambda F(v, w)$

For such a map F, we say it is *symmetric* if $\forall v, w \in V : F(v, w) = F(w, v)$; and we say it is *non-degenerate* if $\forall v \in V : F(v, w) = 0 \implies w = 0$.
When $\mathbb{F} = \mathbb{R}$, we say F is *positive definite* if $\forall v \in V \setminus \{0\} : F(v, v) > 0$. A positive definite bilinear form is always non-degenerate.

Definition 7.2 (Sesquilinear form)

Let V be a vector space over $\mathbb{C}$. A *sesquilinear form* on V is a map $F : V \times V \to \mathbb{C}$ such that for all $u, v, w \in V$ and $\lambda \in \mathbb{C}$,

(i) $F(u+v, w) = F(u, w) + F(v, w)$

(ii) $F(u, v+w) = F(u, v) + F(u, w)$

(iii) $F(\overline{\lambda} v, w) = F(v, \lambda w) = \lambda F(v, w)$

For such a map F, we say it is *conjugate symmetric* if $\forall v, w \in V : F(v, w) = \overline{F(w, v)}$. If so, then $F(v, v) \in \mathbb{R}$ and we further say it is *positive definite* if $\forall v \in V \setminus \{0\} : F(v, v) > 0$.

Note that the "bi" in bilinear form suggests F is linear (w.r.t. scaling) on both arguments, while sesquilinear is linear on one argument, but anti-linear (hence the conjugate) on the other argument.

Definition 7.3 (Inner product space)

A real vector space V endowed with a **bilinear**, **symmetric**, and **positive definite** (and hence non-degenerate) form $\langle \cdot, \cdot \rangle$ is an *inner product space*.
A complex vector space V endowed with a **sesquilinear**, **conjugate symmetric** (and hence non-degenerate) form $\langle \cdot, \cdot \rangle$ is an *inner product space*.

Definition 7.4 (Orthogonality)

Given a (real or complex) inner product space V, the elements $\{w_1, \ldots, w_n\} \subset V$ are mutually *orthogonal* if $\langle w_i, w_j \rangle = 0$ whenever $i \neq j$. Furthermore, if $\langle w_i, w_i \rangle = 1$ for all i, then it is also *orthonormal*.

Proposition 7.1

A set of non-zero, orthogonal elements of an inner product space is also linearly independent.

Proof. Let V be an inner product space over $\mathbb{F}$ (which is $\mathbb{R}$ or $\mathbb{C}$), and $\{w_1, \ldots, w_n\} \subset V$ be non-zero and mutually orthogonal. Suppose there exists some $a_1, \ldots, a_n \in \mathbb{F}$ such that $a_1 w_1 + \ldots a_n w_n = 0$.
We have that for each j,

$$\langle w_j, \sum_i a_i w_i \rangle = \sum_i a_i \langle w_j, w_i \rangle = a_j \langle w_j, w_j \rangle$$

But each $\langle w_j, \sum_i a_i w_i \rangle = 0$[8], and $w_j \neq 0$ so by non-degeneracy $\langle w_j, w_j \rangle \neq 0$, hence $a_j = 0$ for all j. Therefore $\{w_1, \ldots, w_n\}$ are linearly independent. □

[8] WHY?

7.1 Gram–Schmidt Orthonormalization

We aim to find an orthonormal basis for some inner product space $\mathbb{F}^n$ where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$.

Theorem 7.2 (Gram–Schmidt process)

Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$ and $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. Compute

$$\begin{aligned} w_1 &= v_1 \\ w_2 &= v_2 - \frac{\langle w_1, v_2 \rangle}{\langle w_1, w_1 \rangle} w_1 \\ &\vdots \\ w_k &= v_k - \sum_{i=1}^{k-1} \frac{\langle w_i, v_k \rangle}{\langle w_i, w_i \rangle} w_i \\ &\vdots \end{aligned}$$

Then, compute $u_i = \frac{w_i}{\|w_i\|}$. Then $\{u_1, \ldots, u_n\}$ is an orthonormal basis of V.

Proof. We first want to show $\{w_1, \ldots, w_n\}$ is orthogonal. $\{w_1\}$ is certainly orthogonal. Now assume $\{w_1, \ldots, w_k\}$ is orthogonal, then for any $j \leqslant k$,

$$\langle w_j, w_{k+1} \rangle = \langle w_j, v_{k+1} \rangle - \left\langle w_j, \sum_{i=1}^{k} \frac{\langle w_i, v_{k+1} \rangle}{\langle w_i, w_i \rangle} w_i \right\rangle = \langle w_j, v_{k+1} \rangle - \frac{\langle w_j, v_{k+1} \rangle}{\langle w_j, w_j \rangle} \langle w_j, w_j \rangle = 0$$

Similarly, for spanning, first we certainly have $\langle w_1 \rangle = \langle v_1 \rangle$. Now assume $\langle w_1, \ldots, w_k \rangle = \langle v_1, \ldots, v_k \rangle$ spans the same set. Then

$$\langle w_1, \ldots, w_k, w_{k+1} \rangle = \langle w_1, \ldots, w_k, v_{k+1} \rangle = \langle v_1, \ldots, v_k, v_{k+1} \rangle$$

So $\{w_1, \ldots, w_n\}$ is both orthogonal and spanning. They are all non-zero as we started from a basis, so by Proposition 7.1, it is also linearly independent, and hence forms a orthogonal basis. Finally,

$$\langle u_i, u_i \rangle = \left\langle \frac{w_i}{\|w_i\|}, \frac{w_i}{\|w_i\|} \right\rangle = \frac{1}{\sqrt{\langle w_i, w_i \rangle}^2} \langle w_i, w_i \rangle = 1$$

Therefore $\{u_1, \ldots, u_n\}$ is an orthonormal basis. □

7.2 Complements and Duals

Note that in both bilinear and sesquilinear forms, the second argument is always linear, so $\langle v, \cdot \rangle$ is a linear map.

Theorem 7.3

The map $\phi : V \to V'$, $v \mapsto \langle v, \cdot \rangle$ is a natural injective $\mathbb{R}$-linear map. When V is finite-dimensional, it is an isomorphism.

Definition 7.5

Let V be an inner product space and $U \leqslant V$. The *orthogonal complement* of U is

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

Proposition 7.4

Let V be an inner product space and $U, W \leqslant V$. Then,

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

(ii) $U \oplus U^{\perp} = V$ if V is finite-dimensional

(iii) $(U \cap W)^{\perp} \supseteq U^{\perp} + W^{\perp}$ with equality if V is finite-dimensional

(iv) $U \subseteq (U^{\perp})^{\perp}$ with equality if V is finite-dimensional

Proposition 7.5

7.3 Adjoint of Maps

Definition 7.6

Given a linear map $T : V \to W$, another linear map $T^* : V \to V$ is its *adjoint* if for all $v, w \in V$,

$$\langle v, T(w) \rangle = \langle T^*(v), w \rangle$$

Adjoint maps are unique.
If $T^* = T$, then we say T is *self-adjoint*.
If $T^* = -T$, then we say T is *skew-adjoint*.

Proposition 7.6

Let V be finite-dimensional over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). For any $S, T : V \to V$ linear maps and $\lambda \in \mathbb{F}$,

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

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

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

(iv) $T^{**} = T$

(v) $m_{T^*} = \overline{m_T}$ where m_T is the minimal polynomial of T

Theorem 7.7

Let V be finite-dimensional and $T : V \to V$ be linear. Then the adjoint of T exists, is unique, and is linear.

Proof. □

Proposition 7.8

Let V be finite-dimensional and $T : V \to V$ be linear, and let $\mathcal{B}$ be an orthonormal basis for V. Then,

$$_{\mathcal{B}}[T^*]_{\mathcal{B}} = (\overline{_{\mathcal{B}}[T]_{\mathcal{B}}})^T$$

Proof. □

Proposition 7.9

If T is self-adjoint and $U \leqslant V$ is T-invariant, then $U^\perp$ is also T-invariant.

Theorem 7.10

Let V be finite-dimensional and $T : V \to V$ be self-adjoint. Then there exists an orthonormal basis of eigenvectors of T.

7.4 Orthogonal and Unitary Transformations

Definition 7.7

Let V be a finite-dimensional inner product space over $\mathbb{F}$ and $T : V \to V$ be a linear map. If $T^* = T^{-1}$, then T is *orthogonal* when $\mathbb{F} = \mathbb{R}$ and is *unitary* when $\mathbb{F} = \mathbb{C}$.

Theorem 7.11

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

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

(iii) T preserves lengths: $\forall v \in V : \|v\| = \|Tv\|$

In fact, the norm determines the inner product.

Definition 7.8

- *Orthogonal group* $O(n) := \left\{A \in \mathcal{M}_{n\times n}(\mathbb{R}) : A^T A = I\right\}$
- *Special orthogonal group* $SO(n) := \{A \in O(n) : \det A = 1\}$
- *Unitary group* $U(n) := \left\{A \in \mathcal{M}_{n\times n}(\mathbb{C}) : \overline{A}^T A = I\right\}$
- *Special unitary group* $SU(n) := \{A \in U(n) : \det A = 1\}$

Lemma 7.12

If λ is an eigenvalue of an orthogonal or unitary linear map $T : V \to V$, then $|\lambda| = 1$.
This implies that an orthogonal or unitary matrix A has $\det A = 1$.

Lemma 7.13

Let V be finite-dimensional and $T : V \to V$ is such that $T^*T = I$. If U is T-invariant, then $U^\perp$ is also T-invariant.

Theorem 7.14

Let V be finite-dimensional and $T : V \to V$ be unitary. Then there exists an orthonormal basis of eigenvectors of T.

The above theorem tells us that all unitary (orthogonal doesn't suffice!) matrices are diagonalizable.

Theorem 7.15

Let $T : V \to V$ be orthogonal and V be a finite dimensional real vector space. Then there exists an orthonormal basis $\mathcal{B}$ such that

$$ {}_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{bmatrix} I & & & & \\ & -I & & & \\ & & R_{\theta_1} & & \\ & & & \ddots & \\ & & & & R_{\theta_\ell} \end{bmatrix}, \ \theta_i \notin \{0, \pi\} $$

where $R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.