Algebra I

Timothy Hosgood

February 1, 2018

Abstract

These notes are based entirely on lectures given by Ulrike Tillmann to second-year undergraduates at the University of Oxford in the year 2013/14.

Contents

1 Vector spaces

Let $\mathbb{F}$ denote a field. Then both $(\mathbb{F}, +, 0)$ and $(\mathbb{F} \setminus \{0\}, \times, 1)$ are abelian groups, and the distribution law holds:

$$(a+b)c = ac + bc$$

Definition 1.1 (Characteristic)**.** The smallest integer p such that

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

is called the *characteristic of* $\mathbb{F}$. If no such p exists then we define the characteristic of $\mathbb{F}$ to be 0.

Definition 1.2 (Vector space)**.** A *vector space* V over a field $\mathbb{F}$ is an abelian group $(V, +, 0)$, together with scalar multiplication $\mathbb{F} \times V \to V$, such that, for all $a, b \in \mathbb{F}$ and $v, w \in V$,

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

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

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

(iv) $1v = v$

Definition 1.3 (Linear independence)**.** A set $S \subseteq V$ is *linearly independent* if, whenever $a_1 s_1 + \ldots + a_n s_n = 0$, with $a_i \in \mathbb{F}$ and $s_i \in S$, we have that $a_i = 0$ for all $1 \leq i \leq n$.

Definition 1.4 (Spanning)**.** A set $S \subseteq V$ is *spanning* if, for all $v \in V$, there exist $a_1, \ldots, a_n \in \mathbb{F}$ and $s_1, \ldots, s_n \in S$ such that $v = a_1 s_1 + \ldots + a_n s_n$.

Definition 1.5 (Basis)**.** A set $S \subseteq V$ is a *basis* of V if it is both spanning and linearly independent.

Definition 1.6 (Linear transformation)**.** Suppose V and W are vector spaces over $\mathbb{F}$. A map $T : V \to W$ is a *linear transformation* if, for all $a \in \mathbb{F}$ and $v, w \in V$,

$$T(av + w) = aT(v) + T(w)$$

Definition 1.7 (Isomorphism)**.** An *isomorphism* is a bijective linear transformation.

Remark 1.8. Every linear map $T : V \to W$ is determined solely by its action on a basis $\mathcal{B}$ of V (as $\mathcal{B}$ is spanning). Vice versa, any linear map $T : \mathcal{B} \to W$ can be extended to a linear transformation $T' : V \to W$ (as $\mathcal{B}$ is linearly independent).

Definition 1.9. Let $\hom(V, W)$ be the set of linear transformations from V to W. For $a \in \mathbb{F}$, $v \in V$, and $S, T \in \hom(V, W)$, define

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

Lemma 1.10. *With the operations defined as in Definition 1.9,* $\hom(V, W)$ *is a vector space over* $\mathbb{F}$.

Definition 1.11 (Matrix of a linear transformation)**.** Assume that V and W are finite dimensional, and let $\mathcal{B} = \{e_1, \ldots, e_m\}$ and $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ be bases for V and W respectively. Let $T : V \to W$. Define

$${}_{\mathcal{B}'}[T]_{\mathcal{B}} = (a_{ij})_{ij}$$

where a_{ij} are such that

$$a_{1j}e'_1 + \ldots + a_{nj}e'_n = T(e_j)$$

That is, the j^{th} column of ${}_{\mathcal{B}'}[T]_{\mathcal{B}}$ is the coordinate vector of $T(e_j)$ in terms of the basis $\mathcal{B}'$.

Lemma 1.12. *Let* $T, S \in \hom(V, W)$*,* $a \in \mathbb{F}$*, and* $\mathcal{B}, \mathcal{B}'$ *be bases for* V *and* W *respectively. Then*

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

Further, if U *is a finite-dimensional vector space with basis* $\mathcal{B}''$*, and* $R \in \hom(W, U)$*, then*

$${}_{\mathcal{B}''}[R \circ T]_{\mathcal{B}} = ({}_{\mathcal{B}''}[R]_{\mathcal{B}'})({}_{\mathcal{B}'}[T]_{\mathcal{B}})$$

Theorem 1.13. *The map* $T \mapsto {}_{\mathcal{B}'}[T]_{\mathcal{B}}$ *is an isomorphism of vector spaces, from* $\hom(V, W)$ *to* $n \times m$ *matrices, which is compatible with composition.*

Definition 1.14 (Change of basis of a matrix)**.** If V is a finite-dimensional vector space, with two bases, $\mathcal{B}$ and $\mathcal{B}'$, and $T : V \to V$, then

$${}_{\mathcal{B}'}[T]_{\mathcal{B}'} = ({}_{\mathcal{B}'}[\mathrm{Id}]_{\mathcal{B}})({}_{\mathcal{B}}[T]_{\mathcal{B}})({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}'})$$

Note, in particular, that

$$({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}'})({}_{\mathcal{B}'}[\mathrm{Id}]_{\mathcal{B}}) = ({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}}) = I$$

2 Polynomials

Proposition 2.1 (Division algorithm for polynomials)**.** *Let* $f(x), g(x) \in \mathbb{F}[x]$*, with* $g(x) \neq 0$*. Then there exist* $q(x), r(x) \in \mathbb{F}[x]$ *such that*

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

and $\deg r(x) < \deg g(x)$

Proof. If $\deg f(x) < \deg g(x)$ then simply take $g(x) = 0$ and $r(x) = f(x)$. Otherwise, $\deg f(x) \geq \deg g(x)$, and write

$$\begin{aligned} f(x) &= a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0 \\ g(x) &= b_m x^m + b_{m-1}x^{m-1} + \ldots + b_0 \end{aligned}$$

where $n \geq m$.

Now introduce $d = n - m$. For $d = 0$, $n = m$, and see that

$$f(x) = \underbrace{\left(\frac{a_n}{b_n}\right)}_{q(x)} g(x) + \underbrace{\left(f(x) - \frac{a_n}{b_n} g(x)\right)}_{r(x)}$$

where $q(x)$ is well defined thanks to the fact that $b_n \neq 0$, and $\deg r(x) < n$. So we proceed by induction on d, returning to $f_1(x)$. Assuming that the theorem is true for $d < k$, for some k, now consider $d = k$, so that $n = m + k$.

Then define

$$f_1(x) = f(x) - \frac{a_n}{b_m} x^{n-m} g(x)$$

and note that $\deg f_1(x) < \deg f(x)$, so by our inductive step, there exist $q_1(x)$ and $r(x)$ such that

$$f_1(x) = q_1(x)g(x) + r(x)$$

with $\deg r(x) < \deg g(x)$. Thus

$$f(x) = f_1(x) + \frac{a_n}{b_m} x^{n-m} g(x) = \underbrace{\left(q_1(x) + \frac{a_n}{b_m} x^{n-m}\right)}_{q(x)} g(x) + r(x)$$

and we are done. □

Corollary 2.2. *Let $f(x) \in \mathbb{F}[x]$ and $a \in \mathbb{F}$. If $f(a) = 0$ then $(x - a) \mid f(x)$.*

Proof. By the division algorithm we have that there exist $q(x), r(x) \in \mathbb{F}[x]$ with $\deg r(x) < \deg(x - a) = 1$ such that

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

but note that $\deg r(x) < 1$ means that $r(x)$ is a constant.

Evaluating f at $x = a$ gives

$$0 = f(a) = q(x)(a - a) + r = 0 + r = r$$

and so

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

□

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

Proof. Apply induction on Corollary 2.2. □

Definition 2.4 (Algebraically closed)**.** A field $\mathbb{F}$ is *algebraically closed* if every polynomial in $\mathbb{F}[x]$ has a root in $\mathbb{F}$.

Definition 2.5 (Algebraic closure)**.** Let $\mathbb{F}$ be a field which is not algebraically closed. If $\overline{\mathbb{F}}$ is an algebraically closed field with the property that $\mathbb{F} \subset \overline{\mathbb{F}}$, and there is no algebrically closed field $\mathbb{L}$ such that $\mathbb{F} \subset \mathbb{L} \subsetneq \overline{\mathbb{F}}$, then $\overline{\mathbb{F}}$ is an *algebraic closure* of $\mathbb{F}$.

That is, $\overline{\mathbb{F}}$ is the 'smallest' algebraically closed field containing $\mathbb{F}$.

Theorem 2.6. *Every field* $\mathbb{F}$ *has an algebraic closure* $\overline{\mathbb{F}}$*.*

Definition 2.7 (Polynomials of matrices)**.** Let $A \in \mathcal{M}_n(\mathbb{F})$, the space of $n \times n$ matrices over $\mathbb{F}$, and $f(x) = a_k x^k + \ldots + a_0 \in \mathbb{F}[x]$. Define

$$f(A) = (a_k A^k + \ldots + a_0 I) \in \mathcal{M}_{n\times n}(\mathbb{F})$$

Remark 2.8. Since $A^q A^p = A^p A^q$ and $\lambda A = A\lambda$ for $\lambda \in \mathbb{F}$ and $A \in \mathcal{M}_n(\mathbb{F})$, for all $f(x), g(x) \in \mathbb{F}[x]$,

$$f(A)g(A) = g(A)f(A)$$

Also, if $Av = \lambda v$ for some $v \in \mathbb{F}^n$, then $f(A)v = f(\lambda)v$.

Lemma 2.9. *For every* $A \in \mathcal{M}_n(\mathbb{F})$*, there exists a polynomial* $f(x) \in \mathbb{F}[x]$ *such that* $f(A) = 0$*.*

Proof. Note that $\dim \mathcal{M}_n(\mathbb{F}) = n^2 < \infty$. Hence the set $\{I, A, A^2, \ldots, A^k\}$ for $k > n^2$ is linearly dependent. Thus there exists $a_i \in \mathbb{F}$ such that

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

□

Definition 2.10 (Minimal polynomial)**.** The *minimal polynomial* $m_A(x)$ of A is the monic polynomial $p(x)$ of least degree such that $p(A) = 0$.

Theorem 2.11. *If* $f(A) = 0$ *then* $m_A(x) \mid f(x)$*. Further,* $m_A(x)$ *is unique.*

Proof. By the division algorithm, there exist $q(x), r(x) \in \mathbb{F}[x]$ with $\deg r(x) < \deg m_A(x)$ such that

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

Evaluating this at A gives $r(A) = 0$. Thus by the minimality of m_A, $r \equiv 0$.

For uniqueness, if m is 'another' minimal polynomial, then $m(A) = 0$. So by the above, $m_A(x) \mid m(x)$. Thus $m(x) = am_A(x)$ for some $a \in \mathbb{F}$, and by the fact that the minimal polynomial is monic we have that $a = 1$. □

Definition 2.12 (Characteristic polynomial)**.** The *characteristic polynomial* $\chi_A(x)$ is defined by

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

Lemma 2.13.

$$\chi_A(x) = (-1)^n x^n + (-1)^{n-1}\mathrm{tr}A + (\textit{intermediary terms}) + \det A$$

Definition 2.14 (Eigenvalues)**.** Let $A \in \mathcal{M}_n(\mathbb{F}$. Then λ is an *eigenvalue* of A if there exists some non-zero $v \in \mathbb{F}^n$ such that $Av = \lambda v$. Then also v is an *eigenvector* of A.

Theorem 2.15. *The following 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, (i) $\iff$ (ii):

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

Next, (i) $\implies$ (iii):

$$\begin{aligned} \exists v \neq 0 \text{ such that } Av = \lambda v &\implies m_A(\lambda)v = m_A(A)v = 0 \\ &\implies m_A(\lambda) = 0 \text{ (as } v \neq 0) \end{aligned}$$

Finally, (iii) $\implies$ (i):

$$m_A(\lambda) = 0 \implies m_A(x) = (x - \lambda)g(x)$$

for some $g(x)$. By minimality of m_A, we have that $g(A) \neq 0$. Hence there exists some $w \in \mathbb{F}^n$ such that $g(A)w \neq 0$. Let $v = g(A)w$, then

$$(A - \lambda I)v = m_A(A)w = 0$$

□

3 Quotient spaces

Let V be a vector space over a field $\mathbb{F}$, and $U \subseteq V$ a subspace.

Definition 3.1 (Quotient spaces)**.** The set of cosets

$$V/U = \{v + U \mid v \in V\}$$

with the operations defined, for $v, w \in U$ and $a \in \mathbb{F}$, by

$$\begin{aligned}(v + U) + (w + U) &= (v + w) + U \\ a(v + U) &= av + U\end{aligned}$$

form a vector space, called the *quotient space*.

Remark 3.2. It remains to prove that the operations are well defined. The fact that they satisfy the vector space axioms follows immediately from the fact that the operations in V satisfy them. Our concern is instead that two different representations of the same coset might lead to different results.

Assume that $v + U = v' + U$ and $w + U = w' + U$. Then $v = v' + \hat{u}$ and $w = v' + \tilde{u}$ for some $\hat{u}, \tilde{u} \in U$. We can then show that $(v + U) + (w + U) = (v' + U) + (w' + U)$ and $a(v + U) = a(v' + U)$.

Proposition 3.3. *Let $\mathcal{E}$ be a basis of U and $\mathcal{B}$ a basis of V such that $\mathcal{E} \subseteq \mathcal{B}$. Define*

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

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

Proof. Let $v + U \in V/U$. Then there exist some $k, n \in \mathbb{N}$, $a_i \in \mathbb{F}$, $e_1, \ldots, e_k \in \mathcal{E}$, and $e_{k+1}, \ldots, e_n \in \mathcal{B} \setminus \mathcal{E}$ such that

$$v = a_1 e_1 + \ldots + a_k e_k + a_{k+1} e_{k+1} + \ldots + a_n e_n$$

and thus

$$v + U = (a_{k+1} e_{k+1} + \ldots + a_n e_n) + U = a_{k+1}(e_{k+1} + U) + \ldots + a_n(e_n + U)$$

hence $\overline{\mathcal{B}}$ is spanning.

Now assume that, for some $a_i \in \mathbb{F}$ and $e_i \in \mathcal{B} \setminus \mathcal{E}$,

$$\begin{aligned}a_1(e_1 + U) + \ldots + a_n(e_n + U) = U &\implies a_1 e_1 + \ldots + a_n e_n \in U \\ &\implies a_1 e_1 + \ldots + a_n e_n = b_1 e_1' + \ldots + b_k e_k' \\ &\quad \text{(for some } b_i \in \mathbb{F} \text{ and } e_i' \in \mathcal{E}\text{)} \\ &\implies a_1 = \ldots = a_n = -b_1 = \ldots = -b_k = 0 \\ &\text{(as } \mathcal{B} \text{ is linearly independent)}\end{aligned}$$

hence $\overline{\mathcal{B}}$ is linearly independent. ☐

Corollary 3.4. *If V is finite dimensional then* $\dim V = \dim U + \dim V/U$.

Theorem 3.5 (First Isomorphism Theorem for vector spaces)**.** *Let $T : V \to W$ be a linear map of vector spaces over $\mathbb{F}$. Then $\overline{T} : V/\ker T \to \mathrm{im}T$ given by*

$$v + \ker T \mapsto T(v)$$

is a linear isomorphism.

Proof. It follows from the First Isomorphism Theorem for groups that $\overline{T}$ is an isomorphism of abelian groups. We can prove it in greater detail, by showing that $\overline{T}$ is well defined, linear, and bijective. □

Corollary 3.6 (Rank-Nullity Theorem)**.** *If $T : V \to W$ is a linear transformation and V is finitely dimensional, then* $\dim V = \dim\ker T + \dim \mathrm{im}T$*.*

Proof. We have that $\dim V = \dim U + \dim V/U$. Let $U = \ker T$, then by the First Isomorphism Theorem, we have that $\dim V/U = \dim \mathrm{im}T$. □

For what follows, let $T : V \to W$ be a linear transformation, and let $A \subseteq V$, $B \subseteq W$ be linear subspaces.

Lemma 3.7. *The formula $\overline{T}(v+A) = T(v)+B$ defines a linear map of quotients $\overline{T} : V/A \to W/B$ iff $T(A) \subseteq B$.*

Proof. Assume that $T(A) \subseteq B$. Then $\overline{T}$ will be linear if it is well defined, which we can show by letting $v + A = v' + A$, for some $v, v' \in A$. Then $v = v' + a$, for some $a \in A$. So

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

Now, if there exists some $a \in A$ with $T(a) \notin B$, and we assume that $\overline{T}$ is a linear map of quotients, then

$$B = 0 + B = \overline{T}(A) = \overline{T}(a + A) = T(a) + B$$

which is a contradiction, as $B \neq T(a) + B$ by our assumption. □

Now (as before) let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be a basis for V, with $\mathcal{B} \setminus \mathcal{E} = \{e_1, \ldots, e_k\}$ a basis for A. Similarly, let $\mathcal{B}' = \{e'_1, \ldots, e'_m\}$ be a basis for W, with $\mathcal{B}' \setminus \mathcal{E}' = \{e'_1, \ldots, e'_\ell\}$ a basis for B.

Then induced bases for V/A and W/B are given by $\overline{\mathcal{B}} = \{e_{k+1} + A, \ldots, e_n + A\}$ and $\overline{\mathcal{B}'} = \{e'_{\ell+1} + B, \ldots, e'_m + B\}$, respectively.

Let the matrix for ${}_{\mathcal{B}'}[T]_{\mathcal{B}}$ be given (as before) by $(a_{ij})_{1\le i\le m, 1\le j\le n}$, where the a_{ij} satisfy $T(e_j) = a_{1j}e'_1 + \ldots + a_{mj}e'_m$.

Lemma 3.8. *The matrix ${}_{\overline{\mathcal{B}'}}[\overline{T}]_{\overline{\mathcal{B}}}$ is given by $(a_{ij})_{\ell+1\le i\le m, k+1\le j\le n}$.*

Proof.

$$\begin{aligned}\overline{T}(e_j + A) &= T(e_j) + B \\ &= a_{1j}e'_1 + \ldots + a_{mj}e'_m + B \\ &= a_{(\ell+1)j}(e_{\ell+1} + B) + \ldots + a_{mj}(e'_m + B)\end{aligned}$$

□

As $T(A) \subseteq B$, we can restrict T to a linear map $T|_A : A \to B$, with $T|_A(v) = T(v)$ for $v \in A$. Then, in summary, we have the block matrix decomposition:

$$_{\mathcal{B}'}[T]_{\mathcal{B}} = \left(\begin{array}{c|c} _{\mathcal{E}'}[T|_A]_{\mathcal{E}} & * \\ \hline 0 & _{\overline{\mathcal{B}'}}[\overline{T}]_{\overline{\mathcal{B}}} \end{array}\right)$$

4 Triangular form and Cayley-Hamilton

Definition 4.1 (T-invariance)**.** Let $T : V \to V$ be a linear transformation. A subspace $U \subseteq V$ is *T-invariant* if $T(U) \subseteq U$

Lemma 4.2. *Let U be a T- and S-invariant subspace. Then U is also invariant under*

(i) The zero map

(ii) The identity map

(iii) aT, for all $a \in \mathbb{F}$

(iv) $S + T$

(v) $S \circ T$

In particular, U is invariant under any polynomial $p(x)$ evaluated at T. Hence $p(T)$ restricts to U, and also induces a map $\overline{p(T)} : V/U \to V/U$.

Proposition 4.3.

$$\chi_T(x) = \chi_{T|_U}(x) \cdot \chi_{\overline{T}}(x)$$

Proof. This follows from the block matrix decomposition at the end of the previous section. □

Remark 4.4. Note that Proposition 4.3 is not necessarily true for the minimal polynomial.

Definition 4.5 (Upper-triangular matrices)**.** A $n \times n$ matrix A is *upper triangular* if $a_{ij} = 0$ for $i > j$.

Theorem 4.6. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$, and let $T : V \to V$ be a linear map such that its characteristic polynomial is a product of linear factors. Then there exists some basis $\mathcal{B}$ of V such that the matrix of T with respect to this basis is upper triangular.*

Remark 4.7. If $\mathbb{F}$ is an algebraically closed field, then the characteristic polynomial will always satisfy the hypothesis.

Proof. We proceed by induction on n, and by using the block matrix decomposition from the last section. First note that when $n = 1$, the proof is trivial.

In general, χ_T has a root λ, and hence there exists some non-zero $v_1 \in V$ such that $T(v_1) = \lambda v_1$. Let $U = \operatorname{span}\{v_1\}$, and consider $\overline{T} : V/U \to V/U$. By Proposition 4.3, $\chi_{\overline{T}}$ is a product of linear factors, so by the induction hypothesis there exists some $\overline{\mathcal{B}} = \{v_2+U, \dots, v_n+U\}$ such that the matrix of $\overline{T}$ with respect to this basis is upper triangular.

Set $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$, then

$$
{}_{\mathcal{B}}[T]_{\mathcal{B}} = \left(\begin{array}{c|c} \lambda & * \\ \hline 0 & \\ \vdots & {}_{\overline{\mathcal{B}}}[\overline{T}]_{\overline{\mathcal{B}}} \\ 0 & \end{array}\right)
$$

is upper triangular. $\square$

Corollary 4.8. *If A is an $n \times n$ matrix with characteristic polynomial that is a product of linear factors, then there exists some invertible $n \times n$ matrix P such that $P^{-1}AP$ is upper triangular.*

Proposition 4.9. *Let A be an upper-triangular matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. Then*

$$
(A - \lambda_1 I) \dots (A - \lambda_n I) = 0
$$

Proof. Let $e_1, \dots, e_n$ be the standard basis vectors for $\mathbb{F}^n$. Then $(A - \lambda_n I)v \in \operatorname{span}\{e_1, \dots, e_{n-1}\}$ for all $v \in \mathbb{F}^n$. More generally, for all $w \in \operatorname{span}\{e_1, \dots, e_i\}$,

$$
(A - \lambda_i I)w \in \operatorname{span}\{e_1, \dots, e_{i-1}\}
$$

Hence, for all $v \in \mathbb{F}^n$,

$$
\underbrace{(A - \lambda_1 I) \underbrace{\dots \underbrace{(A - \lambda_{n-1} I) \underbrace{(A - \lambda_n I)}_{\in \operatorname{span}\{e_1, \dots, e_{n-1}\}}}_{\in \operatorname{span}\{e_1, \dots, e_{n-2}\}}}_{\in \operatorname{span}\{e_1\}}}_{\in 0} = 0
$$

$\square$

Theorem 4.10 (Cayley-Hamilton Theorem)**.** *If V is a finite-dimensional vector space over a field $\mathbb{F}$, and $T : V \to V$ is a linear transformation, then $\chi_T(T) = 0$. In particular, $m_T(x) \mid \chi_T(x)$.*

Proof. We work over the algebraic closure $\overline{\mathbb{F}}$. Hence $\chi_T(x) = (x-\lambda_1)\dots(x-\lambda_n)$ for some $\lambda_i \in \overline{\mathbb{F}}$. By Theorem 4.6, for some basis $\mathcal{B}$, we have that $A =_{\mathcal{B}} [T]_{\mathcal{B}}$ is upper triangular. Hence $\chi_T(T) = \chi_T(A) = 0$, by Proposition 4.9.

As the minimal polynomial divides all polynomials $p(x)$ with $p(T) = 0$, then in particular $m_T(x) \mid \chi_T(x)$. □

5 Primary Decomposition Theorem

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

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

Proof. Without loss of generality, we can assume that $\deg a \geq \deg b$, and that $\gcd(a, b) = 1$. We proceed by induction on $\deg a + \deg b$.

By the division algorithm, there exist some $q, r \in \mathbb{F}[x]$, with $\deg r < \deg b$, such that

$$a(x) = q(x)b(x) + c(x)$$

Then $\deg r + \deg b > \deg a + \deg b$, and $\gcd(b, r) = 1$ as $\gcd(a, b) = 1$.

Now, if $r(x) \equiv 0$, then $b(x) = \lambda$ (as $\gcd(a, b) = 1$) and

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

and we are done. So assume that $r \neq 0$. Then, by the induction hypothesis, there exist $s', t' \in \mathbb{F}[x]$ such that

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

and hence

$$\begin{aligned} 1 &= s'b + t'(a - qb) \\ &= t'a + (s' - q't)b \end{aligned}$$

and we are done. □

Lemma 5.2. *Let V be finite dimensional, and $T : V \to V$ a linear transformation. Let $W_1, \dots, W_r$ be T-invariant subspaces of V, such that $V = W_1 \oplus \dots \oplus W_r$. Let $\mathcal{B}_i$ be a basis for W_i, and then $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ is a basis for V. Then*

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

where $A_i = {}_{\mathcal{B}_i}[T|_{W_i}]_{\mathcal{B}_i}$.

Futher,

$$\chi_T(x) = \chi_{T|_{W_1}}(x) + \ldots + \chi_{T|_{W_r}}(x)$$

Proposition 5.3. *Assume that* $f(x) = a(x)b(x)$*, 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. As $\gcd(a,b) = 1$, there exist some $s, t \in \mathbb{F}[x]$ with $as + bt = 1$. Then $a(T)s(T) + b(T)t(T) = \mathrm{Id}$, and so

$$v = a(T)s(T)v + b(T)t(T)v, \quad \forall v \in V \qquad (*)$$

Now note that $a(T)\big(b(T)t(T)v\big) = f(T)t(T)v = 0$, and so $b(T)t(T) \in \ker a(T)$, and similarly for $a(T)s(T) \in \ker b(T)$. Thus

$$v = \ker a(T) + \ker b(T)$$

Assume that $v \in \ker a(T) \cap \ker b(T)$, then, by $(*)$, $v = 0 + 0 = 0$. Thus $v = \ker a(T) \oplus \ker b(T)$.

Finally, for $v \in \ker a(T)$, we have that

$$a(T)T(v) = T(a(T)v) = T(0) = 0$$

and similarly for $v \in \ker b(T)$. Hence the decomposition is T-invariant. □

Addendum 5.4. *If* $f(x) = m_T(x)$ *is the minimal polynomial in Proposition 5.3, then, furthermore,*

$$m_{T|_{\ker a(T)}}(x) = a(x) \quad \textit{and} \quad m_{T|_{\ker b(T)}}(x) = b(x)$$

Proof. Let $m_1 = m_{T|_{\ker a(T)}}(x)$ and $m_2 = m_{T|_{\ker b(T)}}(x)$. Then $m_1 \mid a$, as $a(T)|_{\ker a(T)} = 0$, and similarly for $m_2 \mid b$. As any $v \in V$ can be written as $v = w_1 + w_2$, for $w_1 \in \ker a(T)$ and $w_2 \in \ker b(T)$, we have that, for all $v \in V$,

$$\begin{aligned} m_1(T)m_2(T)v &= m_2(T)\big(m_1(T)w_1\big) + m_1(T)\big(m_2(T)w_2\big) \\ &= 0 + 0 = 0 \end{aligned}$$

and thus $m \mid m_1 m_2$. Hence, for reasons of degree, $m_1 = a$ and $m_2 = b$. □

Theorem 5.5 (Primary Decomposition Theorem)**.** *Assume that the minimal polynomial has the form* $m_T(x) = f_1(x)^{m_1} \ldots f_r(x)^{m_r}$*, where the* f_i *are distinct, irreducible, monic polynomials. Let* $W_i = \ker f_i(T)^{m_i}$*. Then*

1. W_i *is* T*-invariant*
2. $V = W_1 \oplus \ldots \oplus W_r$
3. $m_{T|_{W_i}} = f_i^{m_i}$

Further, $\chi_T = f_1^{n_1} \ldots f_r^{n_r}$*, where* $n_i \geq m_i$.

Proof. Put $a = f_1 \dots f_{r-1}$ and $b = f_r$, and proceed by induction on r. The proof of the statement about χ_T has been left, lovingly, as an exercise for the reader (and can be found as an answer to one of the questions on one of the problem sheets). $\square$

Remark 5.6. We have so far proved that

$$\begin{aligned} T \text{ is triangularisable} &\iff \chi_T \text{ factors as a product of linear polynomials} \\ &\iff \text{each } f_i \text{ is linear} \\ &\iff m_T \text{ factors as a product of linear polynomials} \end{aligned}$$

Corollary 5.7. *Let $f_1, \dots, f_r$ be distinct, irreducible, monic polynomials. Then $m_T(x) = f_1(x)^{m_1} \dots f_r(x)^{m_r}$, with $m_i > 0$, iff $\chi_T(x) = f_1(x)^{n_1} \dots f_r(x)^{n_r}$, with $n_i \geq m_i$.*

Proof. By the Cayley-Hamilton Theorem, $m_T \mid \chi_T$. Hence $n_i \geq m_i$, and $\chi_T(x) = f_1(x)^{n_1} \dots f_r(x)^{n_r} b(x)$, with b coprime to $a = f_1^{n_1} \dots f_r^{n_r}$. By Proposition 5.3, $V = \ker a(T) \oplus \ker b(T)$. (Note that $V = \ker m_T(T) \subseteq \ker a(T)$ and $\ker b(T) = 0$). But also, by Addendum 5.4, $b(x) = \chi_{T|_{\ker b(T)}}$, and $\deg b(x) = \dim \ker b(T)$. Hence $b(x) \equiv 1$. $\square$

Theorem 5.8. *Let $T : V \to V$ be a linear transformation on a finite-dimensional vector space V. Then T is diagonalisable iff*

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

for some distinct $\lambda_i \in \mathbb{F}$.

Proof. By the Primary Decomposition Theorem, $V = \ker(T - \lambda_1 I) \oplus \dots \oplus \ker(T - \lambda_r I) = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_r}$, where E_{λ_i} is the λ_i-eigenspace. Let $\mathcal{B}_i$ be a basis for E_{λ_i}, then $\mathcal{B} = \bigcup_{i=1}^r \mathcal{B}_i$ is a basis of eigenvectors of T for V, and

$${}_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & & \\ & & & \lambda_2 & & \\ & & & & \ddots & \\ & & & & & \lambda_r \end{pmatrix}$$

is diagonal.

Vice versa, if T is diagonalisable then there is a basis $\mathcal{B} = \bigcup_{i=1}^n \mathcal{B}_i$ of eigenvectors, as above, and $V = E_{\lambda_1} + \oplus \dots \oplus E_{\lambda_r}$, with distinct λ_i. Let $f(x) = (x - \lambda_1) \dots (x - \lambda_r)$. Then, for any $v = v_1 + \dots + v_r$, with $v_i \in E_{\lambda_i}$,

$$f(T)v = \sum_{i=1}^r \left[\left(\prod_{i \neq j} (T - \lambda_j I) \right) (T - \lambda_i I) v_i \right] = 0$$

and thus $m_T \mid f$. Recall that, if $T(v) = \lambda v$, then for any polynomial $g(x)$, we have that $g(T)v = g(\lambda)v$. Hence here, every λ_i is a root of m_T. Thus $m_T(x) = f(x) = (x - \lambda_1) \dots (x - \lambda_r)$ $\square$

6 Jordan Normal Form

Definition 6.1 (Nilpotent transformations)**.** Let V be finite dimensional and $T : V \to V$ a linear transformation. If $T^m = 0$ for some $m > 0$ then T is *nilpotent.*

Theorem 6.2. *If T is nilpotent and $m_T(x) = x^m$, then there exists a basis $\mathcal{B}$ of V such that*

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

(that is, zeros everywhere except just above the leading diagonal) where $* \in \{0, 1\}$

Proof. First, we note that $0 \subsetneq \ker T \subsetneq \ker T^2 \subsetneq \ldots \subsetneq \ker T^m = V$. Now let $\mathcal{B}_i$ be such that

$$\overline{\mathcal{B}_i} := \{w + \ker T^{i-1} \mid w \in \mathcal{B}_i\} \text{ is a basis for } \frac{\ker T^i}{\ker T^{i-1}}$$

Now we make, and prove, two claims:

(i) $\mathcal{B} = \bigcup_{i=1}^m \mathcal{B}_i$ *is a basis for V*: This follows by induction from Proposition 3.3

(ii) $\{Tw + \ker T^{i-1} \mid w \in \mathcal{B}_{i+1}\}$ *is linearly independent in* $\ker T^i / \ker T^{i-1}$: Assume that we have $\sum_s (a_s T(w_s) + \ker T^{i-1}) = \ker T^{i-1}$. Then

$$\begin{aligned} \sum a_s T(w_s) \in \ker T^{i-1} &\implies T(\sum a_s w_s) \in \ker T^{i-1} \\ &\implies \sum a_s w_s \in \ker T^i \\ &\implies \sum a_s w_s + \ker T^i = \ker T^i \\ &\implies a_s = 0 \forall s \quad \text{by the definition of } \mathcal{B} \end{aligned}$$

So, inductively find $\mathcal{E}_i = \{w_1^i, \ldots, w_{k_i}^i\}$ such that, for $\mathcal{B}_i = \mathcal{E}_i \sqcup T(\mathcal{B}_{i+1})$, $\overline{\mathcal{B}_i}$ is a basis for $\ker T^i / \ker T^{i-1}$. Then

$$\mathcal{B} = \bigcup \mathcal{B}_i = \bigcup_{w \in \mathcal{E}_m} \{T^{m-1}w, \ldots, T(w), w\} \ldots \bigcup_{w \in \mathcal{E}_1} \{w\}$$

is a basis of V, and

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

is block diagonal, with $|\mathcal{E}_i| = k_i$ many Jordan blocks of size j, and where a Jordan block is the $i \times i$ matrix

$$J_i = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

□

Corollary 6.3. *Let $T : V \to V$ be a linear transformation for some finite-dimensional V. Assume that $m_T(x) = (x - \lambda)^m$ for some m. Then there exists a basis $\mathcal{B}$ of V such that ${}_{\mathcal{B}}[T]_{\mathcal{B}}$ is block diagonal, with blocks of the form $\lambda I + J_i$.*

Proof. $T - \lambda I$ is nilpotent and is of the form described in Theorem 6.2. So there exists a basis $\mathcal{B}$ such that ${}_{\mathcal{B}}[T - \lambda I]_{\mathcal{B}}$ is block diagonal, with blocks J_i. Finally, ${}_{\mathcal{B}}[T]_{\mathcal{B}} ={}_{\mathcal{B}} [T - \lambda I] + \lambda I$. □

In the appendix there are two examples of putting a matrix into its Jordan normal form, in an aim to elucidate the method of the proof of the theorem.

Lemma 6.4. *Let $v_n = (v_n^1, \dots, v_n^k)$ and $J_k(\lambda)$ the $j \times j$ Jordan block with λ on the leading diagonal. Consider $v_n = J_k(\lambda) v_{n-1} = (J_k(\lambda))^n v_0$. 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. Proceed by induction: the case is true for $n = 0$. Then

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

as $\binom{n-1}{j} + \binom{n-1}{j-1} = \binom{n}{j+1}$. □

7 Dual spaces

Definition 7.1 (Dual spaces)**.** Let V be a vector space over $\mathbb{F}$. The *dual*, V', is the vector space of linear maps from V to $\mathbb{F}$. That is, $V' = \hom(V, \mathbb{F})$, and its elements are called *linear functionals*.

Theorem 7.2. *Let V be finite dimensional, and $\mathcal{B} = \{e_1, \dots, e_n\}$ be a basis for V. Define the* dual, *e_i' of e_i (relative to $\mathcal{B}$) by*

$$e_i'(e_j) = \delta_{ij}$$

Then $\{e_1', \dots, e_n'\}$ is a basis for V', the dual basis*. In particular, the assignment $e_i \mapsto e_i'$ defines an isomorphism of vector spaces.*

Proof. Assume that $\sum a_i e_i' = 0$, then, for all j, $0 = \sum a_i e_i'(e_j) = a_j$. So we have linear independence.

To show that the set is spanning, assume that $f \in V'$. Let $a_i = f(e_i)$, then $f = \sum a_i e_i'$, as both f and the sum evaluate to a_i on e_i, and any linear map is determined by the values it takes on a basis. □

Theorem 7.3. *Let V be a finite-dimensional vector space. Then $V \to V''$ defined by $v \mapsto E_v$ is a natural linear isomorphism, where E_v is the evaluation map at v. That is, $E_v(f) = f(v)$.*

Remark 7.4. Here 'natural' means independent of a choice of basis. In contrast, the isomorphism $V \cong V'$ is dependent on the choice of basis for V.

Proof. Clearly E_v is a linear map, so it remains to show that it is both injective and surjective.

(i) *Injective*: Assume that $E_v(f) \equiv 0$. Then $E_v(f) = f(v) = 0$ for all $f \in V'$, and hence $v = 0$. (For if $v \neq 0$ then let $e_1 = v$, extend this to a basis for V, and for $f = e_1'$, we would have $E_v(e_1') = 1$, which is a contradiction).

(ii) *Surjective*: This follows from the fact that $\dim V = \dim V' = \dim(V')'$, and from injectivity and the Rank-Nullity Theorem.

□

Definition 7.5 (Annihilators)**.** Let $U \subseteq V$. The *annihilator* of U is defined to be

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

Proposition 7.6. *U^0 is a subspace of V'.*

Proof. Let $f, g \in U^0$ and $\lambda \in \mathbb{F}$. Then, for all $u \in U$,

$$(f + \lambda g)(u) = f(u) + \lambda g(u) = 0$$

and hence $f + \lambda g \in U^0$. Also note that $0 \in U^0$, so $U^0 \neq \varnothing$. □

Theorem 7.7. *Let V be finite dimensional. Then $\dim U^0 = \dim V - \dim U$.*

Proof. Let $\{e_1, \ldots, e_m\}$ be a basis for U, and extend it to a basis $\{e_1, \ldots, e_n\}$ for V. Let $\{e_1', \ldots, e_n'\}$ be the dual basis, and $f \in U^0$. Then there exists some $a_i \in \mathbb{F}$ such that $f = \sum a_i e_i'$. For $i = 1, \ldots, m$, we have that $f(e_i) = a_i = 0$, as $e_i \in U$. For $j = m+1, \ldots, n$, we have that $e'j \in U^0$, as, for $i = 1, \ldots, m$ we have that $e_j'(e_i) = 0$. Hence $\{e_{m+1}', \ldots, e_n'\}$ span U^0, but as a subset of the dual basis it is also linearly independent, and hence a basis. Thus $\dim U^0 = n - m = \dim V - \dim U$. □

Theorem 7.8. *Let $U, W \subseteq V$. Then*

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

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

(iii) $(U \cap W)^0 = U^0 + W^0$ *(if V is finite dimensional)*

Proof. (i)

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

(ii)

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

(iii)

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

Then

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

□

Theorem 7.9. *Let $U \subseteq V$, and V be finite dimensional. Under the natural idendification $V \cong V''$, given by $v \mapsto E_v$, we have that $U = U^{00}$*

Proof. $E_x \in U^{00}$ iff $E_x(f) = f(x) = 0$ for all $f \in U^0$. Hence, if $x \in U$ then $E_x \in U^{00}$, and so $U \subseteq U^{00}$. But also

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

thus $U = U^{00}$. □

Theorem 7.10. *Let $U \subseteq V$, and V be finite dimensional. Then there exists a natural isomorphism $U' \cong V'/U^0$*

Proof. Consider $\psi : V' \to U'$ given by $f \mapsto f|_U$. Then ψ is clearly linear. Further,

$$f \in \ker\psi \iff f|_U = 0 \iff f \in U^0$$

and applying the First Isomorphism Theorem gives

$$\psi : \frac{V'}{U^0} \xrightarrow{\sim} \mathrm{im}\psi \subseteq U'$$

As V is finite dimensional, any basis $\{e_1, \ldots, e_k\}$ of U can be extended to a basis $\{e_1, \ldots, e_n\}$ of V. Then any $g \in U'$ is the image of $\tilde{g} \in V'$, defined by

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

□

Definition 7.11 (Dual maps)**.** Let $T : V \to W$ be a linear transformation. Define the *dual map*:

$$T' : W' \to V', \text{ given by } f \mapsto f \circ T$$

Note that $f \circ T : V \to W \to \mathbb{F}$ is linear, and thus $f \circ T \in V'$.

Proposition 7.12. *T' is a linear map.*

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

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

□

Proposition 7.13. *The map* $\hom(V, W) \to \hom(W', V')$ *given by* $T \mapsto T'$ *is linear.*

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

$$\begin{aligned} \big((T + \lambda S)'(f)\big)(v) &= f(T + \lambda S)(v) \\ &= f(Tv + \lambda Sv) \\ &= f(Tv) + \lambda f(Sv) \\ &= T'(fv) + \lambda S'(fv) \\ &= (T' + \lambda S')(f)(v) \end{aligned}$$

and thus $(T + \lambda S)' = T' + \lambda S'$. □

Theorem 7.14. *Let V, W be finite dimensional. Then $T \mapsto T'$ defines a natural isomorphism between* $\hom(V, W)$ *and* $\hom(W', V')$.

Proof. Assume that $T' = 0$. Then $T'(f)(v) = f(Tv) = 0$, for all $f \in W'$ and $v \in V$. But then $Tv = 0$ for all $v \in V$, and hence $T = 0$. Thus $T \mapsto T'$ is injective.

Further,

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

and so the map is also surjective. □

Remark 7.15. In the above proof we used the fact that V' seperates the elements of V in the sense that

$$f(v) = 0\ \forall f \in V' \implies v = 0$$

and

$$v \neq 0 \implies \exists f \in V' \colon f(v) \neq 0$$

Theorem 7.16. *Let V and W be finite dimensional, with respective bases $\mathcal{B}_V$ and $\mathcal{B}_W$. 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. Let $\mathcal{B}_V = \{e_1, \ldots, e_n\}$ and $\mathcal{B}_W = \{x_1, \ldots, x_m\}$. Write ${}_{\mathcal{B}_W}[T]_{\mathcal{B}_V} = A = (a_{ij})_{ij}$. Then $T(e_j) = \sum_i a_{ij}x_i$, and $x_i'(T(e_j)) = a_{ij}$. Now let ${}_{\mathcal{B}_{V'}}[T']_{\mathcal{B}_{W'}} = B = (b_{ij})_{ij}$. Then $T'(x_i') = \sum_j b_{ji}e_j'$, and $T'(x_i')(e_j) = b_{ji}$. Thus $a_{ij} = b_{ji}$. □

8 Bilinear forms and inner products

Definition 8.1 (Bilinear forms)**.** 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) $F(\lambda v, w) = \lambda F(v, w) = F(v, \lambda w)$

Further, F is

(i) *symmetric* if, for all $v, w \in V$, $F(v, w) = F(w, v)$

(ii) *non degenerate* if $F(v, w) = 0\ \forall v \in V \implies w = 0$

(iii) *positive definite* if, for all $v \neq 0$, $F(v, v) > 0$

Note that non degeneracy follows from positive definiteness.

Definition 8.2 (Sesquilinear forms)**.** 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 $u, v, w \in \mathbb{C}$ 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) $F(\bar{\lambda} v, w) = \lambda F(v, w) = F(v, \lambda w)$

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

Definition 8.3 (Inner product spaces)**.** A real (complex) vector space V with a bilinear (sesquilinear), symmetric (conjugate-symmetric), positive-definition form $F = \langle , \rangle$ in an *inner product space.* Then $\{w_1, \dots, w_n\}$ are *mutually orthogonal* if $\langle w_i, w_j \rangle = 0$ for all $i \neq j$. Further, $\{w_1, \dots, w_n\}$ are *orthonormal* if they are orthogonal and $\langle w_i, w_i \rangle = 1$ for all i.

Proposition 8.4. *Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$, and $\{w_1, \dots, w_n\}$ an orthogonal set with $w_i \neq 0$ for all i. Then $\{w_1, \dots, w_n\}$ is linearly independent.*

Proof. Assume that $\sum_i \lambda_i w_i = 0$ for some $\lambda_i \in \mathbb{F}$. Then, for all j,

$$\begin{aligned} \langle w_j, \sum_i \lambda_i w_i \rangle = 0 &\implies \langle w_j, \lambda_j w_j \rangle = \lambda_j \langle w_j, w_j \rangle = 0 \\ &\implies \lambda_j = 0 \end{aligned}$$

□

Theorem 8.5 (Gram-Schmidt orthonormalisation process)**.** *Let $\{v_1, \dots, v_n\}$ be a basis of the inner product space V. Set*

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

Then $\{w_1, \dots, w_n\}$ is an orthonormal basis of V.

Proof. Prove by induction on $\{w_1, \dots, w_k\}$ that this set spans V, and then use Proposition 8.4 to show linear independence. □

Theorem 8.6. *Let V be an inner product space over $\mathbb{R}$. Then the map $v \mapsto \langle v, \_\rangle$ is a natural injective linear map $\phi : V \to V'$, which is an isomorphism if V is finite dimensional.*

Proof. Firstly, for all $v \in V$, the map $\langle v, \_\rangle : V \to \mathbb{R}$ is a linear functional, as $\langle , \rangle$ is linear in the second argument. That is, ϕ is linear.

As $\langle , \rangle$ is non degenerate, $\langle v, \_\rangle = \langle *, v\rangle = 0$ iff $v = 0$. Hence ϕ is injective.

If V is finite dimensional then $\dim V = \dim V'$, and hence $\mathrm{im}\phi = V'$. □

Definition 8.7 (Orthogonal complements)**.** Let $U \subseteq V$ be a finite-dimensional subspaces of the inner product space V. The *orthogonal complement* of U is defined as

$$U^\perp := \{V \in V \mid \langle u, v\rangle = 0 \ \forall u \in U\}$$

Proposition 8.8. *$U^\perp$ is a subspace of V.*

Proof. Let $v, w \in U^\perp$ and $\lambda \in \mathbb{C}$. Then, for all $u \in U$,

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

□

Proposition 8.9. *Let V be an inner product space, and $U \subseteq V$. Then*

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

(ii) $U \oplus U^\perp = V$*, if* $\dim V < \infty$

(iii) $\dim U^\perp = \dim V - \dim U$*, if* $\dim V < \infty$

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

(v) $U^\perp + W^\perp \subseteq (U \cap W)^\perp$ *(with equality if* $\dim V < \infty$*)*

(vi) $U \subseteq (U^\perp)^\perp$ *(with equailty if* $\dim V < \infty$*)*

Proof. (i) Let $u \in U \cap U^\perp$. Then $\langle u, u\rangle = 0$, so $u = 0$.

(ii) If $\dim V < \infty$ then there exists some orthonormal basis $\{e_1, \dots, e_n\}$ for V such that $\{e_1, \dots, e_k\}$ is a basis for U. Assume that $v = \sum_i a_i e_i \in U^\perp$. Then $\langle e_i, v\rangle = a_i = 0$ for $i = 1, \dots, k$. Hence $v \in \mathrm{span}\{e_{k+1}, \dots, e_n\}$.

Vice versa, $e_j \in U^\perp$, for $j = k+1, \dots, n$, and hence $U^\perp = \mathrm{span}\{e_{k+1}, \dots, e_n\}$.

(iii) See problem sheets

(iv) See problem sheets

(v) Let $u_0 \in U$. Then, for all $w \in U^\perp$, $\langle u_0, w\rangle = \overline{\langle w, u_0\rangle} = 0$, and hence $\langle w, u_0\rangle = 0$. Thus $u_0 \in (U^\perp)^\perp$.

If $\dim V < \infty$ then $\dim V - \dim U^\perp = \dim U$, and so $U = (U^\perp)^\perp$.

□

Proposition 8.10. *Let V be a finite-dimensional vector space over $\mathbb{R}$. Then under the isomorphism $\phi : V \to V'$, given by $v \mapsto \langle v, \_\rangle$, for $U \subseteq V$, $U^\perp \mapsto U^0$.*

Proof. Let $v \in U^\perp$. Then, for all $u \in U$, $\langle u, v\rangle = \langle v, u\rangle = 0$, and thus $\langle v, \_\rangle \in U^0$. Further, $\dim U^\perp = \dim V - \dim U = \dim U^0$. $\square$

9 Adjoint maps

Let V be an inner product space over $\mathbb{K}$, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 9.1 (Adjoints)**.** 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 9.2. *If T^* exists, then it is unique.*

Proof. Let $\tilde{T}$ be another map satisfying the adjoint property. Then, for all $v, w \in V$,

$$\langle T^*v - \tilde{T}v, w\rangle = \langle T^*v, w\rangle - \langle \tilde{T}v, w\rangle = \langle v, Tw\rangle - \langle v, Tw\rangle = 0$$

but as $\langle, \rangle$ is non degenerate, we have that $T^*v - \tilde{T}v = 0$ for all $v \in V$, and thus $T^* = \tilde{T}$. $\square$

Theorem 9.3. *Let $T : V \to V$ be linear, and $\dim V < \infty$. Then the adjoint exists and is linear.*

Proof. Fix $v \in V$ and consider the map $V \to \mathbb{K}$ given by $w \mapsto \langle v, Tw\rangle$. Note that $\langle v, T\_\rangle$ is a linear functional, as T is linear, and so is $\langle, \rangle$ in the second argument. As V is finite dimensional, $\phi : V \to V'$ given by $u \mapsto \langle u, \_\rangle$ is a linear isomorphism for $\mathbb{K} = \mathbb{R}$, and injective if $\mathbb{K} = \mathbb{C}$. Thus there exists some $u \in V$ such that $\langle v, T\_\rangle = \langle u, \_\rangle = \langle T^*v, \_\rangle$, where we define $T^*v = u$.

Then, for all $v_1, v_2, w \in V$, $\lambda \in \mathbb{K}$,

$$\begin{aligned}\langle T^*(v_1 + \lambda v_2), w\rangle &= \langle v_1 + \lambda v_2, Tw\rangle \\ &= \langle v_1, T2\rangle + \bar{\lambda}\langle v_2, Tw\rangle \\ &= \langle T^*v_1, w\rangle + \bar{\lambda}\langle T^*v_2, w\rangle \\ &= \langle T^*v_1 + \lambda T^*v_2, w\rangle\end{aligned}$$

and as $\langle, \rangle$ is non degenerate, we have that $T^*(v_1 + \lambda v_2) = T^*v_1 + \lambda T^*v_2$. $\square$

Proposition 9.4. *Let $T : V \to V$ be linear, and $\mathcal{B} = \{e_1, \ldots, e_n\}$ be an orthonormal basis for V. Then*

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

Proof. Let $A =_{\mathcal{B}} [T]_{\mathcal{B}}$ and $B =_{\mathcal{B}} [T^*]_{\mathcal{B}}$. Then $a_{ij} = \langle e_i, Te_j \rangle$, and

$$b_{ij} = \langle e_i, T^* e_j \rangle = \overline{\langle T^* e_j, e_i \rangle} = \overline{\langle e_j, Te_i \rangle} = \bar{a}_{ji}$$

and so $B = \bar{A}^t$. □

Remark 9.5. For $\mathbb{K} = \mathbb{R}$, under the isomorphism $\phi : V \to V'$, given by $v \mapsto \langle v, \_ \rangle$, T^* is identified with the dual map T', and if $\mathcal{B}'$ is the dual basis of some orthonormal basis $\mathcal{B}$ for V, then

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

Proposition 9.6. *Let $S, T : V \to V$ be linear transformations, $\lambda \in \mathbb{K}$, and $\dim V < \infty$. Then*

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

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

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

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

(v) $m_T^* = \overline{m_T}$

Definition 9.7 (Self-adjoint operators)**.** A linear map $T : V \to V$ is *self adjoint* if $T^* = T$. This is equivalent to saying that the matrix of T is *Hermitian.* That is, that $\bar{A}^t = A$.

Lemma 9.8. *Let λ be an eigenvalue of a self-adjoint operator. Then $\lambda \in \mathbb{R}$.*

Proof. Assume that $w \neq 0$ and $Tw = \lambda w$. Then

$$\begin{aligned} \lambda \langle w, w \rangle &= \langle w, \lambda w \rangle = \langle w, Tw \rangle = \langle T^* w, w \rangle \\ &= \langle Tw, w \rangle = \langle \lambda w, w \rangle = \bar{\lambda} \langle w, w \rangle \end{aligned}$$

and hence, as $\langle w, w \rangle \neq 0$, we have that $\lambda = \bar{\lambda}$. □

Lemma 9.9. *Let T be self adjoint, and $U \subseteq V$ be T-invariant. Then $U^\perp$ is T-invariant.*

Proof. Let $w \in U^\perp$ and $u \in U$. Then

$$\langle u, Tw \rangle = \langle T^* u, w \rangle = \langle Tu, w \rangle = 0$$

and so $Tw \in U^\perp$. □

Theorem 9.10. *Let V be finite dimensional, and $T : V \to V$ self adjoint. Then there exists an orthonormal basis of eigenvectors of T for V.*

Proof. The characteristic polynomial of T has a root over $\mathbb{C}$, but by Lemma 9.8 we have that this root, λ, is real. Thus, by induction, any $n \times n$ self-adjoint matrix over $\mathbb{K}$ has n eigenvalues (maybe not all distinct). Find v_1 such that $Tv_1 = \lambda v_1$, and define $V_1 = (\text{span}\{v_1\})^{\perp}$. Consider the restriction of T to this subspace: by Lemma 9.9, $T|_{V_1} : V_1 \to V_1$. Further, $T_|V_1$ is still self adjoint, so by induction on $\dim V$, there exists an orthonormal basis $\{e_2, \ldots, e_n\}$ of eigenvectors of $T|_{V_1}$ for V_1.

Set $e_i = v_i/\|v_i\|$, and then $\{e_1, \ldots, e_n\}$ is an orthonormal basis of eigenvectors of T for V. □

Corollary 9.11. *Any $n \times n$ matrix with $A = \bar{A}^t$ is diagonalisable by orthogonal matrices. That is, there exists some orthogonal P such that $P^{-1}AP$ is diagonal.*

10 Orthogonal and unitary transformations

Definition 10.1 (Orthogonal and unitary transformations)**.** Let V be a finite-dimensional vector space over $\mathbb{K}$, and $T : V \to V$ be a linear transformation. If $T^* = T^{-1}$ then T is called *orthogonal* if $\mathbb{K} = \mathbb{R}$, or *unitary* if $\mathbb{K} = \mathbb{C}$.

Remark 10.2. Let $\mathcal{B}$ be an orthonormal basis for $\mathbb{K}^n$ (under the usual inner product for $\mathbb{K}^n$), and $\mathcal{E}$ be the standard basis (which is also orthonormal under the usual inner product). Then a matrix whose columns entries are the coordinate representations of the $e_i \in \mathcal{B}$ with respect to $\mathcal{E}$ is orthogonal (unitary).

Theorem 10.3. *The following are equivalent:*

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

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

(iii) *T preserves length:* $\|v\| = \|Tv\|$

Corollary 10.4. *Orthogonal (unitary) linear transformations are isometries. That is,* $\mathrm{d}(v, w) = \|v - w\| = \|Tv - Tw\| = \mathrm{d}(Tv, Tw)$.

Proof.

$(i) \implies (ii)$: $\langle v, w\rangle = \langle \mathrm{Id}v, w\rangle = \langle T^*Tv, w\rangle = \langle Tv, Tw\rangle$

$(ii) \implies (iii)$: $\|v\|^2 = \langle v, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2$

$(ii) \implies (i)$: $\langle v, w\rangle = \langle Tv, Tw\rangle = \langle T^*Tv, w\rangle$. Thus $T^*Tw = w$, and by the non degeneracy of $\langle , \rangle$, we have that $T^*T = \mathrm{Id}$.

$(iii) \implies (i)$: By Proposition 10.5 (below). □

Proposition 10.5. *The length function determines the inner product. That is, for all $v, w \in V$,*

$$\langle v, v\rangle_1 = \langle v, v\rangle_2 \iff \langle v, w\rangle_2 = \langle v, w\rangle_2$$

Proof. The 'only if' direction is clear. For the other direction, note that, when $\mathbb{K} = \mathbb{R}$, $\langle v+w, v+w\rangle = \langle v,v\rangle + \langle v,w\rangle + \overline{\langle v,v\rangle} + \langle w,w\rangle$, and hence

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

When $\mathbb{K} = \mathbb{C}$, consider $\langle v+iw, v+iw\rangle = \langle v,v\rangle + i\langle v,w\rangle - i\overline{\langle v,v\rangle} + \langle w,w\rangle$, and hence

$$\begin{aligned}\mathrm{Re}\langle v,w\rangle &= \frac{1}{2}(\|v+w\|^2 - \|v\|^2 - \|w\|^2)\\ \mathrm{Im}\langle v,w\rangle &= \frac{1}{2}(\|v+iw\|^2 - \|v\|^2 - \|w\|^2)\end{aligned}$$

□

Definition 10.6. We define now some groups of matrices:

$$\begin{array}{rcll} O(n) &=& \{A \in \mathcal{M}_n(\mathbb{R}) \mid A^t A = \mathrm{Id}\} & \textit{orthogonal group} \\ SO(n) &=& \{A \in O_n \mid \det A = 1\} & \textit{special orthogonal group} \\ U(n) &=& \{A \in \mathcal{M}_n(\mathbb{C}) \mid \bar{A}^t A = \mathrm{Id}\} & \textit{unitary group} \\ SU(n) &=& \{A \in U_n \mid \det A = 1\} & \textit{special unitary group} \end{array}$$

Lemma 10.7. *Let λ be an eigenvalue of an orthogonal (unitary) linear transformation $T : V \to V$. Then $|\lambda| = 1$.*

Proof. Let $v \neq 0 \in V$ be a λ-eigenvector. Then $\langle v,v\rangle = \langle Tv,Tv\rangle = \langle \lambda v, \lambda v\rangle = \bar{\lambda}\lambda\langle v,v\rangle$. Thus $|\lambda|^2 = 1$. □

Corollary 10.8. *Let A be an orthogonal (unitary) matrix. Then $|\det A| = 1$.*

Proof. Working over $\mathbb{C}$, we know that $\det A$ is the product of eigenvalues of A. Then $|\det A| = |\lambda_1 \dots \lambda_k| = |\lambda_1| \dots |\lambda_k| = 1$. □

Lemma 10.9. *Let V be finite dimensional, $T : V \to V$ be an orthogonal (unitary) matrix, and $U \subseteq V$ be a T-invariant subspace.*

Proof. Let $u \in U$ and $w \in U^\perp$. Then $\langle u, Tw\rangle = \langle T^* u, w\rangle$, but $T^* = T^{-1} : U \to U$, as U is T-invariant. Hence $T^* u \in U$, and $\langle T^* u, w\rangle = 0$. Thus $Tw \in U^\perp$. □

Theorem 10.10. *Let V be finite dimensional, and $T : V \to V$ be unitary. (Thus here, V is a vector field over $\mathbb{C}$). Then there exists an orthonormal basis of eigenvectors of T for V.*

Proof. As we are working in $\mathbb{C}$, there exists some λ and $v \neq 0$ such that $Tv = \lambda v_1$. Define $U_1 = \mathrm{span}\{v_1\}$ and consider $T|_{U_1^\perp}$. By induction, there exists $\{e_2, \dots, e_n\}$ orthonormal basis of eigenvectors of $T|_{U_1^\perp}$ for $U_1^\perp$. Set $e_i = v_i/\|v_i\|$. Then $\{e_1, \dots, e_n\}$ is an orthonormal basis of eigenvectors of T for V. □

Corollary 10.11. *Let $A \in U(n)$. Then there exists $P \in U(n)$ such that $P^{-1}AP$ is diagonal.*

Remark 10.12. Note that $O(n) \subset U(n)$, so if $A \in O(n)$ then the above still holds, but the diagonal matrix might not have real entries. That is, orthogonal matrices are diagonalisable, but not necessarily over the reals.

Proposition 10.13. *Every $A \in O(n)$ can either be written as R_θ or S_θ, for some $\theta \in \mathbb{R}$, where*

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad S_\theta = \begin{pmatrix} \sin\theta & \cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix}$$

If $\det A = 1$ then $A = R_\theta$ and corresponds to a rotation, and if $\det A = -1$ then $A = S_\theta$ and corresponds to a reflection.

Theorem 10.14. *Let V be a finite-dimensional, real vector space, and $T : V \to V$ be orthogonal. Then there exists an orthonormal basis $\mathcal{B}$ such that*

$${}_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} I & & & & \\ & -I & & & \\ & & R_{\theta_1} & & \\ & & & \ddots & \\ & & & & R_{\theta_k} \end{pmatrix}$$

where $\theta_i \neq 0, \pi$.

N.B. The following proof is lacking in rigour. For an alternative proof see 'Elementary Geometry' by Roe.

Proof. Let $S = T + T^{-1} = T + T^*$. Then $S^* = T^* + T = S$, and thus S is self adjoint. Hence there is some orthonormal basis of eigenvectors of S for V, and $V = V_1 \oplus \ldots \oplus V_k$ decomposes into orthogonal eigenspaces (where $\lambda_i \neq \lambda_j$). Note that each V_i is T-invariant, and so we may restrict ourselves to $T|_{V_i}$.

On V_i, $(T+T^{-1})v = (T+T^*)v = \lambda_i v$, and hence $T^2 - \lambda_i T + I = 0$. Consider first the case that $\lambda_i = \pm 2$, and then $(T \pm I)^2 = 0$ on these V_i. This gives rise to the $\pm I$ in the matrix representation.

The other possibility is that $\lambda_i \neq \pm 2$ on some V_i, and hence $T \neq \pm I$ on these V_i. But since T is orthogonal, ± 1 are the only possible eigenvalues (which would need $T = \pm I$), and hence T has no eigenvalues on these V_i. In particular then, $\{v, Tv\}$ is linearly independent for each non-zero v in this V_i. Consider $W = \text{span}\{v, Tv\}$, which is T-invariant (as $Tv \mapsto T^2 v = \lambda_i Tv - v$), and hence $W^\perp$ is T-invariant also.

By induction, V_i splits into two-dimensional subspaces, and on each of these, by Proposition 10.13, is in the form R_θ for some θ. $\square$

11 Normal operators

Definition 11.1 (Normal operators)**.** Let V be a finite-dimensional complex inner product space, and $T : V \to V$ be a linear transformation. Then T is *normal* if it commutes with its adjoint:

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

Lemma 11.2. *Let T be normal and $\lambda \in \mathbb{C}$. Then*

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

(ii) $T - \lambda I$ *is normal*

(iii) $Tv = \lambda v \implies T^*v = \bar{\lambda} v$

(iv) $Tv = \lambda v,\ Tw = \mu w,\ \lambda \neq \mu \implies \langle v, w\rangle = 0$

Proof. (i) $\langle Tv, Tv\rangle = \langle v, T^*Tv\rangle = \langle v, TT^*v\rangle = \langle T^*v, T^*v\rangle$

(ii) $(T - \lambda I)^* = T^* - \bar{\lambda} I$, and this commutes with $T - \lambda I$.

(iii) $(T - \lambda I)v = 0 \iff (T^* - \bar{\lambda})v = 0$ by the previous two parts.

(iv) $\lambda\langle v, w\rangle = \langle \bar{\lambda} v, w\rangle = \langle T^*v, w\rangle = \langle v, Tw\rangle = \mu\langle v, w\rangle$, and $\lambda \neq \mu$.

□

Theorem 11.3. *Let T be normal (and V is assumed to be finite, as per our definition of normal). Then there exists an orthonormal basis of eigenvectors of T for V.*

Proof. As V is complex, there exists some $\lambda \in \mathbb{C}, v \in V$, such that $\|v\| = 1$ and $Tv = \lambda v$. Let $U_1 = \operatorname{span}\{v\}$. Then U_1 is T-invariant and T^*-invariant. Thus $U_1^\perp$ is T- and T^*-invariant, as, for all $u \in U$, $w \in U^\perp$,

$$\begin{aligned} \langle u, Tw\rangle &= \langle T^*u, w\rangle = 0 \\ \langle u, T^*w\rangle &= \langle Tu, w\rangle = 0 \end{aligned}$$

Once more, proceed by induction, similar to the previous (weaker) statements of this theorem. □

Now we can reformalise all the diagonalisation theorems that we have stated up to this point into one larger theorem:

Theorem 11.4 (Spectral Theorem for normal operators)**.** *Let $T : V \to V$ be a normal (symmetric) transformation on a complex (real) finite-dimensional vector space. Then there exist orthogonal projections $E_1, \ldots, E_r$ on V, and scalars $\lambda_1, \ldots, \lambda_r$, such that*

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

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

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

12 Simultaneous diagonalisation

Remark 12.1. If $\mathcal{B}$ is a basis with respect to which S and T are diagonal, then $ST = TS$. This can be seen by seeing that the matrix of ST splits into the product of S and T (as per usual), but then these matrices commute, as they are diagonal.

Theorem 12.2. *Let $S, T : V \to V$ be normal (symmetric) operators with $ST = TS$, and V be finite dimensional. Then there exists an orthonormal basis of eigenvectors of both S and T simultaneously for V.*

Proof. V decomposes into λ-eigenspaces for S: $V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_r}$. Let $v \in V_{\lambda_i}$. Then

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

and hence Tv is a λ_i-eigenvector for S, and V_{λ_i} is T-invariant.

Let $\mathcal{B}_{\lambda_i}$ be an orthonormal basis of eigenvectors of $T|_{V_{\lambda_i}}$ (which is still normal (symmetric)). Then $\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \ldots \cup \mathcal{B}_{\lambda_r}$ is an orthonormal basis of eigenvectors of S and T simultaneously for V. $\square$

A Examples of JNF

Example A.1 (Eigenvalues being zero)**.** Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by $T(\boldsymbol{x}) = A\boldsymbol{x}$, with

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

First note that $A^2 = 0$, and thus $m_A(x) = x^2$, and $\chi_A(x) = x^3$. We have that $0 \subset \ker T \subset \ker T^2 = \mathbb{R}^3$. It is straightforward to find that

$$\ker T = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

As $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \notin \ker T$, we can write

$$\frac{\ker T^2}{\ker T} = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \ker T \right\}$$

So

$$\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}, \qquad \mathcal{B}_1 = T(\mathcal{B}_2) \cup \mathcal{E}_1 = \left\{ \begin{pmatrix} -2 \\ 14 \\ 10 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

and $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ gives

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

Example A.2 (Eigenvalues being non zero)**.** Let $T : V \to V$ be given by $T(\boldsymbol{x}) = A\boldsymbol{x}$, where

$$ A = \begin{pmatrix} 3 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & 2 \end{pmatrix} $$

Then $\chi_T(x) = \det(A - xI) = \ldots = (2-x)^3$. By calculation, $m_T(x) = (x-2)^3$. We have also that $0 \subset \ker(A-2I) \subset \ker(A-2I)^2 \subset \ker(A-2I)^3 = \mathbb{R}^3$.

So

$$ \mathcal{B}_3 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\} \quad \text{as } (A-2I)^2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \neq 0 $$

$$ \mathcal{B}_2 = (A-2I)\mathcal{B}_3 = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\} $$

$$ \mathcal{B}_1 = (A-2I)\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\} $$

Let $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3$. Then

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