Part IB — Linear Algebra

Based on lectures by Prof. P. Raphael and notes by thirdsgames.co.uk

Michaelmas 2022

Contents

1 Vector spaces and linear dependence **4**
- 1.1 Vector spaces 4
- 1.2 Subspaces 5
- 1.3 Sum of subspaces 6
- 1.4 Quotients 7
- 1.5 Span 7
- 1.6 Dimensionality 8
- 1.7 Linear independence 9
- 1.8 Bases 9
- 1.9 Steinitz exchange lemma 11
- 1.10 Consequences of Steinitz exchange lemma 11
- 1.11 Dimensionality of sums 12
- 1.12 Direct sums 13

2 Linear maps **16**
- 2.1 Linear maps 16
- 2.2 Isomorphism 17
- 2.3 Kernel and image 18
- 2.4 Rank and nullity 19
- 2.5 Space of linear maps 21
- 2.6 Matrices 21
- 2.7 Linear maps as matrices 22
- 2.8 Change of basis 25
- 2.9 Equivalent matrices 27
- 2.10 Column rank and row rank 29
- 2.11 Conjugation and similarity 31
- 2.12 Elementary operations 31
- 2.13 Gauss' pivot algorithm 33
- 2.14 Representation of square invertible matrices 33

3 Dual spaces **35**
3.1 Dual spaces . . . 35
3.2 Annihilators . . . 37
3.3 Dual maps . . . 38
3.4 Properties of the dual map . . . 40
3.5 Double duals . . . 41

4 Bilinear Forms **44**
4.1 Introduction . . . 44
4.2 Change of basis for bilinear forms . . . 47

5 Determinant and Traces **49**
5.1 Trace . . . 49
5.2 Permutations and transpositions . . . 50
5.3 Determinant . . . 50
5.4 Volume forms . . . 51
5.5 Multiplicative property of determinant . . . 54
5.6 Singular and non-singular matrices . . . 55
5.7 Determinants of linear maps . . . 56
5.8 Determinant of block-triangular matrices . . . 57

6 Adjugate Matrices **59**
6.1 Column and row expansions . . . 59
6.2 Adjugates . . . 61
6.3 Cramer's rule . . . 62

7 Eigenvectors and Eigenvalues **64**
7.1 Eigenvalues . . . 64
7.2 Elementary facts about polynomials . . . 65
7.3 Characteristic polynomials . . . 66
7.4 Polynomials for matrices and endomorphisms . . . 68
7.5 Sharp criterion of diagonalisability . . . 69
7.6 Simultaneous diagonalisation . . . 72
7.7 Minimal polynomials . . . 73
7.8 Cayley-Hamilton theorem . . . 73
7.9 Algebraic and geometric multiplicity . . . 75
7.10 Characterisation of diagonalisable complex endomorphisms . . . 77

8 Jordan Normal Form **77**
8.1 Definition . . . 77
8.2 Similarity to Jordan normal form . . . 77
8.3 Direct sum of eigenspaces . . . 78

9 Properties of bilinear forms **82**
9.1 Changing basis . . . 82

9.2 Quadratic forms . . . 82
9.3 Diagonalisation of symmetric bilinear forms . . . 84
9.4 Sylvester's law . . . 85
9.5 Kernels of bilinear forms . . . 88
9.6 Sesquilinear forms . . . 88
9.7 Hermitian forms . . . 89
9.8 Polarisation identity . . . 90
9.9 Hermitian formulation of Sylvester's law . . . 90
9.10 Skew-symmetric forms . . . 91
9.11 Skew-symmetric formulation of Sylvester's law . . . 91

§7 Eigenvectors and Eigenvalues

§7.1 Eigenvalues

Let V be an F-vector space. Let $\dim_F V = n < \infty$, and let α be an endomorphism of V.

Question

Can we find a basis B of V such that, in this basis, $[\alpha]_B \equiv [\alpha]_{B,B}$ has a simple (e.g. diagonal, triangular) form?

Recall that if B' is another basis and P is the change of basis matrix, $[\alpha]_{B'} = P^{-1}[\alpha]_B P$. Equivalently, given a square matrix $A \in M_n(F)$ we want to conjugate it by a matrix P such that the result is 'simpler'.

Definition 7.1 (Diagonalisable)

Let $\alpha \in L(V)$ be an endomorphism. We say that α is **diagonalisable** if there exists a basis B of V such that the matrix $[\alpha]_B$ is diagonal.

Definition 7.2 (Triangulable)

We say that α is **triangulable** if there exists a basis B of V such that $[\alpha]_B$ is triangular.

Remark 33. We can express this equivalently in terms of conjugation of matrices.

Definition 7.3 (Eigenvalue, Eigenvector and Eigenspace)

A scalar $\lambda \in F$ is an **eigenvalue** of an endomorphism α if and only if there exists a vector $v \in V \setminus \{0\}$ such that $\alpha(v) = \lambda v$. Such a vector is an **eigenvector** with eigenvalue λ.
$V_\lambda = \{v \in V : \alpha(v) = \lambda v\} \leq V$ is the **eigenspace** associated to λ.

Lemma 7.1

Let $\alpha \in L(V)$ and $\lambda \in F$.
λ is an eigenvalue iff $\det(\alpha - \lambda I) = 0$.

Proof. If λ is an eigenvalue, there exists a nonzero vector v such that $\alpha(v) = \lambda v$, so $(\alpha - \lambda I)(v) = 0$. So the kernel is non-trivial. So $\alpha - \lambda I$ is not injective, so it is not surjective by the rank-nullity theorem. Hence this matrix is not invertible, so it has zero determinant. □

Remark 34. If $\alpha(v_j) = \lambda_j v_j$ ($v_j \neq 0$) for $j \in \{1, \dots, m\}$, we can complete the family v_j into a basis $(v_1, \dots, v_n)$ of V. Then in this basis, the first m columns of the matrix α has diagonal entries λ_j.

§7.2 Elementary facts about polynomials

Recall the following facts about polynomials on a field F, for instance

$$f(t) = a_n t^n + \dots + a_1 t + a_0, \quad a_i \in F$$

We say that the degree of f, written $\deg f$ is n. The degree of $f + g$ is at most the maximum degree of f and g. $\deg(fg) = \deg f + \deg g$.
Let $F[t]$ be the vector space of polynomials with coefficients in F.
λ is a root of $f(t) \iff f(\lambda = 0)$.

Lemma 7.2

If λ is a root of f then $(t - \lambda)$ divides F. I.e. $f(t) = (t - \lambda)g(t)$ where $g(t) \in F[t]$.

Proof.

$$f(t) = a_n t^n + \dots + a_1 t + a_0$$

Hence,

$$f(\lambda) = a_n \lambda^n + \dots + a_1 \lambda + a_0 = 0$$

which implies that

$$f(t) = f(t) - f(\lambda) = a_n(t^n - \lambda^n) + \dots + a_1(t - \lambda)$$

But note that, for all n,

$$t^n - \lambda^n = (t - \lambda)(t^{n-1} + \lambda t^{n-2} + \dots + \lambda^{n-2} t + \lambda^{n-1})$$

□

Remark 35. We say that λ is a root of **multiplicity** k if $(t - \lambda)^k$ divides f but $(t - \lambda)^{k+1}$ does not.

Corollary 7.1

A nonzero polynomial of degree n has at most n roots, counted with multiplicity.

Proof. Induction on the degree. Left as an exercise. □

Corollary 7.2

If f_1, f_2 are two polynomials of degree less than n such that $f_1(t_i) = f_2(t_i)$ for $i \in \{1, \dots, n\}$ and t_i distinct, then $f_1 \equiv f_2$.

Proof. $f_1 - f_2$ has degree less than n, but has n roots. Hence it is zero. □

Theorem 7.1

Any polynomial $f \in \mathbb{C}[t]$ of positive degree has a complex root. When counted with multiplicity, f has a number of roots equal to its degree.

Corollary 7.3

Any polynomial $f \in \mathbb{C}[t]$ can be factorised into an amount of linear factors equal to its degree. $f(t) = c\prod_{i=1}^{r}(t - \lambda_i)^{\alpha_i}$, with $c \in \mathbb{C}$, $\lambda_i \in \mathbb{C}$, $\alpha_i \in \mathbb{N}$.

Proved in Complex Analysis.

§7.3 Characteristic polynomials

Definition 7.4 (Characteristic polynomials)

Let α be an endomorphism. The **characteristic polynomial** of α is

$$\chi_\alpha(t) = \det(A^{a} - tI)$$

[a] $A = [\alpha]_B$ for any basis B, we will see it's well defined below.

Remark 36.

1. χ_α is a polynomial because the determinant is defined as a polynomial in the terms of the matrix.
2. Note further that conjugate matrices have the same characteristic polynomial, so the above definition is well defined in any basis. Indeed, $\det(P^{-1}AP - \lambda I) = \det(P^{-1}(A - \lambda I)P) = \det(A - \lambda I)$.

Theorem 7.2

Let $\alpha \in L(V)$. α is triangulable iff χ_α can be written as a product of linear factors over F. I.e. $\chi_\alpha(t) = c\prod_{i=1}^n (t - \lambda_i)$[a]

[a] λ_i need not be distinct.

Corollary 7.4

In particular, all complex matrices are triangulable.

Proof. ($\implies$): Suppose α is triangulable. Then for a basis B, $[\alpha]_B$ is triangulable with diagonal entries a_i. Then

$$\chi_\alpha(t) = (a_1 - t)(a_2 - t)\cdots(a_n - t)$$

($\impliedby$): We argue by induction on $n = \dim V$. True for $n = 1$.
By assumption, let $\chi_\alpha(t)$ be the characteristic polynomial of α with a root λ. Then, $\chi_\alpha(\lambda) = 0$ implies λ is an eigenvalue. Let V_λ be the corresponding eigenspace. Let $(v_1, \dots, v_k)$ be the basis of this eigenspace, completed to a basis $(v_1, \dots, v_n)$ of V. Let $W = \operatorname{span}\{v_{k+1}, \dots, v_n\}$, and then $V = V_\lambda \oplus W$. Then

$$[\alpha]_B = \begin{pmatrix} \lambda I & \star \\ 0 & C \end{pmatrix}$$

where $\star$ is arbitrary, and C is a block of size $(n-k)\times(n-k)$.
Then α induces an endomorphism $\overline{\alpha}\colon V/V_\lambda \to V/V_\lambda$ with $C = [\overline{\alpha}]_{\overline{B}}$ and $\overline{B} = (v_{k+1} + V_\lambda, \dots, v_n + V_\lambda)$.

Then (block product)

$$\begin{aligned} \det([\alpha]_B - tI) &= \det\begin{pmatrix} (\lambda - t)I & \star \\ 0 & C - tI \end{pmatrix} \\ &= (\lambda - t)^k \det(C - tI) \\ \text{We know } \det([\alpha]_B - tI) &= c\prod_{i=1}^n (t - a_i) \\ \implies \det(C - tI)^{a} &= c\prod_{k+1}^n (t - \tilde{a}_i) \end{aligned}$$

By induction on the dimension, we can find a basis $(w_{k+1}, \dots, w_n)$ of W for which $[C]_W$ has a triangular form. Then the basis $(v_1, \dots, v_k, w_{k+1}, \dots, w_n)$ is a basis for which α is triangular. $\square$

[a] As $\det(C - tI)$ is a polynomial

Lemma 7.3

Let $n = \dim V$, and V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. Let α be an endomorphism on V. Then

$$\chi_\alpha(t) = (-1)^n t^n + c_{n-1}t^{n-1} + \cdots + c_0$$

with

$$c_0 = \det A; \quad c_{n-1} = (-1)^{n-1} \operatorname{tr} A$$

Proof.

$$\chi_\alpha(t) = \det(\alpha - tI) \implies \chi_\alpha(0) = \det(\alpha)$$

Further, for $\mathbb{R}, \mathbb{C}$ we know that α is triangulable over $\mathbb{C}$. Hence $\chi_\alpha(t)$ is the determinant of a triangular matrix;

$$\chi_\alpha(t) = \prod_{i=1}^{n}(a_i - t)$$

Hence

$$c_{n-1} = (-1)^{n-1} a_i$$

Since the trace is invariant under a change of basis, this is exactly the trace as required. $\square$

§7.4 Polynomials for matrices and endomorphisms

Let $p(t)$ be a polynomial over F. We will write

$$p(t) = a_n t^n + \cdots + a_0$$

For a matrix $A \in M_n(F)$, we write

$$p(A) = a_n A^n + \cdots + a_0 \in M_n(F)$$

For an endomorphism $\alpha \in L(V)$,

$$p(\alpha) = a_n \alpha^n + \cdots + a_0 I \in L(V); \quad \alpha^k \equiv \underbrace{\alpha \circ \cdots \circ \alpha}_{k \text{ times}}$$

§7.5 Sharp criterion of diagonalisability

Theorem 7.3

Let V be a vector space over F of finite dimension n. Let α be an endomorphism of V. Then α is diagonalisable if and only if there exists a polynomial p which is a product of *distinct* linear factors, such that $p(\alpha) = 0$. In other words, there exist distinct $\lambda_1, \dots, \lambda_k$ such that

$$p(t) = \prod_{i=1}^{n}(t - \lambda_i) \implies p(\alpha) = 0$$

Proof. Suppose α is diagonalisable in a basis B. Let $\lambda_1, \dots, \lambda_k$ be the $k \leq n$ *distinct* eigenvalues. Let

$$p(t) = \prod_{i=1}^{k}(t - \lambda_i)$$

Let $v \in B$. Then $\alpha(v) = \lambda_i v$ for some i. Then, since the terms in the following product commute,

$$(\alpha - \lambda_i I)(v) = 0 \implies p(\alpha)(v) = \left[\prod_{i=1}^{k}(\alpha - \lambda_i I)\right](v) = 0$$

So for all basis vectors, $p(\alpha)(v)$. By linearity, $p(\alpha) = 0$.

Conversely, suppose that $p(\alpha) = 0$ for some polynomial $p(t) = \prod_{i=1}^{k}(t - \lambda_i)$ with distinct λ_i. Let $V_{\lambda_i} = \ker(\alpha - \lambda_i I)$. We claim that

$$V = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

Consider the polynomials

$$q_j(t) = \prod_{i=1, i \neq j}^{k} \frac{t - \lambda_i}{\lambda_j - \lambda_i}$$

These polynomials evaluate to one at λ_j and zero at λ_i for $i \neq j$. Hence $q_j(\lambda_i) = \delta_{ij}$. We now define the polynomial

$$q = q_1 + \dots + q_k$$

The degree of q is at most $(k-1)$. Note, $q(\lambda_i) = 1$ for all $i \in \{1, \dots, k\}$. The only polynomial that evaluates to one at k points with degree at most $(k-1)$ is exactly given by $q(t) = 1$. Consider the endomorphism

$$\pi_j = q_j(\alpha) \in L(V)$$

These are called the 'projection operators'. By construction,

$$\sum_{j=1}^{k} \pi_j = \sum_{j=1}^{k} q_j(\alpha) = I$$

So the sum of the π_j is the identity. Hence, for all $v \in V$,

$$I(v) = v = \sum_{j=1}^{k} \pi_j(v) = \sum_{j=1}^{k} q_j(\alpha)(v)$$

So we can decompose any vector as a sum of its projections $\pi_j(v)$. Now, by definition of q_j and p,

$$\begin{aligned}
(\alpha - \lambda_j I) q_j(\alpha)(v) &= \frac{1}{\prod_{i \neq j}(\lambda_j - \lambda_i)} (\alpha - \lambda_j I) \left[\prod_{i \neq j} (t - \lambda_i) \right] (\alpha) \\
&= \frac{1}{\prod_{i \neq j}(\lambda_j - \lambda_i)} \prod_{i=1}^{k} (\alpha - \lambda_i I)(v) \\
&= \frac{1}{\prod_{i \neq j}(\lambda_j - \lambda_i)} p(\alpha)(v)
\end{aligned}$$

By assumption, this is zero. For all v, we have $(\alpha - \lambda_j I) q_j(\alpha)(v)$. Hence,

$$(\alpha - \lambda_j I)\pi_j(v) = 0 \implies \pi_j(v) \in \ker(\alpha - \lambda_j I) = v_j$$

We have then proven that, for all $v \in V$,

$$v = \sum_{j=1}^{k} \underbrace{\pi_j(v)}_{\in V_j}$$

Hence,

$$V = \sum_{j=1}^{k} V_j$$

It remains to show that the sum is direct. Indeed, let

$$v \in V_{\lambda_j} \cap \left(\sum_{i \neq j} V_{\lambda_i} \right)$$

We must show $v = 0$. Applying π_j,

$$\pi_j(v) = q_j(\alpha)(v) = \prod_{i \neq j} \frac{(\alpha - \lambda_i I)(v)}{\lambda_j - \lambda_i}$$

Since $\alpha(v) = \lambda_j v$,

$$\pi_j(v) = \prod_{i \neq j} \frac{(\lambda_j - \lambda_i)v}{\lambda_j - \lambda_i} = v$$

Hence π_j really projects onto V_{λ_j}. However, we also know $v \in \sum_{i \neq j} V_{\lambda_i}$. So we can write $v = \sum_{i \neq j} w_i$ for $w \in V_{\lambda_i}$. Thus,

$$\pi_j(w_i) = \prod_{m \neq j} \frac{(\alpha - \lambda_m I)(v)}{\lambda_m - \lambda_j}$$

Since $\alpha(w_i) = \lambda_i w_i$, one of the factors will vanish, hence

$$\pi_j(w_i) = 0$$

So

$$v = \sum_{i \neq j} w_i \implies \pi_j(v) = \sum_{i \neq j} \pi_j(w_i) = 0$$

But $v = \pi_j(v)$ hence $v = 0$. So the sum is direct. Hence, $B = (B_1, \dots, B_k)$ is a basis of V, where the B_i are bases of V_{λ_i}. Then $[\alpha]_B$ is diagonal. □

Remark 37. We have shown further that if $\lambda_1, \dots, \lambda_k$ are distinct eigenvalues of α, then

$$\sum_{i=1}^{k} V_{\lambda_i} = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

Therefore, the only way that diagonalisation fails is when this sum is not direct, so

$$\sum_{i=1}^{k} V_{\lambda_i} < V$$

Example 7.1

Let $F = \mathbb{C}$. Let $A \in M_n(F)$ such that A has finite order; there exists $m \in \mathbb{N}$ such that $A^m = I$. Then A is diagonalisable. This is because

$$t^m - 1 = p(t) = \prod_{j=1}^{m} (t - \xi_m^j); \quad \xi_m = e^{2\pi i/m}$$

and $p(A) = 0$.

§7.6 Simultaneous diagonalisation

Theorem 7.4

Let α, β be endomorphisms of V which are diagonalisable. Then α, β are *simultaneously diagonalisable* (there exists a basis B of V such that $[\alpha]_B, [\beta]_B$ are diagonal) if and only if α and β commute.

Proof. Two diagonal matrices commute. If such a basis exists, $\alpha\beta = \beta\alpha$ in this basis. So this holds in any basis. Conversely, suppose $\alpha\beta = \beta\alpha$. We have

$$V = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

where $\lambda_i, \dots, \lambda_k$ are the k distinct eigenvalues of α. We claim that $\beta\left(V_{\lambda_j}\right) \leq V_{\lambda_j}$. Indeed, for $v \in V_{\lambda_j}$,

$$\alpha\beta(v) = \beta\alpha(v) = \beta(\lambda_j v) = \lambda_j \beta(v) \implies \alpha(\beta(v)) = \lambda_j \beta(v)$$

Hence, $\beta(v) \in V_{\lambda_j}$. By assumption, β is diagonalisable. Hence, there exists a polynomial p with distinct linear factors such that $p(\beta) = 0$. Now, $\beta\left(V_{\lambda_j}\right) \leq V_{\lambda_j}$ so we can consider $\beta|_{V_{\lambda_j}}$. This is an endomorphism of V_{λ_j}. We can compute

$$p\left(\beta\Big|_{V_{\lambda_j}}\right) = 0$$

Hence, $\beta|_{V_{\lambda_j}}$ is diagonalisable. Let B_i be the basis of V_{λ_i} in which $\beta|_{V_{\lambda_j}}$ is diagonal. Since $V = \bigoplus V_{\lambda_i}$, $B = (B_1, \dots, B_k)$ is a basis of V. Then the matrices of α and β in V are diagonal. $\square$

§7.7 Minimal polynomials

Recall from IB Groups, Rings and Modules the Euclidean algorithm for dividing polynomials. Given a, b polynomials over F with b nonzero, there exist polynomials q, r over F with $\deg r < \deg b$ and $a = qb + r$.

Definition 7.5

Let V be a finite dimensional F-vector space. Let α be an endomorphism on V. The *minimal polynomial* m_α of α is the nonzero polynomial with smallest degree such that $m_\alpha(\alpha) = 0$.

Remark 38. If $\dim V = n < \infty$, then $\dim L(V) = n^2$. In particular, the family $\left\{I, \alpha, \dots, \alpha^{n^2}\right\}$ cannot be free since it has $n^2 + 1$ entries. This generates a polynomial in α which evaluates to zero. Hence, a minimal polynomial always exists.

Lemma 7.4

Let $\alpha \in L(V)$ and $p \in F[t]$ be a polynomial. Then $p(\alpha) = 0$ if and only if m_α is a factor of p. In particular, m_α is well-defined and unique up to a constant multiple.

Proof. Let $p \in F[t]$ such that $p(\alpha) = 0$. If $m_\alpha(\alpha) = 0$ and $\deg m_\alpha < \deg p$, we can perform the division $p = m_\alpha q + r$ for $\deg r < \deg m_\alpha$. Then $p(\alpha) = m_\alpha(\alpha)q(\alpha) + r(\alpha)$. But $m_\alpha(\alpha) = 0$. But $\deg r < \deg m_\alpha$ and m_α is the smallest degree polynomial which evaluates to zero for α, so $r \equiv 0$ so $p = m_\alpha q$. In particular, if m_1, m_2 are both minimal polynomials that evaluate to zero for α, we have m_1 divides m_2 and m_2 divides m_1. Hence they are equivalent up to a constant. □

Example 7.2

Let $V = F^2$ and

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

We can check $p(t) = (t-1)^2$ gives $p(A) = p(B) = 0$. So the minimal polynomial of A or B must be either $(t-1)$ or $(t-1)^2$. For A, we can find the minimal polynomial is $(t-1)$, and for B we require $(t-1)^2$. So B is not diagonalisable, since its minimal polynomial is not a product of distinct linear factors.

§7.8 Cayley-Hamilton theorem

Theorem 7.5

Let V be a finite dimensional F-vector space. Let $\alpha \in L(V)$ with characteristic polynomial $\chi_\alpha(t) = \det(\alpha - tI)$. Then $\chi_\alpha(\alpha) = 0$.

Two proofs will provided; one more physical and based on $F = \mathbb{C}$ and one more algebraic.

Proof. Let $B = \{v_1, \dots, v_n\}$ be a basis of V such that $[\alpha]_B$ is triangular. This can be done when $F = \mathbb{C}$. Note, if the diagonal entries in this basis are a_i,

$$\chi_\alpha(t) = \prod_{i=1}^{n}(a_i - t) \implies \chi_\alpha(\alpha) = (\alpha - a_1 I)\dots(\alpha - a_n I)$$

We want to show that this expansion evaluates to zero. Let $U_j = \operatorname{span}\{v_1, \dots, v_j\}$. Let $v \in V = U_n$. We want to compute $\chi_\alpha(\alpha)(v)$. Note, by construction of the triangular matrix.

$$\begin{aligned}\chi_\alpha(\alpha)(v) &= (\alpha - a_1 I)\dots\underbrace{(\alpha - a_n I)(v)}_{\in U_{n-1}} \\ &= (\alpha - a_1 I)\dots\underbrace{(\alpha - a_{n-1} I)(\alpha - a_n I)(v)}_{\in U_{n-2}} \\ &= \dots \\ &\in U_0\end{aligned}$$

Hence this evaluates to zero. □

The following proof works for any field where we can equate coefficients, but is much less intuitive.

Proof. We will write

$$\det(tI - \alpha) = (-1)^n \chi_\alpha(t) = t^n + a_{n-1}t^{n-1} + \dots + a_0$$

For any matrix B, we have proven $B \operatorname{adj} B = (\det B)I$. We apply this relation to the matrix $B = tI - A$. We can check that

$$\operatorname{adj} B = \operatorname{adj}(tI - A) = B_{n-1}t^{n-1} + \dots + B_1 t + B_0$$

since adjugate matrices are degree $(n-1)$ polynomials for each element. Then, by applying $B \operatorname{adj} B = (\det B)I$,

$$(tI - A)[B_{n-1}t^{n-1} + \dots + B_1 t + B_0] = (\det B)I = (t^n + \dots + a_0)I$$

Since this is true for all t, we can equate coefficients. This gives

$$\begin{aligned}
t^n &: & I &= B_{n-1} \\
t^{n-1} &: & a_{n-1}I &= B_{n-2} - AB_{n-1} \\
&\vdots & &\vdots \\
t^0 &: & a_0 I &= -AB_1
\end{aligned}$$

Then, substituting A for t in each relation will give, for example, $A^n I = A^n B_{n-1}$. Computing the sum of all of these identities, we recover the original polynomial in terms of A instead of in terms of t. Many terms will cancel since the sum telescopes, yielding

$$A^n + a_{n-1}A^{n-1} + \cdots + a_0 I = 0$$

□

§7.9 Algebraic and geometric multiplicity

Definition 7.6

Let V be a finite dimensional F-vector space. Let $\alpha \in L(V)$ and let λ be an eigenvalue of α. Then

$$\chi_\alpha(t) = (t - \lambda)^{a_\lambda} q(t)$$

where $q(t)$ is a polynomial over F such that $(t - \lambda)$ does not divide q. a_λ is known as the *algebraic multiplicity* of the eigenvalue λ. We define the *geometric multiplicity* g_λ of λ to be the dimension of the eigenspace associated with λ, so $g_\lambda = \dim\ker(\alpha - \lambda I)$.

Lemma 7.5

If λ is an eigenvalue of $\alpha \in L(V)$, then $1 \leq g_\lambda \leq a_\lambda$.

Proof. We have $g_\lambda = \dim\ker(\alpha - \lambda I)$. There exists a nontrivial vector $v \in V$ such that $v \in \ker(\alpha - \lambda I)$ since λ is an eigenvalue. Hence $g_\lambda \geq 1$. We will show that $g_\lambda \leq a_\lambda$. Indeed, let $v_1, \dots, v_{g_\lambda}$ be a basis of $V_\lambda \equiv \ker(\alpha - \lambda I)$. We complete this into a basis $B \equiv (v_1, \dots, v_{g_\lambda}, v_{g_\lambda+1}, \dots, v_n)$ of V. Then note that

$$[\alpha]_B = \begin{pmatrix} \lambda I_{g_\lambda} & \star \\ 0 & A_1 \end{pmatrix}$$

for some matrix A_1. Now,

$$\det(\alpha - tI) = \det \begin{pmatrix} (\lambda - t)I_{g_\lambda} & \star \\ 0 & A_1 - tI \end{pmatrix}$$

By the formula for determinants of block matrices with a zero block on the off diagonal,

$$\det(\alpha - tI) = (\lambda - t)^{g_\lambda} \det(A_1 - tI)$$

Hence $g_\lambda \leq a_\lambda$ since the determinant is a polynomial that could have more factors of the same form. □

Lemma 7.6

Let V be a finite dimensional F-vector space. Let $\alpha \in L(V)$ and let λ be an eigenvalue of α. Let c_λ be the multiplicity of λ as a root of the minimal polynomial of α. Then $1 \leq c_\lambda \leq a_\lambda$.

Proof. By the Cayley-Hamilton theorem, $\chi_\alpha(\alpha) = 0$. Since m_α is linear, m_α divides χ_α. Hence $c_\lambda \leq a_\lambda$. Now we show $c_\lambda \geq 1$. Indeed, λ is an eigenvalue hence there exists a nonzero $v \in V$ such that $\alpha(v) = \lambda v$. For such an eigenvector, $\alpha^P(v) = \lambda^P v$ for $P \in \mathbb{N}$. Hence for $p \in F[t]$, $p(\alpha)(v) = [p(\lambda)](v)$. Hence $m_\alpha(\alpha)(v) = [m_\alpha(\lambda)](v)$. Since the left hand side is zero, $m_\alpha(\lambda) = 0$. So $c_\lambda \geq 1$. □

Example 7.3

Let

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

The minimal polynomial can be computed by considering the characteristic polynomial

$$\chi_A(t) = (t-1)^2(t-2)$$

So the minimal polynomial is either $(t-1)^2(t-2)$ or $(t-1)(t-2)$ We check $(t-1)(t-2)$. $(A - I)(A - 2I)$ can be found to be zero. So $m_A(t) = (t-1)(t-2)$. Since this is a product of distinct linear factors, A is diagonalisable.

Example 7.4

Let A be a Jordan block of size $n \geq 2$. Then $g_\lambda = 1$, $a_\lambda = n$, and $c_\lambda = n$.

§7.10 Characterisation of diagonalisable complex endomorphisms

Lemma 7.7

Let $F = \mathbb{C}$. Let V be a finite-dimensional $\mathbb{C}$-vector space. Let α be an endomorphism of V. Then the following are equivalent.

1. α is diagonalisable;
2. for all λ eigenvalues of α, we have $a_\lambda = g_\lambda$;
3. for all λ eigenvalues of α, $c_\lambda = 1$.

Proof. First, the fact that (i) is true if and only if (iii) is true has already been proven. Now let us show that (i) is equivalent to (ii). Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of α. We have already found that α is diagonalisable if and only if $V = \bigoplus V_{\lambda_i}$. The sum was found to be always direct, regardless of diagonalisability. We will compute the dimension of V in two ways;

$$n = \dim V = \deg \chi_\alpha; \quad n = \dim V = \sum_{i=1}^k a_{\lambda_i}$$

since χ_α is a product of $(t - \lambda_i)$ factors as $F = \mathbb{C}$. Since the sum is direct,

$$\dim \left(\bigoplus_{i=1}^k V_{\lambda_i} \right) = \sum_{i=1}^k g_{\lambda_i}$$

α is diagonalisable if and only if the dimensions are equal, so

$$\sum_{i=1}^k g_{\lambda_i} = \sum_{i=1}^k a_{\lambda_i}$$

Conversely, we have proven that for all eigenvalues λ_i, we have $g_{\lambda_i} \leq a_{\lambda_i}$. Hence, $\sum_{i=1}^k g_{\lambda_i} = \sum_{i=1}^k a_{\lambda_i}$ holds if and only if $g_{\lambda_i} = a_{\lambda_i}$ for all i. $\square$