

Further Linear Algebra

Kexing Ying

May 15, 2020

Contents

1	Introduction	1
1.1	Matrices	1
1.2	Geometry	2
2	More on Linear Algebra	3
2.1	Algebraic & Geometric Multiplicities of Eigenvalues	3
2.2	Direct Sums	5
2.3	Quotient Spaces	6
2.4	Triangularisation Theorem	7

1 Introduction

As we have learnt from last year, linear algebra is a important subject regarding matrices, vector spaces, linear maps, and this year we will also take a look at some geometrical interpretations of these concepts.

1.1 Matrices

Definition 1.1 (Similar matrices). Let $A, B \in \mathbb{F}^{n \times n}$ for some field \mathbb{F} . We say A is similar to B if and only if there exists some $P \in \mathbb{F}^{n \times n}$ such that,

$$B = P^{-1}AP.$$

We recall that *similar* is an equivalence relation and similar matrices shares many useful properties such as

- same determinant
- same characteristic polynomial
- same Eigenvalues
- same rank

and many more. As similar matrices share so many properties, one major aim in linear algebra is to find a “nice” matrix B given any arbitrary square matrix A such that A and B

are similar. We first saw a version of this question last year through the *diagonalisation* of matrices. However, as we have seen, not all matrices are diagonalisable, therefore, in this course, we will take a look at some *weaker* versions that are more general.

A version of our aim is the triangular theorem which states that; given $A \in \mathbb{C}^{n \times n}$ (note that this theorem is not true for arbitrary field), there exists (and not uniquely) some upper triangular matrix $B \in \mathbb{C}^{n \times n}$ such that A is similar to B .

Another version of this aim is the *Jordan Canonical Form* theorem. It turns out if $A \in \mathbb{C}^{n \times n}$, then A is similar to a *unique* matrix in the Jordan canonical form. This theorem is powerful due to the canonical nature of this theorem. One immediate result of this theorem is that we can check whether two matrices are similar to each other by checking where or not they have the same *JCF* (which is computationally easy to do).

However, we see that neither of the above version are theorems over arbitrary fields. The *Rational Canonical Form* attempts to solve this.

Definition 1.2 (Companion matrix). Given an arbitrary field \mathbb{F} , $p \in \mathbb{F}[X]$ such that p is monic (i.e. the coefficient of the highest term of p is 1) and $\deg p = k$, the companion matrix of p is the $k \times k$ matrix

$$C(p) = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ & & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{k-1} \end{bmatrix}$$

where a_i is the coefficient of p of the term X^i in \mathbb{F} .

The companion matrix is a nice matrix and it we can in fact show that the characteristic polynomial of the companion matrix of some p is p .

Theorem 1. Let $A \in \mathbb{F}^{n \times n}$ with characteristic polynomial p . Then, there exists a polynomial factorisation such that $p = \prod_{i=1}^k p_i$ and

$$A \sim \begin{bmatrix} C(p_1) & 0 & \cdots & 0 \\ 0 & C(p_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C(p_k) \end{bmatrix}.$$

Furthermore, it turns out this factorisation is unique under certain assumptions which we will take a look at in the course.

1.2 Geometry

Recall the dot product on \mathbb{R}^n where given $u, v \in \mathbb{R}^n$, $\langle u, v \rangle = \sum_{i=1}^n u_i v_i$. Furthermore, recall we also took a look at *orthogonal* and *symmetric* matrices last year. All of these, of course, has geometric interpretations and we will in this part of the course generalise and axiomatise these to the theory of *inner product spaces* of V over \mathbb{R} . We will also extend this theory to arbitrary fields \mathbb{F} - the *Theory of Bilinear Forms*.

2 More on Linear Algebra

From this point forward, we write $\sum \mu S$ as a shorthand for $\sum_{s \in S} \mu_s s$ for some suitable set S and indexed value μ .

2.1 Algebraic & Geometric Multiplicities of Eigenvalues

We recall some basic definitions and properties of Eigenvectors.

Definition 2.1. Let V be some vector space, $T : V \rightarrow V$ a linear map and λ an Eigenvalue of T . Then the λ -Eigenspace of T is the subspace of V ,

$$E_\lambda := \{v \in V \mid (\lambda I_V - T)v = \mathbf{0}\}.$$

We see that this is a subspace as it is the kernel of the linear map $\lambda I_V - T$.

Theorem 2. Let V be some vector space, $T : V \rightarrow V$ a linear map. Suppose that $\{v_1, \dots, v_k\}$ are Eigenvectors corresponding to distinct Eigenvalues $\lambda_1, \dots, \lambda_k$, then it is linearly independent.

Proof. We will prove by contrapositive. Suppose that $\{v_1, \dots, v_k\}$ are Eigenvectors that are linearly independent. Then by definition, there exists a minimal set of $\{\mu_i \mid i \in I\}$, such that $\sum_{i \in I} \mu_i v_i = 0$ (we see that $\mu_i \neq 0$ for all i as otherwise it is not minimal). Now, let $j \in I$, then by rewriting, we have $v_j = \sum_{i \neq j} \mu'_i v_i$. Thus,

$$\lambda_j \sum_{i \neq j} \mu'_i v_i = \lambda_j v_j = T(v_j) = T\left(\sum_{i \neq j} \mu'_i v_i\right) = \sum_{i \neq j} \mu'_i T(v_i) = \sum_{i \neq j} \mu'_i \lambda_i v_i.$$

So, by rearranging, $0 = \sum_{i \neq j} (\lambda_i - \lambda_j) \mu'_i v_i$. Now, if for all $i \neq j$, $\lambda_i \neq \lambda_j$, we have found a smaller subset of $\{v_1, \dots, v_k\}$ that is linearly dependent, contradicting our assumption, so there must be some i such that $\lambda_i = \lambda_j$. \square

Corollary 2.1. Let V be a n -dimensional vector space. Then if the characteristic polynomial of the linear map $T : V \rightarrow V$ has n distinct roots, then T is diagonalisable.

We define *algebraic* and *geometric* multiplicity for Eigenvalues.

Definition 2.2 (Algebraic and Geometric Multiplicity). Let $T : V \rightarrow V$ be a linear map with characteristic polynomial χ_T , such that $\chi_T(\lambda) = 0$ (i.e. λ is an Eigenvalue of T).

The algebraic multiplicity of λ is the number $a(\lambda)$ such that

$$\chi_T(x) = (x - \lambda)^{a(\lambda)} q(x),$$

for some polynomial $q(x)$ where $q(\lambda) \neq 0$.

The geometric multiplicity of λ is

$$g(\lambda) = \dim E_\lambda.$$

Proposition 2.1. Let $T : V \rightarrow V$ be a linear map with an Eigenvalue λ , then $g(\lambda) \leq a(\lambda)$.

Proof. Let $r = g(\lambda) = \dim E_\lambda$, then there exists linearly independent vectors v_1, \dots, v_r which forms a basis of E_λ . Suppose we extend this to a basis of V ,

$$B = \{v_1, \dots, v_r, w_1, \dots, w_s\},$$

then by working out $T(b)$ for all $b \in B$, we find $T(v_i) = \lambda v_i$, and $T(w_i) = \sum \mu_i B$ so,

$$[T]_B = \left[\begin{array}{cccc|cccc} \lambda & 0 & \cdots & 0 & \mu_1(v_1) & \mu_2(v_1) & \cdots & \mu_s(v_1) \\ 0 & \lambda & \cdots & 0 & \mu_1(v_2) & \mu_2(v_2) & \cdots & \mu_s(v_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & \mu_1(v_r) & \mu_2(v_r) & \cdots & \mu_s(v_r) \\ \hline & & & & \mu_1(w_1) & \mu_2(w_1) & \cdots & \mu_s(w_1) \\ & & & & \mu_1(w_2) & \mu_2(w_2) & \cdots & \mu_s(w_2) \\ & & & & \vdots & \vdots & \ddots & \vdots \\ & & & & \mu_1(w_s) & \mu_2(w_s) & \cdots & \mu_s(w_s) \end{array} \right].$$

We will refer to the four quadrants as $[\lambda]$, A , $\mathbf{0}$ and C respectively.

Thus, by considering the characteristic polynomial of this, we have

$$\chi_{[T]_B} = \det(xI - [T]_B) = (x - \lambda)^r \det(xI - C),$$

implying the algebraic multiplicity of λ is at least r . □

Theorem 3. Let $\dim V = n$ and $T : V \rightarrow V$ be a linear map with distinct Eigenvalues $\lambda_1, \dots, \lambda_r$. Suppose that the characteristic polynomial of T is

$$\chi_T = \prod_i (x - \lambda_i)^{a(\lambda_i)},$$

(so, $\sum_i a(\lambda_i) = n$). Then the following are equivalent,

1. T is diagonalisable;
2. $\sum_i g(\lambda_i) = n$;
3. for all i , $g(\lambda_i) = a(\lambda_i)$.

Proof. 2 \iff 3 is trivial so let us consider the other cases.

1 \implies 2. Suppose T is diagonalisable, then there exists some B , a basis of V consisting of Eigenvectors of T . Then, we can partition B into $F_{\lambda_i} := \{v \in B \mid T(v) = \lambda_i v\}$ for all Eigenvalues of T . By noting that the subspace induced by F_{λ_i} is a subspace of the λ_i Eigenspace E_{λ_i} , we have,

$$\sum_i g(\lambda_i) = \sum_i \dim E_{\lambda_i} \geq \sum_i \dim F_{\lambda_i} = n.$$

Now, as $\sum_i g(\lambda_i) \leq \sum_i a(\lambda_i) = n$ by the previous proposition, it follows $\sum_i g(\lambda_i) = n$.

2 \implies 1. Suppose $\sum_i g(\lambda_i) = n$. Let B_i be the basis of E_{λ_i} for all λ_i an Eigenvalue and let $B = \bigcup B_i$. We can straight away see that $|B| = n$ so it suffices to show that B is linearly independent. Suppose otherwise, then there exists an index set $I \subseteq \{1, \dots, r\}$,

$$\sum_{i \in I} \sum \mu_i B_i = 0$$

where $\sum \mu_i B_i \neq 0$ for all i . Now as $\sum \mu_i B_i \in E_{\lambda_i}$, this is a sum of Eigenvectors with distinct Eigenvalues. However, by theorem 2, these Eigenvectors are therefore linearly independent, so they must be zero. $\#$ \square

2.2 Direct Sums

Recall that we can add subspaces of a vector space together forming another subspace, that is, given $U_1, U_2 \leq V$, $U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\} \leq V$. Direct sums is a particular case of this and is closely linked to *block-diagonal* matrices.

Definition 2.3 (Direct Sum of Vector Space). Let V be a vector space with subspaces U_1, \dots, U_k . We write

$$V = \bigoplus_{i=1}^k U_i$$

for the direct sum of subspaces if every $v \in V$, there exists unique $u_i \in U_i$ for all i such that $v = \sum u_i$.

Proposition 2.2. Let V be a vector space with subspaces V_1, V_2 , then $V = V_1 \oplus V_2$ if and only if $V_1 \cap V_2 = \{0_V\}$ and $\dim V_1 + \dim V_2 = \dim V$.

Proof. Follow your nose. \square

Proposition 2.3. Let V be a vector space with subspaces V_1, \dots, V_k , then $V = \bigoplus_{i=1}^k V_i$ if and only if $\sum_{i=1}^k \dim V_i = \dim V$ and if B_i is a basis of V_i then $\bigcup_{i=1}^k B_i$ is a basis of V .

Proof. (\implies). Suppose $V = \bigoplus_{i=1}^k V_i$, then for all i, j , $V_i \cap V_j = \{0_V\}$, thus $B_i \cap B_j = \emptyset$ for all $i \neq j$, implying $\sum_{i=1}^k \dim V_i = \dim V$ and $|\bigcup B_i| = \dim V$ so it suffices to show that $\bigcup B_i$ is linearly independent. However, this is trivial as if $\bigcup B_i$ is linearly dependent, then there are two distinct ways of writing 0 as a sum of vectors in V_i . $\#$

(\impliedby). Suppose $\sum_{i=1}^k \dim V_i = \dim V$ and $\bigcup_{i=1}^k B_i$ is a basis of V , then it follows,

$$V = V_1 + V_2 + \dots + V_k.$$

Suppose for contradiction there is two representations of $v \in V$ where $v = \sum v_i = \sum v'_i$. Then, $v = \sum_i \sum \mu_i B_i = \sum_i \sum \mu'_i B_i$, and thus, $0 = \sum_i \sum \mu_i B_i - \sum_i \sum \mu'_i B_i = \sum_i \sum (\mu_i - \mu'_i) B_i$. By rewriting, $0 = \sum (\mu_i - \mu'_i) \bigcup B_i$, implying $\bigcup B_i$ is linearly dependent. $\#$ \square

Definition 2.4 (Invariant Subspace). Let V be a vector space with subspace W and let $T : V \rightarrow V$ be a linear map. We say W is T -invariant if and only if

$$T(W) \subseteq W.$$

We write $T_W : W \rightarrow W$ as the restriction of T to W .

A example of an invariant subspace is the Eigenspace of a linear map since $T(E_\lambda) = \{T(v) \mid v \in E_\lambda\} = \{\lambda v\} \subseteq E_\lambda$.

Proposition 2.4. *Let $T : V \rightarrow V$ be a linear map and suppose $V = \bigoplus V_i$ where for all i , V_i is T -invariant. Let B_i be a basis of V_i , and $A_i = [T_{V_i}]_{B_i}$, then*

$$[T]_{\bigcup B_i} = \text{diag}(A_1, A_2, \dots, A_k).$$

Proof. Follows directly from the T -invariant property of V_i . \square

From the proposition above, we see the close link between direct sums and block diagonal matrices. To further highlight the fact, from this point forward, we write $\bigoplus_{i=1}^k A_i = \text{diag}(A_1, A_2, \dots, A_k)$ where A_i are block matrices.

2.3 Quotient Spaces

Just like other algebraic graphs we have can construct a quotient structure on vector spaces.

Let V be a vector space and $W \leq V$, then let $\sim_W : V \rightarrow V \rightarrow \text{Prop}$ be the binary relation such that

$$v_1 \sim_W v_2 \iff v_1 + W = v_2 + W,$$

where $v + W = \{v + w \mid w \in W\}$ for all $v \in V$.

By manually checking, we find this is an equivalence relation and the set V/\sim_W equipped with the natural addition and scalar multiplication form a vector space. We will write V/W for this quotient space.

Definition 2.5. Given a quotient space V/W , there exists a linear map

$$q_W : V \rightarrow V/W : v \mapsto v + W.$$

Proposition 2.5. *Let V be a finite dimensional vector space with the subspace W , then $\dim V/W = \dim V - \dim W$.*

Proof. Let B_W be a basis of W and B_V the extension basis of V from B_W . Then we easily see that $V/W \subseteq \text{sp}(q_W(B_V \setminus B_W))$ as for all $v \in V$, $v = \sum \mu B_V$, so $v + W = q_W(v) = q_W(\sum \mu B_V) = \sum \mu q_W(B_W) + \sum \mu q_W(B_V \setminus B_W) = \sum 0_{V/W} + \sum \mu q_W(B_V \setminus B_W) \in \text{sp}(q_W(B_V \setminus B_W))$.

Now suppose $q_W(B_V \setminus B_W)$ is not linearly independent in V/W , then, there exists μ , $0 = \sum \mu q_W(B_V \setminus B_W) = q_W(\sum \mu (B_V \setminus B_W))$, so $\sum \mu (B_V \setminus B_W) \in \ker q_W = W$. If $\sum \mu (B_V \setminus B_W) = 0_V$, then $B_V \setminus B_W$ is not linearly dependent, a contradiction so, $\sum \mu (B_V \setminus B_W) \neq 0_V$. Now, as $\sum \mu (B_V \setminus B_W) \in W$, there is some λ , $\sum \mu (B_V \setminus B_W) = \sum \lambda B_W$, so $\sum \mu (B_V \setminus B_W) - \sum \lambda B_W = 0$ implying B_V is not linearly independent. $\#$ \square

With the above proposition, we have found a method to find a basis of a quotient space V/W by extending the basis of W .

Let us now consider quotient spaces' relation with linear maps.

Definition 2.6 (Quotient Map). Let V be a vector space and W a subspace of V . Suppose $T : V \rightarrow V$ is a linear map and W is T -invariant. Then there is an induced quotient map

$$\bar{T} : V/W \rightarrow V/W : q_W(v) \mapsto q_W(T(v)).$$

To see that this is well defined, let $u, v \in V$, $q_W(u) = q_W(v)$, then $u - v \in W$ implying $T(u - v) \in W$ as W is T -invariant. Thus, $0_{V/W} = q_W(T(u - v)) = q_W(T(u) - T(v)) = q_W(T(u)) - q_W(T(v))$ implying $\bar{T}(u) = \bar{T}(v)$.

Theorem 4. Let V be a vector space W a subspace that is T -invariant for some $T : V \rightarrow V$ a linear map. Let B_W be a basis of W , B the extended basis of V from B_W , and \bar{B} the basis of V/W as constructed by proposition 2.5. Then

$$[T]_B = \left[\begin{array}{c|c} [T_W]_{B_W} & A \\ \hline \mathbf{0} & [\bar{T}]_{\bar{B}} \end{array} \right],$$

where A is some matrix.

Proof. Consider where $T(v)$ lands whenever $u \in B_w \subseteq W$, and where $\bar{T}(v)$ lands for the rest of the basis vectors. \square

Corollary 4.1. Let $T : V \rightarrow V$ be a linear map and $W \leq V$ is T -invariant, then $\chi_T = \chi_{T_W} \chi_{\bar{T}}$ where χ_f denotes the characteristic polynomial of the linear map f .

2.4 Triangularisation Theorem

We have now arrived at the first major theorem of this course, that under certain conditions we can always triangularise matrices. We will in general work with upper triangular matrices when referring to triangular matrices.

Proposition 2.6. Let $A = [a_{i,j}], B = [b_{i,j}] \in M_n(\mathbb{F})$ be triangular, then

- $\chi_A(x) = \prod_{i=1}^n (x - a_{i,i})$;
- $\det A = \prod_{i=1}^n a_{i,i}$;
- AB is also triangular with diagonal $a_{i,i}b_{i,i}$.

The Triangularisation theorem states:

Theorem 5. Let V be a finite dimensional vector space over some field \mathbb{F} , and let $T : V \rightarrow V$ be a linear map. Suppose the the characteristic polynomial of T , χ_T factorises into a product of linear factors, i.e. there exists $\lambda_i \in \mathbb{F}$,

$$\chi_T(x) = \prod (x - \lambda_i),$$

then, there exists a basis B of V such that $[T]_B$ is upper triangular.

Straight away, we see a version of this in terms of matrices instead of linear maps in which the matrix is *similar* to a triangular matrix. We also note that, for some fields, such as the complex numbers \mathbb{C} , we can always triangularise any matrix (by *FTA*). This might not be the case for other fields such as the real numbers.

Proof. We induct on the dimension of V . The theorem is trivial when $\dim V = 1$, so let us consider the case when $\dim V = k + 1$ under the inductive hypothesis.

As χ_T factorises, T has an Eigenvalue λ and some Eigenvector $v \in V$. Let $W = \text{sp}(v)$ be a T -invariant subspace of V . Then, by proposition 2.5, V/W has dimension k and we have the induced quotient map $\bar{T} : V/W \rightarrow V/W$. Now, by corollary 4.1, $\prod (x - \lambda_i) = \chi_T(x) = \chi_{\bar{T}}(x)\chi_{T_W}(x) = \chi_{T_W}(x)(x - \lambda)$. So, $\chi_{T_W}(x)$ is a polynomial of degree k which factorises. Then by our inductive hypothesis, there exists a basis \bar{B} such that $[\bar{T}]_{\bar{B}}$ is triangular. Then by theorem 4, we have found a basis B , $[T]_B$ is triangular. \square

Corollary 5.1. *Let $A \in M_n(\mathbb{C})$ with Eigenvalues λ_i . Then $\sum g(\lambda_i)\lambda_i = \text{tr}(A)$.*

Proof. By the triangularisation theorem, $A = PQP^{-1}$ where Q is triangular. As A and Q have the same Eigenvectors, it suffices to show that $\text{tr}(A) = \text{tr}(Q)$. But this follows as $\text{tr}(A) = \text{tr}(PQP^{-1}) = \text{tr}(P^{-1}PQ) = \text{tr}(Q)$. \square