

Unit 5(b) Lecture Notes for MAT224

Xuanqi Wei 1009353209

21-23 March 2023

§1 5.2 Vector Spaces over a Field

Give an example of a complex vector space V .

One example of a complex vector space V is the space of all $n \times n$ complex matrices, denoted by $M_{n \times n}(\mathbb{C})$.

The addition and scalar multiplication operations are defined as follows:

Given $A, B \in M_{n \times n}(\mathbb{C})$ and $c \in \mathbb{C}$,

$(A + B)i, j = Ai, j + Bi, j$, and $(cA)i, j = cAi, j$.

It can be shown that $M_{n \times n}(\mathbb{C})$ satisfies all the axioms of a complex vector space, including the existence of an additive identity, additive inverses, and distributive laws.

Your vector space V from above is also a real vector space. Give an example of a basis for V as a real vector space (with scalars from \mathbb{R}) and another basis for V as a complex vector space (with scalars from \mathbb{C}).

As a real vector space, a basis for V can be obtained by considering the standard basis for $M_{n \times n}(\mathbb{C})$ as a real vector space.

Let E_{ij} be the $n \times n$ matrix with a 1 in the ij -th position and 0 elsewhere. Then, the set $E_{ij} : 1 \leq i, j \leq n$ is a basis for $M_{n \times n}(\mathbb{C})$ as a real vector space.

As a complex vector space, a natural basis for V is the set of matrices with a single 1 in one position and 0 elsewhere. More formally, let E_{ij} be the $n \times n$ matrix with a 1 in the ij -th position and 0 elsewhere. Then, the set $E_{ij} : 1 \leq i, j \leq n$ is a basis for $M_{n \times n}(\mathbb{C})$ as a complex vector space.

In general, is there a relationship between $\dim_{\mathbb{R}}(V)$ and $\dim_{\mathbb{C}}(V)$? These are the size of a basis for V when linear combinations can use real scalars and complex scalars respectively.

Yes, there is a relationship between $\dim_{\mathbb{R}}(V)$ and $\dim_{\mathbb{C}}(V)$ for any complex vector space V . In fact, we have the following result:

If V is a complex vector space of finite dimension n , then $\dim_{\mathbb{C}}(V) = 2 \dim_{\mathbb{R}}(V)$.

To see why, let v_1, v_2, \dots, v_m be a basis for V as a real vector space, where $m = \dim_{\mathbb{R}}(V)$. Since V is a complex vector space, each v_i can be written as a linear combination of complex basis vectors where u_1, u_2, \dots, u_n is a basis for V as a complex vector space and $a_{ij} \in \mathbb{C}$.

Note that we have m real equations of the form $v_i = \sum_{j=1}^n a_{ij} u_j$ for each $i = 1, 2, \dots, m$. Since u_1, u_2, \dots, u_n is a basis for V as a complex vector space, the coefficients a_{ij} are uniquely determined by these equations for each $i = 1, 2, \dots, m$. Therefore, we have a total of $m \times n$ complex coefficients, which shows that $\dim_{\mathbb{C}}(V) = m \times n = 2 \dim_{\mathbb{R}}(V)$.

§2 5.3 Geometry of a Complex Vector Space

What is the definition of a Hermitian Inner Product?

A Hermitian inner product, also known as a conjugate-symmetric inner product, is a generalization of the standard dot product to complex vector spaces. Let V be a complex vector space. A Hermitian inner product on V is a function $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ that satisfies the following properties for all $u, v, w \in V$ and $\alpha \in \mathbb{C}$:

1. $\langle u, v \rangle = \overline{\langle v, u \rangle}$ (conjugate symmetry).
2. $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0$ if and only if $u = 0$ (positive definiteness).
3. $\langle \alpha u + v, w \rangle = \alpha \langle u, w \rangle + \langle v, w \rangle$ (linearity in the first argument).

In the above definition, the bar notation \bar{z} denotes the complex conjugate of $z \in \mathbb{C}$. Property (1) is a generalization of symmetry in the standard dot product. Property (2) ensures that the inner product of a vector with itself is non-negative and vanishes only when the vector is zero. Property (3) generalizes linearity of the dot product.

How can you use a Hermitian inner product to calculate the length of a vector v ?

Choose an Hermitian inner product and use it to calculate the length of $\begin{pmatrix} 1 \\ i \\ 2 + 3i \end{pmatrix}$.

To calculate the length of a vector v using a Hermitian inner product, we can use the following formula:

$$|v| = \sqrt{\langle v, v \rangle}$$

where $\langle \cdot, \cdot \rangle$ is a Hermitian inner product on the vector space containing v .

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$\begin{aligned} \left\| \begin{pmatrix} 1 \\ i \\ 2 + 3i \end{pmatrix} \right\| &= \sqrt{\left\langle \begin{pmatrix} 1 \\ i \\ 2 + 3i \end{pmatrix}, \begin{pmatrix} 1 \\ i \\ 2 + 3i \end{pmatrix} \right\rangle} = \sqrt{1 \cdot \bar{1} + i \cdot \bar{i} + (2 + 3i) \cdot \overline{(2 + 3i)}} \\ &= \sqrt{1 - 1 + 13} \\ &= \sqrt{13}. \end{aligned}$$

Therefore, the length of $\begin{pmatrix} 1 \\ i \\ 2 + 3i \end{pmatrix}$ with respect to the standard Hermitian inner product on \mathbb{C}^3 is $\sqrt{13}$.

What is another name for the conjugate transpose of a matrix?

The conjugate transpose of a matrix is also known as the Hermitian transpose or the adjoint. It is denoted by A^\dagger or A^* , and can be obtained by taking the transpose of the matrix and then taking the complex conjugate of all its entries.

What is the adjoint M^* of matrix M if

$$M = \begin{bmatrix} 2-2i & 4-i & 0 \\ 7+2i & i & 8i \\ 8 & -5 & -4i \end{bmatrix}$$

Taking the complex conjugate of each entry in M^\top gives us

$$\overline{M^\top} = \begin{bmatrix} 2+2i & 4+i & 0 \\ 7-2i & -i & -8i \\ 8 & 5 & 4i \end{bmatrix}.$$

Therefore, the adjoint of M is

$$M^* = \begin{bmatrix} 2+2i & 4+i & 0 \\ 7-2i & -i & -8i \\ 8 & 5 & 4i \end{bmatrix}.$$

Proposition 5.3.8 says: Let V be a finite dimensional Hermitian inner product space. The adjoint of $T : V \rightarrow V$ satisfies $\langle T(v), w \rangle = \langle v, T^*(w) \rangle$ for all $v, w \in V$.

What does this proposition tell us about the matrix $[T]_\alpha^\alpha$ of a self-adjoint (Hermitian) linear transformation $T : V \rightarrow V$ with α any basis of V ?

Proposition 5.3.8 tells us that the adjoint T of a linear transformation $T : V \rightarrow V$ on a finite dimensional Hermitian inner product space V satisfies the following property for all $v, w \in V$:

$$\langle T(v), w \rangle = \langle v, T(w) \rangle.$$

If T is a self-adjoint linear transformation, then $T^*=T$ and we have:

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

for all $v, w \in V$. Let $\alpha = v_1, \dots, v_n$ be a basis of V . Then, for any $i, j \in 1, \dots, n$, we have:

$$[T(v_j)] = [T][v_j]$$

and

$$[vi][T(vj)] = \langle vi, T(vj) \rangle = \langle T(vi), vj \rangle = [T(vi)][vj]$$

, where the second equality follows from the self-adjointness of T and the third equality follows from Proposition 5.3.8. Therefore, we have:

$$[vi][T][vj] = [T][vi][vj]$$

Since this holds for all $i, j \in 1, \dots, n$, we conclude that $[T]_{\alpha}^{\alpha}$ is a Hermitian (self-adjoint) matrix.

It turns out that, like symmetric transformations, self-adjoint transformations are diagonalizable (Theorem 5.3.12) and have real eigenvalues (Theorem 5.3.13). What does that tell you about the eigenvectors of a self-adjoint transformation? The fact that self-adjoint transformations are diagonalizable and have real eigenvalues tells us that they have a basis consisting of eigenvectors, which are orthogonal to each other with respect to the Hermitian inner product.

More specifically, Theorem 5.3.12 states that if T is a self-adjoint linear transformation on a finite-dimensional Hermitian inner product space V , then T has an orthonormal basis consisting of eigenvectors. That is, there exists an orthonormal basis v_1, \dots, v_n of V such that each v_i is an eigenvector of T . Moreover, the corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ are all real.

Theorem 5.3.13 states that if T is a self-adjoint linear transformation on a finite-dimensional real inner product space V , then T has a basis consisting of eigenvectors, which are orthogonal to each other with respect to the real inner product. That is, there exists an orthogonal basis v_1, \dots, v_n of V such that each v_i is an eigenvector of T . Moreover, the corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ are all real.

Therefore, the eigenvectors of a self-adjoint transformation are not only guaranteed to exist, but they also have some special properties. They are orthogonal (or orthonormal) with respect to the Hermitian (or real) inner product, and the corresponding eigenvalues are real. This makes self-adjoint transformations very useful in applications where real eigenvalues and orthogonal eigenvectors are important, such as in physics, engineering, and statistics.

§3 6.1 Triangular Form

What is a T -invariant subspace? Give an example of a transformation T and a subspace W which is T -invariant.

Let $T: V \rightarrow V$ be a linear transformation on a vector space V . A subspace $W \subseteq V$ is said to be T -invariant if $T(w) \in W$ for all $w \in W$. That is, the subspace W is preserved by the action of the linear transformation T .

To give an example, let $V = \mathbb{R}^3$ be the vector space of 3-dimensional column vectors, and let $T: V \rightarrow V$ be the linear transformation given by the matrix

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

Consider the subspace $W = \text{span}(1, 0, 0)^T, (0, 1, 0)^T$, which is the xy -plane in \mathbb{R}^3 . We claim that W is T -invariant.

To see this, note that any vector in W can be written as a linear combination $w = a(1, 0, 0)^T + b(0, 1, 0)^T$ for some scalars $a, b \in \mathbb{R}$. Then we have

$$T(w) = T(a(1, 0, 0)^T + b(0, 1, 0)^T) = aT(1, 0, 0)^T + bT(0, 1, 0)^T$$

and using the matrix $[T]$, we can compute that

$$T(1, 0, 0)^T = (2, 1, 0)^T, T(0, 1, 0)^T = (1, 2, 1)^T$$

$$T(w) = a(2, 1, 0)^T + b(1, 2, 1)^T = (2a + b, a + 2b, b)^T$$

We see that $T(w)$ is always a linear combination of $(1, 0, 0)^T$, $(0, 1, 0)^T$, and $(0, 0, 1)^T$, which span \mathbb{R}^3 . Therefore, $T(w) \in W$ for any $w \in W$, so W is T -invariant.

How is the upper-triangular form of a matrix helpful in identifying T -invariant subspaces? You can illustrate with an example, rather than explain in words. The upper-triangular form of a matrix is helpful in identifying T -invariant subspaces because it allows us to easily see which subspaces are preserved under the action of the linear transformation T . Specifically, if a subspace is spanned by the first k standard basis vectors, then it is T -invariant if and only if the $(k + 1)$ st row and all subsequent columns of the upper-triangular form of the matrix $[T]_\beta^\beta$ are zero. As an example, consider the linear transformation $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ given by the matrix

$$[T] = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{pmatrix}$$

with respect to the standard basis $\beta = (1, 0, 0)^T, (0, 1, 0)^T, (0, 0, 1)^T$.

$$[T] = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{pmatrix} \sim [T] = \begin{pmatrix} 2 & \frac{1}{2} & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{pmatrix}$$

Since the first two columns of the upper-triangular form are nonzero, the subspace spanned by the first standard basis vector $(1, 0, 0)^T$ is not T -invariant. However, since the third column of the upper-triangular form is zero, the subspace spanned by the first two standard basis vectors $(1, 0, 0)^T, (0, 1, 0)^T$ is T -invariant. This can be verified directly by checking that $T((1, 0, 0)^T)$ and $T((0, 1, 0)^T)$ are both in the subspace $\text{span}((1, 0, 0)^T, (0, 1, 0)^T)$.

How can you tell from the characteristic polynomial of linear transformation $T : V \rightarrow V$, whether or not T is triangularizable? characteristic polynomial splits over the field F and each eigenvalue of T has an eigenvector in V . In other words, T is triangularizable if and only if its characteristic polynomial factors completely into linear factors over F and the algebraic multiplicity of each eigenvalue equals its geometric multiplicity.

Therefore, if the characteristic polynomial of T splits over F but some eigenvalue of T does not have a corresponding eigenvector in V , then T is not triangularizable. For example, suppose $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a linear transformation with characteristic polynomial $(x - 2)^2(x - 3)$, and suppose that T has only one eigenvector corresponding to the eigenvalue 2. Then T is not triangularizable over \mathbb{R} , since the algebraic multiplicity of the eigenvalue 2 is 2, but its geometric multiplicity is 1.

The Cayley-Hamilton Theorem (6.1.12) says: Let $T : V \rightarrow V$ be a linear mapping on a finite-dimensional vector space V , and let $p(t) = \det(T - tI)$ be its characteristic polynomial. Assume that $p(t)$ has $\dim(V)$ roots in the field F over which V is defined. Then $p(T) = 0$ (the zero mapping on V).

What does the theorem say about transformation T , which is defined by

$$T\vec{v} = \begin{bmatrix} 2i & 1 \\ 0 & -2i \end{bmatrix} \vec{v}$$

Let $V = \mathbb{C}^2$ be the vector space over \mathbb{C} on which T is defined as in the question, and let $p(t)$ be the characteristic polynomial of T . We have:

$$\begin{aligned} p(t) &= \det(T - tI) \\ &= \det \begin{pmatrix} 2i - t & 1 \\ 0 & -2i - t \end{pmatrix} \\ &= (2i - t)(-2i - t) - 0 \cdot 1 \\ &= t^2 + 4 \\ &= (t + 2i)(t - 2i). \end{aligned}$$

Since $p(t)$ has two distinct roots $2i$ and $-2i$ in the field \mathbb{C} , which is the field over which V is defined, the Cayley-Hamilton theorem implies that $p(T) = 0$, i.e., the linear transformation T satisfies the polynomial equation $T^2 + 4I = 0$, where I is the 2×2 identity matrix.

§4 6.2 A Canonical Form for Nilpotent Mappings

Recall: A **nilpotent** linear transformation N is a transformation such that N^k is the zero transformation for some $k \in \mathbb{Z}$, $k \geq 1$.

What do we know about the span of a cycle (see 6.2.1) defined by a nilpotent transformation (as opposed to a transformation that is not nilpotent)?

$$\text{span}(N^i v | i \in \mathbb{Z}) = \text{span}(v, Nv, \dots, N^{k-1}v, N^k v, N^{k+1}v, \dots).$$

Let N be a nilpotent linear transformation on a finite-dimensional vector space V , and let \vec{v} be a cyclic vector for N such that $N^i \vec{v} | i = 0, 1, \dots, k-1$ is a basis for the cyclic subspace $W = \text{span}(\vec{v}, N\vec{v}, \dots, N^{k-1}\vec{v})$. Then, the span of the cycle is given by

Since N^k is the zero transformation, it follows that $N^{k+i}\vec{v} = N^i(N^k\vec{v}) = \vec{0}$ for all $i \geq 0$, so $\text{span}(N^i \vec{v} | i \in \mathbb{Z}) = W$. In other words, the span of the cycle is equal to the cyclic subspace generated by \vec{v} .

On the other hand, if N is not nilpotent, then every vector in V is cyclic for N , and the span of the cycle generated by any vector \vec{v} is equal to the whole space V .

Let $N = (T - \lambda I)$ for some eigenvalue λ of linear transformation T . How do you create a tableau to depict the cycles of T . Include an example. Section 6.2 contains many of them. You can use one of those here.

Note: This section merely states that a canonical basis exists without showing you how to find one. For now, you do not have to concern yourself with finding x to generate the cycles. You simply need to know that x exists.

To create a tableau to depict the cycles of T associated with an eigenvalue λ , we start by writing down the generalized eigenspace $W_\lambda(T)$, which is the kernel of $(T - \lambda I)^n$ for some n such that $(T - \lambda I)^n \neq 0$ but $(T - \lambda I)^{n+1} = 0$. We then choose a basis $\beta = \vec{v}_1, \dots, \vec{v}_k$ for $W_\lambda(T)$, and for each \vec{v}_i , we write down the sequence of vectors $\vec{v}_i, (T - \lambda I)\vec{v}_i, \dots, (T - \lambda I)^{n_i-1}\vec{v}_i$ until we reach a vector \vec{w}_j that is already in the span of the previous vectors. We repeat this process for all \vec{v}_i until all vectors in $W_\lambda(T)$ have been accounted for.

To illustrate with an example, consider the linear transformation $T : \mathbb{C}^3 \rightarrow \mathbb{C}^3$ given by the matrix

$$[T]_{\beta}^{\beta} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

where $\beta = (1, 0, 0), (0, 1, 0), (0, 0, 1)$ is the standard basis for \mathbb{C}^3 . The characteristic polynomial of T is $(t-2)^2(t-1)$, so T has one eigenvalue 2 with multiplicity 2, and another eigenvalue 1.

Let's focus on the cycles of T associated with the eigenvalue $\lambda = 2$. We first find the generalized eigenspace $W_2(T)$, which is the kernel of $(T - 2I)^2$. We have

$$(T - 2I)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so $n = 2$ is the smallest integer such that $(T - 2I)^n = 0$. We compute

$$(T - 2I)\vec{v} = \vec{0} \leftrightarrow \vec{v} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

where $a, b \in \mathbb{C}$. Thus, $W_2(T) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$.