LECTURE 15

Endomorphisms, Automorphisms, and Change of Basis

We now specialize to the situation where a vector space homomorphism (a.k.a, linear transformation) maps a vector space to itself.

Definition 15.1. *Let V be a vector space over a field $\mathbb{F}$. A vector space homomorphism that maps V to itself is called an **endomorphism** of V. The set of all endomorphisms of V will be denoted by $L(V,V)$. A vector space isomorphism that maps V to itself is called an **automorphism** of V. The set of all automorphisms of V will be denoted $Aut(V)$.*

We have a natural notions of scalar multiplication and vector addition for elements of $L(V,V)$. If T in $L(V,V)$ and $\lambda \in \mathbb{F}$, then $\lambda \cdot T$ is defined as the mapping whose value at a point $v \in V$ is $\lambda \cdot (T(v))$. $\lambda \cdot T$ is in fact another linear transformation in $L(V,V)$ since for all $\alpha, \beta \in \mathbb{F}$ and all $v, w \in V$

$$\begin{aligned} (\lambda \cdot T)(\alpha v + \beta w) &= \lambda (T\alpha v + \beta w) \\ &= \lambda (\alpha T(v) + \beta T(w)) \\ &= \alpha (\lambda T(v) + \beta (\lambda T(w)) \\ & \alpha (\lambda \cdot T)(v) + \beta (\lambda \cdot T)(w) \end{aligned}$$

If $T, T' \in L(V,V)$, then $T + T'$ is the mapping whose value at $v \in V$ is $T(v) + T'(v)$. This is also another linear transformation since for all $\alpha, \beta \in \mathbb{F}$ and all $v, w \in V$

$$\begin{aligned} (T + T')(\alpha v + \beta w) &= T(\alpha v + \beta w) + T(\alpha v + \beta w) \\ &= \alpha T(v) + \beta T(w) + \alpha T'(v) + \beta T'(w) \\ &= \alpha (T + T')(v) + \beta (T + T')(w) \end{aligned}$$

In fact, it is (tedious but) not to hard to show that

Lemma 15.2. *If V is a vector space over a field $\mathbb{F}$, then $L(V,V)$ endowed with the scalar multiplication and vector addition defined above is a vector space over $\mathbb{F}$. Moreover, $Aut(V)$ is a subspace of $L(V,V)$.*

It is also true (for exactly the same reasons as above) that if $L(V,W)$ is the set of vector space homomorphisms from one vector space V to another W, then $L(V,W)$ can be given the structure of a vector space over the underlying field $\mathbb{F}$. However, what makes $L(V,V)$ (and $Aut(V)$) more interesting is that there is yet another kind of multiplication that can be defined on $L(V,V)$. Indeed, if T and T' are two endomorphisms in $L(V,V)$, then so are their compositions

$$\begin{aligned} T \circ T' &: V \to V \quad , \quad v \longmapsto T(T'(v)) \\ T' \circ T &: V \to V \quad , \quad v \longmapsto T'(T(v)) \end{aligned}$$

Remark 15.3. This multiplicative structure on $L(V,V)$ together with its natural (vector) addition gives $L(V,V)$ the structure of a (non-commutative) ring.

Definition 15.4. *A **ring** (with identity) is a set R endowed with two operations, "addition" and "multiplication" such that the following axioms are satisfied for all $a, b, c, d \in R$:*

(1) $a + b = b + a$ *(commutativity of addition)*
(2) $(a + b) + c = a + (b + c)$ *(associativity of addition)*

(3) *There is an element* $0_R \in R$ *such that* $a + 0 = a$ *for all* $a \in R$ *(additive identity)*
(4) *There is an element* 1_R *such that* $1_R \cdot a = a \cdot 1_R = a$ *for all* $a \in R$ *(multiplicative identity)*
(5) *For each* $a \in R$ *there is an element* $-a \in R$ *such that* $a + (-a) = 0_R$ *(additive inverses)*
(6) $(ab)\,c = a\,(bc)$ *(associativity of multiplication)*
(7) $a\,(b + c) = (ab) + (bc)$ *and* $(a + b)\,c = (ac) + (bc)$ *(distibutivity of multiplication over addition)*
Note that we do not require $ab = ba$.

Here are some more examples of rings:

- Any field is also a ring (in fact, a commutative ring)
- The set $Mat_{n,n}(\mathbb{F})$ of $n \times n$ matrices with entries in $\mathbb{F}$.
- The set of differential operators (acting on functions) on $\mathbb{R}$.

The set $Aut(V)$ of automorphisms of V has yet another special structure. Since the composition of two isomorphism is another isomorphism, and because each isomorphism $T : V \to V$ has an inverse mapping $T^{-1} : V \to V$ that is another element of $Aut(V)$, the set $Aut(V)$ qualifies as a **group**; as per the following defintion.

Definition 15.5. *A* ***group*** *is a set* G *with a notion of multiplication such that*

(1) *There exists an element* 1_G *with the property that* $1_G \cdot g = g \cdot 1_G$ *for all* $g \in G$.
(2) *For each element* $g \in G$*, there exists an element* $g^{-1} \in G$ *such* $gg^{-1} = 1_G = g^{-1}g$.

Here are a few more examples of groups.

- The set $\mathbb{R}$ where the group multiplication is taken to coincide with the usual notion of addition in $\mathbb{R}$.
- A vector space V where the group multiplication is taken to coincide with vector addition in V.
- The set $\mathbb{F}^\times = \{a \in \mathbb{F} \mid a \neq 0_\mathbb{F}\}$ where $\mathbb{F}$ is a field and the group multiplication is taken to coincide with multiplication in $\mathbb{F}$.

The examples above are examples of commutative groups (where $gg' = g'g$ for all $g, g' \in G$). It turns out that the group $Aut(V)$ is commutative only if V is one-dimensional. Here are some other non-commutative groups.

- $GL(n, \mathbb{F})$ the group of invertible $n \times n$ matrices with entries in a field $\mathbb{F}$.
- $O(n)$, the group of rotations in an n-dimensional Euclidean space.
- $SO(3, 1)$, the Lorentz group of Minkowski space time.

Remark 15.6. I have introduced rings and groups not so much because they help us understand Linear Algebra better; but rather because rings and groups are so prominent in modern mathematics and because Linear Algebra is so vital to the study of rings and groups.

1. Changes of Basis

By now, I hope we all are cognizant of the distinction between $\mathbb{F}^n$ and a more general vector space V over $\mathbb{F}$ of dimension n, and yet aware that so long as we have a basis B for V, we can reduce questions about V to computations in $\mathbb{F}^n$. One difficulty that remains, however, is that there is not a unique choice for B. Moreover, some choices of B may be good for some computations, but other choices of B may be more convenient for other computations. The question will now address is how do connect results based on different choices of bases.

Let me put this another way. Think of a basis of V as a way of acribing coordinates in $\mathbb{F}^n$ to elements in V. More explicitly, given a vector $\mathbf{v} \in V$ and a basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, then there is a unique choice of coefficients $a_1, \ldots, a_n \in \mathbb{F}$ such that

$$\mathbf{v} = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_n\mathbf{b}_n \quad . \tag{1}$$

Collecting these coefficients together we can represent $\mathbf{v} \in V$ as the element

$$\mathbf{v}_B = [a_1, \ldots, a_n] \in \mathbb{F}^n$$

and we refer to $\mathbf{v}_B$ as the *coordinate vector for* $\mathbf{v}$ *with respect to the basis* B. Although a bit tautological, another way of writing (1) is thus

$$\mathbf{v} = (\mathbf{v}_B)_1\,\mathbf{b}_1 + (\mathbf{v}_B)_2\,\mathbf{b}_2 + \cdots + (\mathbf{v}_B)_n\,\mathbf{b}_n. \tag{2}$$

A different choice of basis $B' = \{\mathbf{b}'_1, \ldots, \mathbf{b}'_n\}$ then amounts to choosing a different coordinate system for V. What we need to figure out is how to carry out a change of coordinates directly in terms of the corresponding coordinate vectors. That is to say, given that a vector $\mathbf{v} \in \mathbf{V}$ has coordinates $\mathbf{v}_B$ with respect to B, what is its coordinate vector $\mathbf{v}_{B'}$ giving its coordinates with respect to the basis B'?

$$\begin{array}{ccl} & & \mathbf{v}_B \in \mathbb{F}^n \\ & \nearrow & \downarrow \\ V \ni v & & \downarrow ? \\ & \searrow & \downarrow \\ & & \mathbf{v}_{B'} \in \mathbb{F}^n \end{array}$$

So let $B = \{v_1, \ldots, v_n\}$ and $B' = \{u_1, \ldots, u_n\}$ be two bases for V. Then since every vector in V can be expressed as an expansion with respect to either the vectors in B or the vectors in B' we have in particular,

$$v_i = \alpha_1^{(i)}u_1 + \cdots + \alpha_n^{(i)}u_n \quad , \quad i = 1, \ldots n \tag{3}$$

$$u_j = \beta_1^{(j)}v_1 + \cdots + \beta_n^{(j)}v_n \quad , \quad j = 1, \ldots, n \tag{4}$$

for some coefficients $\alpha_j^{(i)}$, $\beta_i^{(j)} \in \mathbb{F}$.[1]

Now suppose a vector $v \in V$ has coordinate vector $\mathbf{v}_B = [a_1, \ldots, a_n] \in \mathbb{F}^n$ with respect to the basis B. Then, as in (2)

$$\begin{aligned} v &= a_1v_1 + \cdots + a_nv_n \\ &= \sum_{j=1}^n a_jv_j \\ &= \sum_{j=1}^n a_j\left(\sum_{i=1}^n \alpha_i^{(j)}u_i\right) \end{aligned}$$

or, after reversing the order of summations,

$$v = \sum_{i=1}^n\left(\sum_{i=1}^n a_j\alpha_i^{(j)}\right)u_i \tag{5}$$

On the other hand, if $\mathbf{v}_{B/} = [b_1, \ldots, b_n]$ is the coordinate vector of v with respect to the basis B', we also have

$$v = \sum_{i=1}^n b_iu_i \tag{6}$$

[1]Note that $\left[a_1^{(i)}, \ldots, \alpha_n^{(i)}\right]$ is the coordinate vector of $u_i \in V$ with respect to the basis B and $\left[\beta_1^{(j)}, \ldots, \beta_n^{(j)}\right]$ is the coordinate vector of v_j with respect to the basis B'.

Comparing (5) and (6) (and using the fact that the vectors u_i are linearly independent) we can conclude

(7)
$$b_i = \sum_{j=1}^{n} a_j \alpha_i^{(j)} \qquad , \quad j = 1, 2, \ldots, n$$

Let me now reformulate this result as a matrix transformation. Form a matrix $\mathbf{C}_{B\to B'}$ by using using the coordinate vector $\left[\alpha_1^{(j)}, \ldots, \alpha_n^{(j)}\right]$ of v_j with respect to the basis B' as its i^{th} column

$$(\mathbf{C}_{B\to B'})_{ij} = \alpha_i^{(j)} \equiv (v_j)_{B'}$$

then (7) reads

$$b_i = \sum_{j=1}^{n} (\mathbf{C}_{B\to B'})_{ij} \, a_j$$

which, upon writing $\mathbf{v}_B = [a_1, \ldots, a_n]$ and $\mathbf{v}_{B'} = [b_1, \ldots, b_n]$ and as column vectors would be equivalent to

$$\mathbf{v}_{B'} = \mathbf{C}_{B\to B'} \mathbf{v}_B$$

Thus, multiplying the coordinate vector $\mathbf{v}_B$ by the $\mathbf{C}_{B\to B'}$ will yield the coordinate vector $\mathbf{v}_{B'}$.

In summary,

THEOREM 15.7. *Let $B = \{v_1, \ldots, v_n\}$ and $B' = \{u_1, \ldots, u_n\}$ be two bases for a vector space V. Let $\{\mathbf{v}_{1,B'}, \ldots, \mathbf{v}_{n,B'}\}$ be the coordinatization of the vectors $\{v_1, \ldots, v_n\}$ with respect to the basis B'. Form an $n \times n$ matrix $C_{B\to B'}$ by using the coordinate vectors $\mathbf{v}_{1,B'}, \ldots, \mathbf{v}_{n,B'}$ as columns*

$$\mathbf{C}_{B\to B'} = \begin{pmatrix} | & & | \\ \mathbf{v}_{1,B'} & \cdots & \mathbf{v}_{n,B'} \\ | & & | \end{pmatrix}$$

Then if $w \in V$ has coordinate vector $\mathbf{w}_B$ with respect to the basis B, it has coordinate vector

$$\mathbf{w}_{B'} = \mathbf{C}_{B\to B'} \mathbf{w}_B$$

with respect to the basis B'.

This explains how to go from coordinates w.r.t. B to coordinates w.r.t. B'. But how now to go from B' to B? Well, if you know the coordinate vectors $\mathbf{u}_{i,B}$ the same procedure works - just reversing the roles of primed and unprimed basis vectors.

However, there is another alternative. Since the vectors $\{v_1, \ldots, v_n\}$ and $\{u_1, , \ldots, u_n\}$ are linearly independent, the rank of the matrix $\mathbf{C}_{B\to B'}$ will be n. Thus, $\mathbf{C}_{B\to B'}$ will be an invertible matrix and the matrix $\mathbf{C}_{B\to B'}^{-1}$ can be used to transform coordinate vectors w.r.t. B' to coordinate vectors w.r.t. B. That is to say,

(8)
$$\mathbf{C}_{B'\to B} = (\mathbf{C}_{B\to B'})^{-1} \quad .$$

REMARK 15.8. It might seem that the need to change coordinates is only a remote possibility. However, soon we will be concerned with the solution of a variety of problems via the diagonalization of some matrix. This diagonalization process will then be understood as a change of basis (from a natural basis to one where the basis vectors are eigenvectors of a matrix).

2. Calculating Change of Bases Matrices for $\mathbb{F}^n$

Suppose $B_1 = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $B_2 = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ are two bases for a finite-dimensional vector space V over $\mathbb{F}$. We can calculate the matrix $\mathbf{C}_{B_1 B_2}$ that sends coordinate vectors with respect to B_1 directly to coordinate vectors with respect to B_2 once we know the coordinate vector $\mathbf{v}_{i,B}$ of each vector $\mathbf{v}_i$ with respect to the basis B_2. But here's the rub. To find $\mathbf{v}_{i,B'}$, one has to solve a system of n equations in n unknowns

$$\mathbf{v}_i = x_1 \mathbf{u}_1 + \cdots x_n \mathbf{u}_n$$

and one has do this for all n basis vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$.

Luckily there's an easier way. What we can do for $\mathbb{F}^n$ is exploit the existence of the standard basis

$$\begin{aligned} B_{std} &= \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\} \\ &= \{[1, 0, \dots, 0], [0, 1, 0, \dots, 0], \dots, [0, \dots, 0, 1]\} \end{aligned}$$

Suppose $\mathbf{v}_B$ is a coordinate vector for a vector $\mathbf{v}$ with respect to the (general) basis $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. According to the result of the preceding section, if we can find the coordinate vector for each $\mathbf{v}_i$ with respect to the standard basis, then we can convert $\mathbf{v}_B$ to its coordinate vector with respect to $B_{std.}$. But in fact, each basis vector $\mathbf{v}_i$ is an element of $\mathbb{F}^n$, and so an ordered list of elements of $\mathbb{F}$ and so

$$\mathbf{v}_j = \left[(v_j)_1, \dots, (v_j)_n \right] = (v_j)_1 \mathbf{e}_1 + \cdots + (v_j)_n \mathbf{e}_n \quad \Rightarrow \quad (\mathbf{v}_{j,B_{std}})_i = (v_j)_i$$

Thus,

$$(\mathbf{C}_{B \to B_{std}})_{ij} = (v_j)_i$$

Put another way, the matrix $\mathbf{C}_{B \to B_{std}}$ that maps coordinate vectors with respect to the general basis B to the corresponding coordinate vectors with respect to the standard basis can be formed by simply using the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ as the columns of $\mathbf{C}_{B \to B_{sstd}}$:

$$\mathbf{C}_{B \to B_{std}} = \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{pmatrix}$$

What about the matrix $\mathbf{C}_{B_{std} \to B}$ that converts standard coordinate vectors to coordinate vectors with respect to the basis B. Well that will be, in view of (8)

$$\mathbf{C}_{B_{std} \to B} = (\mathbf{C}_{B \to B_{std}})^{-1} = \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{pmatrix}^{-1}$$

Okay, now we can convert standard vectors to coordinate vectors with respect to a more general basis B. What about going from coordinate vectors with respect to one general basis $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ to coordinate vectors with respect to another general basis $B' = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$?

Well, we'll do this in two steps. First, we re-express the coordinates $\mathbf{v}_B$ of a vector with respect to B to its standard coordinates using $\mathbf{C}_{B \to B_{std}}$ and then we can go from the standard coordinates to coordinates with respect to B' using $\mathbf{C}_{B_{std} \to B'} = (\mathbf{C}_{B' \to B_{std}})^{-1}$. Thus,

$$\mathbf{C}_{B \to B'} = (\mathbf{C}_{B' \to B_{std}})^{-1} \mathbf{C}_{B \to B_{std}} = \begin{pmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_n \\ | & & | \end{pmatrix}^{-1} \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{pmatrix}$$