M7210 Lecture 6. Vector Spaces Part 3, Linear Maps Friday August 31, 2012

In the previous section, we saw how to associate with an $m \times n$ matrix with entries from $\mathbb{F}$ a function from $\mathbb{F}^n$ to $\mathbb{F}^m$. In the present section, we introduce the general idea of a linear map and then examine the relationships between bases, matrices and linear maps. This mostly concerns notation. The notational conventions take a good deal of care and attention to understand, remember and employ effectively, but there is a big payoff. When we can use notation effectively, it ceases to be a constraint and becomes a powerful tool.

Definition. Let V and W be vector spaces. A linear mapping from V to W is a function $L : V \to W$ such that $L(a_1v_1 + a_2v_2) = a_1L(v_1) + a_2L(v_2)$ for all $a_1, a_2 \in \mathbb{F}$ and all $v_1, v_2 \in V$.

Examples.

i) Let $a, b \in \mathbb{R}$. The function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = ax + by$ is a linear map. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = a\,x + b$ is not a linear map, despite the fact that every American high-school teacher calls this a "linear function." (Purists say that $f(x) = a\,x + b$ is "affine linear".)

ii) Suppose X is a set and $Y \subseteq X$. The the map that sends a function $g \in \mathbb{F}^X$ to its restriction $g|_Y$ in $\mathbb{F}^Y$ is linear.

iii) Differentiation is a linear map from the $\mathbb{R}$-vector space $C^1(\mathbb{R})$ of all continuously differentiable functions on $\mathbb{R}$ to $C(\mathbb{R})$, the vector space of all continuous functions on $\mathbb{R}$. (This map is surjective but not injective.)

Proposition 3.1. *Suppose V has a basis $\{v_1, \ldots, v_n\}$. Let $w_1, \ldots, w_n$ be any elements whatsoever of W. Then there is a unique linear map $L : V \to W$ such that $L(v_i) = w_i$ for $i = 1, \ldots, n$.*

Exercise. Prove Proposition 3.1. *Comment.* Proposition 3.1 is an example of a *universal mapping property*. We will see many more as we proceed through this course.

Using Matrices to Represent Linear Maps. In the next several paragraphs, we will show that if bases are fixed then there is a natural one-to one correspondence between matrices and linear maps.

Suppose V has an ordered basis $\Delta = \{v_1, \ldots, v_n\}$ and W has a ordered basis $\Omega = \{w_1, \ldots, w_m\}$. The next two propositions refer to these bases.

Proposition 3.2. *Let $A = \{a_{ij}\}$ be an $m \times n$ matrix with entries from $\mathbb{F}$. Then there is a unique linear map $L^A_{\Omega\Delta} : V \to W$ such that*

$$L^A_{\Omega\Delta}(v_j) = \sum_{i=1}^{m} a_{ij} w_i. \tag{$*$}$$

Proposition 3.2 is an immediate consequence of Proposition 3.1. The important idea is that once a bases for both V and W have been chosen, then any $\dim W \times \dim V$ matrix determines a linear map from V to W by the rule expressed in $(*)$.

Caution. Take note of the fact that the sum on the right hand side of $(*)$ is a linear combination of vectors. It bears a slight resemblance to the sum we would write if we wanted to express the i-th entry of the matrix product Ac, with c being the column having entries $c_1, \ldots, c_n$:

$$(A\,c)_i = \sum_{j=1}^{n} a_{ij} c_j. \tag{$**$}$$

But note that the summation in$(*)$ has a running index corresponding to the dimension of the codomain W, while the running index in $(**)$ corresponds to the dimension of the domain V. Also, the summands are numbers in $(**)$ but vectors in $(*)$.

We represent linear maps with matrices by letting them operate on column vectors from the left, i.e., (matrix)·(column). Some other authors (e.g., Jacobson, *Basic Algebra I*) let matrices operate on row vectors from the right, i.e., (row)·(matrix). This has advantages in some situations, but it is less common. The convention that we adopt means that the columns of A can be viewed as elements of the codomain, expressed with respect to a specified basis on the codomain.

Fact to Remember. *When representing linear maps my (matrix)·(column), the j-th column of the matrix, $\{a_{ij}\}_{i=1}^{m}$, lists the coefficients that appear when the image of the j-th basis vector of V is written as a linear combination of the basis elements of W.*

Elaboration. If an element $v \in V$ is represented as a column vector c_Δ^v whose n entries record the coefficients used to express it as a linear combination of elements of Δ, then the product $A\,c_\Delta^v$ is a column that lists the coefficients that appear when the image $L_{\Omega\Delta}^{A}(v)$ of v is written as a linear combination of the elements of Ω. In other words, if $w = L_{\Omega\Delta}^{A}(v)$, then

$$A\,c_\Delta^v = c_\Omega^w.$$

Let us emphasize that it is necessary to specify bases for V and W in order to interpret matrices as linear maps. In the case that $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$, we have the canonical bases. But often we must deal with vector spaces for which no bases are given or use bases for $\mathbb{F}^n$ that are other than the canonical one.

Proposition 3.3. *Suppose V has an ordered basis $\Delta = \{v_1, \ldots, v_n\}$ and W has a ordered basis $\Omega = \{w_1, \ldots, w_m\}$. Let $L : V \to W$ be a linear map. There is a unique $m \times n$ matrix $A_{\Omega\Delta}^{L} = \{a_{ij}\}$ such that*

$$L(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Proof. This is an immediate consequence of Proposition 1.1, which says that every element of W has a unique representation as a linear combination of elements of Ω. /////

Let us review. We have seen that the choice of a basis for a vector space provides way to uniquely represent its elements as tuples. We have also seen that the choice of bases for the domain and codomain of a linear map provides a way of representing that map by a matrix.

What happens if we change bases? To examine this, it is convenient to have notation that keeps track of the bases that are being used. We introduced a version of this, but the textbook has a nice alternative; see page 44 and following. (All vector spaces we talk about will be assumed to be finite-dimensional.)

Knapp's Notation. If U is a vector space with ordered basis $\Gamma = (u_1, \ldots, u_n)$ and $u = c_1u_1 + \cdots + c_nu_n \in U$, we let

$$\begin{pmatrix} u \\ \Gamma \end{pmatrix} := \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

If $L : U \to V$ is a linear map, and V has ordered basis $\Delta = (v_1, \ldots, v_k)$, and $L(u_j) = \sum_{i=1}^{k} a_{ij}v_j$ (cf. equation $(*)$) then

$$\begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix} := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{k1} & \cdots & a_{kn} \end{pmatrix}.$$

According to the "Fact to Remember":

> The j-th column of $\begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix}$ lists the coefficients that appear when the L-image of the j-th element of Γ is written as a linear combination of the elements of Δ.

We have

$$\begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix} \begin{pmatrix} u \\ \Gamma \end{pmatrix} = \begin{pmatrix} L(u) \\ \Delta \end{pmatrix}. \tag{3.1}$$

If $M : V \to W$, and Ω is an ordered basis for W:

$$\begin{pmatrix} M \\ \Omega\Delta \end{pmatrix} \begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix} = \begin{pmatrix} ML \\ \Omega\Gamma \end{pmatrix}. \tag{3.2}$$

On the left of (3.2), we have the matrix product of matrices $\begin{pmatrix} M \\ \Omega\Delta \end{pmatrix}$ and $\begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix}$. On the right, we have the matrix representing the composite map, ML:

$$U \xrightarrow{L} V \xrightarrow{M} W.$$

All of these things follow from the propositions proved above and in the previous lecture. Yet more detail can be found in the textbook.

The above is useful when bases have been chosen. However, it is very often the case that a statement or proof is obscure or difficult when expressed in terms of one basis or pair of bases, but clear and simple in another. So, we ought to have a way to transform representations made with respect to one basis to representations with respect to another. You will be delighted to learn that the same notation that we have employed for representing elements and maps with respect to selected bases also works to perform such transformations. The key idea is to use the identity map on a given space V, represented in terms of two *different* bases.

Change of basis. Suppose $\Gamma = (u_1, \ldots, u_n)$ and $\Delta = (v_1, \ldots, v_n)$ are two different ordered bases for a single vector space V. Let $I : V \to V$ be the identity map. By applying (3.1), we see that $\begin{pmatrix} I \\ \Delta\Gamma \end{pmatrix}$ is a matrix that changes a representation of $v \in V$ with respect to Γ to a representation of V with respect to Δ:

$$\begin{pmatrix} I \\ \Delta\Gamma \end{pmatrix}\begin{pmatrix} v \\ \Gamma \end{pmatrix} = \begin{pmatrix} v \\ \Delta \end{pmatrix}.$$

Applying (3.2), we have

$$\begin{pmatrix} I \\ \Delta\Gamma \end{pmatrix}\begin{pmatrix} I \\ \Gamma\Delta \end{pmatrix} = \begin{pmatrix} I \\ \Delta\Delta \end{pmatrix} = \text{identity matrix},$$

so we conclude that

$$\begin{pmatrix} I \\ \Delta\Gamma \end{pmatrix}^{-1} = \begin{pmatrix} I \\ \Gamma\Delta \end{pmatrix}.$$

In other words, the matrix that changes basis from Δ to Γ is the inverse of the matrix which changes basis from Γ to Δ.

Example. Let $V = \mathbb{F}^2$, and let $\Omega = (e_1, e_2)$ be the standard basis. Let $\Gamma = (u_1, u_2)$, where $u_1 = 3e_1 + 5e_2$ and $u_2 = e_1 + 2e_2$. Let $\Delta = (v_1, v_2)$, where $v_1 = 7e_1 + 5e_2$ and $v_2 = 4e_1 + 3e_2$. Our goal is to find the matrix $\begin{pmatrix} I \\ \Delta\Gamma \end{pmatrix}$. According to the Scholium,

$$\begin{pmatrix} I \\ \Omega\Gamma \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}; \begin{pmatrix} I \\ \Omega\Delta \end{pmatrix} = \begin{pmatrix} 7 & 4 \\ 5 & 3 \end{pmatrix}.$$

Thus

$$\begin{pmatrix} I \\ \Delta\Gamma \end{pmatrix} = \begin{pmatrix} I \\ \Delta\Omega \end{pmatrix}\begin{pmatrix} I \\ \Omega\Gamma \end{pmatrix} = \begin{pmatrix} 3 & -4 \\ -5 & 7 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} -11 & -5 \\ 20 & 9 \end{pmatrix}.$$

This tells us that $u_1 = -11v_1 + 20v_2$ and $u_2 = -5v_1 + 9v_2$. Let us check:

$$-11(7e_1 + 5e_2) + 20(4e_1 + 3e_2) = 3e_1 + 5e_2 = u_1,$$
$$-5(7e_1 + 5e_2) + 9(4e_1 + 3e_2) = e_1 + 2e_2 = u_2.$$

Example. Suppose $L : V \to V$ is an endomorphism (i.e., a linear transformation from V to itself) and Γ and Δ are two bases of V. Let $A = \begin{pmatrix} L \\ \Gamma\Gamma \end{pmatrix}$ and $B = \begin{pmatrix} L \\ \Delta\Delta \end{pmatrix}$. Then $B = C^{-1}AC$, where $C = \begin{pmatrix} I \\ \Gamma\Delta \end{pmatrix}$. The j-th column of C lists the coefficients that are used when the j-th element of Δ is written as a linear combination of elements of Γ.

Homework. **Do** the exercise after Proposition 3.1. **Do** page 82: 4, 5. **Do** the following exercises are from S. Lang, *Algebra, 3rd ed.*, Springer 2002, page 546:[1]

8. Let N be a strictly upper triangular $n \times n$ matrix, That is $N = (a_{ij})$ and $a_{ij} = 0$ if $i \geq j$. Show that $N^n = 0$.
9. Let E be a vector space over k [k = a field], of dimension n. Let $T : E \to E$ be a linear map such that T is nilpotent, that is $T^m = 0$ for some positive integer m. Show that there exists a basis of E over k such the matrix of T with respect to this basis is strictly upper triangular.

[1] If you have trouble, try doing these exercises for $n = 2$ and $n = 3$, then see if you can generalize.