260 Notes 10/17/18

David L. Meretzky

October 17th, 2018

Matrices

Definition 1 (3.30)**.** Let m and n denote positive integers. An m-by-n matrix A is a rectangular array of elements of $\mathbf{F}$ with m rows and n columns:

$$A = \begin{pmatrix} A_{(1,1)} & ... & A_{(1,n)} \\ \vdots & & \vdots \\ A_{(m,1)} & ... & A_{(m,n)} \end{pmatrix} \tag{1}$$

Each $A_{(j,k)} \in \mathbf{F}$ and denotes the entry in row j and column k.

Definition 2 (3.35)**.** The sum of two matricies of the same size is defined by adding corresponding entries in the matrix.

$$\begin{pmatrix} A_{(1,1)} & ... & A_{(1,n)} \\ \vdots & & \vdots \\ A_{(m,1)} & ... & A_{(m,n)} \end{pmatrix} + \begin{pmatrix} C_{(1,1)} & ... & C_{(1,n)} \\ \vdots & & \vdots \\ C_{(m,1)} & ... & C_{(m,n)} \end{pmatrix} = \begin{pmatrix} A_{(1,1)} + C_{(1,1)} & ... & A_{(1,n)} + C_{(1,n)} \\ \vdots & & \vdots \\ A_{(m,1)} + C_{(m,1)} & ... & A_{(m,n)} + C_{(m,n)} \end{pmatrix} \tag{2}$$

Definition 3 (3.37)**.** The product of a scalar and a matrix is the matrix defined by multiplying each entry in the matrix by the scalar:

For $\lambda \in \mathbf{F}$

$$\lambda \begin{pmatrix} A_{(1,1)} & ... & A_{(1,n)} \\ \vdots & & \vdots \\ A_{(m,1)} & ... & A_{(m,n)} \end{pmatrix} = \begin{pmatrix} \lambda A_{(1,1)} & ... & \lambda A_{(1,n)} \\ \vdots & & \vdots \\ \lambda A_{(m,1)} & ... & \lambda A_{(m,n)} \end{pmatrix} \tag{3}$$

You should check that the way we have defined addition and scalar multiplication make the collection of m-by-n matrices into a vector space. We call this vector space $\mathbf{F}^{m \times n}$.

Review Cartesian Product

Let n and m be sets as follows: $n = \{1, 2, 3, .., n\}$ and $m = \{1, 2, 3, ..., m\}$. The set $n \times m$ is called the cartesian product of n and m and is the collection of pairs (x, y) where x is an element in n and y is an element of m. For instance as sets, $5 \times 3 = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), ..., (5, 1), (5, 2), (5, 3)\} =$

$$\left\{ \begin{matrix} (1,1) & (1,2) & (1,3) \\ (2,1) & (2,2) & (2,3) \\ \vdots & & \vdots \\ (5,1) & (5,2) & (5,3) \end{matrix} \right\}$$

What do sets see?

We return now to a discussion of the notation $\mathbf{F}^S$. Recall the special case $\mathbf{F}^2 = \{(x, y)|\ x, y \in \mathbf{F}\}$. In this case $S = 2 = \{1, 2\}$. As we said in class, this is the collection of functions $\{1, 2\} \to \mathbf{F}$. That is to say:

(x, y) is a function $(x, y) : \{1, 2\} \to \mathbf{F}$ where $(x, y)(1) = x$ and $(x, y)(2) = y$

and a function $f : \{1, 2\} \to \mathbf{F}$ is a pair $(f(1), f(2))$.

In a sense, when the set $\{1, 2\}$ looks at $\mathbf{F}$ it "sees" lists of length 2 (x, y).

What do products of sets see?

The question can be rephrased as: what is $\mathbf{F}^{m\times n}$?

Let A be a function from 5×3 into $\mathbf{F}$, $A : 5 \times 3 \to \mathbf{F}$. Let's take the first element in 5×3, $(1, 1)$. Applying A we obtain $A(1, 1)$, a scalar in $\mathbf{F}$.

The function A is entirely determined by it's value on each of the 15 elements of 5×3. Thus all of the data is contained in the following matrix:

$A : 5 \times 3 \to \mathbf{F}$

$$A = \begin{pmatrix} A(1,1) & A(1,2) & A(1,3) \\ A(2,1) & A(2,2) & A(2,3) \\ \vdots & & \vdots \\ A(5,1) & A(5,2) & A(5,3) \end{pmatrix} \tag{4}$$

Writing the function $A(j, k)$ as $A_{(j,k)}$ we obtain a matrix! So when a product of sets like 5×3 looks at $\mathbf{F}$, it "sees" 5-by-3 matrices. Thus the collection of m-by-n matrices is $\mathbf{F}^{m\times n}$.

Vectors in $\mathbf{F}^m$

As you may have guessed $\mathbf{F}^m$ is the "same" as $\mathbf{F}^{m\times 1}$. Recall $\mathbf{F}^m$ is the collection of lists of length m of elements in $\mathbf{F}$ while $\mathbf{F}^{m\times 1}$ is the collection of m-by-1 matrices, with m. Something deeper is true however:

Theorem 1. Let V be a vector space of dimension m, a positive integer. For a basis $B = e_1, e_2, e_3, ..., e_m$ of V there is an injective and surjective linear map

$$[\ \ ]_B : V \to \mathbf{F}^{m\times 1}$$

For $v \in V$ we write $[\ \ ]_B(v)$ as $[v]_B$ and the map is defined as follows:

First represent v in terms of the basis B, $v = A_1e_1 + A_2e_2 + ... + A_me_m$, let

$$[v]_B = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix} \tag{5}$$

It is a simple matter to show this map is injective and surjective. It follows directly from the fact that the basis and therefore every vector can be written as a linear combination of the basis in exactly one way. In fact this is our notion of sameness of two vector spaces. Informally, A vector space V is the same as a vectorspace W if there exists an injective and surjective linear map from V to W. We will discuss this in depth next class.

Example 1. For $\mathbf{F}^n$ with the standard basis, this map is simply rotates the list

$$[(1, 2, 3, ..., n)]_B = \begin{pmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ n \end{pmatrix} \tag{6}$$

Example 2. For $\mathcal{P}_3(\mathbf{F})$. The situation is slightly different. Pick the standard basis of polynomials $B = 1, x, x^2, x^3$. Let $p(x) = 2 - 4x + 7x^3$.

$$[2 - 4x + 7x^3]_B = \begin{pmatrix} 2 \\ -4 \\ 0 \\ 7 \end{pmatrix} \tag{7}$$

$\mathscr{L}(V, W)$ and $\mathbf{F}^{m\times n}$

Let $\mathscr{L}(V, W)$ denote the vector space of linear maps from a vector space V of dimension n to a vector space W of dimension m. We will show that these vector spaces are "the same".

Theorem 2. Let V and W be vector spaces of dimension n and m respectively. For a basis $B = v_1, v_2, v_3, ..., v_n$ of V and $B' = w_1, ...w_m$ there is an injective and surjective linear map

$$[\ \ ]^B_{B'} : \mathscr{L}(V, W) \to \mathbf{F}^{m\times n}$$

For $T \in \mathscr{L}(V, W)$ we write $[\ \ ]^B_{B'}(T)$ as $[T]^B_{B'}$ and the map is defined as follows:

$[T]^B_{B'}$ is the matrix who's entries $A_{(j,k)}$ are defined by $Tv_k = A_{(1,k)}w_1 + ... + A_{(m,k)}w_m$, that is, the matrix who's columns are $[Tv_k]_{B'}$

$$[T]^B_{B'} = \begin{pmatrix} A_{(1,1)} & ... & A_{(1,n)} \\ \vdots & & \vdots \\ A_{(m,1)} & ... & A_{(m,n)} \end{pmatrix} \tag{8}$$

The elements of $\mathscr{L}(V, W)$ are linear maps from V to W. As we just saw V is the same as $\mathbf{F}^{n\times 1}$, W is the same as $\mathbf{F}^{m\times 1}$ and $\mathscr{L}(V, W)$ is the same as $\mathbf{F}^{m\times n}$. Hence an element of $\mathbf{F}^{m\times n}$ should be a linear map from $\mathbf{F}^{n\times 1}$ to $\mathbf{F}^{m\times 1}$.

If T maps the vector v in V to the vector Tv in W, How does $[T]^B_{B'}$ map $[v]_B$ to $[Tv]_{B'}$?

We have the following diagram of vectorspaces and linear maps. Notice here that we are conjecturing that $[T]^B_{B'}$ is a linear map between $\mathbf{F}^{n\times 1}$ and $\mathbf{F}^{m\times 1}$ although we have not explicitly defined how it acts on vectors yet.

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "{[\ ]_B}"'] & W \arrow[d, "{[\ ]_{B'}}"] \\ \mathbf{F}^{n\times 1} \arrow[r, "{[T]^B_{B'}}"] & \mathbf{F}^{m\times 1} \end{tikzcd}$$

Taking a closer look, starting with a vector $v \in V$ we can apply T and come up with a vector $Tv \in W$. We can then apply $[\ \ ]_{B'}$ and obtain the column vector $[Tv]_{B'}$. This is shown by following the top and right arrows of the following diagram. Following the other pair of arrows below, we could first go downwards, appling $[\ \ ]_B$ to v to obtain the column vector $[v]_B$. The diagram would suggest then that we should be able to apply $[T]^B_{B'}$ to $[v]_B$, written $[T]^B_{B'}[v]_B$. We should define $[T]^B_{B'}[v]_B$ in such a way that it is equal to $[Tv]_{B'}$. The fraction in the diagram is not a fraction, it notation showing that if you follow the top (T) and then right ($[\ \ ]_{B'}$) arrows you obtain the "numerator", $[Tv]_{B'}$. If you follow the left ($[\ \ ]_B$) and then bottom ($[T]^B_{B'}$) arrows you obtain the "denominator" $[T]^B_{B'}[v]_B$.

$$\begin{tikzcd} v \arrow[r, "T"] \arrow[d, "{[\ ]_B}"'] & Tv \arrow[d, "{[\ ]_{B'}}"] \\ {[v]_B} \arrow[r, "{[T]^B_{B'}}"] & \dfrac{[Tv]_{B'}}{[T]^B_{B'}[v]_B} \end{tikzcd}$$

To reiterate, we would like to define $[T]^B_{B'}[v]_B$ so that

$$[Tv]_{B'} = [T]^B_{B'}[v]_B \tag{9}$$

Let us first try to follow the bottom route: Suppose we take any vector $v \in V$, v can be expressed in terms of the basis as

$$v = x_1v_1 + x_2v_2 + ... + x_nv_n \tag{10}$$

This is exactly the data of $[v]_B$.

$$[v]_B = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{11}$$

At this point we are stuck because we do not know how to apply $[T]^B_{B'}$.

Let us try to follow the other route: Applying T to both sides of (10) and using the linearity of T we obtain the vector $Tv \in W$

$$Tv = T(x_1v_1 + x_2v_2 + ... + x_nv_n) = x_1Tv_1 + x_2Tv_2 + ... + x_nTv_n \tag{12}$$

Then we would like to express Tv in terms of the basis B' so we can apply $[\ \ ]_{B'}$ and obtain $[Tv]_{B'}$. Since B' is a basis there exists coefficients $A_1, ..., A_m$ such that

$$Tv = A_1w_1 + A_2w_2 + ... + A_mw_m$$

Then

$$[Tv]_{B'} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix} \tag{13}$$

However, since $[\ \ ]_{B'}$ is a linear map, we can apply it to equation (12) as

$$[Tv]_{B'} = x_1[Tv_1]_{B'} + x_2[Tv_2]_{B'} + ... + x_n[Tv_n]_{B'} \tag{14}$$

Since B' is a basis we can represent each Tv_k in terms of B'. That is there exist coefficients $A_{(j,k)}$ for j in the range from 1 to m and k in the range from 1 to n such that

$$Tv_k = A_{(1,k)}w_1 + A_{(2,k)}w_2 + ... + A_{(m,k)}w_m \tag{15}$$

Then

$$[Tv_k]_{B'} = \begin{pmatrix} A_{(1,k)} \\ A_{(2,k)} \\ \vdots \\ A_{(m,k)} \end{pmatrix} \tag{16}$$

Notice that the column vector in the above equation is exactly the way we defined the entries of $[T]^B_{B'}$. See Theorem 2 above and (8).

Written in column notation, equations (14) and (16) become

$$\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix} = x_1 \begin{pmatrix} A_{(1,1)} \\ A_{(2,1)} \\ \vdots \\ A_{(m,1)} \end{pmatrix} + x_2 \begin{pmatrix} A_{(1,2)} \\ A_{(2,2)} \\ \vdots \\ A_{(m,2)} \end{pmatrix} + ... + x_n \begin{pmatrix} A_{(1,n)} \\ A_{(2,n)} \\ \vdots \\ A_{(m,n)} \end{pmatrix} \tag{17}$$

Notice that this is eerily close to equation (9)

$$[Tv]_{B'} = [T]^B_{B'}[v]_B$$

which can be written in column and matrix notation as:

$$\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix} = \begin{pmatrix} A_{(1,1)} & A_{(1,2)} & \dots & A_{(1,n)} \\ A_{(2,1)} & A_{(2,2)} & \dots & A_{(2,n)} \\ \vdots & \vdots & & \vdots \\ A_{(m,1)} & A_{(m,2)} & \dots & A_{(m,n)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{18}$$

In order to make equation (18) hold, we will simply define

$$\begin{pmatrix} A_{(1,1)} & A_{(1,2)} & \dots & A_{(1,n)} \\ A_{(2,1)} & A_{(2,2)} & \dots & A_{(2,n)} \\ \vdots & \vdots & & \vdots \\ A_{(m,1)} & A_{(m,2)} & \dots & A_{(m,n)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} A_{(1,1)} \\ A_{(2,1)} \\ \vdots \\ A_{(m,1)} \end{pmatrix} + \dots + x_n \begin{pmatrix} A_{(1,n)} \\ A_{(2,n)} \\ \vdots \\ A_{(m,n)} \end{pmatrix} \tag{19}$$

In bracket notation, not matrix notation, we have defined $[T]^B_{B'}[v]_B$ as

$$[T]^B_{B'}[v]_B = x_1[Tv_1]_{B'} + x_2[Tv_2]_{B'} + \dots + x_n[Tv_n]_{B'} \tag{20}$$

In bracket notation equations (14) and (20) together give us equation (9). Said in matrix notation, equations (19) and (17) give us equation (18).

In other words, the diagrams just above equation (9) commute, meaning one can follow either path and get the same result. Concretely, this means that if one has a vector v and a linear transformation T, one can apply T and then represent Tv as a column vector $[Tv]_{B'}$ with respect to the basis, B' of the output space, *or* one can first represent v as a column vector $[v]_B$ with respect to the basis of the input space B, then apply the matrix of T, $[T]^B_{B'}$ using definition (19) and obtain $[T]^B_{B'}[v]_B$, and by (9) either route is the same.

Example 3. Pick a polynomial $p(x) \in \mathcal{P}_3(\mathbf{F})$ and use the linear map D which differentiates polynomials. We saw in class that with respect to the standard bases of $\mathcal{P}_3(\mathbf{F})$ and $\mathcal{P}_2(\mathbf{F})$, $B = 1, x, x^2, x^3$ and $B' = 1, x, x^2$ respectively,

$$[D]^B_{B'} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \tag{21}$$

Draw the 2 diagrams before equation (9) in the context of this example and check that both routes result in the same column vector.

Example 4. Repeat the preceding exercise but use different bases: $C = 1, x - 1, (x-1)^2, (x-1)^3$ and $C' = 1, x-1, (x-1)^2$. You should still find $[D]^C_{C'}$.