Math 201 lecture for Monday, Week 6

Linear transformations and matrices I

Our next goal is to encode linear functions by matrices. We first treat the special case of linear functions of the form $F^n \to F^m$. Next, we consider linear functions $V \to W$ between general finite-dimensional vector spaces. If $\dim V = n$ and $\dim W = m$ we saw last time that a choices of bases give isomorphisms $V \simeq F^n$ and $W \simeq F^m$, which reduces the problem to the special case.

Matrices for linear functions $F^n \to F^m$. The *dot product* of vectors $(a_1, \dots, a_n)$ and $(b_1, \dots, b_n)$ in F^n is defined by

$$(a_1, \dots, a_n) \cdot (b_1, \dots, b_n) := \sum_{i=1}^{n} a_i b_i = a_1 b_1 + \cdots + a_n b_n.$$

From now on we make adopt the convention of identifying vectors $(a_1, \dots, a_n) \in F^n$ with $n \times 1$ matrices, also called *column vectors*:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

If $A \in M_{m\times n}(F)$ and $x = (x_1, \dots, x_n) \in F^n$, we define $Ax \in F^m$ to be the element of F^m whose i-th component $(Ax)_i$ is the dot product of the i-th row of A with x:

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

$$= (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \dots, a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n).$$

The latter equals sign is just making the identification of column vectors with elements of F^m. Equivalently,

$$Ax := x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

We could similarly, convert the above notation into a statement about a linear combination of m-tuples in F^m instead of using column vectors.

Definition. Let $A \in M_{m\times n}(F)$. The linear associated with A is

$$\begin{aligned} L_A \colon F^n &\to F^m \\ x &\mapsto Ax. \end{aligned}$$

Exercise. The reader should perform the routine check that L_A is a linear function: $L_A(x+\lambda y) = L_A(x) + \lambda L_A(y)$.

Examples.

(1) The matrix

$$A = \begin{pmatrix} 2 & -5 & 4 \\ 3 & 0 & 2 \end{pmatrix}$$

has corresponding linear mapping

$$\begin{aligned} L_A \colon F^3 &\to F^2 \\ (x,y,z) &\mapsto \begin{pmatrix} 2x-5y+4z \\ 3x+2z \end{pmatrix}. \end{aligned}$$

Recall that we are identifying To save space, we could we will write this as

$$\begin{aligned} L_A \colon F^3 &\to F^2 \\ (x,y,z) &\mapsto (2x-5y+4z, 3x+2z). \end{aligned}$$

(2) Note that if you were given the linear function L_A, you could easily recover the matrix: just read off the coefficients of each component of $L_A(x)$ to find the rows of A. (We will see another way of recovering A below.) For example, find the matrix corresponding to the linear function $\phi \colon F^3 \to F^2$ defined by $\phi(u,v) = (4u-3v, 6u+2v, 3v)$.

Solution. Reading off the coefficients of each component of ϕ gives our matrix. Defining

$$A := \begin{pmatrix} 4 & -3 \\ 6 & 2 \\ 0 & 3 \end{pmatrix},$$

it is easy to check that $\phi = L_A$.

(3) Here are some important special cases of this correspondence between linear functions and matrices:

$$\begin{aligned} L_A(x) = (2x, 5x, 7x) \qquad &\leftrightsquigarrow \qquad A = \begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix} \\ L_B(w,x,y,z) = w+2x-4y+z \qquad &\leftrightsquigarrow \qquad B = \begin{pmatrix} 1 & 2 & -4 & 1 \end{pmatrix} \\ L_C(t) = 8t \qquad &\leftrightsquigarrow \qquad C = \begin{pmatrix} 8 \end{pmatrix}. \end{aligned}$$

We have formally defined the linear mapping L_A associated with a matrix A, and from the examples above, it may be clear how to go in the other direction to find the matrix of a given linear function. Here is the formal definition:

Definition. The matrix associated with the linear function $L\colon F^n \to F^m$ is the element $A \in M_{m\times n}(F)$ whose j-th column is $L(e_j)$ where e_j is the j-th standard basis vector for F^n.

Examples. Consider the first two examples given above.

(1) Consider the linear function $L\colon F^3 \to F^2$ given by $L(x, y, z) = (2x - 5y + 4z, 3x + 2z)$. Evaluate L at the three standard basis vectors for F^3:

$$\begin{aligned} L(e_1) &= L(1,0,0) = (2,3) \\ L(e_2) &= L(0,1,0) = (-5,0) \\ L(e_3) &= L(0,0,1) = (4,2). \end{aligned}$$

Use these three vectors to form a matrix:

$$A = \begin{pmatrix} 2 & -5 & 4 \\ 3 & 0 & 2 \end{pmatrix}.$$

Thus, $L = L_A$.

(2) Consider the linear function $\phi\colon F^3 \to F^2$ given by $\phi(u, v) = (4u - 3v, 6u + 2v, 3v)$. Then,

$$\phi(1,0) = (4,6,0) \quad \text{and} \quad \phi(0,1) = (-3,2,3).$$

Place these vectors as columns to get the matrix

$$\begin{pmatrix} 4 & -3 \\ 6 & 2 \\ 0 & 3 \end{pmatrix}.$$

We have, thus, created a bijective correspondence between linear function $F^n \to F^m$ and matrices in $M_{m\times n}(F)$.

Matrices for linear functions $V \to W$.

Let V and W be vector spaces with ordered bases $\mathcal{B} = \langle v_1, \ldots, v_n\rangle$ and $\mathcal{D} = \langle w_1, \ldots, w_m\rangle$, respectively. Taking coordinates with respect to these bases yields isomorphisms $\phi_{\mathcal{B}} : V \to F^n$ and $\phi_{\mathcal{D}}\colon W \to F^m$. For instance, if $v \in V$, we write $v = \sum_{i=1}^n a_i v_i$, and then $\phi_{\mathcal{B}}(v) := (a_1, \ldots, a_n)$. Now suppose we have a linear function $f\colon V \to W$. So up to now we have three mappings we are considering:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\phi_{\mathcal{B}}"', "\wr"] & W \arrow[d, "\phi_{\mathcal{D}}", "\wr"'] \\ F^n & F^m \end{tikzcd}$$

We now describe how to use this diagram to create a linear function $L\colon F^n \to F^m$. Since $\phi_{\mathcal{B}}$ is an isomorphism, we can invert it and then define L by starting at F^n, applying $\phi_{\mathcal{B}}^{-1}$ to go up the left-hand side of the diagram arriving at V, then applying f to go to W, and finally using $\phi_{\mathcal{D}}$ to go from W to F^m. More succinctly, define:

$$L := \phi_{\mathcal{D}} \circ f \circ \phi_{\mathcal{B}}^{-1}.$$

In sum, we have the following important *commutative diagram*:

$$\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, "\phi_{\mathcal{B}}"', "\wr"] & W \arrow[d, "\phi_{\mathcal{D}}", "\wr"'] \\
F^n \arrow[r, "L"] & F^m
\end{tikzcd}$$

Saying the diagram is *commutative* means the no matter which path we take from V to F^m, we arrive at the same place, i.e.,

$$L \circ \phi_{\mathcal{B}} = \phi_{\mathcal{D}} \circ f.$$

Now L is a mapping between tuples and, thus, has a matrix, as discussed at the beginning of this lecture. To keep track of all of the input data, we use the following, necessarily complicated, notation for this matrix:

$$[f]_{\mathcal{B}}^{\mathcal{D}} := \text{ matrix corresponding to } L.$$

How do we compute this matrix? The algorithm for computing $[f]_{\alpha}^{\beta}$ is summarized in the diagram below:

$$\begin{tikzcd}
v_j \arrow[rrr] \arrow[ddd, dashed] & & & f(v_j) \arrow[ddd, "\text{take coords. wrt. } \mathcal{D}"] \\
& V \arrow[r, "f"] \arrow[d, "\phi_{\mathcal{B}}"', "\wr"] & W \arrow[d, "\phi_{\mathcal{D}}", "\wr"'] & \\
& F^n \arrow[r, "{[f]_{\mathcal{B}}^{\mathcal{D}}}"] & F^m. & \\
e_j \arrow[rrr, dashed] & & & j\text{-th col. of } [f]_{\mathcal{B}}^{\mathcal{D}}
\end{tikzcd}$$

In words: since $[f]_{\mathcal{B}}^{\mathcal{D}}$ is a matrix, its j-th column is given by $[f]_{\mathcal{B}}^{\mathcal{D}}(e_j)$. By definition,

$$[f]_{\mathcal{B}}^{\mathcal{D}}(e_j) = \phi_{\mathcal{D}} \circ f \circ \phi_{\mathcal{B}}^{-1}(e_j) = \phi_{\mathcal{D}}\left(f\left(\phi_{\mathcal{B}}^{-1}(e_j)\right)\right).$$

We have $\phi_{\mathcal{B}}(v_j) = e_j$. Hence, $\phi_{\mathcal{B}}^{-1}(e_j) = v_j \in V$. So,

$$[f]_{\mathcal{B}}^{\mathcal{D}}(e_j) = \phi_{\mathcal{D}}\left(f\left(\phi_{\mathcal{B}}^{-1}(e_j)\right)\right) = \phi_{\mathcal{D}}\left(f(v_j)\right).$$

So here is the algorithm for computing $[f]_{\mathcal{B}}^{\mathcal{D}}$:

To find the j-th column of $[f]_{\mathcal{B}}^{\mathcal{D}}$ compute the coordinates of $f(v_j)$ with respect to $\mathcal{D} = \langle w_1, \ldots, w_m \rangle$ for each $v_j \in \mathcal{B} = \langle v_1, \ldots, v_n \rangle$.

Example. Consider linear function $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ given by the matrix

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}.$$

Thus, $f(x, y) = (x + 4y, 2x + 3y$. Using the notation above, we are letting $V = W = \mathbb{R}^2$. Take the same ordered basis for both V and W given by

$$\mathcal{B} = \mathcal{D} = \langle (1, 1), (-2, 1) \rangle.$$

Find the matrix representing f with respect to this choice of bases for domain and codomain.

Solution. To conform with our earlier notation, we take $v_1 = (1, 1)$ and $v_2 = (-2, 1)$. First apply f to each of the basis vectors for V:

$$\begin{aligned} f(v_1) &= (5, 5) \\ f(v_2) &= (2, -1). \end{aligned}$$

Next, take the coordinates of these vectors with respect to the basis $\mathcal{B}$ for W:

$$\begin{aligned} (5, 5) &= 5v_1 + 0 \cdot v_2 \\ (2, -1) &= 0 \cdot v_1 - v_2. \end{aligned}$$

Hence,

$$\begin{aligned} \phi_{\mathcal{B}}(v_1) &= (5, 0) \\ \phi_{\mathcal{B}}(v_2) &= (0, -1). \end{aligned}$$

These are the columns for our matrix:

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}.$$

We arrive at the commutative diagram:

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "f"] \arrow[d, "\phi_{\mathcal{B}}"', "\wr"] & \mathbb{R}^2 \arrow[d, "\phi_{\mathcal{B}}", "\wr"'] \\ \mathbb{R}^2 \arrow[r, "L"] & \mathbb{R}^2 \end{tikzcd}$$

Where L is the linear function corresponding to $[f]_{\mathcal{B}}^{\mathcal{B}}$, i.e.,

$$L(x, y) = (5x, -y).$$

Example. Consider the linear mapping

$$\begin{aligned} f\colon \mathbb{R}[x]_{\leq 2} &\to \mathbb{R}[x]_{\leq 3} \\ p &\mapsto xp. \end{aligned}$$

Thus, f consists of multiplying a polynomial by x. Choose bases $\mathcal{B} = \langle 1, x, x^2 \rangle$ for the domain and $\mathcal{D} = \langle 1, x, x^2, x^3 \rangle$ for the codomain. Thus, $\phi_{\mathcal{B}}(a+bx+cx^2) = (a, b, c)$ and $\phi_{\mathcal{D}}(a+bx+cx^2+dx^3) = (a, b, c, d)$. To find $[f]_{\mathcal{B}}^{\mathcal{D}}$, compute the images of the elements in $\mathcal{B}$ and express them as linear combinations of elements of $\mathcal{D}$:

$$\begin{aligned} f(1) &= x = 0 \cdot 1 + 1 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\ f(x) &= x^2 = 0 \cdot 1 + 0 \cdot x + 1 \cdot x^2 + 0 \cdot x^3 \\ f(x^2) &= x^3 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 1 \cdot x^3. \end{aligned}$$

Therefore,

$$\begin{aligned} [f(1)]_{\mathcal{D}} &= (0, 1, 0, 0) \\ [f(x)]_{\mathcal{D}} &= (0, 0, 1, 0) \\ [f(x^2)]_{\mathcal{D}} &= (0, 0, 0, 1). \end{aligned}$$

These vectors are the columns for our matrix:

$$[f]_{\mathcal{B}}^{\mathcal{D}} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$