Math 201 lecture for Wednesday, Week 6

Linear transformations and matrices II

Recall from last time that given a linear mapping $f\colon V \to W$ and ordered bases $\mathcal{B} = \langle v_1, \dots, v_n\rangle$ and $\mathcal{D} = \langle w_1, \dots, w_m\rangle$ for V and W, respectively, we have a commutative diagram

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

where

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

The matrix representing L is denoted $[f]_{\mathcal{B}}^{\mathcal{D}}$ and its calculation is displayed in the following diagram:

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

We compute $[f]_{\mathcal{B}}^{\mathcal{D}}$ by finding each of its columns: To find the j-th column of $[f]_{\mathcal{B}}^{\mathcal{D}}$ compute the coordinates of $f(v_j)$ with respect to $\mathcal{D} = \langle w_1, \dots, w_m\rangle$ for each $v_j \in \mathcal{B} = \langle v_1, \dots, v_n\rangle$. Commutativity of the diagram says that for each $v \in V$

$$[f]_{\mathcal{B}}^{\mathcal{D}}(\phi_{\mathcal{B}}(v)) = \phi_{\mathcal{D}}(f(v)).$$

Recall our notation for the coordinates of a vector with respect to a ordered basis, we can rewrite that above as

$$[f]_{\mathcal{B}}^{\mathcal{D}}[v]_{\mathcal{B}} = [f(v)]_{\mathcal{D}}.$$

Example. Consider the linear function

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

Choose ordered bases $\mathcal{B} = \langle 1, x, x^2\rangle$ and $\mathcal{D} = \langle 1, x, x^2, x^3\rangle$ for the domain and codomain, respectively. Find the matrix representing f with respect to these bases, and use the matrix to computer $f(3 + 2x + x^2)$.

Solution. Compute the images of the basis vectors in $\mathcal{B}$:

$$\begin{aligned}
f(1) &= x \cdot 1 + 2(1)' = x \\
f(x) &= x \cdot x + 2(x)' = x^2 + 2 \\
f(x^2) &= x \cdot x^2 + 2(x^2)' = x^3 + 4x.
\end{aligned}$$

Next, find the coordinates of each of these with respect to $\mathcal{D}$:

$$\begin{aligned}
[x]_{\mathcal{D}} &= (0,1,0,0) \\
[x^2+2]_{\mathcal{D}} &= (2,0,1,0) \\
[x^3+4]_{\mathcal{D}} &= (0,4,0,1).
\end{aligned}$$

Therefore,

$$[f]_{\mathcal{B}}^{\mathcal{D}} = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Here is a helpful way to think about this matrix:

$$\begin{array}{c c}
 & \begin{array}{ccc} f(1) & f(x) & f(x^2) \end{array} \\
\begin{array}{c} 1 \\ x \\ x^2 \\ x^3 \end{array} & \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{array}$$

The columns are labeled by the images of the basis vectors of the domain, and the rows are labeled by basis vectors of codomain.

To find $f(3+2x+x^2)$, we first do the calculation using coordinates:

$$\begin{aligned}
[f(3+2x+x^2)]_{\mathcal{D}} &= [f]_{\mathcal{B}}^{\mathcal{D}}[3+2x+x^2]_{\mathcal{B}} \\
&= \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 4 \\ 7 \\ 2 \\ 1 \end{pmatrix}.
\end{aligned}$$

It follows that

$$f(3+2x+x^2) = 4 + 7x + 2x^2 + x^3.$$

Check using the definition of f:

$$\begin{aligned} f(3+2x+x^2) &= x(3+2x+x^2) + 2(3+2x+x^2)' \\ &= (3x+2x^2+x^3) + 2(2+2x) \\ &= 4+7x+2x^2+x^3. \end{aligned}$$

We would next like to prove that the rank of a linear function is equal to the rank of any matrix representative of that function. Recall that the rank of a linear function is the dimension of its image, and the rank of a matrix is the dimension of its column space (which we saw is equal to the dimension of its row space—it is the number of pivot columns in the reduced row echelon form of the matrix).

Proposition. Let V and W be finite-dimensional vector spaces with ordered bases $\mathcal{B}$ and $\mathcal{D}$, respectively. Let $f\colon V \to W$ be a linear transformation. Then

$$\operatorname{rank}(f) = \operatorname{rank}([f]_{\mathcal{B}}^{\mathcal{D}}).$$

Proof. We first consider the special case of a linear mapping $L_A\colon F^n \to F^m$ where $A \in M_{m\times n}$. Thus, $L_A(x) = Ax$. We saw last time that the image of L_A is the span of the column of A, i.e., the column space of A. Thus, the result holds in this case:

$$\operatorname{rank}(L_A) := \dim(\operatorname{im}(L_A)) = \dim(\operatorname{colspace}(A)) = \operatorname{rank}(A).$$

Now consider the general case. We have the commutative diagram

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

where, in this case, $A = [f]_{\mathcal{B}}^{\mathcal{D}}$. Since $\phi_{\mathcal{B}}$ and $\phi_{\mathcal{D}}$ are isomorphism and the diagram commutes,

$$\operatorname{rank}(L_A) := \dim(\operatorname{im}(L_A)) = \dim(\operatorname{im}(L_A \circ \phi_{\mathcal{B}})) = \dim(\operatorname{im}(\phi_{\mathcal{D}} \circ f)) = \dim(\operatorname{im}(f)) =: \operatorname{rank}(f).$$

We have seen that $\operatorname{rank}(L_A) = \operatorname{rank}(A)$. So the result follows, in general. $\square$

Corollary. With notation as above, let $A = [f]_{\mathcal{B}}^{\mathcal{D}} \in M_{m\times n}(F)$.

(a) f is surjective if and only if $\operatorname{rank}(A) = m = \dim(W)$.
(b) f is injective if and only if $\operatorname{rank}(A) = n = \dim(V)$.
(c) f is an isomorphism if and only if $\operatorname{rank}(A) = m = n$.

Proof.

(a) The function f being surjective means that $\operatorname{im}(f) = W$, which is equivalent to saying that $\dim(\operatorname{im}(f)) = \dim(W)$, or that $\operatorname{rank}(f) = m$, and we have just seen that $\operatorname{rank}(f) = \operatorname{rank}(A)$.

(b) We know that f is injective if and only if $\dim(\ker(f)) = 0$. By the rank-nullity theorem,

$$n = \dim V = \dim(\operatorname{im}(f)) + \dim(\ker(f)) = \operatorname{rank}(f) + \dim(\ker(f)).$$

From the Proposition, we have $\operatorname{rank}(f) = \operatorname{rank}(A)$. Therefore, $\dim(\ker f) = 0$ if and only if $\operatorname{rank}(A) = n$.

(c) This part follows from the previous two.

□

Composition of linear functions. Consider the linear functions

$$\begin{aligned} f\colon \mathbb{R}^4 &\to \mathbb{R}^2 \\ (x, y, z, w) &\mapsto (2x - z + 3w, x - y + 4z) \end{aligned}$$

and

$$\begin{aligned} g\colon \mathbb{R}^2 &\to \mathbb{R}^3 \\ (s, t) &\mapsto (5s - t, 2t, -3s). \end{aligned}$$

Let's compute the composition $g \circ f\colon \mathbb{R}^4 \to \mathbb{R}^3$:

$$\begin{aligned} (g \circ f)(x, y, z, w) &= g(\underbrace{2x - z + 3w}_{s}, \underbrace{x - y + 4z}_{t}) \\ &= (5(2x - z + 3w) - (x - y + 4z), 2(x - y + 4z), -3(2x - z + 3w)) \\ &= (9x + y - 9z + 15w, 2x - 2y + 8z, -6x + 3z - 9w). \end{aligned}$$

The matrices associated with f and g (with respect to the standard bases) are, respectively,

$$\begin{pmatrix} 2 & 0 & -1 & 3 \\ 1 & -1 & 4 & 0 \end{pmatrix}, \quad \begin{pmatrix} 5 & -1 \\ 0 & 2 \\ -3 & 0 \end{pmatrix}, \quad \begin{pmatrix} 9 & 1 & -9 & 15 \\ 2 & -2 & 8 & 0 \\ -6 & 0 & 3 & -9 \end{pmatrix}.$$

What is the relation among these matrices? We will take up this question next time.