Math 201 lecture for Friday, Week 6

Linear transformations and matrices III

The goal today is to formally define the algebraic structure for matrices (linear structure and multiplication). Multiplication of matrices corresponds with composition of their corresponding linear transformations.

Composition of linear functions.

Proposition. Let $f\colon V \to W$ and $g\colon W \to U$ be linear functions. Their the composition $g \circ f\colon V \to U$ is a linear function.

Proof. Let $u, v \in V$ and $\lambda \in F$. Then, since f and g are linear,

$$\begin{aligned}(g \circ f)(u + \lambda v) &:= g(f(u + \lambda v))\\ &= g\left(f(u) + \lambda f(v)\right)\\ &= g(f(u)) + \lambda g(f(v))\\ &= (g \circ f)(u) + \lambda (g \circ f)(v).\end{aligned}$$

□

Let $f\colon V \to W$ and $g\colon W \to U$ be a linear functions. We are interested in a matrices representing the composition

$$g \circ f\colon V \xrightarrow{f} W \xrightarrow{g} U.$$

Fix ordered bases $\mathcal{B} = \langle v_1, \dots, v_n\rangle$ for V, $\mathcal{C} = \langle w_1, \dots, w_\ell\rangle$ for W, and $\mathcal{D} = \langle u_1, \dots, u_m\rangle$ for U. Let

$$P := [g]_{\mathcal{C}}^{\mathcal{D}} \quad \text{and} \quad Q = [f]_{\mathcal{B}}^{\mathcal{C}}.$$

Thus, $P \in M_{m\times\ell}(F)$ and $Q \in M_{\ell\times n}$. The relevant commutative diagram is

$$\begin{array}{ccccc} V & \xrightarrow{f} & W & \xrightarrow{g} & U \\ {\scriptstyle \phi_{\mathcal{B}}}\downarrow\wr & & \wr\downarrow{\scriptstyle \phi_{\mathcal{C}}} & & \wr\downarrow{\scriptstyle \phi_{\mathcal{D}}} \\ F^n & \xrightarrow{Q} & F^\ell & \xrightarrow{P} & F^m \end{array}$$

Let's compute $[g \circ f]_{\mathcal{B}}^{\mathcal{D}}$. To find its j-th column, we find the coordinates of $(g \circ f)(v_j)$ with respect to the ordered basis $\mathcal{D}$:

$$\begin{aligned}
(g \circ f)(v_j) &= g(f(v_j)) \\
&= g\left(\sum_{k=1}^{\ell} Q_{kj} w_k\right) && (\text{ j-th column of } Q) \\
&= \sum_{k=1}^{\ell} Q_{kj} g(w_k) \\
&= \sum_{k=1}^{\ell} Q_{kj} \left(\sum_{i=1}^{m} P_{ik} u_i\right) && (\text{ k-th column of } P) \\
&= \sum_{i=1}^{m} \left(\sum_{k=1}^{\ell} P_{ik} Q_{kj}\right) u_i.
\end{aligned}$$

So the j column of $[g \circ f]_{\mathcal{B}}^{\mathcal{D}}$ is given by the coefficients of the u_i in the above some. That means that the (i, j)-th entry of the matrix $[g \circ f]_{\mathcal{B}}^{\mathcal{D}}$, i.e., the entry in its i-row and j-th column is

$$\left([g \circ f]_{\mathcal{B}}^{\mathcal{D}}\right)_{ij} = \sum_{k=1}^{m} P_{ik} Q_{kj}.$$

Definition. (Multiplication of matrices) Let $P \in M_{m \times \ell}(F)$ and $Q \in M_{\ell \times n}(F)$, then the product $PQ \in M_{m \times n}(F)$ is defined by

$$(PQ)_{ij} = \sum_{k=1}^{\ell} P_{ik} Q_{kj}.$$

Note: The formula says that the (i, j)-th entry of the product PQ is the dot product of the i-th row of P with the j-th column of Q. That's what one thinks about when performing the calculation of PQ in practice.

Example. Here is an example of the product of two matrices. For instance, to find the $(2, 3)$-entry of the product, we take the dot product of the second row of the first matrix with the third column of the second:

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

Recall the relevance of this computation: the first two matrices encode linear functions g and f, and their product is a matrix encoding the composition $g \circ f$.

Proposition. Let $f\colon V \to W$ and $g\colon W \to U$ be a linear functions, and fix ordered bases $\mathcal{B} = \langle v_1, \ldots, v_n\rangle$ for V, $\mathcal{C} = \langle w_1, \ldots, w_\ell\rangle$ for W, and $\mathcal{D} = \langle u_1, \ldots, u_m\rangle$ for U. Then we have

$$[g \circ f]_{\mathcal{B}}^{\mathcal{D}} = [g]_{\mathcal{C}}^{\mathcal{D}}[f]_{\mathcal{B}}^{\mathcal{C}}.$$

Proof. The proof is exactly the motivation we just gave for the definition of the matrix product. □

We summarize some basic properties of matrix algebra.

Proposition. Let A be an $m \times n$ matrix, B an $n \times r$ matrix, both over a field F, and $\lambda \in F$.

(a) $\lambda(AB) = (\lambda A)B = A(\lambda B)$.

(b) $A(BC) = (AB)C$ for all $r \times s$ matrices C.

(c) $A(B + C) = AB + AC$ for all $n \times r$ matrices C.

(d) $(C + D)A = CA + DA$ for all $r \times m$ matrices C and D.

Proof. We will just prove part (b), associativity of multiplication. So let C be an $r \times s$ matrix. We have

$$\begin{aligned}
(A(BC))_{ij} &= \sum_{k=1}^{n} A_{ik}(BC)_{kj} \\
&= \sum_{k=1}^{n} \left(A_{ik} \left(\sum_{\ell=1}^{r} B_{k\ell}C_{\ell j} \right) \right) \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{r} A_{ik}(B_{k\ell}C_{\ell j}) \\
&= \sum_{\ell=1}^{r} \sum_{k=1}^{n} A_{ik}(B_{k\ell}C_{\ell j}) \\
&= \sum_{\ell=1}^{r} \sum_{k=1}^{n} (A_{ik}B_{k\ell})C_{\ell j} \\
&= \sum_{\ell=1}^{r} \left(\sum_{k=1}^{n} A_{ik}B_{k\ell} \right) C_{\ell j} \\
&= \sum_{\ell=1}^{r} (AB)_{i\ell}C_{\ell j} \\
&= ((AB)C)_{ij}.
\end{aligned}$$

□

Warning. Matrix multiplication is not commutative, in general. First of all, if the dimensions aren't right, multiplication for both AB and BA might not make sense. For instance, if

$$A = \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 4 \end{pmatrix},$$

then AB is defined, but not BA.

However, even if AB and BA are both defined, it is usually not the case that $AB = BA$. Try just about any example with 2×2 matrices to see this.

Definition. Let V and W be vector spaces over a field F. The set of linear transformations (homomorphisms) from V to W is denoted $\mathcal{L}(V, W)$ or $\mathrm{Hom}(V, W)$. It forms a vector space with operations defined as follows: for $f, g \in \mathrm{Hom}(V, W)$ and $\lambda \in F$,

$$(f + g)(v) = f(v) + g(v) \quad \text{and} \quad (\lambda f)(v) = \lambda f(v)$$

for all $v \in V$.

Proposition. Let V and W be vectors spaces over F of dimension n and m, respectively. Then there is an isomorphism of vector spaces

$$\mathrm{Hom}(V, W) \to M_{m \times n}(F).$$

Sketch of proof. Choose ordered bases $\mathcal{B}$ and $\mathcal{D}$ for V and W, respectively. Then an isomorphism is given by

$$\begin{gathered} \mathrm{Hom}(V, W) \to M_{m \times n}(F) \\ f \mapsto [f]_{\mathcal{B}}^{\mathcal{D}}. \end{gathered}$$

This isomorphism will change to a different isomorphism if different bases are chosen. □