MATRICES

ROHAN RAMCHAND

We will start by defining the concept of a matrix.

Definition 1. *Let F be a field and let $m, n \in \mathbb{N}^{\geq 0}$. Then an $m \times n$ F-matrix is a table consisting of m rows and n columns with elements in F.*

We defined in previous sections (see Linear Transformations) the **matrix space** as

$$(\mathrm{Mat}_{m\times n}(F), +, \alpha),$$

or the set of all $m \times n$ F-matrices, along with matrix addition and scalar multiplication. Furthermore, $\dim \mathrm{Mat}_{m\times n}(F) = mn$. The proof of this statement follows.

Proof. We begin by defining the **standard basis of** $\mathrm{Mat}_{m\times n}(F)$ as the set of matrices $E = \{E_{0,0}..., E_{m,n}\}$, where $E_{m,n} = (e_{i,j})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$ such that $e_{i,j} = \delta_{m,n}(i,j)$. Then $A = \sum_{i,j} \alpha_{i,j} E_{i,j}$, where $\alpha_{i,j} = A_{i,j}$; therefore, E spans A. Furthermore, if $\sum_{i,j} \alpha_{i,j} E_{i,j} = 0$, then $\alpha_{i,j} = 0$ and E is linearly independent. Therefore, E is a basis with length mn; then $\dim \mathrm{Mat}_{m\times n} = mn$ and the proof is complete. $\square$

We now define one final operation on matrices: matrix multiplication.

Definition 2. *Let $A \in \mathrm{Mat}_{m\times r}(F)$ and $B \in \mathrm{Mat}_{r\times n}(F)$. Then the product of A and B, $C = A \cdot B \in \mathrm{Mat}_{m\times n}(F)$, is defined as*

$$c_{i,j} = \sum_{k=1}^{r} a_{i,k} b_{k,j}.$$

Matrices and Linear Transformations

The following two properties of matrices are used in the proofs that follow and are stated without proof.

Distributivity: Let $A \in \mathrm{Mat}_{m\times r}(F)$ and $B_1, B_2 \in \mathrm{Mat}_{r\times n}(F)$. Then

$$A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2.$$

Scalar Multiplication: Let $A \in \mathrm{Mat}_{m\times r}(F)$, $B \in \mathrm{Mat}_{r\times n}(F)$, and $\lambda \in F$. Then

$$A \cdot (\lambda \cdot B) = \lambda \cdot (A \cdot B).$$

We now define the following map.

Definition 3. *Let $A \in \mathrm{Mat}_{m\times n}(F)$. Define*

$$T_A : \mathrm{Mat}_{n\times r} \to \mathrm{Mat}_{m\times r} : T_A(B) = A \cdot B.$$

Theorem 1. *T_A is a linear transformation.*

Proof. Both properties of matrix multiplication above satisfy the properties of a linear transformation. □

We will, in particular, study the case $r = 1$, defined below.

Definition 4. *Let $A \in \mathrm{Mat}_{m\times r}$ and let $r = 1$. Then A is a* **column vector**. *In particular, the set of m-dimensional column vectors is referred to as $\mathrm{col}_m(F)$. Furthermore, $\mathrm{col}_m(F)$ can be identified with F^m, the former being a column vector and the latter being (canonically) a row vector.*

Therefore, we redefine T_A as $T_A : \mathrm{Mat}_{m\times 1} \to \mathrm{Mat}_{n\times 1} : F^n \to F^m$.

We will now revisit the concept of matrix multiplication as composition of linear transformations. Let

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

and let

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Then

$$\begin{aligned}
T_A(\vec{x}) &= \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ a_n \end{pmatrix} \\
&= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \end{pmatrix} \\
&= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + ... + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}
\end{aligned}$$

This is one of the more important results of this course: the product of a matrix and a vector is a linear combination of the columns of the matrix, where the coefficients in the combination are the elements of the vector. This implies that the result of matrix multiplication is a linear transformation in and of itself, and therefore T_A, as defined above, sends $\mathrm{Mat}_{m\times n}(F)$ to $\mathrm{Hom}(F^n, F^m)$.

We now state an important theorem regarding T_A.

Theorem 2. *T_A is a bijective linear transformation; in other words, $\mathrm{Mat}_{m\times n}(F)$ and $\mathrm{Hom}(F^n, F^m)$ are isomorphic.*

Proof. Let $A, A_1, A_2 \in \mathrm{Mat}_{m\times n}(F)$ and $\lambda \in F$.

$$\begin{aligned} T_{A_1+A_2}(\vec{x}) &= (A_1 + A_2)\vec{x} \\ &= A_1\vec{x} + A_2\vec{x} \\ &= T_{A_1}(\vec{x}) + T_{A_2}(\vec{x}) \\ T_{\lambda A}(\vec{x}) &= (\lambda A)(\vec{x}) \\ &= \lambda(A\vec{x}) \\ &= \lambda T_A(\vec{x}) \end{aligned}$$

Therefore, T_A is a linear transformation.

Let $M = T^{-1} : \mathrm{Hom}(F^n, F^m) \to \mathrm{Mat}_{m\times n}(F)$. For a linear transformation $A : F^n \to F^m$, M_A is a matrix, composed of columns $\vec{m}_1, ..., \vec{m}_n$, such that

$$M_A \cdot \vec{x} = A(\vec{x}) = \sum_{i=1}^{n} x_i \vec{m}_i.$$

By definition, the basis $(e_1, e_2, ..., e_n)$, with $e_i = (\delta_i(1), \delta_i(2), ..., \delta_i(n))$ spans any vector space V: therefore,

$$A(\vec{x}) = A \cdot (x_1 e_1 + ... + x_n e_n) = \sum_{i=1}^{n} x_i A(e_i).$$

Therefore,

$$M_A = (A(e_1), ..., A(e_n))$$

and M_A is defined for all A; therefore, T_A has an inverse for all linear transformations A and is therefore an isomorphism. □

Note that we did not actually prove that M is a linear transformation; the proof of this statement is left as an exercise.

This concept is illustrated with the following example.

Example. *Let*

$$A : \mathbb{R}^2 \to \mathbb{R}^2 : A(x, y) = \begin{pmatrix} x + y \\ x - y \end{pmatrix}.$$

Then

$$M_A = (A(1,0), A(0,1)) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

It has been stated on multiple occasions that composition of linear transformations is equivalent to matrix multiplications; this theorem is restated formally and proven below.

Theorem 3. *Let $A \in \mathrm{Mat}_{m\times r}(F), B \in \mathrm{Mat}_{r\times n}(F)$, with $AB \in \mathrm{Mat}_{m\times n}(F)$. Then let $T_A : F^r \to F^m, T_B : F^n \to F^r$, with $T_{AB} : F^n \to F^m$. Then*

$$T_{AB} = T_A \circ T_B.$$

Proof. As a convenience measure, we will denote the ith column of a matrix M as M_i.

$$\begin{aligned} T_{AB}(\vec{x}) &= (A \cdot B) \cdot \vec{x} \\ &= \sum_{i=1}^{n} x_i (A \cdot B)_i \\ &= \sum_{i=1}^{n} x_i (A \cdot B_i) && ((A \cdot B)_i = A \cdot B_i) \\ &= \sum_{i=1}^{n} A(x_i B_i) \\ &= A \cdot \sum_{i=1}^{n} x_i B_i \\ &= A \cdot (B \cdot \vec{x}) \\ &= (T_A \circ T_B)(\vec{x}) \end{aligned}$$

□

Armed with this theorem, we can prove another theorem of matrix multiplication.

Theorem 4. *Let $A \in \mathrm{Mat}_{m\times r}(F), B \in \mathrm{Mat}_{r\times l}(F), C \in \mathrm{Mat}\, l \times n(F)$ be matrices. Then*

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Proof.

$$T_{(A\cdot B)\cdot C} = T_{A\cdot(B\cdot C)} \tag{1}$$

$$T_{A\cdot B} \circ T_C = T_A \circ T_{B\cdot C} \tag{2}$$

$$(T_A \circ T_B) \circ T_C = T_A \circ (T_B \circ T_C) \tag{3}$$

By associativity of composition of functions, (3) is true and the proof is complete. □