Matrices Representing Transformations

Matrix multiplication is useful for defining linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. Let us review this central idea.

Suppose $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear transformation. Then the linearity of $T$ permits us to completely determine its values by knowing what it does to the vectors in a basis for the domain $\mathbb{R}^n$. Suppose that $\{E_1, E_2, \ldots, E_m\}$ is the canonical basis for $\mathbb{R}^m$ and that $\{F_1, F_2, \ldots, F_n\}$ is the canonical basis for $\mathbb{R}^n$. Then we can express each of the vectors $T(F_1), T(F_2), \ldots, T(F_n)$ as a linear combination of the $E$’s:

\[
T(F_1) = a_{11}E_1 + a_{21}E_2 + \cdots + a_{m1}E_m,
\]
\[
T(F_2) = a_{12}E_1 + a_{22}E_2 + \cdots + a_{m2}E_m,
\]
\[
\vdots
\]
\[
T(F_n) = a_{1n}E_1 + a_{2n}E_2 + \cdots + a_{mn}E_m
\]

and since every $X \in \mathbb{R}^n$ is a linear combination of the $F$’s, so that we can write

\[
X = x_1F_1 + x_2F_2 + \cdots + x_nF_n,
\]
then

\[ T(\mathbf{X}) = T(x_1 \mathbf{F}_1 + x_2 \mathbf{F}_2 + \cdots + x_n \mathbf{F}_n) \]
\[ = x_1 T(\mathbf{F}_1) + x_2 T(\mathbf{F}_2) + \cdots + x_n T(\mathbf{F}_n) \]
\[ = x_1 (a_{11} \mathbf{E}_1 + a_{21} \mathbf{E}_2 + \cdots + a_{m1} \mathbf{E}_m) \]
\[ + x_2 (a_{12} \mathbf{E}_1 + a_{22} \mathbf{E}_2 + \cdots + a_{m2} \mathbf{E}_m) \]
\[ + \cdots + x_n (a_{1n} \mathbf{E}_1 + a_{2n} \mathbf{E}_2 + \cdots + a_{mn} \mathbf{E}_m) \]
\[ = (x_1 a_{11} + x_2 a_{12} + \cdots + x_n a_{1n}) \mathbf{E}_1 \]
\[ + (x_1 a_{21} + x_2 a_{22} + \cdots + x_n a_{2n}) \mathbf{E}_2 \]
\[ + \cdots + (x_1 a_{m1} + x_2 a_{m2} + \cdots + x_n a_{mn}) \mathbf{E}_m \]

so that, if we represent the vector \( \mathbf{X} \) in terms of its coordinates relative to the basis of \( \mathbf{F}'s \) as the \( n \times 1 \) (column) matrix

\[
\mathbf{X}_F = \begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{bmatrix},
\]

then the coordinates of \( T(\mathbf{X}) \) relative to the basis of \( \mathbf{E}'s \) can be represented as the \( m \times 1 \) (column) matrix
where the $m \times n$ matrix $A$ is seen to be independent of $X$. That is, the image of $X$ under $T$ can be determined by a matrix multiplication. We call $A$ the **matrix of $T$ relative to the ordered bases** \{$E_1, E_2, \ldots, E_m$\} and \{$F_1, F_2, \ldots, F_n$\}. (The bases are ordered because the list of $E$’s fixes the ordering of the rows of $A$ and the list of $F$’s fixes its columns.) While we described this idea of the matrix of a transformation in the context of a transformation carrying vectors from $\mathbb{R}^n$ to $\mathbb{R}^m$, there is nothing in this discussion that referred to features of the domain and codomain vector spaces other than that they are finite dimensional.

That is, we can just as easily form the matrix of a linear transformation $T: \mathcal{V} \rightarrow \mathcal{W}$ between any two finite dimensional vector spaces by choosing an
ordered basis \( \{ \mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n \} \) for \( \mathcal{V} \) and an ordered basis \( \{ \mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_m \} \) for \( \mathcal{W} \); the images of the basis vectors \( T(\mathbf{V}_1), T(\mathbf{V}_2), \ldots, T(\mathbf{V}_n) \) can be represented in terms of the \( \mathbf{W} \)'s in the form

\[
T(\mathbf{V}_j) = \sum_{i=1}^{m} a_{ij} \mathbf{W}_i, \quad j = 1, 2, \ldots, n.
\]

Then, expressing any \( \mathbf{X} \in \mathcal{V} \) in terms of the \( \mathbf{V} \)'s as \( \mathbf{X} = \sum_{j=1}^{n} x_j \mathbf{V}_j \) allows us to represent it as a column vector (an \( n \times 1 \) matrix)

\[
\mathbf{X}_\mathcal{V} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix};
\]

linearity of \( T \) then permits \( T(\mathbf{X}) \) to be written in terms of the \( \mathbf{W} \)'s as a column vector (an \( m \times 1 \) matrix)

\[
T(\mathbf{X})_{\mathcal{W}} = \begin{bmatrix} x_1 a_{11} + x_2 a_{12} + \cdots + x_n a_{1n} \\ x_1 a_{21} + x_2 a_{22} + \cdots + x_n a_{2n} \\ \vdots \\ x_1 a_{m1} + x_2 a_{m2} + \cdots + x_n a_{mn} \end{bmatrix}.
\]
which is more simply realized as the matrix product

\[
T(\mathbf{X})_W = \mathbf{A}\mathbf{X}_V = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} & x_1 \\
a_{21} & a_{22} & \cdots & a_{2n} & x_2 \\
\vdots & \vdots & & \vdots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn} & x_n
\end{bmatrix}
\]

where \( \mathbf{A} \) is the \( m \times n \) matrix \((a_{ij})\).

Note again that the representation of the linear transformation \( T : \mathcal{V} \to \mathcal{W} \) in terms of the \( m \times n \) matrix \( \mathbf{A} \) depends on the choice of ordered bases for \( \mathcal{V} \) and \( \mathcal{W} \): different bases for these spaces (or even different orderings of the same bases) can provide different matrix representations for the same transformation \( T \).
The correspondence between linear transformations $T: \mathcal{V} \to \mathcal{W}$ and their $m \times n$ matrix representations $A$ in terms of a choice of ordered basis for each of the spaces $\mathcal{V}$ and $\mathcal{W}$ leads to this natural

**Theorem** The map $M: \mathcal{L}(\mathcal{V}, \mathcal{W}) \to \text{Mat}_{m,n}$ that sends the linear transformation $T: \mathcal{V} \to \mathcal{W}$ to its matrix representation $A$ in terms of some fixed choice of ordered basis for each of the spaces $\mathcal{V}$ and $\mathcal{W}$ is in fact an isomorphism of vector spaces.

**Proof** $M$ is a linear map because if $T, U \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, and $T$ has matrix $A$ and $U$ matrix $B$ with respect to these ordered bases, then the transformation $T + U$ is represented by matrix $A + B$ (why?). Also, the transformation $rT$, where $r$ is any scalar, is represented by the matrix $rA$ in terms of these same bases (why?).

Further, given any $A \in \text{Mat}_{m,n}$, we can define a linear transformation $T: \mathcal{V} \to \mathcal{W}$ by means of the formula $T(X)_W = AX_V$ where $V$ and $W$ refer to some pair of fixed ordered bases for $\mathcal{V}$ and $\mathcal{W}$ (why is this linear?). It follows that we can define another map $L: \text{Mat}_{m,n} \to \mathcal{L}(\mathcal{V}, \mathcal{W})$ between these vector spaces that sends the matrix $A$ to the linear
transformation $T(X)_W = AX_V$.

Because of how we have defined $L$, the matrix of the transformation $T = L(A)$ is $A$ itself. That is, $M(L(A)) = A$. On the other hand, given any $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, if $M(T)$ is the matrix $A$, then the linear transformation $L(A) = L(M(T))$ satisfies

$$[L(M(T))(X)]_W = AX_V = T(X)_W$$

for every vector $X$ in $\mathcal{V}$. Thus $L(M(T))$ is identical with the transformation $T$.

This shows that $M \circ L$ is the identity map on $\text{Mat}_{m,n}$ and that $L \circ M$ is the identity map on $\mathcal{L}(\mathcal{V}, \mathcal{W})$. So $M$ is an isomorphism of vector spaces with inverse $L$. //

**Corollary** $\dim(\mathcal{L}(\mathcal{V}, \mathcal{W})) = mn = \dim(\mathcal{V}) \cdot \dim(\mathcal{W})$. //
There is yet additional algebraic structure that links the behavior of linear transformations in $\mathcal{L}(\mathcal{V}, \mathcal{W})$ and of matrices in $\text{Mat}_{m,n}$. Suppose that $S: \mathcal{V} \to \mathcal{W}$ is a linear transformation in $\mathcal{L}(\mathcal{V}, \mathcal{W})$ and that $T: \mathcal{U} \to \mathcal{V}$ is a linear transformation in $\mathcal{L}(\mathcal{U}, \mathcal{V})$ (where we will take $\mathcal{U}$ to be a $p$-dimensional vector space). Then the composition

$$S \cdot T: \mathcal{U} \xrightarrow{T} \mathcal{V} \xrightarrow{S} \mathcal{W}$$

is a linear transformation in $\mathcal{L}(\mathcal{U}, \mathcal{W})$. If we choose in each of these vector spaces $\mathcal{U}$, $\mathcal{V}$, $\mathcal{W}$ an ordered basis, then with respect to these bases we can express the transformations $S$ and $T$ in terms of matrices. What can we say about the matrix representing $S \cdot T$?

**Proposition** Suppose $\mathcal{U}$, $\mathcal{V}$, $\mathcal{W}$ are vector spaces of dimensions $p$, $n$, and $n$, respectively, and fix ordered bases for each. If $A \in \text{Mat}_{m,n}$ is the matrix representing $S \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ and $B \in \text{Mat}_{n,p}$ is the matrix representing $T \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, then the matrix representing $S \cdot T \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ is $AB \in \text{Mat}_{m,p}$. That is, the matrix representing the composition of transformations is the product of the matrices that represent the transformations being composed.
Proof Suppose that it is the ordered bases 
\{U_1, U_2, \ldots, U_p\}, \{V_1, V_2, \ldots, V_n\}, \{W_1, W_2, \ldots, W_m\}
for \(\mathcal{U}, \mathcal{V}, \mathcal{W}\), respectively, with respect to which the transformations \(S\) and \(T\) have been represented by the \(m \times n\) matrix \(A = (a_{ij})\) and the \(n \times p\) matrix \(B = (b_{kl})\). Then \(S\) is the linear extension of the map

\[ S(V_j) = \sum_{i=1}^{m} a_{ij} W_i, \quad j = 1, 2, \ldots, n, \]

and \(T\) is the linear extension of the map

\[ T(U_l) = \sum_{k=1}^{n} b_{kl} V_k, \quad l = 1, 2, \ldots, p. \]

So the composition map satisfies

\[ S \circ T(U_l) = S(T(U_l)), \quad l = 1, 2, \ldots, p. \]

Using the formulas above and the linearity of the maps, we can manipulate this to obtain

\[ S \circ T(U_l) = S(\sum_{k=1}^{n} b_{kl} V_k) \]

\[ = \sum_{k=1}^{n} b_{kl} S(V_k) \]
\[
\begin{align*}
&= \sum_{k=1}^{n} b_{kl} \left( \sum_{i=1}^{m} a_{ik} W_i \right) \\
&= \sum_{k=1}^{n} \sum_{i=1}^{m} a_{ik} b_{kl} W_i \\
&= \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} b_{kl} W_i \\
&= \sum_{i=1}^{m} \left( \sum_{k=1}^{n} a_{ik} b_{kl} \right) W_i
\end{align*}
\]

which we recognize as being represented by the matrix $AB$ with respect to the bases for $U$ and $W$.

---

**Corollary** $T: \mathcal{V} \rightarrow \mathcal{V}$ is an **automorphism** (an isomorphism between a vector space and itself) if and only if the matrix $M(T) = B$ that represents it with respect to any pair of ordered bases for $\mathcal{V}$ is invertible.

**Proof** Take $U = \mathcal{V} = \mathcal{W}$ in the theorem above; then $m = n = p$ as well. Fix two ordered bases $\{U_1, U_2, \ldots, U_n\}$ and $\{V_1, V_2, \ldots, V_n\}$ and let $M(T) = B$ be the matrix that represents $T$ with respect to these bases.
If $T$ is an isomorphism, then there is an inverse map $S: V \to V$ that satisfies $S \circ T(X) = X$ for every vector $X \in V$. Let $S$ be represented by the matrix $A$ in terms of the ordered bases $\{V_1, V_2, \ldots, V_n\}$ and $\{U_1, U_2, \ldots, U_n\}$ (in the context of the theorem, we are taking the basis of $W$’s to be identical with the basis of $U$’s). Then $AB$ must represent the identity map on $V$ in terms of the ordered basis $\{U_1, U_2, \ldots, U_n\}$ (both for the domain and codomain spaces). Clearly then, $AB = I$, the $n \times n$ identity matrix. In addition, $T \circ S(Y) = Y$ for every vector $Y \in V$, whence $BA$ represents the identity map on $V$ in terms of the ordered basis $\{V_1, V_2, \ldots, V_n\}$ (both for domain and codomain spaces), and thus, $BA = I$ as well. It follows that $B$ is invertible, with inverse $A$.

Conversely, suppose $M(T) = B$ is invertible, with inverse $A = (a_{ij})$, where $B$ represents $T$ with respect to the ordered bases $\{U_1, U_2, \ldots, U_n\}$ and $\{V_1, V_2, \ldots, V_n\}$ for $V$. Then define the linear transformation $S: V \to V$ to be the extension of

$$S(V_j) = \sum_{i=1}^{n} a_{ij} U_i, \quad j = 1, 2, \ldots, n$$
It is straightforward to check that $S \circ T(X) = X$ for every vector $X \in \mathcal{V}$ (since $AB = I$), and that $T \circ S(Y) = Y$ for every vector $Y \in \mathcal{V}$ (since $BA = I$), whence $T$ is an isomorphism. //

**Corollary** The isomorphism $M: \mathcal{L}(\mathcal{V}, \mathcal{V}) \rightarrow \text{Mat}_{n,n}$ of vector spaces that sends the linear transformation $S: \mathcal{V} \rightarrow \mathcal{V}$ to the matrix $A$ that represents it with respect to a particular ordered basis $\{V_1, V_2, \ldots, V_n\}$ and sends the linear transformation $T: \mathcal{V} \rightarrow \mathcal{V}$ to the matrix $B$ that represents it with respect to $\{V_1, V_2, \ldots, V_n\}$ also satisfies $M(S \circ T) = AB = M(S) \cdot M(T)$. That is, $M$ is an isomorphism of algebras. (In particular, $\mathcal{L}(\mathcal{V}, \mathcal{V})$ is an algebra over its field of scalars, where vector multiplication corresponds to composition of maps.) //

**Corollary** If $A, B \in \text{Mat}_{n,n}$ satisfy $AB = I$, then $BA = I$.

**Proof** Let $S, T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the linear transformations defined by the formulas $S(X) = AX$, $T(X) = BX$. (That is, the matrix representing $S$ in terms of the canonical basis for $\mathbb{R}^n$ in both domain and codomain is given to be $A$. Similarly, the
matrix representing \( T \) in terms of the canonical basis for \( \mathbb{R}^n \) in both domain and codomain is \( B \).

Then \( S \circ T(X) = ABX = X \), that is, \( S \circ T \) is the identity map and \( S \) and \( T \) are inverse functions: if we put \( Y = BX = T(X) \), then \( S(Y) = AY = ABX = X \).

It follows that \( T \circ S(Y) = Y \), that is, \( T \circ S \) is the identity map also, so it is represented by the identity matrix \( I \). But \( T \circ S(Y) = T(AY) = BAY \) shows that \( T \circ S \) is also represented by the matrix \( BA \), so \( BA = I \). //