Matrix multiplication and composition of linear transformations

September 12, 2007

Let $B \in M_{nq}$ and let $A \in M_{pm}$ be matrices. Note that $q$ is the number of columns of $B$ and is also the length of the rows of $B$, and that $p$ is the number of rows of $A$ and is also the length of the columns of $A$.

**Definition 1** If $B \in M_{nq}$ and $A \in M_{pm}$, the matrix product $BA$ is defined if $q = p$. In this case it is the element of $M_{nm}$ whose $ij$th entry is given by

$$(BA)_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{ip}B_{pj}.$$ 

Thus the matrix product is an operation:

$$M_{np} \times M_{pm} \rightarrow M_{nm}.$$ 

**Formulas:**

1. If $m = 1$, multiplication by $B$ is a map $\mathbb{R}^p = M_{p1} \rightarrow \mathbb{R}^n = M_{n1}$. It sends a column vector $X = \begin{pmatrix} x_1 \\ \cdots \\ x_p \end{pmatrix}$ to

$$BX = x_1C_1(B) + x_2C_2(B) + \cdots + x_pC_p(B),$$

where $C_j(B)$ is the $j$th column of $B$.

2. If $m = 1$, $Be_j = C_j(B)$, where $e_j$ is the $j$th element of the standard frame for $\mathbb{R}^p$. 

1
**Theorem 2** Let \( T : \mathbb{R}^p \to \mathbb{R}^n \) be any function. Then the following are equivalent:

1. There exists an \( n \times p \) matrix \( B \) such that \( T(X) = T_B(X) \) for all \( X \in \mathbb{R}^p \).

2. \( T \) satisfies the principle(s) of superposition:
   
   (a) \( T(X + X') = T(X) + T(X') \) for all \( X \) and \( X' \) in \( \mathbb{R}^p \), and
   
   (b) \( T(cX) = cT(X) \) for all \( X \in \mathbb{R}^p \) and \( c \in \mathbb{R} \).

**Proof:** The proof that (1) implies (2) was given earlier and is in the book; it is a consequence of the properties of matrix multiplication. Let us prove that (2) implies (1). First note that if (2) is true and if \( X_1, \ldots X_r \) are vectors and \( c_1, \ldots c_r \) are numbers, then \( T(c_1 X_1 + \cdots + c_r X_r) = c_1 T(X_1) + \cdots + c_r T(X_r) \). Now if \( T \) satisfies (2), we let \( B \) the \( n \times p \) matrix whose \( j \)th column is the vector \( T(e_j) \), for \( j = 1, \ldots p \). We must prove that \( T(X) = T_B(X) \) for all \( X \in \mathbb{R}^p \). But if \( X \in \mathbb{R}^p \), \( X = x_1 e_1 + \cdots + x_p e_p \), where the \( x_i \)'s are the coordinates of \( X \). Hence

\[
T(X) = T(x_1 e_1 + \cdots x_p e_p) \\
= x_1 T(e_1) + \cdots + x_p T(e_p) \\
= x_1 C_1(B) + \cdots + x_p C_p(B) \\
= BX \\
= T_B(X)
\]

\[\square\]

**Corollary 3** A linear transformation is uniquely determined by its effect on the standard frame. More precisely, if \( T \) and \( T' \) are linear transformations from \( \mathbb{R}^p \) to \( \mathbb{R}^n \) and if \( T(e_j) = T'(e_j) \) for all \( j \), then \( T(X) = T'(X) \) for all \( X \).

**Corollary 4** The composite of two linear transformations is linear.
Proof: Suppose $S$ and $T$ are linear and composable. Then
\[
S \circ T(X + X') = S(T(X + X')) = S(T(X)) + S(T(X')) = S \circ T(X) + S \circ T(X').
\]
The proof of the compatibility with scalar multiplication is similar. \qed

**Theorem 5** If $B \in M_{np}$ and $A \in M_{pm}$, then the composite $T_B \circ T_A$ is $T_{BA}$.

**Proof:** We know that $T_B \circ T_A$ and $T_{BA}$ are linear. To prove that they are equal it suffices to check that they have the same effect on each $e_j$. We compute
\[
T_B \circ T_A(e_j) = T_B(T_A(e_j)) = T_B(C_j(A)) = BC_j(A) = C_j(BA) = T_{BA}(e_j).
\]
\[\square\]

**Corollary 6** Matrix multiplication is associative. That is if $C, B$ and $A$ are matrices with the correct dimensions, then $(CB)A = C(BA)$.

**Theorem 7** If $A$ and $B$ are $n \times n$ matrices such that $BA = I_n$ (the identity matrix), then $B$ and $A$ are invertible, and $B = A^{-1}$.

**Proof:** Suppose that $BA = I_n$. Let us prove that the rank of $A$ is $n$. To do this it suffices to check that $T_A$ is injective. But $T_B \circ T_A = T_{BA} = T_{I_n}$ is the identity transformation. Thus if $T_A(X) = T_A(X')$, then $X = T_B(T_A(X)) = T_B(T_A(X'))$, so $X = X'$, i.e., $T_A$ is injective, $A$ has rank $n$, and hence $A$ is invertible. Let $A^{-1}$ be its inverse. We have $I_n = BA$, so
\[
A^{-1} = I_n A^{-1} = (BA)A^{-1} = B(AA^{-1}) = BI_n = B.
\]
\[\square\]