Section 1.3 Matrix Products

A linear combination is a sum of scalars times quantities. Such expressions arise quite frequently and have the form

\[(\text{scalar } #1)(\text{quantity } #1) + (\text{scalar } #2)(\text{quantity } #2) + \ldots + (\text{scalar } #n)(\text{quantity } #n)\]

This pattern of a sum of products arises so frequently we give it a special name.

**Definition** The **dot product** or **inner product** of two \(n\)-vectors of scalars \(a = [a_1, a_2, \ldots, a_n]\) and \(b = [b_1, b_2, \ldots, b_n]\) is

\[a \cdot b = a_1b_1 + a_2b_2 + \ldots + a_nb_n = \sum_{j=1}^{n} a_jb_j.\]
Example  An instructor gives four exams worth 20%, 25%, 25% and 30% respectively. A student gets the following scores out of 100; 78, 65, 85, 82. **Determine the course average using dot products.**

Example  Find the dot product of each of the following pairs of vectors.

a) \( \mathbf{v} = \begin{bmatrix} 7 & -2 & 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 3 & 6 & 0 \end{bmatrix} \)

b) \( \mathbf{v} = \begin{bmatrix} 3 & -4 & 1 & 2 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 5 \\ 0 \\ -1 \\ -2 \end{bmatrix} \)
Every linear combination of vectors can be written as a product in a special way.

Many times we will be using a linear combination of columns like

\[
\begin{bmatrix}
2 \\
3 \\
-2
\end{bmatrix} + \begin{bmatrix}
5 \\
7 \\
0
\end{bmatrix}
\]

Let’s show how to write this as a special kind of product.

**Definition**  Let \( A \) be an \( m \times n \) matrix and \( c \) a column vector with \( n \) entries, then the **matrix-vector product** \( Ac \) is the linear combination

\[
c_1 \text{col}_1(A) + c_2 \text{col}_2(A) + \cdots + c_n \text{col}_n(A)
\]

and the entries of \( Ac \) can be computed directly as **dot products** using

\[
Ac = \begin{bmatrix}
\text{row}_1(A) \cdot c \\
\text{row}_2(A) \cdot c \\
\vdots \\
\text{row}_n(A) \cdot c
\end{bmatrix}
\]
Example Let $A = \begin{bmatrix} 4 & 5 & -2 \\ 3 & -6 & 1 \end{bmatrix}$ and $c = \begin{bmatrix} 8 \\ 12 \\ 7 \end{bmatrix}$.

Write $Ac$ as a linear combination of columns of $A$.

Write $Ac$ in terms of dot products.
Every system of equations can be written as a matrix equation involving a matrix-vector product.

\[ \begin{align*}
5x_1 - 3x_2 + x_3 &= 1 \\
-4x_1 + 2x_2 + 6x_3 &= -8
\end{align*} \]
We can form the product of a pair of matrices under certain circumstances.

**Definition** Let \( A = [a_{ij}] \) be an \( m \times n \) matrix and \( B = [b_{ij}] \) an \( n \times p \) matrix. Then the **product** of matrices \( A \) and \( B \), denoted \( AB \), is the \( m \times p \) matrix \( C = [c_{ij}] \) where

\[
c_{ij} = \text{row}_i(A) \cdot \text{col}_j(B) = [\begin{array}{cccc} a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{i2} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{in} & a_{in} & \cdots & a_{in} \\ \end{array}] \cdot \left[ \begin{array}{ccc} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{array} \right] = \sum_{k=1}^{n} a_{ik} b_{kj}
\]

for \( 1 \leq i \leq m \) and \( 1 \leq j \leq p \). (**Product** \( AB \) **is defined only when** the number **of columns of** \( A \) **equals** the number **of rows of** \( B \).)

Figure 1 illustrates a matrix product \( AB \), where \( A \) is \( m \times n \) and \( B \) is \( n \times p \).
Examples: Form the following products, if possible.

\[
\begin{bmatrix}
3 & 4 & 2 \\
-1 & 1 & 0
\end{bmatrix} \quad \begin{bmatrix}
1 & 0 & -1 \\
2 & 1 & 1 \\
1 & -1 & 3
\end{bmatrix} \quad \begin{bmatrix}
3 & 2 & -1 & 4
\end{bmatrix} \quad \begin{bmatrix}
1 & 3 \\
0 & 5 \\
-1 & -2 \\
0 & 1
\end{bmatrix} \quad \begin{bmatrix}
2 & 1
\end{bmatrix} \quad \begin{bmatrix}
1 & -2 \\
3 & 1 \\
-1 & 2
\end{bmatrix}
\]

AB

BF

AD

EA
Products of Matrices vs. Products of Real Numbers

Example 6 in the text illustrates several major distinctions between multiplication of matrices and multiplication of scalars.

1. Multiplication of matrices is not defined for all pairs of matrices. The number of columns of the first matrix must equal the number of rows of the second matrix.

2. Matrix multiplication is not commutative; that is, $AB$ need not equal $BA$.

3. The product of two matrices can be zero when neither of the matrices consists of all zeros.
Special Terminology

The definition of a matrix product includes row-by-column products as a $1 \times n$ matrix times an $n \times 1$ matrix and it also includes matrix-vector products as an $m \times n$ matrix times an $n \times 1$ matrix. Thus we will only refer to matrix products from here on, except when we want to emphasize the nature of a product as one of these special cases.

In addition note that

$$C = AB = A[\text{col}_1(B) \quad \text{col}_2(B) \quad \cdots \quad \text{col}_p(B)] = [A\text{col}_1(B) \quad A\text{col}_2(B) \quad \cdots \quad A\text{col}_p(B)]$$

which says that $\text{col}_j(AB)$ is a linear combination of the columns of $A$ with scalars from $\text{col}_j(B)$ since

$$A\text{col}_j(B) = b_{1j}\text{col}_1(A) + b_{2j}\text{col}_2(A) + \cdots + b_{nj}\text{col}_n(A)$$

Hence linear combinations are the fundamental building blocks of matrix products.
Matrix Powers

If $A$ is a square matrix, then products of $A$ with itself like $AA$, $AAA$, ... are defined and we denote them as powers of the matrix $A$ in the form $AA = A^2$, $AAA = A^3$, ...

If $p$ is a positive integer, then we define

$$A^p = AA \ldots A \overset{p \text{ factors}}{\cdots}$$

For an $n \times n$ matrix $A$ we define $A^0 = I_n$, the $n \times n$ identity matrix. For nonnegative integers $p$ and $q$, matrix powers obey familiar rules for exponents, namely

$$A^p A^q = A^{p+q}$$

and

$$(A^p)^q = A^{pq}.$$ 

However, since matrix multiplication is not commutative, for general square matrices

$$(AB)^p \neq A^p B^p.$$
Example A government employee at an obscure warehouse keeps busy by shifting portions of stock between two rooms. Each week two-thirds of the stock from room R1 is shifted to room R2 and two-fifths of the stock from R2 to R1. The following transition matrix $T$ provides a model for this activity:

$$T = \begin{bmatrix} 1/3 & 2/5 \\ 2/3 & 3/5 \end{bmatrix}$$

Entry $t_{ij}$ is the fraction of stock that is shipped to $R_i$ from $R_j$ each week. If R1 contains 3000 items and R2 contains 9000 items, at the end of the week the stock in a room is the amount that remains in the room plus the amount that is shipped in from the other room. We have

$$\frac{1}{3}(\text{stock in R1}) + \frac{2}{5}(\text{stock in R2}) = \frac{1}{3}(3000) + \frac{2}{5}(9000) = 4600$$

$$\frac{2}{3}(\text{stock in R1}) + \frac{3}{5}(\text{stock in R2}) = \frac{2}{3}(3000) + \frac{3}{5}(9000) = 7400$$

So the new stock in each room is a linear combination of the stocks at the start of the week and this can be computed as the product of the transition matrix and $S_0 = \begin{bmatrix} 3000 \\ 9000 \end{bmatrix}$, the original stock vector. Thus the stock vector at the end of the first week is $TS_0 = S_1 = \begin{bmatrix} 4600 \\ 7400 \end{bmatrix}$. The stock vector at the end of the second week is computed as $TS_1 = T^2S_0 = \begin{bmatrix} 4493 \\ 7507 \end{bmatrix} = S_2$ and it follows that $S_3 = TS_2 = T^3S_0 = \begin{bmatrix} 4500 \\ 7500 \end{bmatrix}$.

(We have rounded the entries of the stock vectors to whole numbers.) The table is a portion of the records that the employee keeps.

<table>
<thead>
<tr>
<th>Original stock</th>
<th>Stock at end of week #1</th>
<th>Stock at end of week #2</th>
<th>Stock at end of week #3</th>
<th>Stock at end of week #4</th>
<th>Stock at end of week #5</th>
</tr>
</thead>
<tbody>
<tr>
<td>3000</td>
<td>4600</td>
<td>4493</td>
<td>4500</td>
<td>4500</td>
<td>4500</td>
</tr>
<tr>
<td>9000</td>
<td>7400</td>
<td>7507</td>
<td>7500</td>
<td>7500</td>
<td>7500</td>
</tr>
</tbody>
</table>
We see that in just a few weeks the employee can claim to do the shifting of the stock without actually performing the task since the number of items in each room does not change after the third week. We say the process has reached a **steady state**. This type of iterative process is called a **Markov chain** or **Markov process**. The basic question that we pose for such a process is to determine the limiting behavior of the sequence of vectors $S_0, S_1, S_2, S_3, \ldots$. We show later that the ultimate behavior depends upon the information contained in the transition matrix and we show how to determine the steady state, if it exists.
Properties of Matrix Products

Property of Scalar Multiplication and Matrix Products.
Let $r$ be a scalar and $A$ and $B$ be matrices. Then

$$r(AB) = (rA)B = A(rB). \iff$$

Properties of Matrix Products and Linear Combinations.
Let $A$, $B$, and $C$ be matrices, $x$ and $y$ be vectors, and $r$ and $s$ be scalars. Then

$$A(B + C) = AB + AC \quad \text{and} \quad A(rx + sy) = rAx + sAy.$$ 

Each of these properties is called a **distributive rule**.

Property of Matrix Products and the Transpose.
If $A$ and $B$ are matrices, then

$$(AB)^T = B^TA^T.$$
The Angle Between 2-vectors.

The dot product of a pair of n-vectors provided us with an algebraic tool for computing matrix-vector products and matrix products. However, the dot product is more than just an algebraic construction, it arises naturally when we determine the angle between ‘geometric’ n-vectors; see Section 1.1. Here we show how $\mathbf{a} \cdot \mathbf{b}$ is related to the angle between the 2-vectors $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$.

Using the geometric representation for 2-vectors given in Section 1.1 we can view vectors $\mathbf{a}$ and $\mathbf{b}$ as shown in Figure 2 where $\alpha$ is the angle between vector $\mathbf{a}$ and the horizontal axis and $\beta$ is the angle between vector $\mathbf{b}$ and the horizontal axis. Hence the angle between vectors $\mathbf{a}$ and $\mathbf{b}$ shown in Figure 2 is $\alpha - \beta$.

![Figure 2.](image)

From basic trigonometric relationships we have

$$\cos(\alpha) = \frac{a_1}{\sqrt{a_1^2 + a_2^2}} \quad \cos(\beta) = \frac{b_1}{\sqrt{b_1^2 + b_2^2}}$$

$$\sin(\alpha) = \frac{a_2}{\sqrt{a_1^2 + a_2^2}} \quad \sin(\beta) = \frac{b_2}{\sqrt{b_1^2 + b_2^2}}$$

From the trigonometric identity $\cos(\alpha - \beta) = \cos(\alpha) \cos(\beta) + \sin(\alpha) \sin(\beta)$ it follows that
Thus the cosine of the angle between 2-vectors $a$ and $b$ has been expressed in terms of dot products. Later we generalize this result to a pair of n-vectors.
Application: Path Lengths

In Section 1.1 we introduced an incidence matrix as a mathematical representation of a graph. The graph in Figure 3 is represented by the incidence matrix

\[
A = \begin{bmatrix}
0 & 1 & 1 & 0 \\
1 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 \\
0 & 0 & 1 & 0
\end{bmatrix}.
\]

We say there is a path between nodes \( P_i \) and \( P_j \) if we can move along edges of the graph to go from \( P_i \) to \( P_j \). For example, there is a path between \( P_1 \) and \( P_4 \) in Figure 3; traverse the edge from \( P_1 \) to \( P_3 \) and then from \( P_3 \) to \( P_4 \). However, there is no edge from \( P_1 \) to \( P_4 \). For the graph in Figure 3, there is a path connecting every pair of nodes. This is not always the case; see Figure 4. There is no path between \( P_1 \) and \( P_4 \). The graph in Figure 4 has two separate, or disjoint, pieces.

In some applications that can be modeled by using a graph there can be more than one edge between a pair of nodes. For the graph in Figure 5 there are two edges between \( P_2 \) and \( P_3 \). In this case the incidence matrix is

\[
B = \begin{bmatrix}
0 & 1 & 1 & 0 \\
1 & 0 & 2 & 0 \\
1 & 2 & 0 & 1 \\
0 & 0 & 1 & 0
\end{bmatrix}.
\]

We use the term path of length 1 to mean that there are no intervening nodes of the graph between the ends of the path. (The term length here does not mean a measure of the actual distance between nodes, rather it is a count of the number of edges traversed between the starting node and the ending node.
of a path.) The entries of the incidence matrix of a graph count the number of paths of length 1 between a pair of nodes.

In certain applications it is important to determine the number of paths of a certain length \( k \) between each pair of nodes of a graph. Also, one frequently must calculate the longest path length that will ever be needed to travel between any pair of nodes. Both of these problems can be solved by manipulating the incidence matrix associated with a graph. Our manipulation tools will be matrix addition and multiplication.

Since the incidence matrix \( A \) of a graph is square, its powers \( A^2, A^3, \) etc. are defined. We have

\[
\text{ent}_{ij}(A^2) = \text{row}_i(A) \cdot \text{col}_j(A) = a_{i1}a_{1j} + a_{i2}a_{2j} + \cdots + a_{in}a_{nj}.
\]

We know that

\[
a_{i1} = \text{number of paths of length 1 from } P_i \text{ to } P_1
\]
\[
a_{1j} = \text{number of paths of length 1 from } P_1 \text{ to } P_j
\]

so the only way there could be a path from \( P_i \) to \( P_j \) through node \( P_1 \) is if both \( a_{i1} \) and \( a_{1j} \) are not zero. Thus product \( a_{i1}a_{1j} \) is the number of paths of length two between \( P_i \) and \( P_j \) through \( P_1 \). Similarly, \( a_{i2}a_{2j} \) is the number of paths of length 2 between \( P_i \) and \( P_j \) through \( P_s \). Hence the dot product \( \text{row}_i(A) \cdot \text{col}_j(A) \) gives the total number of paths of length 2 between \( P_i \) and \( P_j \). It follows that matrix \( A^2 \) provides a mathematical model whose entries count the number of paths of length 2 between nodes of the graph. By corresponding arguments, the number of paths of length \( k \) from \( P_i \) to \( P_j \) will be given by the \((i,j)\)-entry of matrix \( A^k \).

In order to calculate the longest path length that will ever be required to travel between any pair of nodes in a graph, we proceed as follows. The \((i,j)\)-entry of matrix \( A + A^2 \) is the number of paths of length less than or equal to 2 from \( P_i \) to \( P_j \). Similarly, the entries of \( A + A^2 + A^3 \) are the number of paths of length less than or equal to 3 between nodes of the graph. We continue this process until we find a value of \( k \) such that all the nondiagonal entries of

\[
A + A^2 + A^3 + \ldots + A^k
\]

are nonzero. This implies that there is a path of length \( k \) or less between distinct nodes of the graph. (Why can the diagonal entries be ignored?) The value of \( k \)
determined in this way is the length of the longest path we would ever need to traverse to go from one node of the graph to any other node of the graph.
True/False Review Questions

1. If \( x \) and \( y \) are both 3x1 matrices then \( x \cdot y \) is the same as \( x^T y \).

2. If \( A \) is 3x2 and \( x \) is 2x1 then \( Ax \) is a linear combination of the columns of \( A \).

3. If \( A \) is 4x2 and \( B \) is 2x4, then \( AB = BA \).

4. If \( A \) is 10x7 and \( B \) is 7x8, then \( \text{ent}_{5,2}(AB) = \text{row}_5(A)\text{col}_2(B) \).

5. Matrix multiplication is commutative.

6. If \( A \) is 3x3 and \( b \) is 3x1 so that \( Ab = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \) then either \( A \) is a zero matrix or \( b \) is a zero column.

7. If \( A \) and \( B \) are the same size, then \( AB = BA \).

8. Powers of a matrix are really successive matrix products.

9. The dot product of a pair of real n-vectors \( x \) and \( y \) is commutative; that is, \( x \cdot y = y \cdot x \).

10. Let \( x = [1 \ 1 \ 1] \), then \( (x^T x)^2 = 3x^T x \).
**Terminology**

<table>
<thead>
<tr>
<th>Dot product; inner product.</th>
<th>Row-by-column-product.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix-vector product.</td>
<td>Matrix equation.</td>
</tr>
<tr>
<td>Matrix product.</td>
<td>Matrix powers.</td>
</tr>
<tr>
<td>Markov chain; steady state.</td>
<td>Angle between 2-vectors.</td>
</tr>
</tbody>
</table>

There is a connection between some of the terms listed above. Understanding and being able to explicitly state such connections is important in later sections. For practice with the concepts of this section formulate responses to the following questions and statements in your own words.

- How are dot products and row-by-column products related?
- Describe how matrix-vector products are computed using dot products.
- Explain how we to compute a matrix-vector product as a linear combination of columns.
- Explain how we formulate every linear system of equations as a matrix equation. (Explicitly describe all matrices used in the equation.)
- Explain how to express the matrix product $\mathbf{AB}$ in terms of matrix-vector products.
- Explain how to express the matrix product $\mathbf{AB}$ in terms of dot products.
- For what type of matrix $\mathbf{A}$ can we compute powers, $\mathbf{A}^2$, $\mathbf{A}^3$, …, etc. Explain why there is such a restriction.
- What do we mean when we say that matrix multiplication distributes over matrix addition?
- Complete the following:
  The transpose of a product of two matrices is ________________.
- Explain what we mean by the steady state of a Markov chain. (Express this relationship in words and using a matrix equation.)