MTHSC 3110 Section 2.1 – Matrix Operations

Kevin James

Notation

Let A be an $m \times n$ matrix, that is, m rows and n columns. We'll refer to the entries of A by their row and column indices. The entry in the i^{th} row and j^{th} column is denoted by a_{ij}, and is called the (i,j)-entry of A.

$$\begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

Notation

The columns of A are vectors in $\mathbb{R}^m$, and are denoted in the book by $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ and in my notes by $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$. In order to focus attention on the columns we write

$$A = [\vec{a}_1 \quad \vec{a}_2 \quad \dots \quad \vec{a}_n]$$

Definition

Suppose that $A = [\vec{a}_1 \quad \vec{a}_2 \quad \dots \quad \vec{a}_n]$ and $B = [\vec{b}_1 \quad \vec{b}_2 \quad \dots \quad \vec{b}_n]$ are both $m \times n$ matrices. Then we define their sum as

$$\begin{aligned} A + B &= [\vec{a}_1 \quad \vec{a}_2 \quad \dots \quad \vec{a}_n] + [\vec{b}_1 \quad \vec{b}_2 \quad \dots \quad \vec{b}_n] \\ &= [(\vec{a}_1 + \vec{b}_1) \quad (\vec{a}_2 + \vec{b}_2) \quad \dots \quad (\vec{a}_n + \vec{b}_n)] \end{aligned}$$

Definition

Suppose that $A = [\vec{a}_1 \quad \vec{a}_2 \quad \dots \quad \vec{a}_n]$ and $B = [\vec{b}_1 \quad \vec{b}_2 \quad \dots \quad \vec{b}_n]$ are both $m \times n$ matrices. Then we define their sum as

$$\begin{aligned} A + B &= [\vec{a}_1 \quad \vec{a}_2 \quad \dots \quad \vec{a}_n] + [\vec{b}_1 \quad \vec{b}_2 \quad \dots \quad \vec{b}_n] \\ &= [(\vec{a}_1 + \vec{b}_1) \quad (\vec{a}_2 + \vec{b}_2) \quad \dots \quad (\vec{a}_n + \vec{b}_n)] \end{aligned}$$

Alternative Definition

Given two $m \times n$ matrices A and B, we can define their sum $C = A + B$ as the $m \times n$ matrix whose entries are $c_{ij} = a_{ij} + b_{ij}$.

Example

$$\begin{pmatrix} 1 & 2 \\ 3 & -1 \\ 2 & 4 \end{pmatrix} + \begin{pmatrix} -1 & 3 \\ 2 & -1 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} & & & \\ & & & \\ & & & \end{pmatrix}$$

Definition

We denote by 0 the matrix all of whose elements are zero.

Definition

We denote by 0 the matrix all of whose elements are zero.

Theorem

Let A, B, C be matrices of the same size, and let r, s be scalars. Then

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $r(A + B) = rA + rB$
5. $(r + s)A = rA + sA$
6. $r(sA) = (rs)A$

Definition

We denote by 0 the matrix all of whose elements are zero.

Theorem

Let A, B, C be matrices of the same size, and let r, s be scalars. Then

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $r(A + B) = rA + rB$
5. $(r + s)A = rA + sA$
6. $r(sA) = (rs)A$

Note

This is very similar to the corresponding theorem for vector addition.

Recall

If f and g are functions and the image of g is contained in the domain of f, then we define the composition of f and g by

$$f \circ g(x) = f(g(x))$$

Recall

If f and g are functions and the image of g is contained in the domain of f, then we define the composition of f and g by

$$f \circ g(x) = f(g(x))$$

Example

Suppose that $f, g : \mathbb{R} \to \mathbb{R}$ are defined by $f(y) = y^2$ and $g(x) = \sin(x)$,

Recall

If f and g are functions and the image of g is contained in the domain of f, then we define the composition of f and g by

$$f \circ g(x) = f(g(x))$$

Example

Suppose that $f, g : \mathbb{R} \to \mathbb{R}$ are defined by $f(y) = y^2$ and $g(x) = \sin(x)$, then
$f \circ g(x) = f(g(x)) = (\sin(x))^2$

Note

1. We will specialize our attention to linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.
2. We have seen that a linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ corresponds to multiplying a vector $\vec{x} \in \mathbb{R}^n$ by an $m \times n$ matrix A.

Note

1. We will specialize our attention to linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.
2. We have seen that a linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ corresponds to multiplying a vector $\vec{x} \in \mathbb{R}^n$ by an $m \times n$ matrix A.
3. We will consider linear transformations

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m,$$

where $U : \mathbb{R}^p \longrightarrow \mathbb{R}^n$, and $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. So, since $U(\vec{x})$ is in the domain of T, we can compute $T(U(\vec{x}))$.

Fact

The composition of two linear transformations is again a linear transformation and thus can be written as multiplication by a matrix.

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.
Let A be the $m \times n$ matrix corresponding to the transformation T.

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.
Let A be the $m \times n$ matrix corresponding to the transformation T.
(-i.e. $U(\vec{x}) = B\vec{x}$, and $T(\vec{v}) = A\vec{v}$).

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.
Let A be the $m \times n$ matrix corresponding to the transformation T.
(-i.e. $U(\vec{x}) = B\vec{x}$, and $T(\vec{v}) = A\vec{v}$).
Then we have

$$T \circ U(\vec{x}) \quad = \quad T(U(\vec{x})) =$$

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.
Let A be the $m \times n$ matrix corresponding to the transformation T.
(-i.e. $U(\vec{x}) = B\vec{x}$, and $T(\vec{v}) = A\vec{v}$).
Then we have

$$T \circ U(\vec{x}) \quad = \quad T(U(\vec{x})) = A(B\vec{x}) =$$

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.
Let A be the $m \times n$ matrix corresponding to the transformation T.
(-i.e. $U(\vec{x}) = B\vec{x}$, and $T(\vec{v}) = A\vec{v}$).
Then we have

$$T \circ U(\vec{x}) \quad = \quad T(U(\vec{x})) = A(B\vec{x}) = A(x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_p\vec{b}_p)$$

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.
Let A be the $m \times n$ matrix corresponding to the transformation T.
(-i.e. $U(\vec{x}) = B\vec{x}$, and $T(\vec{v}) = A\vec{v}$).
Then we have

$$\begin{aligned} T \circ U(\vec{x}) &= T(U(\vec{x})) = A(B\vec{x}) = A(x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_p\vec{b}_p) \\ &= Ax_1\vec{b}_1 + Ax_2\vec{b}_2 + \cdots + Ax_p\vec{b}_p \end{aligned}$$

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.
Let A be the $m \times n$ matrix corresponding to the transformation T.
(-i.e. $U(\vec{x}) = B\vec{x}$, and $T(\vec{v}) = A\vec{v}$).
Then we have

$$\begin{aligned} T \circ U(\vec{x}) &= T(U(\vec{x})) = A(B\vec{x}) = A(x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_p\vec{b}_p) \\ &= Ax_1\vec{b}_1 + Ax_2\vec{b}_2 + \cdots + Ax_p\vec{b}_p \\ &= x_1A\vec{b}_1 + x_2A\vec{b}_2 + \cdots + x_pA\vec{b}_p \end{aligned}$$

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.
Let A be the $m \times n$ matrix corresponding to the transformation T.
(-i.e. $U(\vec{x}) = B\vec{x}$, and $T(\vec{v}) = A\vec{v}$).
Then we have

$$\begin{aligned} T \circ U(\vec{x}) &= T(U(\vec{x})) = A(B\vec{x}) = A(x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_p\vec{b}_p) \\ &= Ax_1\vec{b}_1 + Ax_2\vec{b}_2 + \cdots + Ax_p\vec{b}_p \\ &= x_1A\vec{b}_1 + x_2A\vec{b}_2 + \cdots + x_pA\vec{b}_p \\ &= [A\vec{b}_1 \quad A\vec{b}_2 \quad \ldots \quad A\vec{b}_p]\ \vec{x} \end{aligned}$$

Suppose that

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m.$$

Let B be the $n \times p$ matrix corresponding to the transformation U.
Let A be the $m \times n$ matrix corresponding to the transformation T.
(-i.e. $U(\vec{x}) = B\vec{x}$, and $T(\vec{v}) = A\vec{v}$).
Then we have

$$\begin{aligned} T \circ U(\vec{x}) &= T(U(\vec{x})) = A(B\vec{x}) = A(x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_p\vec{b}_p) \\ &= Ax_1\vec{b}_1 + Ax_2\vec{b}_2 + \cdots + Ax_p\vec{b}_p \\ &= x_1A\vec{b}_1 + x_2A\vec{b}_2 + \cdots + x_pA\vec{b}_p \\ &= [A\vec{b}_1 \quad A\vec{b}_2 \quad \ldots \quad A\vec{b}_p]\ \vec{x} \end{aligned}$$

Thus, $T \circ U(\vec{x}) = [A\vec{b}_1 \quad A\vec{b}_2 \quad \ldots \quad A\vec{b}_p]\ \vec{x}$.

Definition

If A is an $m \times n$ matrix, and if $B = [\vec{b}_1, \vec{b}_2 \ldots, \vec{b}_p]$ is a $n \times p$ matrix, then the matrix product AB is the following $m \times p$ matrix.

$$AB = [A\vec{b}_1 \quad A\vec{b}_2 \quad \ldots \quad A\vec{b}_p]$$

Definition

If A is an $m \times n$ matrix, and if $B = [\vec{b}_1, \vec{b}_2 \ldots, \vec{b}_p]$ is a $n \times p$ matrix, then the matrix product AB is the following $m \times p$ matrix.

$$AB = [A\vec{b}_1 \quad A\vec{b}_2 \quad \ldots \quad A\vec{b}_p]$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix}$ and let $B = \begin{pmatrix} 3 & -1 & 6 \\ 7 & 5 & 3 \end{pmatrix}$. Compute AB.

Row-Column Rule

If A is $m \times n$ and if B is $n \times p$ the (i,j)-entry of AB is given by

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Row-Column Rule

If A is $m \times n$ and if B is $n \times p$ the (i, j)-entry of AB is given by

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Note

$\text{Row}_i(AB) = \text{Row}_i(A) \cdot B$.

Definition

We define the $m \times m$ identity matrix as

$$I_m = [\vec{e}_1, \vec{e}_2, \ldots \vec{e}_m] = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Definition

We define the $m \times m$ identity matrix as

$$I_m = [\vec{e}_1, \vec{e}_2, \dots \vec{e}_m] = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Theorem

With A, B and C appropriately sized matrices and r a scalar

1. $(AB)C = A(BC)$
2. $A(B + C) = AB + AC$
3. $(B + C)A = BA + CA$
4. $r(AB) = (rA)B = A(rB)$
5. $I_m A = A = AI_n$

Properties Not held by Matrix Multiplication

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 5 & 10 \end{pmatrix}$.

Let $B = \begin{pmatrix} 2 & 6 \\ -1 & -3 \end{pmatrix}$.

Compute AB, BA, and $A \cdot 0$.

Properties Not held by Matrix Multiplication

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 5 & 10 \end{pmatrix}$.

Let $B = \begin{pmatrix} 2 & 6 \\ -1 & -3 \end{pmatrix}$.

Compute AB, BA, and $A \cdot 0$.

Note

1. It is **NOT** in general true that $AB = BA$.
2. It is **NOT** in general true when $AC = AB$ that $C = B$.

Definition

We define non negative powers of an $m \times m$ matrix as follows.

$$A^0 = I_m, \qquad A^1 = A, \qquad A^n = A^{n-1}A \qquad \text{for } n > 1.$$

Definition

We define non negative powers of an $m \times m$ matrix as follows.

$$A^0 = I_m, \qquad A^1 = A, \qquad A^n = A^{n-1}A \qquad \text{for } n > 1.$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Compute A^3.

Definition

The transpose of a $m \times n$ matrix A is the matrix A^T having (i, j)-entry a_{ji}. That is,

$$(A^T)_{ij} = a_{ji}.$$

Definition

The transpose of a $m \times n$ matrix A is the matrix A^T having (i,j)-entry a_{ji}. That is,

$$(A^T)_{ij} = a_{ji}.$$

Example

For example, $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ has transpose

Definition

The transpose of a $m \times n$ matrix A is the matrix A^T having (i, j)-entry a_{ji}. That is,

$$(A^T)_{ij} = a_{ji}.$$

Example

For example, $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ has transpose $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

The Transpose of a Matrix

Definition

The transpose of a $m \times n$ matrix A is the matrix A^T having (i,j)-entry a_{ji}. That is,

$$(A^T)_{ij} = a_{ji}.$$

Example

For example, $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ has transpose $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

Note

The rows of A become the columns of A^T and vice versa.

Theorem

Let A and B be matrices whose sizes are appropriate for the following sums and products to be defined

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$.
3. *For any scalar* r, $(rA)^T = rA^T$.
4. $(AB)^T = B^T A^T$

Example

$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, and $B = \begin{pmatrix} 5 & 1 & -1 \\ 1 & 2 & 2 \end{pmatrix}$ then

$$AB = \begin{pmatrix} 7 & 5 & 3 \\ 9 & 11 & 5 \end{pmatrix} \qquad (AB)^T = \begin{pmatrix} 7 & 9 \\ 5 & 11 \\ 3 & 5 \end{pmatrix} = B^T A^T$$

but A^T is 2×2 and B^T is 3×2, so $A^T B^T$ isn't even defined.