Chapter 3 Matrix Algebra

3.1 Matrix Operations

Definition 3.1 Two matrices $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$ and $\mathbf{B} = [b_{ij}] \in \mathbb{R}^{p \times q}$ are ***equal***, written $\mathbf{A} = \mathbf{B}$, if $m = p$ and $n = q$, and $a_{ij} = b_{ij}$, for all i and j.

Two matrices are equal if they are of the same order and each pair of corresponding elements are equal.

Definition 3.2 Given a matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$ and a scalar c, the ***scalar multiplication of matrix*** **A** by the scalar c is given by $c\mathbf{A} = [ca_{ij}] \in \mathbb{R}^{m \times n}$.

The result of scalar multiplication of a matrix is just that matrix with each of its elements multiplied by the scalar.

Example a) $3\begin{bmatrix} 4 & 2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 12 & 6 \\ -3 & 0 \end{bmatrix}$.

b) $0\begin{bmatrix} 1 & 1.6 & 4 \\ 0 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Definition 3.3 The ***addition of two matrices*** $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$ and $\mathbf{B} = [b_{ij}] \in \mathbb{R}^{p \times q}$ is defined only if $m = p$ and $n = q$, and is given by

$$\mathbf{A} + \mathbf{B} = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] \in \mathbb{R}^{m \times n}.$$

The addition of two matrices is defined only when they are of the same order and the result of the addition is a matrix whose each element is the sum of corresponding elements of **A** and **B**.

Example a) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 2 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 5 & -1 \end{bmatrix}$.

b) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ is not defined.

Note that the matrix subtraction $\mathbf{A} - \mathbf{B}$ can be readily defined as $\mathbf{A} + (-1)\mathbf{B}$.

Definition 3.4 The ***multiplication of matrix*** $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$ ***by matrix*** $\mathbf{B} = [b_{ij}] \in \mathbb{R}^{p \times q}$ is defined only if $n = p$, and is given by,

$$\mathbf{AB} = \left[\sum_{k=1}^{n} a_{ik} b_{kj}\right] \in \mathbb{R}^{m\times q}$$

The product of the matrix multiplication **AB** is a matrix whose number of rows equal to that of **A** and number of columns equal to that of **B**. The $(i,j)^{\text{th}}$ element of **AB** is the sum of the products of elements in i^{th} row of **A** multiplied by corresponding elements in the j^{th} column of **B**.

Example The matrix multiplication $\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}1 & 4 & 3\\ 0 & -1 & 2\end{bmatrix} = \begin{bmatrix}1 & 2 & 7\\ 3 & 8 & 17\end{bmatrix}$, but $\begin{bmatrix}1 & 4 & 3\\ 0 & -1 & 2\end{bmatrix}\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$ is not defined.

The example shows that the order of the matrices in the multiplication is important. In general, $\mathbf{AB} \neq \mathbf{BA}$.

Problem If $\mathbf{AB} = \mathbf{0}$, does it mean that at least one of **A** and **B** has to be a zero matrix?

2. Matrix Representation of System of Linear Equations

We can use the matrix multiplication to concisely represent a system of linear equation as,

$$\mathbf{Ax} = \mathbf{b} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

where **A** is the coefficient matrix and **b** is the RHS vector and **x** is the vector whose elements are the variables.

Definition 3.5 When a row vector is multiplied by a column vector, the product is called ***dot product*** or ***inner product***. That is, if $\mathbf{a} \in \mathbb{R}^{1\times n}$ and $\mathbf{b} \in \mathbb{R}^{n+1}$, then

$$\mathbf{ab} = [a_1 \quad a_2 \quad \cdots \quad a_n]\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{j=1}^{n} a_j b_j \in \mathbb{R},$$

which is just a scalar.

With the dot product defined, we can visualize the result of the matrix multiplication as a matrix of dot products. We can write the matrix **A** as a column vector whose elements are themselves row vectors, and write the matrix $\mathbf{B}$ as a row vector whose elements are columns vectors. That is,

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_{1\cdot} \\ \mathbf{a}_{2\cdot} \\ \vdots \\ \mathbf{a}_{m\cdot} \end{bmatrix} \in \mathbb{R}^{m\times n}, \mathbf{a}_{i\cdot} = [a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}] \in \mathbb{R}^{1\times n}, i = 1,2,\ldots,m$$

$$\mathbf{B} = [\mathbf{b}_{\cdot 1} \quad \mathbf{b}_{\cdot 2} \quad \cdots \quad \mathbf{b}_{\cdot q}] \in \mathbb{R}^{n\times q}, \mathbf{b}_{\cdot j} = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} \in \mathbb{R}^{n\times 1}, j = 1,2,\ldots,q$$

Then, we can write

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_{1\cdot} \\ \mathbf{a}_{2\cdot} \\ \vdots \\ \mathbf{a}_{m\cdot} \end{bmatrix} [\mathbf{b}_{\cdot 1} \quad \mathbf{b}_{\cdot 2} \quad \cdots \quad \mathbf{b}_{\cdot q}] = \begin{bmatrix} \mathbf{a}_{1\cdot}\mathbf{b}_{\cdot 1} & \mathbf{a}_{1\cdot}\mathbf{b}_{\cdot 2} & \cdots & \mathbf{a}_{1\cdot}\mathbf{b}_{\cdot q} \\ \mathbf{a}_{2\cdot}\mathbf{b}_{\cdot 1} & \mathbf{a}_{2\cdot}\mathbf{b}_{\cdot 2} & \cdots & \mathbf{a}_{2\cdot}\mathbf{b}_{\cdot q} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_{m\cdot}\mathbf{b}_{\cdot 1} & \mathbf{a}_{m\cdot}\mathbf{b}_{\cdot 2} & \cdots & \mathbf{a}_{m\cdot}\mathbf{b}_{\cdot q} \end{bmatrix} = [\mathbf{a}_{i\cdot}\mathbf{b}_{\cdot j}]_{m\times q}.$$

With the same idea of writing a matrix as a vector whose elements are themselves vectors, we can write the product

$$\mathbf{AB} = \mathbf{A}[\mathbf{b}_{\cdot 1} \quad \mathbf{b}_{\cdot 2} \quad \cdots \quad \mathbf{b}_{\cdot q}] = [\mathbf{Ab}_{\cdot 1} \quad \mathbf{Ab}_{\cdot 2} \quad \cdots \quad \mathbf{Ab}_{\cdot q}] = \begin{bmatrix} \mathbf{a}_{1\cdot} \\ \mathbf{a}_{2\cdot} \\ \vdots \\ \mathbf{a}_{m\cdot} \end{bmatrix} \mathbf{B} = \begin{bmatrix} \mathbf{a}_{1\cdot}\mathbf{B} \\ \mathbf{a}_{2\cdot}\mathbf{B} \\ \vdots \\ \mathbf{a}_{m\cdot}\mathbf{B} \end{bmatrix}.$$

Similarly, a system of linear equations can be written as

$$\mathbf{b} = \mathbf{Ax} = \begin{bmatrix} \mathbf{a}_{1\cdot} \\ \mathbf{a}_{2\cdot} \\ \vdots \\ \mathbf{a}_{m\cdot} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_{1\cdot}\mathbf{x} \\ \mathbf{a}_{2\cdot}\mathbf{x} \\ \vdots \\ \mathbf{a}_{m\cdot}\mathbf{x} \end{bmatrix}$$

That means each linear equation is an inner product of a row of $\mathbf{A}$ and $\mathbf{x}$. Alternatively, the matrix $\mathbf{A}$ can be written as a row vector whose elements are columns of $\mathbf{A} = [\mathbf{a}_{\cdot 1} \quad \mathbf{a}_{\cdot 2} \quad \cdots \quad \mathbf{a}_{\cdot n}]$, and

$$\mathbf{b} = \mathbf{Ax} = [\mathbf{a}_{\cdot 1} \quad \mathbf{a}_{\cdot 2} \quad \cdots \quad \mathbf{a}_{\cdot n}]\mathbf{x}$$

$$= [\mathbf{a}_{\cdot 1} \quad \mathbf{a}_{\cdot 2} \quad \cdots \quad \mathbf{a}_{\cdot n}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$
$$= x_1\mathbf{a}_{\cdot 1} + x_2\mathbf{a}_{\cdot 2} + \cdots + x_n\mathbf{a}_{\cdot n}.$$

The last term is, as to be defined in Chapter 6, ***a linear combination of columns*** of **A**. The problem of finding a solution to the system of linear equations can be reframed as finding a linear combinations of columns of **A** to be equal to the RHS vector **b**.

Problem Write $\mathbf{A} = [\mathbf{a}_{\cdot 1} \quad \mathbf{a}_{\cdot 2} \quad \cdots \quad \mathbf{a}_{\cdot n}] \in \mathbb{R}^{m\times n}$ and $\mathbf{B} = \begin{bmatrix} \mathbf{b}_{1\cdot} \\ \mathbf{b}_{2\cdot} \\ \vdots \\ \mathbf{b}_{n\cdot} \end{bmatrix} \in \mathbb{R}^{n\times q}$. Find the product **AB** and verify that it is equivalent to all the above representations of $\mathbf{AB} = \begin{bmatrix} \mathbf{a}_{1\cdot} \\ \mathbf{a}_{2\cdot} \\ \vdots \\ \mathbf{a}_{m\cdot} \end{bmatrix} [\mathbf{b}_{\cdot 1} \quad \mathbf{b}_{\cdot 2} \quad \cdots \quad \mathbf{b}_{\cdot q}] \in \mathbb{R}^{m\times q}$.

Problem Johnson, Riess & Arnold [1998] page 59, #60. Let **A** and **B** be matrices such that the product **AB** is defined. Prove each of the following.

a. If **A** has a row of zeros, then so does **AB**
b. If **B** has a column of zeros, then so does **AB**.
c. If **A** has two identical rows, then so does **AB**.
d. If **B** has two identical columns, then so does **AB**.

Definition 3.6 Two vectors are said to be ***orthogonal*** if their dot product is zero. (See **Fraleigh & Beauregard** [1995] page 22-29 for geometrical interpretation of dot product)

3.3 Laws of Matrix Operations

The proofs of the following laws of scalar multiplication, matrix addition and multiplication are left as exercises.

Theorem 3.1 Let and be scalars and $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ be matrices whose orders are such that all operations below are defined. Then,

a) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$; commutative,
b) $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$; associative,
c) $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$; distributive,
d) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$; associative,
e) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$, and
$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$; distributive, and
f) $\alpha\mathbf{AB} = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$.

Proof Exercises.❑

As noted earlier, matrix multiplication does not possess commutative property.

Problem Johnson, Riess & Arnold [1998] page 69, #26, 27, 28, 44, 45.

26. Let **A** and **B** be matrices. Prove or find a counterexample for this statement: $(\mathbf{A} + \mathbf{B})(\mathbf{A} - \mathbf{B}) = \mathbf{A}^2 - \mathbf{B}^2$.
27. Let **A** and **B** be matrices such that $\mathbf{A}^2 = \mathbf{AB}$ and $\mathbf{A} \neq \mathbf{0}$ (see Definition 3.8 for zero matrix) Can we assert that, by cancellation, $\mathbf{A} = \mathbf{B}$? Explain.
28. Let **A** and **B** be as in Exercise 27. Find the flaw in the follwing proof that $\mathbf{A} = \mathbf{B}$. Since $\mathbf{A}^2 = \mathbf{AB}$, $\mathbf{A}^2 - \mathbf{AB} = \mathbf{0}$. Factoring yields $\mathbf{A}(\mathbf{A} - \mathbf{B}) = \mathbf{0}$. Since $\mathbf{A} \neq \mathbf{0}$, if follows that $\mathbf{A} - \mathbf{B} = \mathbf{0}$. Therefore, $\mathbf{A} = \mathbf{B}$.

3.4 Transposition of a Matrix

Another matrix operations that is frequently used is the transposition. It is very useful in matrix manipulation.

Definition 3.7 Let **A** be a matrix in $\mathbb{R}^{m\times n}$. The ***transpose of*** **A** is given by,

$$\mathbf{A}^{\mathrm{T}} = \left[a_{ij}^{\mathrm{T}}\right] \in \mathbb{R}^{n\times m}, \text{ where } a_{ij}^{\mathrm{T}} = a_{ji}.$$

When a matrix is transposed, the rows becomes columns and vice versa.

Example If $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, then $\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

Theorem 3.2 Let **A** and **B** be matrices of the same order. Then

a) $(\mathbf{A}+\mathbf{B})^{\mathrm{T}}=\mathbf{A}^{\mathrm{T}}+\mathbf{B}^{\mathrm{T}}$,
b) $(\mathbf{A}^{\mathrm{T}})^{\mathrm{T}}=\mathbf{A}$, and
c) $(\alpha\mathbf{A})^{\mathrm{T}}=\alpha\mathbf{A}^{\mathrm{T}}$.

Proof The proof is straight forward and is left as exercise.❑

Theorem 3.3 Let $\mathbf{A}\in\mathbb{R}^{m\times n}$ and $\mathbf{B}\in\mathbb{R}^{n\times p}$. Then $(\mathbf{AB})^{\mathrm{T}}=\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$.

Proof We have to show that the dimensions of the matrix on the left of the equation equals that on the right, and that each pair of corresponding elements are equal.

The dimension of $\mathbf{AB}$ is $m\times p$ so the dimension of $(\mathbf{AB})^{\mathrm{T}}$ is $p\times m$. The dimension of $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{B}^{\mathrm{T}}$ are $n\times m$ and $p\times n$ respectively. So the dimension of the product $\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ is also $p\times m$.

Let $\mathbf{C}=\mathbf{AB}$ so that $c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}$. Thus $\mathbf{C}^{\mathrm{T}}=\left[c_{ij}^{\mathrm{T}}\right]=\left[c_{ji}\right]=\left[\sum_{k=1}^{n}a_{jk}b_{ki}\right]$. We need to show that this $(i,j)^{\text{th}}$ element of $\mathbf{C}^{\mathrm{T}}$ is equal to the $(i,j)^{\text{th}}$ element of $\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$, which is $\sum_{k=1}^{n}b_{ik}^{\mathrm{T}}a_{kj}^{\mathrm{T}}$. That is

$$c_{ij}^{\mathrm{T}}=\sum_{k=1}^{n}a_{jk}b_{ki}=\sum_{k=1}^{n}a_{kj}^{\mathrm{T}}b_{ik}^{\mathrm{T}}=\sum_{k=1}^{n}b_{ik}^{\mathrm{T}}a_{kj}^{\mathrm{T}}\ ❑$$

Corollary 3.1 Let $\mathbf{A}_1,\mathbf{A}_2,\dots,\mathbf{A}_k$ be matrices whose dimensions are such that their product $\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k$ is defined. Then,

$$(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k)^{\mathrm{T}}=\mathbf{A}_k^{\mathrm{T}}\mathbf{A}_{k-1}^{\mathrm{T}}\cdots\mathbf{A}_1^{\mathrm{T}}.$$

Proof Prove by induction. By Theorem 3.3, the Corollary is true for $k=2$. We can then state the induction hypothesis that the Corollary holds for any arbitrary k, and show that it also holds for $k+1$. That is,

$$\begin{aligned}(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k\mathbf{A}_{k+1})^{\mathrm{T}}&=\left((\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k)\mathbf{A}_{k+1}\right)^{\mathrm{T}}\\&=\mathbf{A}_{k+1}^{\mathrm{T}}(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k)^{\mathrm{T}}\\&=\mathbf{A}_{k+1}^{\mathrm{T}}\mathbf{A}_k^{\mathrm{T}}\mathbf{A}_{k-1}^{\mathrm{T}}\cdots\mathbf{A}_1^{\mathrm{T}}.\end{aligned}$$

The first equality holds by the associative property of matrix multiplication as stated in Theorem 3.1 (d), the second by Theorem 3.3, and the last by the induction hypothesis.❑

Problem Show that $(\mathbf{A}^k)^{\mathrm{T}} = (\mathbf{A}^k)^{\mathrm{T}}$, where $\mathbf{A}^k$ is matrix $\mathbf{A}$ multiplied by itself k times, for any positive integer k. (Note that for $\mathbf{A}^k$ to be defined, $\mathbf{A}$ has to have the same number of rows and columns. That is, $\mathbf{A}$ has to be square. See Definition 3.9.)

3.5 Special Types of Matrices

Definition 3.8 A matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a ***zero matrix*** if all of its elements are zeroes. We write $\mathbf{A} = \mathbf{0}$.

Definition 3.9 A matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a ***square matrix*** if $m = n$. That is, the number of rows is equal to the number of columns.

Example The matrix $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is square, but $\begin{bmatrix} 1 & 2 & -1 \\ 8 & 1 & 0 \end{bmatrix}$ is not.

Definition 3.10 The ***diagonal elements*** of a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ are the elements $a_{11}, a_{22}, \dots, a_{nn}$.

Example The diagonal elements of the square matrix of the previous example are 1 and 4.

Definition 3.11 A ***diagonal matrix*** is a square matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ where $a_{ij} = 0$ if $i \neq j$.

That is, only the diagonal elements of diagonal matrix may be nonzero.

Example The following matrices are diagonal;

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}, \text{and} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -5 \end{bmatrix}.$$

Problem Leon [1994], page 63, #20.

20. In general, matrix multiplication is not commutative (i.e., $\mathbf{AB} \neq \mathbf{BA}$). However, there are certain special cases where the commutative property does hold. Show that:
    a. If $\mathbf{D}_1$ and $\mathbf{D}_2$ are $n \times n$ diagonal matrices, then $\mathbf{D}_1\mathbf{D}_2 = \mathbf{D}_2\mathbf{D}_1$.
    b. If $\mathbf{A}$ is an $n \times n$ matrix and

$$\mathbf{B} = a_0\mathbf{I} + a_1\mathbf{A} + a_2\mathbf{A}^2 + \cdots + a_k\mathbf{A}^k.$$

where $a_0, a_1, \dots, a_k$ are scalars, then $\mathbf{AB} = \mathbf{BA}$. (See Definition 3.12 below for identity matrix $\mathbf{I}$.)

Definition 3.12 An ***identity matrix*** $\mathbf{I}_n$ is a diagonal matrix of order $n \times n$ whose diagonal elements being ones. An identity matrix is thus necessarily square.

Example The following are identity matrix of order 2×2, 3×3 and 4×4;

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } \mathbf{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The identity matrix is so called because of the following property.

Proposition 3.1 Let $\mathbf{A}$ be a matrix in $\mathbb{R}^{m \times n}$. Then, $\mathbf{I}_m\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_n$.

Proof The identity matrix $\mathbf{I}_m$ can be written as,

$$\mathbf{I}_m = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_m \end{bmatrix}, \text{ where } \mathbf{e}_i = [0 \quad \cdots \quad 0 \quad 1 \quad 0 \quad \cdots \quad 0].$$

The elements of row vector $\mathbf{e}_i$ are zeros except the i^{th} element being one. When $\mathbf{A}$ is premultiplied by $\mathbf{I}_m$, we have $\mathbf{I}_m\mathbf{A} = [\mathbf{e}_i\mathbf{a}_j] = [a_{ij}] = \mathbf{A}$. The second equation can be proved similarly and left as an exercise.❑

When its dimension is unambiguous, the subscript of identity matrix $\mathbf{I}_m$ can be suppressed. We can write $\mathbf{IA} = \mathbf{A}$.

Definition 3.13 A matrix $\mathbf{A}$ is ***upper triangular*** if it is square and $a_{ij} = 0$ if $i > j$ and ***lower triangular*** if it is square and $a_{ij} = 0$ if $i < j$.

Example The matrix $\begin{bmatrix} 4 & 3 \\ 0 & -1 \end{bmatrix}$ is upper triangular and $\begin{bmatrix} 4 & 0 \\ 2 & 0 \end{bmatrix}$ is lower triangular.

A diagonal matrix, identity matrix and even a zero square matrix are both upper and lower triangular. However, a row echelon matrix is not upper triangular. Why?

Problem Suppose that **A** and **B** are $n \times n$ upper triangular matrices. Is **AB** also upper triangular?

Problem Lipschutz [1968], #3.58, page 60. Show that the sum, product and scalar multiple of:

a. lower triangular matrices is lower triangular.
b. diagonal matrices is diagonal.

Definition 3.14 A matrix is ***symmetric*** if $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$. That is, $a_{ij} = a_{ji}$, for all i and j.

A symmetric matrix is thus necessarily square. For the special matrices seen so far, zero, identity and diagonal matrices are symmetric.

Problem Johnson, Riess & Arnold [1998], page 69, #30, 31.

30. Find 2×2 matrices **A** and **B** such that **A** and **B** are symmetric, but **AB** are not symmetric. (*Hint*: $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \mathbf{BA}$)
31. Let **A** and **B** be $n \times n$ symmetric matrices. Give a necessary and sufficient condition for **AB** to be symmetric. (*Hint*: Recall Exercise 30.)

Problem Let **Q** be an $m \times n$ matrix. Show that

a. $\mathbf{QQ}^{\mathrm{T}}$ and $\mathbf{Q}^{\mathrm{T}}\mathbf{Q}$ are symmetric.
b. $\mathbf{x}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\mathbf{x} \geq 0$, for any vector $\mathbf{x} \in \mathbb{R}^n$.

Definition 3.15 A matrix is ***idempotent*** if $\mathbf{A}^2 = \mathbf{A}$.

We can show by induction that a matrix **A** is idempotent if, and only if, for any positive integer k, $\mathbf{A}^k = \mathbf{A}$. Is an idempotent matrix necessarily square?

Example A square zero matrix and identity matrix are idempotent.

Problem Simon & Blume [1994], #8.7, page 162. Show that $\begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}$ and $\begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix}$ are idempotent.

Problem Johnson, Riess & Arnold [1998], page 103, #52. Let **u** be a vector in $\mathbb{R}^n$ such that $\mathbf{u}^{\mathrm{T}}\mathbf{u} =$

1. Let $\mathbf{A} = \mathbf{I} - \mathbf{u}\mathbf{u}^{\mathrm{T}}$, where $\mathbf{I}$ is an $n \times n$ identity matrix. Verify that $\mathbf{A}$ is idempotent, that is, $\mathbf{AA} = \mathbf{A}$.

Definition 3.16 An ***elementary matrix*** is a matrix obtained by performing elementary row operation *exactly once* on an identity matrix.

Example The matrices $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$ are elementary. An identity matrix is elementary but diagonal matrix is not in general. For example, $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ is not elementary.

3.6 Elementary Matrices and Elementary Row Operations

If a matrix $\mathbf{B}$ is obtained from performing an elementary row operations on $\mathbf{A}$, then $\mathbf{B}$ can also be obtained by first performing that exact row operation on an identity matrix to get an elementary matrix $\mathbf{E}$, and then premultiplying matrix $\mathbf{A}$ by $\mathbf{E}$. That is, $\mathbf{B} = \mathbf{EA}$.

Example Let $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix}$, obtained by interchanging the first and second rows. The elementary matrix needed is $\mathbf{E} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. So,

$$\mathbf{B} = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \mathbf{EA}.$$

Theorem 3.4 If matrix $\boldsymbol{B} \in \mathbb{R}^{m \times n}$ is obtained by performing an elementary row operation once on a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, then $\mathbf{B} = \mathbf{EA}$, where $\mathbf{E} \in \mathbb{R}^{m \times m}$ is the elementary matrix obtained from performing that exact elementary row operation on identity matrix.

Proof Write identity matrix as $\mathbf{I}_m = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_m \end{bmatrix}$. By multiplying row i by a constant α, the elementary matrix obtained is $\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \alpha\mathbf{e}_i \\ \vdots \\ \mathbf{e}_m \end{bmatrix}$, and the matrix **B** is the product,

$$\mathbf{EA} = \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \alpha\mathbf{e}_i \\ \vdots \\ \mathbf{e}_m \end{bmatrix} [\mathbf{a}_{\cdot 1} \quad \cdots \quad \mathbf{a}_{\cdot j} \quad \cdots \quad \mathbf{a}_{\cdot n}]$$

$$= \begin{bmatrix} \mathbf{e}_1\mathbf{a}_{\cdot 1} & \cdots & \mathbf{e}_1\mathbf{a}_{\cdot j} & \cdots & \mathbf{e}_1\mathbf{a}_{\cdot n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha\mathbf{e}_i\mathbf{a}_{\cdot 1} & \cdots & \alpha\mathbf{e}_i\mathbf{a}_{\cdot j} & \cdots & \alpha\mathbf{e}_i\mathbf{a}_{\cdot n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{e}_m\mathbf{a}_{\cdot 1} & \cdots & \mathbf{e}_m\mathbf{a}_{\cdot j} & \cdots & \mathbf{e}_m\mathbf{a}_{\cdot n} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha a_{i1} & \cdots & \alpha a_{ij} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \alpha\mathbf{a}_i \\ \vdots \\ \mathbf{a}_m \end{bmatrix}.$$

Similar proofs for other two kinds of elementary row operations are left as exercises.❑

Problem Similar to the elementary row operations, we can define three elementary column operations to be interchanging a pair of columns, multiplying a column by a constant, and multiplying a column by a constant and then add it to another column. Write a similar theorem for elementary column operations and prove it.

The reduction of a matrix into a row echelon form by elementary row operations can then be viewed as premultiplying the matrix by a series of elementary matrices until a row echelon form is obtained.

***Problem** Is $\mathbf{E}^{\mathrm{T}}$ also elementary matrix?

***Problem** **Leon** [1994], page 62, #15. Is the transpose of an elementary matrix an elementary matrix of the same type? Is the product of two elementary matrices an elementary matrix?

Theorem 3.5 Let $\mathbf{A}$ be a matrix in $\mathbb{R}^{m\times n}$. There exists a series of elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k$, and $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_k$ for some finite integer k such that,

a) $\mathbf{E}_k\mathbf{E}_{k-1}\cdots\mathbf{E}_1\mathbf{A} = \mathbf{R}$, and

b) $\mathbf{A} = \mathbf{F}_1\mathbf{F}_2\cdots\mathbf{F}_k\mathbf{R}$.

where **R** is a matrix in row echelon form.

Proof Any matrix can be reduced into row echelon form in a finite number of k elementary row operations, and by Theorem 3.4 we have k elementary matrices. Premultiplying **A** consecutively by $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k$ we have **R**.

For each elementary matrix, we can do an elementary row operation to reverse it back to an identity matrix (See **Lipschutz** [1968], #3.21, page 53). That is, for each $\mathbf{E}_i$, there exists $\mathbf{F}_i$ for $i = 1,2,\dots,k$, such that $\mathbf{F}_i\mathbf{E}_i = \mathbf{I}$. Thus, part **(b)** is obtained by premultiplying both sides of **(a)** by $\mathbf{F}_1\mathbf{F}_2\cdots\mathbf{F}_k$.❑

Corollary 3.2 Let **A** be a square matrix in $\mathbb{R}^{n\times n}$. If **A** is full rank, there exists a series of elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k$, and $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_k$ for some finite integer k such that,

a) $\mathbf{E}_k\mathbf{E}_{k-1}\cdots\mathbf{E}_1\mathbf{A} = \mathbf{I}$, and

b) $\mathbf{A} = \mathbf{F}_1\mathbf{F}_2\cdots\mathbf{F}_k$.

Proof Exercise.❑

Problem Show that $rank\ \mathbf{AB} = rank\ \mathbf{B}$ if the matrix **A** is square and full rank.