2 Matrix Algebra

2.1

MATRIX OPERATIONS

□ Could be replaced by {R}

□ Could be replaced by {~}

MATRIX OPERATIONS

- If A is an $m \times n$ matrix—that is, a matrix with m rows and n columns—then the scalar entry in the ith row and jth column of A is denoted by a_{ij} and is called the (i, j)-entry of A. See the figure below.
- Each column of A is a list of m real numbers, which identifies a vector in $\mathbb{R}^m$.

$$\begin{array}{cc}
 & \begin{array}{ccccc} & & \text{Column} & & \\ & & j & & \end{array} \\
\text{Row } i & \left[\begin{array}{ccccc} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{array}\right] = A \\
 & \begin{array}{ccccc} \uparrow & & \uparrow & & \uparrow \\ \mathbf{a}_1 & & \mathbf{a}_j & & \mathbf{a}_n \end{array}
\end{array}$$

Matrix notation.

MATRIX OPERATIONS

- The columns are denoted by $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and the matrix A is written as

$$A = \begin{bmatrix} \mathrm{a}_1 & \mathrm{a}_2 & \cdots & \mathrm{a}_n \end{bmatrix}.$$

- The number a_{ij} is the ith entry (from the top) of the jth column vector $\mathbf{a}_j$.
- The **diagonal entries** in an $m \times n$ matrix $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ are $\mathrm{a}_{11}, \mathrm{a}_{22}, \mathrm{a}_{33}, \ldots$, and they form the **main diagonal** of A.
- A **diagonal matrix** is a sequence $n \times m$ matrix whose nondiagonal entries are zero.
- An example is the $n \times n$ identity matrix, I_n.

SUMS AND SCALAR MULTIPLES

- An $m \times n$ matrix whose entries are all zero is a **zero matrix** and is written as 0.

- The two matrices are **equal** if they have the same size (*i.e.*, the same number of rows and the same number of columns) and if their corresponding columns are equal, which amounts to saying that their corresponding entries are equal.

- If A and B are $m \times n$ matrices, then the **sum** $A + B$ is the $m \times n$ matrix whose columns are the sums of the corresponding columns in A and B.

SUMS AND SCALAR MULTIPLES

- Since vector addition of the columns is done entrywise, each entry in $A+B$ is the sum of the corresponding entries in A and B.

- The sum $A+B$ is defined only when A and B are the same size.

- **Example 1:** Let $A=\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, B=\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix},$

and $C=\begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$. Find $A+B$ and $A+C$.

SUMS AND SCALAR MULTIPLES

- **Solution:** $A+B=\begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$ but $A+C$ is not defined because A and C have different sizes.

- If r is a scalar and A is a matrix, then the **scalar multiple** rA is the matrix whose columns are r times the corresponding columns in A.

- **Theorem 1:** Let A, B, and C be matrices of the same size, and let r and s be scalars.
  a. $A+B=B+A$

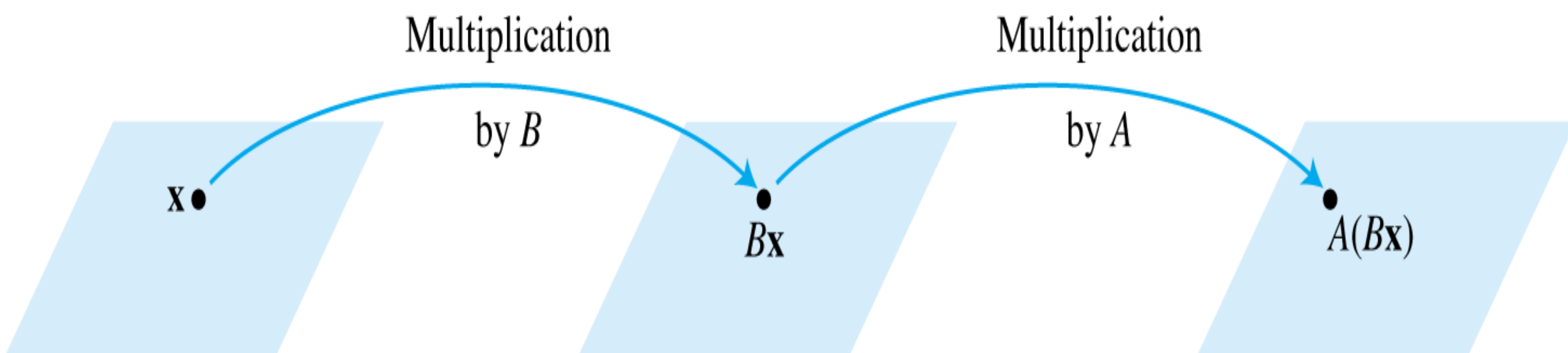
SUMS AND SCALAR MULTIPLES

b. $(A+B)+C = A+(B+C)$
c. $A+0 = A$
d. $r(A+B) = rA+rB$
e. $(r+s)A = rA+sA$
f. $r(sA) = (rs)A$

- Each quantity in Theorem 1 is verified by showing that the matrix on the left side has the same size as the matrix on the right and that corresponding columns are equal.

MATRIX MULTIPLICATION

- When a matrix B multiplies a vector $\mathbf{x}$, it transforms $\mathbf{x}$ into the vector $B\mathbf{x}$.
- If this vector is then multiplied in turn by a matrix A, the resulting vector is $A(B\mathbf{x})$. See the Fig. below.

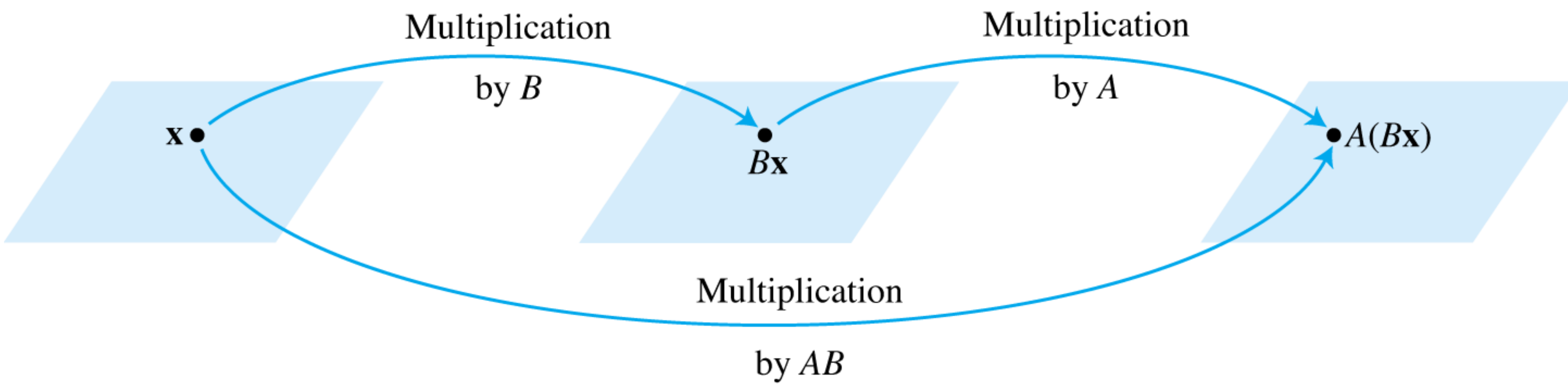


Multiplication by B and then A.

- Thus $A(B\mathbf{x})$ is produced from x by a composition of mappings—the linear transformations.

MATRIX MULTIPLICATION

- Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by AB, so that $A(B\text{x})=(\text{AB})\text{x}$. See the figure below.



Multiplication by AB.

- If A is $m \times n$, B is $n \times p$, and $\mathbf{x}$ is in $\mathbb{R}^p$, denote the columns of B by $\mathbf{b}_1, \ldots, \mathbf{b}p$ and the entries in $\mathbf{x}$ by $\mathbf{x}_1, \ldots, \mathbf{x}_p$.

MATRIX MULTIPLICATION

- Then

$$B\mathrm{x} = x_1\mathrm{b}_1 + \ldots + x_p\mathrm{b}_p$$

- By the linearity of multiplication by A,

$$A(B\mathrm{x}) = A(x_1\mathrm{b}_1) + \ldots + A(x_p\mathrm{b}_p)$$

$$= x_1 A\mathrm{b}_1 + \ldots + x_p A\mathrm{b}_p$$

- The vector A ($B\mathbf{x}$) is a linear combination of the vectors $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$, using the entries in $\mathbf{x}$ as weights.
- In matrix notation, this linear combination is written as

$$A(B\mathrm{x}) = \begin{bmatrix} A\mathrm{b}_1 & A\mathrm{b}_2 & \cdots & A\mathrm{b}_p \end{bmatrix}\mathrm{x}.$$

MATRIX MULTIPLICATION

- Thus multiplication by $\begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}$ transforms $\mathbf{x}$ into $A(B\mathbf{x})$.

- **Definition:** If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$.

- That is,

$$AB = A\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}$$

- *Multiplication of matrices corresponds to composition of linear transformations.*

MATRIX MULTIPLICATION

- **Example 2:** Compute AB, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 9 \\ 1 & -2 & 3 \end{bmatrix}$.

- **Solution:** Write $B = \begin{bmatrix} \mathrm{b}_1 & \mathrm{b}_2 & \mathrm{b}_3 \end{bmatrix}$, and compute:

MATRIX MULTIPLICATION

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}, A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}, A\mathbf{b}_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

- Then

$$AB = A\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{bmatrix} = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$$\begin{matrix} \uparrow & \uparrow & \uparrow \\ A\mathbf{b}_1 & A\mathbf{b}_2 & A\mathbf{b}_3 \end{matrix}$$

MATRIX MULTIPLICATION

- Each column of AB is a linear combination of the columns of A using weights from the corresponding column of B.

- Row—column rule for computing AB
- If a product AB is defined, then the entry in row i and column j of AB is the sum of the products of corresponding entries from row i of A and column j of B.
- If $(AB)_{ij}$ denotes the (i, j)-entry in AB, and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + \ldots + a_{in}b_{nj}.$$

PROPERTIES OF MATRIX MULTIPLICATION

- **Theorem 2:** Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.
  a. $A(BC) = (AB)C$ (associative law of multiplication)
  b. $A(B+C) = AB + AC$ (left distributive law)
  c. $(B+C)A = BA + CA$ (right distributive law)
  d. $r(AB) = (rA)B = A(rB)$ for any scalar r
  e. $I_m A = A = AI_n$ (identity for matrix multiplication)

PROPERTIES OF MATRIX MULTIPLICATION

- **Proof:** Property (a) follows from the fact that matrix multiplication corresponds to composition of linear transformations (which are functions), and it is known that the composition of functions is associative.

- Let $C = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_p \end{bmatrix}$

- By the definition of matrix multiplication,

$$BC = \begin{bmatrix} B\mathbf{c}_1 & \cdots & B\mathbf{c}_p \end{bmatrix}$$

$$A(BC) = \begin{bmatrix} A(B\mathbf{c}_1) & \cdots & A(B\mathbf{c}_p) \end{bmatrix}$$

PROPERTIES OF MATRIX MULTIPLICATION

- The definition of AB makes $A(B\mathbf{x}) = (AB)\mathbf{x}$ for all $\mathbf{x}$, so

$$A(BC) = \begin{bmatrix} (AB)\mathbf{c}_1 & \cdots & (AB)\mathbf{c}_p \end{bmatrix} = (AB)C$$

- The left-to-right order in products is critical because AB and BA are usually not the same.
- Because the columns of AB are linear combinations of the columns of A, whereas the columns of BA are constructed from the columns of B.
- The position of the factors in the product AB is emphasized by saying that A is *right-multiplied* by B or that B is *left-multiplied* by A.

PROPERTIES OF MATRIX MULTIPLICATION

- If $AB = BA$, we say that A and B **commute** with one another.

- **Warnings:**
  1. In general, $AB \neq BA$.
  2. The cancellation laws do *not* hold for matrix multiplication. That is, if $AB = AC$, then it is *not* true in general that $B = C$.
  3. If a product AB is the zero matrix, you *cannot* conclude in general that either $A = 0$ or $B = 0$.

POWERS OF A MATRIX

- If A is an $n \times n$ matrix and if k is a positive integer, then A^k denotes the product of k copies of A:

$$A^k = \underbrace{A \cdots A}_{k}$$

- If A is nonzero and if $\mathbf{x}$ is in $\mathbb{R}^n$, then $A^k\mathbf{x}$ is the result of left-multiplying $\mathbf{x}$ by A repeatedly k times.

- If $k = 0$, then $A^0\mathbf{x}$ should be $\mathbf{x}$ itself.

- Thus A^0 is interpreted as the identity matrix.

THE TRANSPOSE OF A MATRIX

- Given an $m \times n$ matrix A, the **transpose** of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

Theorem 3: Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^T)^T = A$

b. $(A+B)^T = A^T + B^T$

c. For any scalar r, $(rA)^T = rA^T$

d. $(AB)^T = B^T A^T$

THE TRANSPOSE OF A MATRIX

- The transpose of a product of matrices equals the product of their transposes in the *reverse* order.

2 Matrix Algebra

2.1

THE INVERSE OF A MATRIX

MATRIX OPERATIONS

- An $n \times n$ matrix A is said to be invertible if there is an $n \times n$ matrix C such that
$$CA = I \quad \text{and} \quad AC = I$$
where $I = I_n$, the $n \times n$ identity matrix.
- In this case, C is an inverse of A.
- In fact, C is uniquely determined by A, because if B were another inverse of A, then
$$B = BI = B(AC) = (BA)C = IC = C.$$
- This unique inverse is denoted by A^{-1}, so that
$$A^{-1}A = I \text{ and } AA^{-1} = I.$$

MATRIX OPERATIONS

- **Theorem 4:** Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then

  A is invertible and

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

  If $ad - bc = 0$, then A is not invertible.

- The quantity $ad - bc$ is called the determinant of A, and we write $\det A = ad - bc$

- This theorem says that a 2×2 matrix A is invertible if and only if $\det A \neq 0$.

MATRIX OPERATIONS

- **Theorem 5:** If A is an invertible $n \times n$ matrix, then for each **b** in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.
- **Proof:** Take any **b** in $\mathbb{R}^n$.
- A solution exists because if $A^{-1}\mathbf{b}$ is substituted for **x**, then $A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$.
- So $A^{-1}\mathbf{b}$ is a solution.
- To prove that the solution is unique, show that if **u** is any solution, then **u** must be $A^{-1}\mathbf{b}$.
- If $A\mathbf{u} = \mathbf{b}$, we can multiply both sides by A^{-1} and obtain $A^{-1}A\mathbf{u} = A^{-1}\mathbf{b}$, $I\mathbf{u} = A^{-1}\mathbf{b}$, and $\mathbf{u} = A^{-1}\mathbf{b}$.

MATRIX OPERATIONS

- **Theorem 6:**
  a. If A is an invertible matrix, then A^{-1} is invertible and
  $$(A^{-1})^{-1} = A$$
  b. If A and B are $n \times n$ invertible matrices, then so is AB, and the inverse of AB is the product of the inverses of A and B in the reverse order. That is, $$(AB)^{-1} = B^{-1}A^{-1}$$
  c. If A is an invertible matrix, then so is A^T, and the inverse of A^T is the transpose of A^{-1}. That is, $$(A^T)^{-1} = (A^{-1})^T$$

MATRIX OPERATIONS

- **Proof:** To verify statement (a), find a matrix C such that

$$A^{-1}C = I \text{ and } CA^{-1} = I$$

- These equations are satisfied with A in place of C. Hence A^{-1} is invertible, and A is its inverse.
- Next, to prove statement (b), compute:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

- A similar calculation shows that $(B^{-1}A^{-1})(AB) = I$.
- For statement (c), use Theorem 3(d), read from right to left, $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$.
- Similarly, $A^T(A^{-1})^T = I^T = I$.

ELEMENTARY MATRICES

- Hence A^T is invertible, and its inverse is $(A^{-1})T$.
- The generalization of Theorem 6(b) is as follows:

  The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.
- An invertible matrix A is row equivalent to an identity matrix, and we can find A^{-1} by watching the row reduction of A to I.
- An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix.

ELEMENTARY MATRICES

- **Example 1:** Let $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \quad A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Compute E_1A, E_2A, and E_3A, and describe how these products can be obtained by elementary row operations on A.

ELEMENTARY MATRICES

- **Solution:** Verify that

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix}, \; E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix},$$

$$E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}.$$

- Addition of -4 times row 1 of A to row 3 produces E_1A.

ELEMENTARY MATRICES

- An interchange of rows 1 and 2 of A produces E_2A, and multiplication of row 3 of A by 5 produces E_3A.

- Left-multiplication by E_1 in Example 1 has the same effect on any $3 \times n$ matrix.

- Since $E_1 \cdot I = E_1$, we see that E_1 itself is produced by this same row operation on the identity.

ELEMENTARY MATRICES

- Example 1 illustrates the following general fact about elementary matrices.

- If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operation on I_m.

- Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I.

ELEMENTARY MATRICES

- **Theorem 7:** An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.
- **Proof:** Suppose that A is invertible.
- Then, since the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each **b** (Theorem 5), A has a pivot position in every row.
- Because A is square, the n pivot positions must be on the diagonal, which implies that the reduced echelon form of A is I_n. That is, $A \sim I_n$.

ELEMENTARY MATRICES

- Now suppose, conversely, that $A \sim I_n$.
- Then, since each step of the row reduction of A corresponds to left-multiplication by an elementary matrix, there exist elementary matrices $E_1, \ldots, E_p$ such that

$$A \sim E_1 A \sim E_2(E_1 A) \sim \ldots \sim E_p(E_{p-1} \ldots E_1 A) = I_n .$$

- That is, $\qquad E_p \ldots E_1 A = I_n \qquad$ ----(1)
- Since the product $E_p \ldots E_1$ of invertible matrices is invertible, (1) leads to

$$(E_p \ldots E_1)^{-1}(E_p \ldots E_1) A = (E_p \ldots E_1)^{-1} I_n$$

$$A = (E_p \ldots E_1)^{-1} .$$

ALGORITHM FOR FINDING A^{-1}

- Thus A is invertible, as it is the inverse of an invertible matrix (Theorem 6). Also,
$$A^{-1} = \left[(E_p \ldots E_1)^{-1}\right]^{-1} = E_p \ldots E_1.$$

- Then $A^{-1} = E_p \ldots E_1 \cdot I_n$, which says that A^{-1} results from applying E_1, ..., E_p successively to I_n.
- This is the same sequence in (1) that reduced A to I_n.
- Row reduce the augmented matrix $\begin{bmatrix} A & I \end{bmatrix}$. If A is row equivalent to I, then $\begin{bmatrix} A & I \end{bmatrix}$ is row equivalent to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$. Otherwise, A does not have an inverse.

ALGORITHM FOR FINDING A^{-1}

- **Example 2:** Find the inverse of the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}, \text{ if it exists.}$$

- **Solution:**

$$\begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix}$$

ALGORITHM FOR FINDING A^{-1}

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

ALGORITHM FOR FINDING A^{-1}

- Theorem 7 shows, since $A \sim I$, that A is invertible, and

$$A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}.$$

- Now, check the final answer.

$$AA^{-1} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

ANOTHER VIEW OF MATRIX INVERSION

- It is not necessary to check that $A^{-1}A = I$ since A is invertible.

- Denote the columns of I_n by $\mathbf{e}_1,\ldots,\mathbf{e}_n$.

- Then row reduction of $\begin{bmatrix} A & I \end{bmatrix}$ to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$ can be viewed as the simultaneous solution of the n systems

  $$A\mathrm{x} = \mathrm{e}_1,\ A\mathrm{x} = \mathrm{e}_2,\ \ldots,\ A\mathrm{x} = \mathrm{e}_n \qquad \text{----(2)}$$

  where the "augmented columns" of these systems have all been placed next to A to form

  $$\begin{bmatrix} A & \mathrm{e}_1 & \mathrm{e}_2 & \cdots & \mathrm{e}_n \end{bmatrix} = \begin{bmatrix} A & I \end{bmatrix}.$$

ANOTHER VIEW OF MATRIX INVERSION

- The equation $AA^{-1} = I$ and the definition of matrix multiplication show that the columns of A^{-1} are precisely the solutions of the systems in (2).