Chapter 1
Systems of Linear Equations and Matrices
Outline

- 1.1 Introduction to Systems of Linear Equations
- 1.2 Gaussian Elimination
- 1.3 Matrices and Matrix Operations
- 1.4 Inverse; Algebraic Properties of Matrices
- 1.5 Elementary Matrices and a Method for Finding $A^{-1}$
- 1.6 More on Linear Systems and Invertible Matrices
- 1.7 Diagonal, Triangular, and Symmetric Matrices
1.1
Introduction to Systems of Linear Equations
Linear Equations

- Any straight line in $xy$-plane can be represented algebraically by an equation of the form:
  \[ a_1x + a_2y = b \]
- General form: Define a linear equation in the $n$ variables $x_1, x_2, \ldots, x_n$:
  \[ a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \]
  where $a_1, a_2, \ldots, a_n$ and $b$ are real constants.
- The variables in a linear equation are sometimes called unknowns.
Example (Linear Equations)

The equations \( x + 3y = 7, \ y = \frac{1}{2}x + 3z + 1, \) and \( x_1 - 2x_2 - 3x_3 + x_4 = 7 \) are linear

- A linear equation does not involve any products or roots of variables
- All variables occur only to the first power and do not appear as arguments for trigonometric, logarithmic, or exponential functions.

The equations \( x + 3\sqrt{y} = 5, 3x + 2y - z + xz = 4, \) and \( y = \sin x \) are not linear

A solution of a linear equation is a sequence of \( n \) numbers \( s_1, s_2, \ldots, s_n \) such that the equation is satisfied.

The set of all solutions of the equation is called its solution set or general solution (通解) of the equation.
Example

- Find the solution of \( x_1 - 4x_2 + 7x_3 = 5 \)
- Solution:
  - We can assign arbitrary values to any two variables and solve for the third variable
  - For example
    \[
    x_1 = 5 + 4s - 7t, \quad x_2 = s, \quad x_3 = t
    \]
    where \( s, t \) are arbitrary values
Linear Systems

- A finite set of linear equations in the variables $x_1, x_2, \ldots, x_n$ is called a system of linear equations or a linear system (線性系統).
- A sequence of numbers $s_1, s_2, \ldots, s_n$ is called a solution of the system.
- A system has no solution is said to be inconsistent. (矛盾方程組)
- If there is at least one solution of the system, it is called consistent. (相容的)
- Every system of linear equations has either no solutions, exactly one solution, or infinitely many solutions.
- A general system of two linear equations:
  \[
  a_1x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\
  a_2x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\
  \vdots \quad \vdots \quad \vdots \\
  a_mx_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m
  \]
  - Two lines may be parallel – no solution
  - Two lines may be intersect at only one point – one solution
  - Two lines may coincide – infinitely many solutions
Linear Systems

- Example of a linear system of three equations in three unknowns

[Diagrams showing possible solutions for a system of three equations in three unknowns: 1) No solution (three parallel planes), 2) No solution (two parallel planes), 3) No solution (three planes intersecting at a line), 4) No solution (two planes parallel and a third plane intersecting), 5) One solution (three planes intersecting at a point), 6) Infinite solutions (two planes parallel and a third plane coinciding), 7) Infinite solutions (three planes coinciding)].

▲ 図 1.1.2
Example

\[
\begin{align*}
x - y + 2z &= 5 \\
2x - 2y + 4z &= 10 \\
3x - 3y + 6z &= 15
\end{align*}
\]

After elimination \[ x - y + 2z = 5 \]

General solution: \[ x = 5 + r - 2s \]
\[ y = r \]
\[ z = s \]

- The three planes coincide!
Augmented Matrices

- The location of the +’s, the x’s, and the =’s can be abbreviated by writing only the rectangular array of numbers.
- This is called the augmented matrix (增廣矩陣) for the system.
- It must be written in the same order in each equation as the unknowns and the constants must be on the right.

\[
\begin{align*}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\
\vdots & \quad \vdots & \quad \vdots & \quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m
\end{align*}
\]

\[
\begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\
a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
a_{m1} & a_{m2} & \ldots & a_{mn} & b_m
\end{bmatrix}
\]
Elementary Row Operations

- The basic method for solving a system of linear equations is to replace the given system by a new system that has the same solution set but which is easier to solve.

- Since the rows of an augmented matrix correspond to the equations in the associated system, a new system is generally obtained in a series of steps by applying the following three types of operations to eliminate unknowns systematically.

- These are called elementary row operations
  - Multiply an equation through by an nonzero constant
  - Interchange two equations
  - Add a multiple of one equation to another
**Example (Using Elementary Row Operations)**

\[
\begin{align*}
\begin{bmatrix}
1 & 1 & 2 & 9 \\
2 & 4 & -3 & 1 \\
3 & 6 & -5 & 0
\end{bmatrix}
& \rightarrow
\begin{bmatrix}
1 & 1 & 2 & 9 \\
0 & 2 & -7 & -17 \\
3 & 6 & -5 & 0
\end{bmatrix}
& \rightarrow
\begin{bmatrix}
1 & 1 & 2 & 9 \\
0 & 2 & -7 & -17 \\
0 & 3 & -11 & -27
\end{bmatrix}
& \rightarrow
\begin{bmatrix}
1 & 1 & 2 & 9 \\
0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\
0 & 3 & -11 & -27
\end{bmatrix}
\end{align*}
\]

\[
\begin{align*}
x + y + 2z &= 9 \\
y - \frac{7}{2}z &= -\frac{17}{2} \\
-\frac{1}{2}z &= -\frac{3}{2}
\end{align*}
\]

\[
\begin{align*}
x + y + 2z &= 9 \\
y - \frac{7}{2}z &= -\frac{17}{2} \\
z &= 3
\end{align*}
\]

\[
\begin{align*}
x + y + 2z &= 9 \\
y - \frac{7}{2}z &= -\frac{17}{2} \\
x + \frac{11}{2}z &= \frac{35}{2}
\end{align*}
\]

\[
\begin{align*}
x &= 1 \\
y &= 2 \\
z &= 3
\end{align*}
\]
1.2
Gaussian Elimination
Echelon Forms

- A matrix which has the following properties is in **reduced row-echelon form** (as in the previous example) (簡約列-梯型)
  - If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. We call this a **leading 1**. (首項1)
  - If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
  - In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.
  - Each *column* that contains a leading 1 has zeros everywhere else.

- A matrix that has the **first three properties** is said to be in **row-echelon form**. (列-梯型)

- Note: A matrix in reduced row-echelon form is of necessity in row-echelon form, but not conversely.
Example

- Reduced row-echelon form:

\[
\begin{bmatrix}
1 & 0 & 0 & 4 \\
0 & 1 & 0 & 7 \\
0 & 0 & 1 & -1
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
0 & 1 & -2 & 0 & 1 \\
0 & 0 & 0 & 1 & 3 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

- Row-echelon form:

\[
\begin{bmatrix}
1 & 4 & -3 & 7 \\
0 & 1 & 6 & 2 \\
0 & 0 & 1 & 5
\end{bmatrix}
\begin{bmatrix}
1 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
0 & 1 & 2 & 6 & 0 \\
0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]
Example

All matrices of the following types are in **row-echelon form** (any real numbers substituted for the *’s. ):

\[
\begin{bmatrix}
1 & * & * & * \\
0 & 1 & * & * \\
0 & 0 & 1 & * \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

All matrices of the following types are in **reduced row-echelon form** (any real numbers substituted for the *’s. ):

\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]
Example

- Suppose that the augmented matrix for a linear system in the unknowns \( x, y, z \) has been reduced as
  \[
  \begin{bmatrix}
  1 & -5 & 1 & 4 \\
  0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0
  \end{bmatrix}
  \]

  \[
  x - 5y + z = 4 \\
  0x + 0y + 0z = 0
  \]

  \[
  x = 4 + 5y - z
  \]

  General solution:
  \[
  x = 4 + 5s - t \\
  y = s \\
  z = t
  \]

  \( s \) and \( t \) can be arbitrary values
Elimination Methods

- A step-by-step elimination procedure that can be used to reduce any matrix to reduced row-echelon form

\[
\begin{bmatrix}
0 & 0 & -2 & 0 & 7 & 12 \\
2 & 4 & -10 & 6 & 12 & 28 \\
2 & 4 & -5 & 6 & -5 & -1
\end{bmatrix}
\]
Elimination Methods

- Step 1. Locate the leftmost column that does not consist entirely of zeros.

\[
\begin{bmatrix}
0 & 0 & -2 & 0 & 7 & 12 \\
2 & 4 & -10 & 6 & 12 & 28 \\
2 & 4 & -5 & 6 & -5 & -1 \\
\end{bmatrix}
\]

- Step 2. Interchange the top row with another row, to bring a nonzero entry to top of the column found in Step 1

\[
\begin{bmatrix}
2 & 4 & -10 & 6 & 12 & 28 \\
0 & 0 & -2 & 0 & 7 & 12 \\
2 & 4 & -5 & 6 & -5 & -1 \\
\end{bmatrix}
\]
Elimination Methods

- Step 3. If the entry that is now at the top of the column found in Step 1 is $a$, multiply the first row by $1/a$ in order to introduce a leading 1.

\[
\begin{bmatrix}
1 & 2 & -5 & 3 & 6 & 14 \\
0 & 0 & -2 & 0 & 7 & 12 \\
2 & 4 & -5 & 6 & -5 & -1 \\
\end{bmatrix}
\]

The 1st row of the preceding matrix was multiplied by 1/2.

- Step 4. Add suitable multiples of the top row to the rows below so that all entries below the leading 1 become zeros.

\[
\begin{bmatrix}
1 & 2 & -5 & 3 & 6 & 14 \\
0 & 0 & -2 & 0 & 7 & 12 \\
0 & 0 & 5 & 0 & -17 & -29 \\
\end{bmatrix}
\]

-2 times the 1st row of the preceding matrix was added to the 3rd row.
Elimination Methods

- **Step 5.** Now cover the top row in the matrix and begin again with Step 1 applied to the submatrix that remains. Continue in this way until the entire matrix is in row-echelon form.

\[
\begin{bmatrix}
1 & 2 & -5 & 3 & 6 & 14 \\
0 & 0 & -2 & 0 & 7 & 12 \\
0 & 0 & 5 & 0 & -17 & -29 \\
\end{bmatrix}
\]

The 1st row in the submatrix was multiplied by $-1/2$ to introduce a leading 1.

\[
\begin{bmatrix}
1 & 2 & -5 & 3 & 6 & 14 \\
0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\
0 & 0 & 5 & 0 & -17 & -29 \\
\end{bmatrix}
\]

Leftmost nonzero column in the submatrix
Elimination Methods

\[
\begin{bmatrix}
1 & 2 & -5 & 3 & 6 & 14 \\
0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\
0 & 0 & 0 & 0 & \frac{1}{2} & 1 \\
\end{bmatrix}
\]

-5 times the 1st row of the submatrix was added to the 2nd row of the submatrix to introduce a zero below the leading 1.

\[
\begin{bmatrix}
1 & 2 & -5 & 3 & 6 & 14 \\
0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\
0 & 0 & 0 & 0 & \frac{1}{2} & 1 \\
\end{bmatrix}
\]

The top row in the submatrix was covered, and we returned again Step1.

\[
\begin{bmatrix}
1 & 2 & -5 & 3 & 6 & 14 \\
0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\
0 & 0 & 0 & 0 & 1 & 2 \\
\end{bmatrix}
\]

Leftmost nonzero column in the new submatrix

The first (and only) row in the new submatrix was multiplied by 2 to introduce a leading 1.
Elimination Methods

- Step 6. Beginning with the last nonzero row and working upward, add suitable multiplies of each row to the rows above to introduce zeros above the leading 1’s.

\[
\begin{bmatrix}
1 & 2 & -5 & 3 & 6 & 14 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 2
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 2 & -5 & 3 & 6 & 14 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 2
\end{bmatrix}
\]

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

\[
\begin{bmatrix}
1 & 2 & 0 & 3 & 0 & 7 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 2
\end{bmatrix}
\]

7/2 times the third row was added to the second row

-6 times the third row was added to the first row

5 times the second row was added to the first row

The last matrix is in reduced row-echelon form
Elimination Methods

- Step1~Step5: the above procedure produces a row-echelon form and is called **Gaussian elimination**
- Step1~Step6: the above procedure produces a reduced row-echelon form and is called **Gaussian-Jordan elimination**
- Every matrix has a **unique reduced row-echelon** form but a row-echelon form of a given matrix is not unique
- **Back-Substitution**
  - It is sometimes preferable to solve a system of linear equations by using Gaussian elimination to bring the augmented matrix into row-echelon form **without continuing all the way to the reduced row-echelon form**.
  - When this is done, the corresponding system of equations can be solved by a technique called **back-substitution**
Homogeneous Linear Systems

A system of linear equations is said to be homogeneous (齊次的) if the constant terms are all zero; that is, the system has the form:

\[
\begin{align*}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0 \\
\vdots & \quad \vdots & \quad \vdots & \quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0
\end{align*}
\]

Every homogeneous system of linear equation is consistent, since all such system have \(x_1 = 0, x_2 = 0, \ldots, x_n = 0\) as a solution.

- This solution is called the trivial solution. (明顯解)
- If there are another solutions, they are called nontrivial solutions.

There are only two possibilities for its solutions:

- There is only the trivial solution
- There are infinitely many solutions in addition to the trivial solution
Example

- A homogeneous linear system of two equations in two unknowns

\[
\begin{align*}
    a_1 x + b_1 y &= 0 \\
    a_2 x + b_2 y &= 0
\end{align*}
\]
Example (Gauss-Jordan Elimination)

- Solve the homogeneous system of linear equations by Gauss-Jordan elimination
  \[2x_1 + 2x_2 - x_3 + x_5 = 0\]
  \[-x_1 - x_2 + 2x_3 - 3x_4 + x_5 = 0\]
  \[x_1 + x_2 - 2x_3 - x_5 = 0\]
  \[x_3 + x_4 + x_5 = 0\]
  
- The augmented matrix
  \[
  \begin{bmatrix}
  2 & 2 & -1 & 0 & 1 & 0 \\
  -1 & -1 & 2 & -3 & 1 & 0 \\
  1 & 1 & -2 & 0 & -1 & 0 \\
  0 & 0 & 0 & 1 & 0 & 0
  \end{bmatrix}
  \]

- Reducing this matrix to reduced row-echelon form
  \[
  \begin{bmatrix}
  1 & 1 & 0 & 0 & 1 & 0 \\
  0 & 0 & 1 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 & 0
  \end{bmatrix}
  \]

- The general solution is
  \[x_1 = -s - t, x_2 = s\]
  \[x_3 = -t, x_4 = 0, x_5 = t\]

- Note: the trivial solution is obtained when \(s = t = 0\)
Two important points:

- None of the three row operations alters the final column of zeros, so the system of equations corresponding to the reduced row-echelon form of the augmented matrix must also be a homogeneous system.

- If the given homogeneous system has \( m \) equations in \( n \) unknowns with \( m < n \), and there are \( r \) nonzero rows in reduced row-echelon form of the augmented matrix, we will have \( r < n \). It will have the form:

\[
\begin{align*}
\cdots x_{k_1} + \sum 0 &= 0 \\
\cdots x_{k_2} + \sum 0 &= 0 \\
\vdots & \quad \vdots \\
\cdots x_{k_r} + \sum 0 &= 0 \\
\end{align*}
\]

\[
\begin{align*}
x_{k_1} &= -\sum 0 \\
x_{k_2} &= -\sum 0 \\
\vdots \\
x_{k_r} &= -\sum 0 \\
\end{align*}
\]

\[
\begin{align*}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\vdots & \quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{align*}
\]
Theorem

- **Theorem 1.2.1**
  - If a homogeneous system has $n$ unknowns, and if the reduced row echelon form of its augmented matrix has $r$ nonzero rows, then the system has $n-r$ free variables.

- **Theorem 1.2.2**
  - A homogeneous system of linear equations with more unknowns than equations has **infinitely many solutions**.
Remarks

- Every matrix has a unique reduced row echelon form
- Row echelon forms are not unique
- Although row echelon forms are not unique, all row echelon forms of a matrix $A$ have the same number of zero rows, and the leading 1’s always occur in the same positions in the row echelon forms of $A$.

$$A = \begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$
1.3 Matrices and Matrix Operations
Definition and Notation

- A **matrix** is a rectangular array of numbers. The numbers in the array are called the **entries (元素)** in the matrix.
- A general $m \times n$ matrix $A$ is denoted as

$$A = \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} \\
    a_{21} & a_{22} & \cdots & a_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}$$

- The entry that occurs in row $i$ and column $j$ of matrix $A$ will be denoted $a_{ij}$ or $\langle A \rangle_{ij}$. If $a_{ij}$ is real number, it is common to be referred as **scalars (純量)**.
- The preceding matrix can be written as $[a_{ij}]_{m \times n}$ or $[a_{ij}]$.
- A matrix $A$ with $n$ rows and $n$ columns is called a **square matrix of order $n$**.
Examples of Matrices

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

- A matrix with only one row is called a *row matrix* (or a *row vector*).
- A matrix with only one column is called a *column matrix* (or a *column vector*).
Sum, Difference, and Product

- Two matrices are defined to be equal if they have the same size and their corresponding entries are equal. If \( A = [a_{ij}] \) and \( B = [b_{ij}] \) have the same size, then \( A = B \) if and only if \( a_{ij} = b_{ij} \) for all \( i \) and \( j \).

- If \( A \) and \( B \) are matrices of the same size, then the sum \( A + B \) is the matrix obtained by adding the entries of \( B \) to the corresponding entries of \( A \).

- The difference \( A - B \) is the matrix obtained by subtracting the entries of \( B \) from the corresponding entries of \( A \).

- If \( A \) is any matrix and \( c \) is any scalar, then the product \( cA \) is the matrix obtained by multiplying each entry of the matrix \( A \) by \( c \). The matrix \( cA \) is said to be the scalar multiple of \( A \).

- If \( A = [a_{ij}] \), then \( \langle cA \rangle_{ij} = c \langle A \rangle_{ij} = ca_{ij} \).
Example

\[ A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 3 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 2 & 7 \\ -1 & 3 & -5 \end{bmatrix} \]

\[ A + B = \begin{bmatrix} 2 & 5 & 11 \\ 0 & 6 & -4 \end{bmatrix} \quad 2A = \begin{bmatrix} 4 & 6 & 8 \\ 2 & 6 & 2 \end{bmatrix} \]

\[ A - B = \begin{bmatrix} 2 & 1 & -3 \\ 2 & 0 & 6 \end{bmatrix} \]

linear combination:

\[ 2A + 3B = \begin{bmatrix} 4 & 6 & 8 \\ 2 & 6 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 6 & 21 \\ -3 & 9 & -15 \end{bmatrix} = \begin{bmatrix} 4 & 12 & 29 \\ -1 & 15 & -13 \end{bmatrix} \]
Product of Matrices

- If $A$ is an $m \times r$ matrix and $B$ is an $r \times n$ matrix, then the product $AB$ is the $m \times n$ matrix whose entries are determined as follows.

- To find the entry in row $i$ and column $j$ of $AB$, single out row $i$ from the matrix $A$ and column $j$ from the matrix $B$. Multiply the corresponding entries from the row and column together and then add up the resulting products.

  That is, $(AB)_{m \times n} = A_{m \times r} B_{r \times n}$

$$AB = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1r} \\
  a_{21} & a_{22} & \cdots & a_{2r} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{i1} & a_{i2} & \cdots & a_{ir} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mr}
\end{bmatrix} \begin{bmatrix}
  b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\
  b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\
  \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
  b_{r1} & b_{r2} & \cdots & b_{rj} & \cdots & b_{rn}
\end{bmatrix}$$

the entry $\langle AB \rangle_{ij}$ in row $i$ and column $j$ of $AB$ is given by

$$\langle AB \rangle_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{ir}b_{rj}$$
Product of Matrices

\[ A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} \]

\[ AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & \boxed{4} & 3 \\ 0 & -1 & \boxed{3} & 1 \\ 2 & 7 & \boxed{5} & 2 \end{bmatrix} = \begin{bmatrix} \text{\textbullet\textbullet\textbullet\textbullet} \\ \text{\textbullet\textbullet\textbullet\textbullet 26} \end{bmatrix} \]

\[ 2 \times 4 + 6 \times 3 + 0 \times 5 = 26 \]

\[ AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & \boxed{3} \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} \text{\textbullet\textbullet\textbullet\textbullet} \\ \text{\textbullet\textbullet\textbullet\textbullet 13} \end{bmatrix} \]

\[ 1 \times 3 + 2 \times 1 + 4 \times 2 = 13 \]
Example

- Determining whether a product is defined

  \[ A_{3 \times 4} \quad B_{4 \times 7} \quad C_{7 \times 3} \]

- \( AB \) is defined and is a \( 3 \times 7 \) matrix; \( BC \) is defined and is a \( 4 \times 3 \) matrix; and \( CA \) is defined and is a \( 7 \times 4 \) matrix.

- The products \( AC \), \( CB \), and \( BA \) are all undefined.
Example

- If $A = [a_{ij}]$ is a $m \times r$ matrix and $B = [b_{ij}]$ is an $r \times n$ matrix, then the entry $(AB)_{ij}$ is given by

$$AB = \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1r} \\
    a_{21} & a_{22} & \cdots & a_{2r} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{i1} & a_{i2} & \cdots & a_{ir} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{m1} & a_{m2} & \cdots & a_{mr}
\end{bmatrix}
\begin{bmatrix}
    b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\
    b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\
    \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
    b_{r1} & b_{r2} & \cdots & b_{rj} & \cdots & b_{rn}
\end{bmatrix}
$$

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj}$$
Partitioned Matrices

- A matrix can be subdivided or **partitioned** into smaller matrices by inserting horizontal and vertical rules between selected rows and columns.

- For example, three possible partitions of a $3 \times 4$ matrix $A$:
  - The partition of $A$ into four submatrices $A_{11}$, $A_{12}$, $A_{21}$, and $A_{22}$:
    $$ A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} $$
  - The partition of $A$ into its row matrices $r_1$, $r_2$, and $r_3$:
    $$ A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} $$
  - The partition of $A$ into its column matrices $c_1$, $c_2$, $c_3$, and $c_4$:
    $$ A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & c_3 & c_4 \end{bmatrix} $$
Multiplication by Columns and by Rows

- It is possible to compute a particular row or column of a matrix product $AB$ without computing the entire product:
  
  \[ j \text{th column matrix of } AB = A[j \text{th column matrix of } B] \]
  
  \[ i \text{th row matrix of } AB = [i \text{th row matrix of } A]B \]

- If $a_1, a_2, ..., a_m$ denote the row matrices of $A$ and $b_1, b_2, ..., b_n$ denote the column matrices of $B$, then

\[
AB = A\begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = [Ab_1 \ Ab_2 \ \cdots \ Ab_n]
\]

\[
AB = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \quad B = \begin{bmatrix} a_1B \\ a_2B \\ \vdots \\ a_mB \end{bmatrix}
\]
Example

\[
A = \begin{bmatrix}
1 & 2 & 4 \\
2 & 6 & 0
\end{bmatrix} \quad B = \begin{bmatrix}
4 & 1 & 4 & 3 \\
0 & -1 & 3 & 1 \\
2 & 7 & 5 & 2
\end{bmatrix} \quad AB = \begin{bmatrix}
12 & 27 & 30 & 13 \\
8 & -4 & 26 & 12
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 2 & 4 \\
2 & 6 & 0
\end{bmatrix} \begin{bmatrix}
1 \\
-1 \\
7
\end{bmatrix} = \begin{bmatrix}
27 \\
-4
\end{bmatrix}
\]

Second column of \(AB\)

\[
\begin{bmatrix}
1 & 2 & 4 \\
2 & 6 & 0
\end{bmatrix} \begin{bmatrix}
4 & 1 & 4 & 3 \\
0 & -1 & 3 & 1 \\
2 & 7 & 5 & 2
\end{bmatrix} = \begin{bmatrix}
12 & 27 & 30 & 13
\end{bmatrix}
\]

First row of \(AB\)
Matrix Products as Linear Combinations

Let

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\text{ and } \quad x = \begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
\]

Then

\[
Ax = \begin{bmatrix}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n
\end{bmatrix} = x_1 \begin{bmatrix} a_{11} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \end{bmatrix}
\]

The product \(Ax\) of a matrix \(A\) with a column matrix \(x\) is a linear combination of the column matrices of \(A\) with the coefficients coming from the matrix \(x\)
Example

The matrix product

\[
\begin{bmatrix}
-1 & 3 & 2 \\
1 & 2 & -3 \\
2 & 1 & -2
\end{bmatrix}
\begin{bmatrix}
2 \\
-1 \\
3
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
-9 \\
-3
\end{bmatrix}
\]

can be written as the linear combination of column matrices

\[
2 \begin{bmatrix}
-1 \\
1 \\
2
\end{bmatrix} - 1 \begin{bmatrix}
3 \\
2 \\
1
\end{bmatrix} + 3 \begin{bmatrix}
2 \\
-3 \\
-2
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
-9 \\
-3
\end{bmatrix}
\]

The matrix product

\[
[1 \quad -9 \quad -3] 
\begin{bmatrix}
-1 & 3 & 2 \\
1 & 2 & -3 \\
2 & 1 & -2
\end{bmatrix}
= 
\begin{bmatrix}
-16 & -18 & 35
\end{bmatrix}
\]

can be written as the linear combination of row matrices

\[
1[\begin{bmatrix}
-1 & 3 & 2
\end{bmatrix}] - 9[\begin{bmatrix}
1 & 2 & -3
\end{bmatrix}] - 3[\begin{bmatrix}
2 & 1 & -2
\end{bmatrix}]
= 
\begin{bmatrix}
-16 & -18 & 35
\end{bmatrix}
\]
Example (Columns of a Product $AB$ as Linear Combinations)

We showed in Example 5 that

$$AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \\ 2 & 7 & 5 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 \\ 0 & -1 & 3 \\ 2 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

The column matrices of $AB$ can be expressed as linear combinations of the column matrices of $A$ as follows:

$$\begin{bmatrix} 12 \\ 8 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 6 \end{bmatrix} + 2 \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 27 \\ -4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 6 \end{bmatrix} + 7 \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 30 \\ 26 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 6 \end{bmatrix} + 5 \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 13 \\ 12 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 6 \end{bmatrix} + 2 \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$
Matrix Form of a Linear System

Consider any system of $m$ linear equations in $n$ unknowns:

\[ a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \]
\[ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \]
\[ \vdots \]
\[ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \]

\[
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_m
\end{bmatrix}
= 
\begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m
\end{bmatrix}
\]

\[ A\mathbf{x} = \mathbf{b} \]

- The matrix $A$ is called the coefficient matrix of the system.
- The augmented matrix of the system is given by

\[
[A \mid \mathbf{b}]= 
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\
  a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn} & b_m
\end{bmatrix}
\]
Matrices Defining Functions

- We can view $A$ as defining a rule that shows how a given $x$ is mapped into a corresponding $y$.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad x = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$y = Ax = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ -b \end{bmatrix}$$

- The effect of multiplying $A$ by a column vector is to change the sign of the second entry of the column vector.
Matrices Defining Functions

\[ B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad y = Bx = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} b \\ -a \end{bmatrix} \]

- The effect of multiplying \( B \) by a column vector is to interchange the first and second entries of the column vector, also changing the sign of the first entry.
If $A$ is any $m \times n$ matrix, then the transpose of $A$, denoted by $A^T$, is defined to be the $n \times m$ matrix that results from interchanging the rows and columns of $A$

- That is, the first column of $A^T$ is the first row of $A$, the second column of $A^T$ is the second row of $A$, and so forth

If $A$ is a square matrix, then the trace (跡數) of $A$, denoted by tr($A$), is defined to be the sum of the entries on the main diagonal of $A$. The trace of $A$ is undefined if $A$ is not a square matrix.

- For an $n \times n$ matrix $A = [a_{ij}]$, $\text{tr}(A) = \sum_{i=1}^{n} a_{ii}$
Example

\[ A = \begin{bmatrix}
    a_{11} & a_{12} & a_{13} & a_{14} \\
    a_{21} & a_{22} & a_{23} & a_{24} \\
    a_{31} & a_{32} & a_{33} & a_{34}
\end{bmatrix} \quad B = \begin{bmatrix}
    2 & 3 \\
    1 & 4 \\
    5 & 6
\end{bmatrix} \quad C = \begin{bmatrix}
    1 & 3 & 5
\end{bmatrix} \quad D = [4]

\[ A^T = \begin{bmatrix}
    a_{11} & a_{21} & a_{31} \\
    a_{12} & a_{22} & a_{32} \\
    a_{13} & a_{23} & a_{33} \\
    a_{14} & a_{24} & a_{34}
\end{bmatrix} \quad B^T = \begin{bmatrix}
    2 & 1 & 5 \\
    3 & 4 & 6
\end{bmatrix} \quad C^T = \begin{bmatrix}
    1 \\
    3 \\
    5
\end{bmatrix} \quad D^T = [4] \]
Example

\[ A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad B = \begin{bmatrix} -1 & 2 & 7 & 0 \\ 3 & 5 & -8 & 4 \\ 1 & 2 & 7 & -3 \\ 4 & -2 & 1 & 0 \end{bmatrix} \]

\[ tr(A) = a_{11} + a_{22} + a_{33} \quad tr(B) = -1 + 5 + 7 + 0 = 11 \]
1.4
Inverse; Algebraic Properties of Matrices
Properties of Matrix Operations

- For real numbers $a$ and $b$, we always have $ab = ba$, which is called the *commutative law for multiplication*. For matrices, however, $AB$ and $BA$ need not be equal.

- Equality can fail to hold for three reasons:
  - The product $AB$ is defined but $BA$ is undefined.
  - $AB$ and $BA$ are both defined but have different sizes.
  - It is possible to have $AB \neq BA$ even if both $AB$ and $BA$ are defined and have the same size.
Example

\[ A = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} \]

\[ AB = \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix} \quad BA = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix} \]

\[ AB \neq BA \]
Theorem 1.4.1
(Properties of Matrix Arithmetic)

Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid:

- $A + B = B + A$ (commutative law for addition)
- $A + (B + C) = (A + B) + C$ (associative law for addition)
- $A(BC) = (AB)C$ (associative law for multiplication)
- $A(B + C) = AB + AC$ (left distributive law)
- $(B + C)A = BA + CA$ (right distributive law)
- $A(B - C) = AB - AC$, $(B - C)A = BA - CA$
- $a(B + C) = aB + aC$, $a(B - C) = aB - aC$
- $(a+b)C = aC + bC$, $(a-b)C = aC - bC$
- $a(bC) = (ab)C$, $a(BC) = (aB)C = B(aC)$
Proof (d) \[ A(B + C) = AB + AC \]

- We must show that \( A(B+C) \) and \( AB+AC \) have the same size and that corresponding entries are equal.
- To form \( A(B+C) \), the matrices \( B \) and \( C \) must have the same size, say \( m \times n \), and the matrix \( A \) must then have \( m \) columns, so its size must be of the form \( r \times m \). This makes \( A(B+C) \) an \( r \times n \) matrix.
- It follows that \( AB+AC \) is also an \( r \times n \) matrix.
Proof (d)

\[ A(B + C) = AB + AC \]

- Suppose that \( A = [a_{ij}] \), \( B = [b_{ij}] \), and \( C = [c_{ij}] \). We want to show

\[ [A(B + C)]_{ij} = [AB + AC]_{ij} \]

- From the definitions of matrix addition and matrix multiplication, we have

\[
[A(B + C)]_{ij} = a_{i1}(b_{1j} + c_{1j}) + a_{i2}(b_{2j} + c_{2j}) + \cdots + a_{im}(b_{mj} + c_{mj}) \\
= (a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}) + (a_{i1}c_{1j} + a_{i2}c_{2j} + \cdots + a_{im}c_{mj}) \\
= [AB]_{ij} + [AC]_{ij} \\
= [AB + AC]_{ij}
\]
Example

As an illustration of the associative law for matrix multiplication, consider

\[
A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}
\]

Then

\[
AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 5 \\ 20 & 13 \\ 2 & 1 \end{bmatrix}
\]

and

\[
BC = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix}
\]

Thus,

\[
(AB)C = \begin{bmatrix} 8 & 5 \\ 20 & 13 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}
\]

and

\[
A(BC) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}
\]

so \((AB)C = A(BC)\), as guaranteed by Theorem 1.4.1c.
Zero Matrices (零矩陣)

- A matrix, all of whose entries are zero, is called a **zero matrix**
- A zero matrix will be denoted by $0$
- If it is important to emphasize the size, we shall write $0_{m \times n}$ for the $m \times n$ zero matrix.
- In keeping with our convention of using **boldface symbols** for matrices with one column, we will denote a zero matrix with one column by $\mathbf{0}$
- **Theorem 1.4.2 (Properties of Zero Matrices)**
  - Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid
    - $A + 0 = 0 + A = A$
    - $A - A = 0$
    - $0 - A = -A$
    - $A0 = 0; \ 0A = 0$
Cancellation Law

- For real numbers:
  - If $ab=ac$ and $a\neq 0$, then $b = c$
  - If $ab = 0$, then at least one of the factors on the left is 0.

- It fails in matrix operation

\[
A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix} \quad C = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix}
\]

\[
AB = AC = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix} \quad \text{but } B \neq C
\]

\[
A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 7 \\ 0 & 0 \end{bmatrix} \quad AB = 0 \quad \text{but } A \neq 0 \quad \text{and } B \neq 0
\]
Identity Matrices (單位矩陣)

- A square matrix with 1's on the main diagonal and 0's off the main diagonal is called an **identity matrix** and is denoted by $I$, or $I_n$ for the $n \times n$ identity matrix.
- If $A$ is an $m \times n$ matrix, then $AI_n = A$ and $I_mA = A$.
- An identity matrix plays the same role in matrix arithmetic as the number 1 plays in the numerical relationships $a \cdot 1 = 1 \cdot a = a$.
- **Theorem 1.4.3**
  - If $R$ is the reduced row-echelon form of an $n \times n$ matrix $A$, then either $R$ has a row of zeros or $R$ is the identity matrix $I_n$. 
Example

- **Zero matrices**
  
  $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad [0]$

- **Identity matrices**
  
  $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
Proof of Theorem 1.4.3

- Suppose that the reduced row-echelon form of $A$ is

\[
R = \begin{bmatrix}
  r_{11} & r_{12} & \cdots & r_{1n} \\
  r_{21} & r_{22} & \cdots & r_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  r_{n1} & r_{n2} & \cdots & r_{nn}
\end{bmatrix}
\]

- Either the last row in this matrix consists entirely of zeros or it does not.

- If not, the matrix contains no zero rows, and consequently each of the $n$ rows has a leading entry of 1.
Proof of Theorem 1.4.3

Since these leading 1’s occur progressively farther to the right as we move down the matrix, each of these 1’s must occur on the main diagonal.

Since the other entries in the same column as one of these 1’s are zero, $R$ must be $I_n$.

Thus, either $R$ has a row of zeros or $R = I_n$. 

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}$$
Inverse

- If $A$ is a square matrix, and if a matrix $B$ of the same size can be found such that $AB = BA = I$, then $A$ is said to be invertible (可逆的) or nonsingular and $B$ is called an inverse (逆矩阵) of $A$. If no such matrix $B$ can be found, then $A$ is said to be singular. (奇異的)

- Remark:
  - The inverse of $A$ is denoted as $A^{-1}$
  - Not every (square) matrix has an inverse
  - An inverse matrix has exactly one inverse
Example

\[ B = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} \text{ is an inverse of } \quad A = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} \]

\[ AB = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \]

\[ BA = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \]

\( A \) and \( B \) are invertible and each is an inverse of the other.
Example

The matrix \( A = \begin{bmatrix} 1 & 4 & 0 \\ 2 & 5 & 0 \\ 3 & 6 & 0 \end{bmatrix} \) is singular.

Let \( B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \)

The third column of \( BA \) is

\[
\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow BA \neq I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

\( j \)th column matrix of \( BA = B[j \text{th column matrix of } A] \)
Theorems

- **Theorem 1.4.4**
  - If $B$ and $C$ are both inverses of the matrix $A$, then $B = C$

- **Theorem 1.4.5**
  - The matrix
    $$ A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} $$
  - is invertible if $ad - bc \neq 0$, in which case the inverse is given by the formula
    $$ A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} $$

- **Theorem 1.4.6**
  - If $A$ and $B$ are invertible matrices of the same size, then $AB$ is invertible and $(AB)^{-1} = B^{-1}A^{-1}$
Proof of 1.4.4

If $B$ and $C$ are both inverses of the matrix $A$, then $B = C$

- Since $B$ is an inverse of $A$, we have $BA = I$.
- Multiplying both sides on the right by $C$ gives $(BA)C = IC = C$.
- But $(BA)C = B(AC) = BI = B$, so $C = B$. 
Proof of 1.4.6

If $A$ and $B$ are invertible matrices of the same size, then $AB$ is invertible and $(AB)^{-1} = B^{-1}A^{-1}$

- If we can show that $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I$, then we will have simultaneously shown that the matrix $AB$ is invertible and that $(AB)^{-1} = B^{-1}A^{-1}$.
- $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$.
- A similar argument shows that $(B^{-1}A^{-1})(AB) = I$
Example

\[
A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix} \quad AB = \begin{bmatrix} 7 & 6 \\ 9 & 8 \end{bmatrix}
\]

Applying the formula in Theorem 1.4.5, we obtain

\[
A^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \quad B^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix} \quad (AB)^{-1} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}
\]

\[
B^{-1}A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}
\]
Powers of A Matrix

- If $A$ is a square matrix, then we define the nonnegative integer powers of $A$ to be

$$A^0 = I, \quad A^n = A A \cdots A \quad (n > 0)$$

- If $A$ is invertible, then we define the negative integer powers to be

$$A^{-n} = (A^{-1})^n = A^{-1} A^{-1} \cdots A^{-1} \quad (n > 0)$$
Example

\[ A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \quad A^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \]

\[ A^3 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 30 \\ 15 & 41 \end{bmatrix} \]

\[ A^{-3} = (A^{-1})^3 = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix} \]
Theorems

- If $A$ is a square matrix and $r$ and $s$ are integers, then $A^r A^s = A^{r+s}$, $(A^r)^s = A^{rs}$

- Theorem 1.4.7 (Laws of Exponents)
  - If $A$ is invertible and $n$ is a nonnegative integer, then:
    - $A^{-1}$ is invertible and $(A^{-1})^{-1} = A$
    - $A^n$ is invertible and $(A^n)^{-1} = (A^{-1})^n$ for $n = 0, 1, 2, \ldots$
    - For any nonzero scalar $k$, the matrix $kA$ is invertible and $(kA)^{-1} = k^{-1}A^{-1} = (1/k)A^{-1}$
Proof

$A^{-1}$ is invertible and $(A^{-1})^{-1} = A$

- Since $AA^{-1} = A^{-1}A = I$, the matrix $A^{-1}$ is invertible and $(A^{-1})^{-1} = A$.

For any nonzero scalar $k$, the matrix $kA$ is invertible and $(kA)^{-1} = (1/k)A^{-1}$

$$ (kA)(\frac{1}{k}A^{-1}) = \frac{1}{k}(kA)A^{-1} = (\frac{1}{k}k)AA^{-1} = 1I = I $$

Similarly, $\left(\frac{1}{k}A^{-1}\right)(kA) = I$
Polynomial Expressions Involving Matrices

- If $A$ is a square matrix, say $m \times m$, and if
  \[ p(x) = a_0 + a_1x + \ldots + a_nx^n \]
is any polynomial, then we define
  \[ p(A) = a_0I + a_1A + \ldots + a_nA^n \]
where $I$ is the $m \times m$ identity matrix.

- That is, $p(A)$ is the $m \times m$ matrix that results when $A$ is substituted for $x$ in the above equation and $a_0$ is replaced by $a_0I$
Example (Matrix Polynomial)

If

\[ p(x) = 2x^2 - 3x + 4 \quad \text{and} \quad A = \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} \]

then

\[ p(A) = 2A^2 - 3A + 4I = 2\begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}^2 - 3\begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} + 4\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

\[ = \begin{bmatrix} 2 & 8 \\ 0 & 18 \end{bmatrix} - \begin{bmatrix} -3 & 6 \\ 0 & 9 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 9 & 2 \\ 0 & 13 \end{bmatrix} \]
Theorems

- Theorem 1.4.8 (Properties of the Transpose)
  - If the sizes of the matrices are such that the stated operations can be performed, then
    - \((A^T)^T = A\)
    - \((A + B)^T = A^T + B^T\) and \((A - B)^T = A^T - B^T\)
    - \((kA)^T = kA^T\), where \(k\) is any scalar
    - \((AB)^T = B^T A^T\)

- Theorem 1.4.9 (Invertibility of a Transpose)
  - If \(A\) is an invertible matrix, then \(A^T\) is also invertible and \((A^T)^{-1} = (A^{-1})^T\)
Proof

If $A$ is an invertible matrix, then $A^T$ is also invertible and $(A^T)^{-1} = (A^{-1})^T$

- We can prove the invertibility of $A^T$ by showing that

$$A^T(A^{-1})^T = (A^{-1})^T A^T = I$$
$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$$

$$(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$$

$$A = \begin{bmatrix} -5 & -3 \\ 2 & 1 \end{bmatrix} \quad A^T = \begin{bmatrix} -5 & 2 \\ -3 & 1 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 1 & 3 \\ -2 & -5 \end{bmatrix} \quad (A^{-1})^T = \begin{bmatrix} 1 & -2 \\ 3 & -5 \end{bmatrix} \quad (A^T)^{-1} = \begin{bmatrix} 1 & -2 \\ 3 & -5 \end{bmatrix}$$
1.5

Elementary Matrices and a Method for Finding $A^{-1}$
An elementary row operation (sometimes called just a row operation) on a matrix $A$ is any one of the following three types of operations:

- Interchange of two rows of $A$
- Replacement of a row $r$ of $A$ by $cr$ for some number $c \neq 0$
- Replacement of a row $r_1$ of $A$ by the sum $r_1 + cr_2$ of that row and a multiple of another row $r_2$ of $A$

Matrices $A$ and $B$ are row equivalent if either can be obtained from the other by a sequence of elementary row operations.
An $n \times n$ elementary matrix is a matrix produced by applying exactly one elementary row operation to $I_n$

- $E_{ij}$ is the elementary matrix obtained by interchanging the $i$-th and $j$-th rows of $I_n$
- $E_i(c)$ is the elementary matrix obtained by multiplying the $i$-th row of $I_n$ by $c \neq 0$
- $E_{ij}(c)$ is the elementary matrix obtained by adding $c$ times the $j$-th row to the $i$-th row of $I_n$, where $i \neq j$
Example (Elementary Matrices and Row Operations)

Listed below are four elementary matrices and the operations that produce them.

\[
\begin{bmatrix}
1 & 0 \\
0 & -3
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 3 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

- Multiply the second row of \( I_2 \) by \(-3\).
- Interchange the second and fourth rows of \( I_4 \).
- Add 3 times the third row of \( I_3 \) to the first row.
- Multiply the first row of \( I_3 \) by 1.
Theorem 1.5.1 (Elementary Matrices and Row Operations)

Suppose that $E$ is an $m \times m$ elementary matrix produced by applying a particular elementary row operation to $I_m$, and that $A$ is an $m \times n$ matrix. Then $EA$ is the matrix that results from applying that same elementary row operation to $A$.

Remark:

When a matrix $A$ is multiplied on the left by an elementary matrix $E$, the effect is to perform an elementary row operation on $A$. 
Example (Using Elementary Matrices)

Consider the matrix

\[
A = \begin{bmatrix}
1 & 0 & 2 & 3 \\
2 & -1 & 3 & 6 \\
1 & 4 & 4 & 0
\end{bmatrix}
\]

and consider the elementary matrix

\[
E = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
3 & 0 & 1
\end{bmatrix}
\]

which results from adding 3 times the first row of \(I_3\) to the third row. The product \(EA\) is

\[
EA = \begin{bmatrix}
1 & 0 & 2 & 3 \\
2 & -1 & 3 & 6 \\
4 & 4 & 10 & 9
\end{bmatrix}
\]

which is precisely the same matrix that results when we add 3 times the first row of \(A\) to the third row.
Inverse Operations

If an elementary row operation is applied to an identity matrix \( I \) to produce an elementary matrix \( E \), then there is a second row operation that, when applied to \( E \), produces \( I \) back again
Example

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 0 \\
0 & 7
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]

Multiply the second row by 7
Multiply the second row by 1/7

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix} \rightarrow \begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]

Interchange the first and the second rows
Interchange the first and the second rows

\[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 5 \\
0 & 1
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]

Add 5 times the second row to the first
Add -5 times the second row to the first
Theorem 1.5.2 (Elementary Matrices and Nonsingularity)

- Each elementary matrix is nonsingular (is invertible), and its inverse is itself an elementary matrix. More precisely,
  - \( E_{ij}^{-1} = E_{ji} (= E_{ij}) \)
  - \( E_i(c)^{-1} = E_i(1/c) \) with \( c \neq 0 \)
  - \( E_{ij}(c)^{-1} = E_{ij}(-c) \) with \( i \neq j \)
Theorem 1.5.3 (Equivalent Statements)

If $A$ is an $n \times n$ matrix, then the following statements are equivalent, that is, all true or all false

(a) $A$ is invertible

(b) $Ax = 0$ has only the trivial solution

(c) The reduced row-echelon form of $A$ is $I_n$

(d) $A$ is expressible as a product of elementary matrices
Proof

(a) → (b)

- Assume $A$ is invertible and let $x_0$ be any solution of $Ax = 0$ thus $Ax_0 = 0$.
- Multiplying both sides of this equation by the matrix $A^{-1}$ gives $A^{-1}(Ax_0) = A^{-1} \cdot 0$, or $(A^{-1}A)x_0 = 0$, or $Ix_0 = 0$, or $x_0 = 0$.
- Thus, $Ax = 0$ has only the trivial solution.
Proof

(b) → (c)

Let $Ax = 0$ be the matrix form of the system

\[
\begin{align*}
  a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
  a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
  \vdots & \quad \vdots & \quad \vdots & \quad \vdots \\
  a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0
\end{align*}
\]

Assume that the system has only the trivial solution. If we solve by Gauss-Jordan elimination, the reduced row-echelon form of the augmented matrix will be

\[
\begin{align*}
  x_1 &= 0 \\
  x_2 &= 0 \\
  \vdots & \\
  x_n &= 0
\end{align*}
\]
Proof

(b) → (c)

- The augmented matrix

\[
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} & 0 \\
a_{21} & a_{22} & \cdots & a_{2n} & 0 \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn} & 0
\end{bmatrix}
\]

can be reduced to the augmented matrix

\[
\begin{bmatrix}
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & 0
\end{bmatrix}
\]

The reduced row-echelon form of A is \( I_n \)
Proof

(c) → (d)

- Assume that the reduced row-echelon form of $A$ is $I_n$, so that $A$ can be reduced to $I_n$ by a finite sequence of elementary row operations.

- By Theorem 1.5.1, each of these operations can be accomplished by multiplying on the left by an appropriate elementary matrix. Thus we can find elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k \cdots E_2 E_1 A = I_n$$
Proof

$E_k \cdots E_2 E_1 A = I_n$

(c) $\rightarrow$ (d)

- By Theorem 1.5.2, $E_1, E_2, \ldots, E_k$ are invertible. Multiplying both sides on the left successively by $E_k^{-1}, \ldots, E_2^{-1}, E_1^{-1}$ we obtain

\[
E_k^{-1} E_2^{-1} \cdots E_1^{-1} (E_k \cdots E_2 E_1 A) = E_k^{-1} E_2^{-1} \cdots E_1^{-1} I_n
\]

\[
A = E_k^{-1} E_2^{-1} \cdots E_1^{-1} I_n = E_k^{-1} E_2^{-1} \cdots E_1^{-1}
\]

- By Theorem 1.5.2, this equation expresses $A$ as a product of elementary matrices.
A Method for Inverting Matrices

\[ E_k \cdots E_2 E_1 A = I_n \]

- Multiplying on the right by \( A^{-1} \) yields

\[ E_k \cdots E_2 E_1 AA^{-1} = I_n A^{-1} \]

\[ A^{-1} = E_k \cdots E_2 E_1 I_n \]

- \( A^{-1} \) can be obtained by multiplying \( I_n \) successively on the left by the elementary matrices \( E_1, E_2, \ldots, E_k \).

- The sequence of row operations that reduces \( A \) to \( I_n \) will reduce \( I_n \) to \( A^{-1} \).
A Method for Inverting Matrices

To find the inverse of an invertible matrix $A$, we must find a sequence of elementary row operations that reduces $A$ to the identity and then perform this same sequence of operations on $I_n$ to obtain $A^{-1}$.

Remark

- Suppose we can find elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k \ldots E_2 E_1 A = I_n$$

then

$$A^{-1} = E_k \ldots E_2 E_1 I_n$$
Example (Using Row Operations to Find $A^{-1}$)

- Find the inverse of

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$$

- Solution:
  - To accomplish this we shall adjoin the identity matrix to the right side of $A$, thereby producing a matrix of the form $[A \mid I]$
  - We shall apply row operations to this matrix until the left side is reduced to $I$; these operations will convert the right side to $A^{-1}$, so that the final matrix will have the form $[I \mid A^{-1}]$
Example

The computations are as follows:

\[
\begin{bmatrix}
1 & 2 & 3 & | & 1 & 0 & 0 \\
2 & 5 & 3 & | & 0 & 1 & 0 \\
1 & 0 & 8 & | & 0 & 0 & 1 \\
\end{bmatrix}
\]

We added \(-2\) times the first row to the second and \(-1\) times the first row to the third.

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

We added 2 times the second row to the third.

\[
\begin{bmatrix}
1 & 2 & 3 & | & 1 & 0 & 0 \\
0 & 1 & -3 & | & -2 & 1 & 0 \\
0 & 0 & -1 & | & -5 & 2 & 1 \\
\end{bmatrix}
\]
Example

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

\[
\begin{bmatrix}
1 & 2 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
-14 & 6 & 3 \\
13 & -5 & -3 \\
5 & -2 & -1 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
-40 & 16 & 9 \\
13 & -5 & -3 \\
5 & -2 & -1 \\
\end{bmatrix}
\]

Thus,

\[A^{-1} = \begin{bmatrix}
-40 & 16 & 9 \\
13 & -5 & -3 \\
5 & -2 & -1 \\
\end{bmatrix}\]
Example

- **Not every matrix is invertible**

\[
\begin{bmatrix}
1 & 6 & 4 & | & 1 & 0 & 0 \\
2 & 4 & -1 & | & 0 & 1 & 0 \\
-1 & 2 & 5 & | & 0 & 0 & 1 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 6 & 4 & | & 1 & 0 & 0 \\
0 & -8 & -9 & | & -2 & 1 & 0 \\
0 & 8 & 9 & | & 1 & 0 & 1 \\
\end{bmatrix}
\]

We added -2 times the first row to the second and added the first row to the third.

\[
\begin{bmatrix}
1 & 6 & 4 & | & 1 & 0 & 0 \\
0 & -8 & -9 & | & -2 & 1 & 0 \\
0 & 0 & 0 & | & -1 & 1 & 1 \\
\end{bmatrix}
\]

We added the second row to the third.

Since we have obtained a row of zeros on the left side, \( A \) is not invertible.
Example

- Determine whether the given homogeneous system has nontrivial solutions

\[
\begin{align*}
    x_1 + 2x_2 + 3x_3 &= 0 \\
    2x_1 + 5x_2 + 3x_3 &= 0 \\
    x_1 + 8x_3 &= 0
\end{align*}
\]

\[
\begin{align*}
    x_1 + 6x_2 + 4x_3 &= 0 \\
    2x_1 + 4x_2 - x_3 &= 0 \\
    -x_1 + 2x_2 + 5x_3 &= 0
\end{align*}
\]

\[
\begin{bmatrix}
    1 & 2 & 3 \\
    2 & 5 & 3 \\
    1 & 0 & 8
\end{bmatrix}
\] is invertible, and the first system has only trivial solution

\[
\begin{bmatrix}
    1 & 6 & 4 \\
    2 & 4 & -1 \\
    -1 & 2 & 5
\end{bmatrix}
\] is not invertible, and the second system has nontrivial solutions
1.6

More on Linear Systems and Invertible Matrices
Theorems

- **Theorem 1.6.1**
  - Every system of linear equations has either no solutions, exactly one solution, or finitely many solutions.

- **Theorem 1.6.2**
  - If $A$ is an invertible $n \times n$ matrix, then for each $n \times 1$ matrix $b$, the system of equations $Ax = b$ has exactly one solution, namely, $x = A^{-1}b$. 
The proof will be complete if we can show that the system has infinitely many solutions if the system has more than one solution.

Assume that \(Ax = b\) has more than one solution, and let \(x_0 = x_1 - x_2\), where \(x_1\) and \(x_2\) are any two distinct solutions. Because \(x_1\) and \(x_2\) are distinct, \(x_0\) is nonzero.

\[Ax_0 = A(x_1 - x_2) = Ax_1 - Ax_2 = b - b = 0\]

\[A(x_1 + kx_0) = Ax_1 + k(Ax_0) = b + k0 = b + 0 = b.\] This says that \(x_1 + kx_0\) is a solution of \(Ax = b\).

Since \(x_0\) is nonzero and there are infinitely many choices for \(k\), the system \(Ax = b\) has infinitely many solutions.
Example

Consider the system of linear equations

\[ \begin{align*}
    x_1 + 2x_2 + 3x_3 &= 5 \\
    2x_1 + 5x_2 + 3x_3 &= 3 \\
    x_1 + 8x_3 &= 17
\end{align*} \]

In matrix form this system can be written as \( Ax = b \), where

\[
A = \begin{bmatrix}
    1 & 2 & 3 \\
    2 & 5 & 3 \\
    1 & 0 & 8
\end{bmatrix}, \quad
x = \begin{bmatrix}
    x_1 \\
    x_2 \\
    x_3
\end{bmatrix}, \quad
b = \begin{bmatrix}
    5 \\
    3 \\
    17
\end{bmatrix}
\]

In Example 4 of the preceding section we showed that \( A \) is invertible and

\[
A^{-1} = \begin{bmatrix}
    -40 & 16 & 9 \\
    13 & -5 & -3 \\
    5 & -2 & -1
\end{bmatrix}
\]

By Theorem 1.6.2 the solution of the system is

\[
x = A^{-1}b = \begin{bmatrix}
    -40 & 16 & 9 \\
    13 & -5 & -3 \\
    5 & -2 & -1
\end{bmatrix} \begin{bmatrix}
    5 \\
    3 \\
    17
\end{bmatrix} = \begin{bmatrix}
    1 \\
    -1 \\
    2
\end{bmatrix}
\]

or \( x_1 = 1, x_2 = -1, x_3 = 2. \)
Linear Systems with a Common Coefficient Matrix

- To solve a sequence of linear systems, \( Ax = b_1, \ Ax = b_2, \ldots, \ Ax = b_k \), with common coefficient matrix \( A \)

- If \( A \) is invertible, then the solutions \( x_1 = A^{-1}b_1, \ x_2 = A^{-1}b_2, \ldots, \ x_k = A^{-1}b_k \)

- A more efficient method is to form the matrix \([A | b_1 | b_2 | \ldots | b_k]\) by reducing it to reduced row-echelon form

- By reducing it to reduced row-echelon form we can solve all \( k \) systems at once by Gauss-Jordan elimination.
Example

- Solve the systems

\[
\begin{align*}
  x_1 + 2x_2 + 3x_3 &= 4 \\
  2x_1 + 5x_2 + 3x_3 &= 5 \\
  x_1 + 8x_3 &= 9
\end{align*}
\]

\[
\begin{bmatrix}
1 & 2 & 3 & 4 & 1 \\
2 & 5 & 3 & 5 & 6 \\
1 & 0 & 8 & 9 & -6
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & 0 & 1 & 2 \\
0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 1 & -1
\end{bmatrix}
\]

\[
\begin{align*}
  x_1 + 2x_2 + 3x_3 &= 1 \\
  2x_1 + 5x_2 + 3x_3 &= 6 \\
  x_1 + 8x_3 &= -6
\end{align*}
\]

\[
\begin{bmatrix}
1 & 0 & 0 & 1 & 2 \\
0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 1 & -1
\end{bmatrix}
\]

\[
\begin{align*}
  x_1 &= 1, \ x_2 = 0, \ x_3 = 1 \\
  x_1 &= 2, \ x_2 = 1, \ x_3 = -1
\end{align*}
\]
Theorems

- **Theorem 1.6.3**
  - Let $A$ be a square matrix
    - If $B$ is a square matrix satisfying $BA = I$, then $B = A^{-1}$
    - If $B$ is a square matrix satisfying $AB = I$, then $B = A^{-1}$
Proof of Theorem 1.6.3

- Assume that $BA = I$. If we can show that $A$ is invertible, the proof can be completed by multiplying $BA = I$ on both sides by $A^{-1}$ to obtain

$$BA A^{-1} = IA^{-1} \quad BI = IA^{-1} \quad B = A^{-1}$$

- To show that $A$ is invertible, it suffices to show that the system $A\mathbf{x} = 0$ has only the trivial solution.

- Let $\mathbf{x}_0$ be any solution of this system. If we multiply both sides of $A\mathbf{x}_0 = 0$ on the left by $B$, we obtain $BA\mathbf{x}_0 = B\mathbf{0}$ or $I\mathbf{x}_0 = 0$ or $\mathbf{x}_0 = 0$. Thus, the system $A\mathbf{x} = 0$ has only the trivial solution.
Theorem 1.6.4 (Equivalent Statements)

- If $A$ is an $n \times n$ matrix, then the following statements are equivalent
  - $A$ is invertible
  - $Ax = 0$ has only the trivial solution
  - The reduced row-echelon form of $A$ is $I_n$
  - $A$ is expressible as a product of elementary matrices
  - $Ax = b$ is consistent for every $n \times 1$ matrix $b$
  - $Ax = b$ has exactly one solution for every $n \times 1$ matrix $b$
Theorems

- **Theorem 1.6.5**
  - Let $A$ and $B$ be square matrices of the same size. If $AB$ is invertible, then $A$ and $B$ must also be invertible.

- **A fundamental problem**: Let $A$ be a fixed $m \times n$ matrix. Find all $m \times 1$ matrices $b$ such that the system of equations $Ax = b$ is consistent.
Fundamental Problem

- If $A$ is invertible, Theorem 1.6.2 says that $Ax=b$ has the unique solution.
- If $A$ is not square, or if $A$ is square but not invertible
  - The matrix $b$ must usually satisfy certain conditions in order for $Ax=b$ to be consistent
Example

What conditions must $b_1$, $b_2$, and $b_3$ satisfy in order for the systems of equations to be consistent?

\[
\begin{align*}
x_1 + 2x_2 + 3x_3 &= b_1 \\
x_1 + x_3 &= b_2 \\
2x_1 + x_2 + 3x_3 &= b_3
\end{align*}
\]

The augmented matrix is

\[
\begin{bmatrix}
1 & 1 & 2 & b_1 \\
1 & 0 & 1 & b_2 \\
2 & 1 & 3 & b_3
\end{bmatrix}
\]
Example

\[
\begin{bmatrix}
1 & 1 & 2 & b_1 \\
0 & -1 & -1 & b_2 - b_1 \\
0 & -1 & -1 & b_3 - 2b_1 \\
\end{bmatrix}
\]

-1 times the first row was added to the second and -2 times the first row was added to the third.

\[
\begin{bmatrix}
1 & 1 & 2 & b_1 \\
0 & 1 & 1 & b_1 - b_2 \\
0 & -1 & -1 & b_3 - 2b_1 \\
\end{bmatrix}
\]

The second row was multiplied by -1.

\[
\begin{bmatrix}
1 & 1 & 2 & b_1 \\
0 & 1 & 1 & b_1 - b_2 \\
0 & 0 & 0 & b_3 - b_2 - b_1 \\
\end{bmatrix}
\]

The second row was multiplied to the third.
Example

\[
\begin{bmatrix}
1 & 1 & 2 & b_1 \\
0 & 1 & 1 & b_1 - b_2 \\
0 & 0 & 0 & b_3 - b_2 - b_1
\end{bmatrix}
\]

- The system has a solution if and only if \( b_1, b_2, \) and \( b_3 \) satisfy the condition
  \[ b_3 - b_2 - b_1 = 0 \quad \text{or} \quad b_3 = b_2 + b_1 \]

- To express this condition another way, \( Ax = b \) is consistent if and only if \( b \) is a matrix of the form
  \[
  b = \begin{bmatrix}
  b_1 \\
  b_2 \\
  b_1 + b_2
  \end{bmatrix}
  \]
  where \( b_1 \) and \( b_2 \) are arbitrary.
1.7 Diagonal, Triangular, and Symmetric Matrices
Diagonal and Triangular

- A square matrix $A$ is $m \times n$ with $m = n$; the $(i,i)$-entries for $1 \leq i \leq m$ form the main diagonal of $A$.
- A diagonal matrix (對角矩陣) is a square matrix all of whose entries not on the main diagonal equal zero. By diag($d_1, \ldots, d_m$) is meant the $m \times m$ diagonal matrix whose $(i,i)$-entry equals $d_i$ for $1 \leq i \leq m$.
- A $n \times n$ lower-triangular matrix (下三角矩陣) $L$ satisfies $(L)_{ij} = 0$ if $i < j$, for $1 \leq i \leq m$ and $1 \leq j \leq n$.
- A $n \times n$ upper-triangular matrix (上三角矩陣) $U$ satisfies $(U)_{ij} = 0$ if $i > j$, for $1 \leq i \leq m$ and $1 \leq j \leq n$. 
Properties of Diagonal Matrices

- A general \( n \times n \) diagonal matrix \( D \) can be written as:

\[
D = \begin{bmatrix}
d_1 & 0 & \cdots & 0 \\
0 & d_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & d_n
\end{bmatrix}
\]

- A diagonal matrix is invertible if and only if all of its diagonal entries are nonzero:

\[
D^{-1} = \begin{bmatrix}
1/d_1 & 0 & \cdots & 0 \\
0 & 1/d_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1/d_n
\end{bmatrix}
\]

- Powers of diagonal matrices are easy to compute:

\[
D^k = \begin{bmatrix}
d_1^k & 0 & \cdots & 0 \\
0 & d_2^k & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & d_n^k
\end{bmatrix}
\]
Example

\[ A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \]

\[ A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \quad A^5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -243 & 0 \\ 0 & 0 & 32 \end{bmatrix} \quad A^{-5} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{243} & 0 \\ 0 & 0 & \frac{1}{32} \end{bmatrix} \]
Properties of Diagonal Matrices

- Matrix products that involve diagonal factors are especially easy to compute

\[
\begin{bmatrix}
d_1 & 0 & 0 \\
0 & d_2 & 0 \\
0 & 0 & d_3
\end{bmatrix}
\begin{bmatrix}
a_{11} & a_{12} & a_{13} & a_{14} \\
a_{21} & a_{22} & a_{23} & a_{24} \\
a_{31} & a_{32} & a_{33} & a_{34}
\end{bmatrix} =
\begin{bmatrix}
d_1a_{11} & d_1a_{12} & d_1a_{13} & d_1a_{14} \\
d_2a_{21} & d_2a_{22} & d_2a_{23} & d_2a_{24} \\
d_3a_{31} & d_3a_{32} & d_3a_{33} & d_3a_{34}
\end{bmatrix}
\]

\[
\begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33} \\
a_{41} & a_{42} & a_{43}
\end{bmatrix}
\begin{bmatrix}
d_1 & 0 & 0 \\
0 & d_2 & 0 \\
0 & 0 & d_3
\end{bmatrix} =
\begin{bmatrix}
d_1a_{11} & d_2a_{12} & d_3a_{13} \\
d_1a_{21} & d_2a_{22} & d_3a_{23} \\
d_1a_{31} & d_2a_{32} & d_3a_{33} \\
d_1a_{41} & d_2a_{42} & d_3a_{43}
\end{bmatrix}
\]

To multiply a matrix \(A\) on the left by a diagonal matrix \(D\), one can multiply successive rows of \(A\) by the successive diagonal entries of \(D\).
To multiply \(A\) on the right by \(D\), one can multiply successive columns of \(A\) by the successive diagonal entries of \(D\).
Theorem 1.7.1

- The transpose of a lower triangular matrix is upper triangular, and the transpose of an upper triangular matrix is lower triangular.
- The product of lower triangular matrices is lower triangular, and the product of upper triangular matrices is upper triangular.
- A triangular matrix is invertible if and only if its diagonal entries are all nonzero.
- The inverse of an invertible lower triangular matrix is lower triangular, and the inverse of an invertible upper triangular matrix is upper triangular.
Example

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

The matrix \( A \) is invertible, since its diagonal entries are nonzero, but the matrix \( B \) is not.

\[
A^{-1} = \begin{bmatrix}
1 & \frac{-3}{2} & \frac{7}{5} \\
0 & \frac{1}{2} & \frac{-2}{5} \\
0 & 0 & \frac{1}{5}
\end{bmatrix}
\]

The product \( AB \) is also upper triangular.

\[
AB = \begin{bmatrix}
3 & -2 & -2 \\
0 & 0 & 2 \\
0 & 0 & 5
\end{bmatrix}
\]
Proof

The product of lower triangular matrices is lower triangular.

- Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be lower triangular $n \times n$ matrices, and let $C = [c_{ij}]$ be the product $C = AB$.

- We can prove that $C$ is lower triangular by showing that $c_{ij} = 0$ for $i < j$.

\[
c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}
\]

- If we assume that $i < j$, then the terms can be grouped as

\[
c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{i(j-1)}b_{(j-1)j} + a_{ij}b_{jj} + \cdots + a_{in}b_{nj}
\]

The row number of $b$ is less than the column number of $b$

The row number of $a$ is less than the column number of $a$
Proof

The product of lower triangular matrices is lower triangular.

\[ c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{i(j-1)}b_{(j-1)j} + a_{ij}b_{jj} + \cdots + a_{in}b_{nj} \]

- The row number of \( b \) is less than the column number of \( b \)
- The row number of \( a \) is less than the column number of \( a \)

In the first grouping all of the \( b \) factors are zero since \( B \) is lower triangular. In the second grouping all of the \( a \) factors are zero since \( A \) is lower triangular. Thus, \( c_{ij}=0 \).
Symmetric Matrices

Definition
- A (square) matrix $A$ for which $A^T = A$, so that $\langle A\rangle_{ij} = \langle A\rangle_{ji}$ for all $i$ and $j$, is said to be symmetric.

Theorem 1.7.2
- If $A$ and $B$ are symmetric matrices (對稱矩陣) with the same size, and if $k$ is any scalar, then
  - $A^T$ is symmetric
  - $A + B$ and $A - B$ are symmetric
  - $kA$ is symmetric

Theorem 1.7.3
- The product of two symmetric matrices is symmetric if and only if the matrices commute (可交換), i.e., $AB = BA$
Example

- It is not true, in general, that the product of symmetric matrices is symmetric.

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

- If these two matrices commute, the product of two symmetric matrices is symmetric.

\[
\begin{bmatrix}
1 & 2 \\
2 & 3
\end{bmatrix}
\begin{bmatrix}
-4 & 3 \\
3 & -1
\end{bmatrix}
= 
\begin{bmatrix}
2 & 1 \\
1 & 3
\end{bmatrix}
\]

\[
\begin{bmatrix}
-4 & 3 \\
3 & -1
\end{bmatrix}
\begin{bmatrix}
1 & 2 \\
2 & 3
\end{bmatrix}
= 
\begin{bmatrix}
2 & 1 \\
1 & 3
\end{bmatrix}
\]
Theorems

- Theorem 1.7.4
  - If $A$ is an invertible symmetric matrix, then $A^{-1}$ is symmetric.

- Remark:
  - In general, a symmetric matrix needs not be invertible.
  - The products $AA^T$ and $A^TA$ are always symmetric

- Theorem 1.7.5
  - If $A$ is an invertible matrix, then $AA^T$ and $A^TA$ are also invertible
Proof

If $A$ is an invertible symmetric matrix, then $A^{-1}$ is symmetric.

- Assume that $A$ is symmetric and invertible. From Theorem 1.4.9 and the fact that $A = A^T$, we have

$$\left( A^{-1} \right)^T = \left( A^T \right)^{-1} = A^{-1}$$

which proves that $A^{-1}$ is symmetric.

The products $AA^T$ and $A^TA$ are always symmetric

- $(AA^T)^T = (A^T)^T A^T = AA^T$
- $(A^TA)^T = A^T (A^T)^T = A^TA$

Theorem 1.4.9

$(A^{-1})^T = (A^T)^{-1}$
Example

Let $A$ be the $2 \times 3$ matrix

$$A = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix}$$

Then

$$A^T A = \begin{bmatrix} 1 & -2 \\ -2 & 0 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 10 & -2 & -11 \\ -2 & 4 & -8 \\ -11 & -8 & 41 \end{bmatrix}$$

$$A A^T = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ 4 & -5 \end{bmatrix} = \begin{bmatrix} 21 & -17 \\ -17 & 34 \end{bmatrix}$$

Observe that $A^T A$ and $A A^T$ are symmetric as expected.
Proof

If $A$ is an invertible matrix, then $AA^T$ and $A^TA$ are also invertible

- Since $A$ is invertible, so is $A^T$ by Theorem 1.4.9.
- Thus $AA^T$ and $A^TA$ are invertible, since they are the products of invertible matrices.

Theorem 1.4.9

$$(A^{-1})^T = (A^T)^{-1}$$