# 4 Elementary matrices and a method for finding inverse

**Definition 4.1.** A matrix is called an **elementary** matrix if it can be obtained from the identity matrix by performing a single elementary row operation.

**Example 4.1.** Specify elementary matrices that are obtained by applying the following elementary row operation on the identity matrix.

(a) Multiply the second row of $I_2$ by $-3$

(b) Interchange the second and the fourth rows of $I_4$.

(c) Add 3 time the third row of $I_3$ to the first row.

**Theorem 4.1** (Row Operations by Matrix Multiplication)**.** If the elementary matrix $\mathbf{E}$ results from performing a certain row operation on $\mathbf{I}_m$ and if $\mathbf{A}$ is an $m \times n$ matrix, then the product $\mathbf{EA}$ is the matrix that results when this same row operation is performed on $\mathbf{A}$.

**Example 4.2.** Consider the matrix $\mathbf{A}$ and the elementary matrix $\mathbf{E}$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 1 & 4 & 4 & 0 \end{bmatrix}, \qquad \mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10 & \\ 3 & 0 & 1 \end{bmatrix}.$$

Note that the matrix $\mathbf{E}$ results from adding 3 times the first row of $\mathbf{I}_3$ to the third row.
The product

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

is precisely the same matrix that results when we add 3 times the first row of $\mathbf{A}$ to the third row.

**Remark:**

- If an elementary row operation is applied to an identity matrix $\mathbf{I}$ to produce an elementary matrix $\mathbf{E}$, then there is a second row operation that, when applied to $\mathbf{E}$, say $\hat{E}$, produces $\mathbf{I}$ back again.
- The operation that produce $\hat{\mathbf{E}}$ are called the **inverse operations** of the corresponding operation that produce $\mathbf{E}$. I.e.

$$\hat{\mathbf{E}}\mathbf{E}\mathbf{I} = \mathbf{I}.$$

Note:we can show that $\mathbf{E}\hat{\mathbf{E}}\mathbf{I} = \mathbf{I}$ and therefore

$$\hat{\mathbf{E}} = \mathbf{E}^{-1}.$$

**Theorem 4.2.** Every elementary matrix is **invertible**, and the inverse is also an elementary matrix.

**Proof:** If $\mathbf{E}$ is an elementary matrix, then $\mathbf{E}$ results from performing some row operation on $\mathbf{I}$. Let $\hat{\mathbf{E}}$ be the matrix that results when the inverse of this operation is performed on $\mathbf{I}$. Applying the previous theorem and using the fact that inverse row operations cancel the effect of each other, it follows that

$$\hat{\mathbf{E}}\mathbf{E} = \mathbf{I} \quad \text{and} \quad \mathbf{E}\hat{\mathbf{E}} = \mathbf{I}$$

Thus, the elementary matrix $\hat{\mathbf{E}}$ is the inverse of $\mathbf{E}$. ■

The following theorem provides relationships among:
- invertibility,
-homogeneous linear systems,
- reduced row-echelon forms,
- elementary matrices.

**Theorem 4.3** (Equivalent Statements)**.** If $\mathbf{A}$ is an $n \times n$ matrix, then the following statements are equivalent, that is, all true or all false.

(a) $\mathbf{A}$ is invertible.

(b) $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

(c) The reduced row-echelon form of $\mathbf{A}$ is $\mathbf{I}_n$.

(d) $\mathbf{A}$ is expressible as a product of elementary matrices.

**Proof:** We will prove the equivalence by establishing the chain of implications:
$(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a)$.

Proof (Cont'):

### 4.1 A Method for Finding Inverses of Matrices by Elementary Row Operations

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be an invertible matrix. From the previous theorem (in the proof), we have

$$\mathbf{E}_k....\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}_n$$

where $\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_k$ are elementary matrices. Hence,

$$\mathbf{A}^{-1} = \mathbf{E}_k....\mathbf{E}_2\mathbf{E}_1\mathbf{I}_n.$$

That is, **the sequence of row operations that reduces $\mathbf{A}$ to $\mathbf{I}_n$ will reduce $\mathbf{I}_n$ to $\mathbf{A}^{-1}$ .**

**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 to obtain .**

**Example 4.3.** Find the inverse of $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$ if $\mathbf{A}$ is invertible.

**<u>Remark:</u> Not invertible matrix**

- If an $n \times n$ matrix $\mathbf{A}$ is not invertible, then it cannot be reduced to by elementary row operations [part (c) of Theorem 4.3].
- I.e.,the reduced row-echelon form of $\mathbf{A}$ has at **least one row of zeros** when $\mathbf{A}$ is not invertible.
- Thus, if the procedure in the last example is attempted on a matrix that is not invertible, then at some point in the computations a row of zeros will occur on the **left side** of the augmented matrix.

**Example 4.4.** Find the inverse of $\mathbf{A} = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{bmatrix}$ if $\mathbf{A}$ is invertible.

# 5 Invertible Matrix and Linear Systems

**Theorem 5.1.** Every system of linear equations has no solutions, or has exactly one solution, or has infinitely many solutions.

**Proof:**

**Theorem 5.2. Solution of a Linear System Using $\mathbf{A}^{-1}$**
If $\mathbf{A}$ is an invertible matrix, then for each matrix $\mathbf{b}$, the system of equations has exactly one solution, namely, $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

**Proof:**

## 5.1 Properties of Invertible Matrices

Recall that when we have to show that $\mathbf{B}$ is an inverse of an $n \times n$ matrix $\mathbf{A}$ by using definition, we have to show that

$$\mathbf{AB} = \mathbf{I} \qquad \text{and} \qquad \mathbf{BA} = \mathbf{I}$$

The next theorem shows that if we produce an matrix $\mathbf{B}$ satisfying **either condition** above, then the other condition holds automatically.

**Theorem 5.3.** Let A be a square matrix.

(a) If $\mathbf{B}$ is a square matrix satisfying $\mathbf{BA} = \mathbf{I}$ , then $\mathbf{B} = \mathbf{A}^{-1}$.

(b) If $\mathbf{B}$ is a square matrix satisfying $\mathbf{AB} = \mathbf{I}$ , then $\mathbf{B} = \mathbf{A}^{-1}$.

**Proof:**

**Theorem 5.4. Equivalent Statements**
If $\mathbf{A}$ is an $n \times n$ matrix, then the following are equivalent.

(a) $\mathbf{A}$ is invertible.

(b) $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

(c) The reduced row-echelon form of $\mathbf{A}$ is $\mathbf{I}_n$.

(d) $\mathbf{A}$ is expressible as a product of elementary matrices.

(e) $\mathbf{Ax} = \mathbf{b}$ is consistent for every matrix $\mathbf{b} \in \mathbb{R}^n$.

(f) $\mathbf{Ax} = \mathbf{b}$ has exactly one solution for every matrix $\mathbf{b}$.

**Proof:**

### 5.2 General Linear systems $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{A} \in \mathbb{R}^{m\times n}$

- When $\mathbf{A} \in \mathbb{R}^{n\times n}$ is square and invertible, from Theorem 5.2, there is a unique solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ for any $\mathbf{b} \in \mathbb{R}^n$.
- If is **not square**, or if A is square but **not invertible**, then Theorem 5.2 does not apply. In these cases the matrix $\mathbf{b}$ must usually satisfy certain conditions in order for $\mathbf{Ax} = \mathbf{b}$ to be consistent.

  The following example illustrates how the elimination methods can be used to determine such conditions.

**Example 5.1.** Consider the linear systems:

$$\begin{aligned} x_1 + x_2 + 2x_3 &= b_1 \\ x_1 + x_3 &= b_2 \\ 2x_1 + x_2 + 3x_3 &= b_3 \end{aligned}$$

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

# 6 Matrices with special forms

In this section we shall consider certain classes of matrices that have special forms, which are listed below.

- Diagonal matrices
- Triangular marices.
- Symmetric matrices.

These matrices are among the most important kinds of matrices encountered in linear algebra.

## 6.1 Diagonal Matrix

**Definition 6.1.** A square matrix in which all the entries off the main diagonal are zero is called a diagonal matrix.

A general diagonal matrix **A** can be written as

$$\mathbf{D} = \begin{bmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & d_{nn} \end{bmatrix}$$

**Properties of Diagonal Matrix**

- A diagonal matrix **A** is invertible if and only if all of its diagonal entries are nonzero; in this case the inverse of **D** is

$$\mathbf{D}^{-1} = \begin{bmatrix} 1/d_{11} & 0 & \dots & 0 \\ 0 & 1/d_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & 1/d_{nn} \end{bmatrix}$$

  Note: We can verify that $\mathbf{D}\mathbf{D}^{-1} = \mathbf{D}^{-1}\mathbf{D} = \mathbf{I}$

- Powers of diagonal matrices are easy to compute. I.e. if **D** is the diagonal matrix and $k$ is a positive integer, then it can be shown that (exercise)

$$\mathbf{D}^k = \begin{bmatrix} d_{11}^k & 0 & \dots & 0 \\ 0 & d_{22}^k & \dots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & d_{nn}^k \end{bmatrix}.$$

- Left and right multiplication:
  - 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** (scale each row $i$ of **A** by $d_{ii}$).
  - To multiply **A** on the right by **D**, one can multiply successive columns of A by the successive diagonal entries of **D** (scale each column $i$ of **A** by $d_{ii}$).

### 6.2 Triangular matrices

**Definition 6.2.** Upper and lower triangular matrices.

- A square matrix in which all the entries above the main diagonal are zero is called **lower triangular**.

$$\begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}$$

- a square matrix in which all the entries below the main diagonal are zero is called **upper triangular**.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

A matrix that is either upper triangular or lower triangular is called **triangular**.

**Remarks:**

- Diagonal matrices are both upper triangular and lower triangular since they have zeros below and above the main diagonal.
- A square matrix in row-echelon form is upper triangular since it has zeros below the main diagonal.

The following theorem provides some properties of triangular matrices.

**Theorem 6.1. Properties of triangular matrices**

(a) The transpose of a lower triangular matrix is upper triangular, and the transpose of an upper triangular matrix is lower triangular.

(b) The product of lower triangular matrices is lower triangular, and the product of upper triangular matrices is upper triangular.

(c) A triangular matrix is invertible if and only if its diagonal entries are all nonzero.

(d) The inverse of an invertible lower triangular matrix is lower triangular, and the inverse of an invertible upper triangular matrix is upper triangular.

### 6.3 Symmetric Matrices

**Definition 6.3.** A square matrix $\mathbf{A}$ is called **symmetric** if $\mathbf{A} = \mathbf{A}^T$.

Note: Expressed in terms of the individual entries, an $n \times n$ matrix $\mathbf{A} = [a_{ij}]$ is symmetric if and only if $a_{ij} = a_{ji}$ for all $i, j = 1, 2, ..., n$.

The following theorem lists the main algebraic properties of symmetric matrices.

**Theorem 6.2.** If $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices with the same size, and if $k$ is any scalar, then:

(a) $\mathbf{A}^T$ is symmetric.

(b) $\mathbf{A} + \mathbf{B}$ and $\mathbf{A} - \mathbf{B}$ are symmetric.

(c) $k\mathbf{A}$is symmetric.

**Proof:** (Exercise)

**Example 6.1.** Prove or disprove that "**the product of symmetric matrices is symmetric.** "

**Example 6.2.** Prove or disprove the following statements.

(a) If $\mathbf{A}$ is a symmetric matrix, then $\mathbf{A}$ is invertible.

(b) If $\mathbf{A}$ is an invertible symmetric matrix, then $\mathbf{A}^{-1}$ is symmetric.

(c) If $\mathbf{A}$ is an $m \times n$ matrix, then $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$ are always symmetric.