Chapter 1 Matrices and System of Equations

Section 1.5 Elementary Matrices

Definition (Elementary Matrix)

Perform exactly one elementary row operation on the identity matrix I, the resulting matrix is called an **elementary matrix**.

Three types of elementary matrices corresponding to the three types of elementary row operations.

(I) $R_i \leftrightarrow R_j$

(II) $cR_i \rightarrow R_i, c \neq 0$

(III) $cR_j + R_i \rightarrow R_i$

Type I elementary matrix:
A matrix obtained by interchanging two rows of I.

Example Exchanging the i^{th} row and j^{th} row of I_n, we obtain a Type I elementary matrix

$$E = \begin{bmatrix} I_{i-1} & & & & \\ & 0 & & 1 & \\ & & I_{j-i-1} & & \\ & 1 & & 0 & \\ & & & & I_{n-j} \end{bmatrix}_{n \times n}$$

Type II elementary matrix:
A matrix obtained by multiplying a row of I by a nonzero constant, α say.

Example Multiplying the i^{th} row of I_n by a nonzero real number α, we have a type II elementary matrix:

$$E = \operatorname{diag}(1, \ldots, 1, \alpha, 1, \ldots, 1)$$

Type III elementary matrix:
A matrix obtained from I by adding a multiple of one row to another row.

Example Replacing R_i of A by $\alpha R_j + R_i$, we have a Type III elementary matrix

$$E = \begin{bmatrix} I_{i-1} & & & & \\ & 1 & & \alpha & \\ & & I_{j-i-1} & & \\ & & & 1 & \\ & & & & I_{n-j} \end{bmatrix}$$

Theorem Let E be an elementary matrix of size $n \times n$.

1. For any $m \times n$ matrix A, EA is the matrix obtained when the same row operation is performed on A.
2. For any $n \times r$ matrix B, BE is the matrix obtained when the same column operation is performed on B.

Example (Type I)

$$EA = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ a_{31} & a_{32} & \cdots & a_{3n} \end{pmatrix}$$

$$BE = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{pmatrix} b_{12} & b_{11} & b_{13} \\ b_{22} & b_{21} & b_{23} \\ b_{32} & b_{31} & b_{33} \end{pmatrix}.$$

Example (Type II)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \alpha a_{21} & \alpha a_{22} & \alpha a_{23} & \alpha a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}.$$

Example (Type III)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} - 2a_{31} & a_{22} - 2a_{32} & a_{23} - 2a_{33} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Left multiply E onto A = The same Elementary Row Operation on A.

Right multiply E onto A = The same Elementary Column Operation on A.

Extra exercises

Given that $A = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 0 & 1 \\ 0 & -3 & 1 \end{pmatrix}$, and elementary matrices
$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix}$, $E_3 = \begin{pmatrix} 1 & 0 & \beta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Find

1. E_1A and AE_1.
2. E_2A and AE_2.
3. E_1E_3A and E_3E_2A.

Theorem If E is an elementary matrix, then E is nonsingular and E^{-1} is an elementary matrix of the same type.

Proof The inverse of an elementary matrix is constructed by doing the reverse row operation on I. E^{-1} will be obtained by performing the row operation which would carry E back to I.
If E is obtained by switching rows i and j, then E^{-1} is also obtained by switching rows i and j.
If E is obtained by multiplying row i by the scalar α, then E^{-1} is obtained by multiplying row i by the scalar $1/\alpha$.
If E is obtained by adding α times row i to row j, then E^{-1} is obtained by adding $-\alpha$ times row i from row j.

Definition (Row equivalent) A matrix B is **row equivalent** to a matrix A if there exists a finite sequence $E_1, E_2, \ldots, E_k$ of elementary matrices such that

$$B = E_k E_{k-1} \cdots E_1 A.$$

In other words, B is row equivalent to A if B can be obtained from A by a finite number of row operations.

In particular, if two augmented matrices $(A|\mathbf{b})$ and $(B|\mathbf{c})$ are row equivalent, then $A\mathbf{x} = \mathbf{b}$ and $B\mathbf{x} = \mathbf{c}$ are equivalent systems.

Property of row equivalent matrices
I. If A is row equivalent to B, then B is row equivalent to A.
II. If A is row equivalent to B, and B is row equivalent to C, then A is row equivalent to C.

Proof Exercise

Theorem (Equivalent Conditions for Nonsingularity)

Let A be an $n \times n$ matrix. The following are equivalent:

(a) A is nonsingular;

(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{0}$;

(c) A is row equivalent to I.
(Or simply, $\mathrm{rref}(A) = I$. A can be written as a product of elementary matrices.)

(d) ......more in Chap 2-6

A is nonsingular $\quad \rightarrow \quad$ $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{0}$

Proof $(a) \Rightarrow (b)$ If A is nonsingular and $\mathbf{x}_0$ is a solution of $A\mathbf{x} = \mathbf{0}$, then

$$\mathbf{x}_0 = I\mathbf{x}_0 = (A^{-1}A)\mathbf{x}_0 = A^{-1}(A\mathbf{x}_0) = A^{-1}\mathbf{0} = \mathbf{0}.$$

Thus, $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

$A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\quad \rightarrow \quad \mathrm{rref}(A) = I$

Proof $(b) \Rightarrow (c)$ If we use elementary row operations, the system can be transformed into the form $U\mathbf{x} = \mathbf{0}$, where U is in row echelon form. If one of the diagonal elements of U were 0, the last row of U would consist entirely of 0's. But then $A\mathbf{x} = \mathbf{0}$ would be equivalent to a system with more unknowns than equations and, hence, there would have a nontrivial solution. Thus, U must be a strictly triangular matrix with diagonal elements all equal to 1. Hence, I is the reduced row echelon form of A and A is row equivalent to I.

$\text{rref}(A) = I \quad \rightarrow \quad A$ is nonsingular

Proof $(c) \Rightarrow (a)$ If A is row equivalent to I_n, there exist elementary matrices $E_1, E_2, \cdots, E_k$ such that

$$A = E_k E_{k-1} \cdots E_1 I_n = E_k E_{k-1} \cdots E_1$$

Since E_i is invertible, $i = 1, \cdots, k$, the product $E_k E_{k-1} \cdots E_1$ is also invertible. Hence, A is nonsingular and

$$A^{-1} = (E_k E_{k-1} \cdots E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}.$$

$\square$

Corollary The system $A\mathbf{x} = \mathbf{b}$ of n linear equations in n unknowns has a unique solution **if and only if** A is nonsingular.

Proof If A is nonsingular and $\mathbf{x}_0$ is any solution of $A\mathbf{x} = \mathbf{b}$, then $A\mathbf{x}_0 = \mathbf{b}$. Multiplying both sides of this equation by A^{-1}, we must have $\mathbf{x}_0 = A^{-1}\mathbf{b}$.

Conversely, if $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x}_0$, then we claim that A cannot be singular. Indeed, if A were singular, then the equation $A\mathbf{x} = \mathbf{0}$ would have a solution $\mathbf{z} \neq \mathbf{0}$. But this would imply that $\mathbf{y} = \mathbf{x}_0 + \mathbf{z}$ is a second solution of $A\mathbf{x} = \mathbf{b}$, since

$$A\mathbf{y} = A(\mathbf{x}_0 + \mathbf{z}) = A\mathbf{x}_0 + A\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

Therefore, if $A\mathbf{x} = \mathbf{b}$ has a unique solution, then A must be nonsingular.

Steps to compute the inverse of an $n \times n$ matrix A

1. Form the $n \times 2n$ matrix $[A|I_n]$
2. Compute rref$(A|In)$
3. If rref$(A|I_n) = (I_n|C)$, then $A^{-1} = C$. Otherwise, A is singular.

Why it work? If A is nonsingular, then A is row equivalent to I and hence there exist elementary matrices $E_1, \cdots, E_k$ such that

$$
\begin{aligned}
(A|I_n) &\to E_1(A|I_n) = (E_1A|E_1) \\
&\to E_2(E_1A|E_1) = (E_2E_1A|E_2E_1) \\
&\vdots \\
&\to E_k(E_{k-1}\cdots E_1A|E_{k-1}\cdots E_1) = (E_kE_{k-1}\cdots E_1A|E_kE_{k-1}\cdots E_1)
\end{aligned}
$$

If $(E_kE_{k-1}\cdots E_1A|E_kE_{k-1}\cdots E_1) = (I_n|C)$, then $E_kE_{k-1}\cdots E_1A = I_n$ and $E_kE_{k-1}\cdots E_1 = C$, giving $A^{-1} = E_kE_{k-1}\cdots E_1 = C$.

Example Let $A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix}$. Find inverse of A.

Solution

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[-r_1+r_3\to r_3]{-2r_1+r_2\to r_2} \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & -6 & -3 & -2 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{-1/6r_2+r_3\to r_3} \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & -6 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1/2 & -2/3 & -1/6 & 1 \end{array}\right)$$

$$\xrightarrow[-2r_3\to r_3]{-1/6r_2\to r_2}\left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1/2 & 1/3 & -1/6 & 0 \\ 0 & 0 & 1 & 4/3 & 1/3 & -2 \end{array}\right)\xrightarrow[-(1/2)r_3+r_2\to r_2]{-r_3+r_1\to r_1}$$

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 0 & -1/3 & -1/3 & 2 \\ 0 & 1 & 0 & -1/3 & -1/3 & 1 \\ 0 & 0 & 1 & 4/3 & 1/3 & -2 \end{array}\right)\xrightarrow{-3r_2+r_1\to r_1}\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2/3 & 2/3 & -1 \\ 0 & 1 & 0 & -1/3 & -1/3 & 1 \\ 0 & 0 & 1 & 4/3 & 1/3 & -2 \end{array}\right)$$

$$A^{-1}=\begin{pmatrix} 2/3 & 2/3 & -1 \\ -1/3 & -1/3 & 1 \\ 4/3 & 1/3 & -2 \end{pmatrix}$$

Definition (Triangular matrices)

1. A square matrix $U = (u_{ij})$ is *upper triangular* if $u_{ij} = 0$ for $i > j$.
2. A square matrix $L = (l_{ij})$ is *lower triangular* if $l_{ij} = 0$ for $i < j$.
3. A matrix is *triangular* if it is either upper triangular or lower triangular.
4. A matrix is *unit lower (upper respectively) triangular* if it is a lower (upper respectively) triangular matrix with 1's on the diagonal.

Example

$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & 6 \end{pmatrix}$ is upper triangular. $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{pmatrix}$ is lower triangular.

$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}$ is unit upper triangular. $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 5 & 1 \end{pmatrix}$ is unit lower triangular.

Definition (LU factorization) The factorization of the matrix A into a product of a unit lower triangular matrix L times an upper triangular matrix U, i.e.,

$$A_{n\times n} = LU = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \star & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \star & \cdots & \star & 1 \end{bmatrix} \begin{bmatrix} \blacktriangle & \blacktriangle & \cdots & \blacktriangle \\ 0 & \blacktriangle & \ddots & \vdots \\ \vdots & \ddots & \ddots & \blacktriangle \\ 0 & \cdots & 0 & \blacktriangle \end{bmatrix}.$$

How to find L and U?

A square matrix A is row equivalent to an upper triangular matrix U using only elementary matrix of type III that add a multiple of one row to another row **below** it. Thus, there exist a sequence of unit lower triangular elementary matrices $E_1, \cdots, E_k$ s.t.

$$E_k \cdots E_1 A = U,$$

and $L = (E_k \cdots E_1)^{-1} = E_1^{-1} \cdots E_k^{-1}$ is a unit lower triangular matrix.

LU factorization in solving a linear system:

The system $A\mathbf{x} = \mathbf{b}$ becomes $LU\mathbf{x} = \mathbf{b}$. Therefore, we have

$$L(U\mathbf{x}) = \mathbf{b}.$$

Let $\mathbf{y} = U\mathbf{x}$, we can find $\mathbf{x}$ by solving the following two systems of equations

$$L\mathbf{y} = \mathbf{b}, \quad U\mathbf{x} = \mathbf{y}.$$

Example

$$\begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix} \xrightarrow[-1r_1+r_3]{-2r_1+r_2} \begin{pmatrix} 1 & 3 & 1 \\ 0 & -6 & -3 \\ 0 & -1 & -1 \end{pmatrix} \xrightarrow{\frac{-1}{6}r_2+r_3} \begin{pmatrix} 1 & 3 & 1 \\ 0 & -6 & -3 \\ 0 & 0 & \frac{-1}{2} \end{pmatrix}$$

$$\text{Let } E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{6} & 1 \end{pmatrix}$$

Then

$$E_3E_2E_1A = U$$

Since $E_3 E_2 E_1 A = U$, we have $A = E_1^{-1} E_2^{-1} E_3^{-1} U$.
Take $L = E_1^{-1} E_2^{-1} E_3^{-1}$.

$$
\begin{aligned}
L =& E_1^{-1} E_2^{-1} E_3^{-1} \\
=& \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{6} & 1 \end{pmatrix}^{-1} \\
=& \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{6} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & \frac{1}{6} & 1 \end{pmatrix}
\end{aligned}
$$

Remark This method does not work if A cannot be reduced into an upper triangular matrix using III row operation.