Slide 6–Matrix Inverse

MAT2040 Linear Algebra

Example 6.1 Consider the following linear system

$$\begin{aligned} -7x_1 - 6x_2 - 12x_3 &= -33, \\ 5x_1 + 5x_2 + 7x_3 &= 24, \\ x_1 + 4x_3 &= 5. \end{aligned}$$

It can be represented as

$$A\mathbf{x} = \mathbf{b},$$

where

$$A = \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -33 \\ 24 \\ 5 \end{bmatrix}$$

Now define

$$B = \begin{bmatrix} -10 & -12 & -9 \\ \frac{13}{2} & 8 & \frac{11}{2} \\ \frac{5}{2} & 3 & \frac{5}{2} \end{bmatrix}$$

One can check that

$$BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

Apply this to solve the equation

$$\mathbf{x} = I_3\mathbf{x} = BA\mathbf{x} = B\mathbf{b} = \begin{bmatrix} -10 & -12 & -9 \\ \frac{13}{2} & 8 & \frac{11}{2} \\ \frac{5}{2} & 3 & \frac{5}{2} \end{bmatrix} \begin{bmatrix} -33 \\ 24 \\ 5 \end{bmatrix} = \begin{bmatrix} -3 \\ 5 \\ 2 \end{bmatrix}$$

Definition 6.2 (Invertible Matrix) Let $A, B \in \mathbb{R}^{n \times n}$ be such that

$$AB = BA = I_n$$

then A is invertible and B is the inverse of A. We shall write $B = A^{-1}$.

Remark. For a linear system $A\mathbf{x} = \mathbf{b}$ with n equations and n variables, then A is a square matrix with size $n \times n$. If A is invertible, then the system has a unique solution, which is given by $\mathbf{x} = A^{-1}\mathbf{b}$. This is due to $\mathbf{x} = I\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$..

Question:

For any $A \in \mathbb{R}^{n \times n}$, does there exist B, s.t. $BA = AB = I_n$?

Answer: Not all matrices satisfy this condition, e.g. $A = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$.

Example 6.3

(1) Let

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$$

then

$$\begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus,

$$A^{-1} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}$$

(2) Let

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

then

$$\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus,

$$A^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

(3) Let

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

Suppose

$$B = \begin{bmatrix} w & x \\ y & z \end{bmatrix} = A^{-1},$$

then

$$\begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

But

$$\begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 2w+y & 2x+z \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

In this case,

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix} \text{is not invertible.}$$

Remark

(1) The diagonal matrix $A = diag(a_{11}, a_{22}, \cdots, a_{nn})$ is invertible if and only if all $a_{ii} \neq 0, i = 1, \cdots, n$, and $A^{-1} = diag(a_{11}^{-1}, a_{22}^{-1}, \cdots, a_{nn}^{-1})$.

(2) Not all square matrices are invertible.

(3) **Invertible matrices** are sometimes called **nonsingular** or **nondegenerate matrices**. On the other hand, square matrices that are **not invertible** are also called **singular** or **degenerate**.

Theorem 6.4 (**Matrix inverse is unique**) Suppose the square matrix A has an inverse. Then A^{-1} is unique.

Proof. Let B, C be the inverse of A, thus, $AB = BA = I, AC = CA = I$. Then $B = BI = BAC = IC = C$

Property 6.5 (**Matrix Inverse of a Matrix Transpose**) Suppose A is an invertible matrix. Then A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$.
Proof. $A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$ and
$(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$.

Property 6.6 (**Matrix Inverse of a Scalar Multiple**) Suppose A is an invertible matrix and α is a nonzero scalar. Then αA is invertible and $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$.
Proof. $(\alpha A)(\frac{1}{\alpha}A^{-1}) = AA^{-1} = I$, and $(\frac{1}{\alpha}A^{-1})(\alpha A) = A^{-1}A = I$.

Property 6.7 (**Matrix Inverse of a Matrix Inverse**) Suppose A is an invertible matrix. Then A^{-1} is invertible and $(A^{-1})^{-1} = A$.

Proof. By definition, $AA^{-1} = A^{-1}A = I$, thus $(A^{-1})^{-1} = A$.

Theorem 6.8 (**Matrix Inverse of Matrices Product**) Suppose A and B are invertible matrices of size n, then AB is invertible and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof. By assumption, A^{-1} and B^{-1} exist. Thus,

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

$$B^{-1}A^{-1}(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I.$$

Hence, $(AB)^{-1} = B^{-1}A^{-1}$.

Remark. $(A+B)^{-1} \neq A^{-1} + B^{-1}$, can you find an counterexample?

Elementary Matrix and its inverse

Definition 6.9 (**Elementary Matrices**) If we start with the identity matrix, and perform exactly one type of elementary row operations, then the resulting matrix is called elementary matrix.

(1) The elementary matrix corresponding to elementary row operation 1 $(R_i \leftrightarrow R_j)$ is (elementary matrix type I)

$$
E_{R_iR_j} = \begin{array}{cc}
\left[\begin{array}{ccc|c|ccc|c|ccc}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
\hline
& & & 0 & & & & 1 & & & \\
\hline
& & & & 1 & & & & & & \\
& & & & & \ddots & & & & & \\
& & & & & & 1 & & & & \\
\hline
& & & 1 & & & & 0 & & & \\
\hline
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{array}\right]
&
\begin{array}{l}
\\ \\ \\ \textit{ith row} \\ \\ \\ \\ \textit{jth row} \\ \\ \\ \\
\end{array}
\\
\begin{array}{c} i\text{th column} \qquad j\text{th column} \end{array} &
\end{array}
$$

Suppose $E_{R_iR_j} \in \mathbb{R}^{m\times m}$, $A \in \mathbb{R}^{m\times n}$, the result of $E_{R_iR_j}A$ is just to exchange the ith row and jth row of matrix A, $E_{R_iR_j}$ is also called the **row exchange matrix**.

Example

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} y & z \\ w & x \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} \\ a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix}$$

Actually, $E_{R_iR_j}$ is a **permutation matrix** (will be defined later on).

(2) The elementary matrix corresponding to elementary row operation 2 ($R_i \to \alpha R_i (\alpha \neq 0)$) is (elementary matrix type II)

$$E_{\alpha R_i} = \begin{array}{cc} \left[\begin{array}{ccc|c|ccc} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ \hline & & & \alpha & & & \\ \hline & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{array}\right] & \begin{array}{c} \\ \\ \\ ith\ row \\ \\ \\ \\ \end{array} \\ i\text{th column} & \end{array}$$

Suppose $E_{\alpha R_i} \in \mathbb{R}^{m\times m}(\alpha \neq 0)$, the result of $E_{\alpha R_i}A$ is just to multiply each element of ith row of matrix A by α.

Example

$$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} w & x \\ 3y & 3z \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ 2a_{31} & 2a_{32} \end{bmatrix}$$

(3) The elementary matrix corresponding to elementary row operation 3 ($R_j \to \beta R_i + R_j$) is (elementary matrix type III)

$$E_{\beta R_i+R_j} = \begin{array}{cl} \left[\begin{array}{ccc|c|ccc|c|ccc} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ \hline &&& 1 &&&&&&& \\ \hline &&&& 1 &&&&&& \\ &&&&& \ddots &&&&& \\ &&&&&& 1 &&&& \\ \hline &&& \beta &&&& 1 &&& \\ \hline &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{array}\right] & \begin{array}{l} \\ \\ \\ i\text{th row} \\ \\ \\ \\ j\text{th row} \\ \\ \\ \\ \end{array} \\ \begin{array}{c} \qquad\qquad i\text{th column} \qquad j\text{th column} \end{array} & \end{array}$$

Suppose $E_{\beta R_i+R_j} \in \mathbb{R}^{m\times m}$, the result of $E_{\beta R_i+R_j}A(\alpha \neq 0)$ is to multiply each element of ith row of matrix A by α, then add them into the jth row while keeping ith row unchanged.

Example

$$E_{2R_1+R_2}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} w & x \\ 2w+y & 2x+z \end{bmatrix}$$

$$E_{-2R_1+R_3}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ -2a_{11}+a_{31} & -2a_{12}+a_{32} \end{bmatrix}$$

Important property for elementary matrices

For a given matrix A, performing elementary row operation for A is equivalent to premultiplying A by the corresponding elementary matrix.

Theorem 6.10 (Elementary Matrices are Invertible and Their Inverse are also Elementary Matrices)

(1) $E_{R_iR_j}^{-1} = E_{R_iR_j}$, corresponding to the reverse row operation 1: $R_i \leftrightarrow R_j$.

(2) $E_{\alpha R_i}^{-1} = E_{\frac{1}{\alpha}R_i}$ $(\alpha \neq 0)$, corresponding to the reverse row operation 2: $R_i \to \frac{1}{\alpha}R_i$.

(3) $E_{\beta R_i+R_j}^{-1} = E_{-\beta R_i+R_j}$, corresponding to the reverse row operation 3: $R_j \to -\beta R_i + R_j$.

Remark. The inverse of the elementary matrices corresponding to the reverse row operations and belong to the same type of elementary matrices.

Example

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

Definition 6.11 (Permutation matrix)
A permutation matrix is a square matrix that has exactly one entry of 1 in each row and each column and 0s elsewhere.

Remark A permutation matrix can be obtained by reordering the rows of the identity matrix.

Example

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Property 6.12 For a permutation matrix P, it can always be decomposed into a multiplication of finite number of row exchange matrices $E_{R_i R_j}$(corresponding to the row exchange $R_i \leftrightarrow R_j$), i.e.

$$P = E_{R_{i_k} R_{j_k}} \cdots E_{R_{i_2} R_{j_2}} E_{R_{i_1} R_{j_1}}$$

Example

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = E_{R_2 R_4} E_{R_1 R_3}$$

Property 6.13 For a permutation matrix P, $P^{-1} = P^T$, since

$$\begin{aligned} P^{-1} &= (E_{R_{i_k} R_{j_k}} \cdots E_{R_{i_2} R_{j_2}} E_{R_{i_1} R_{j_1}})^{-1} \\ &= E^{-1}_{R_{i_1} R_{j_1}} E^{-1}_{R_{i_2} R_{j_2}} \cdots E^{-1}_{R_{i_k} R_{j_k}} \\ &= E_{R_{i_1} R_{j_1}} E_{R_{i_2} R_{j_2}} \cdots E_{R_{i_k} R_{j_k}} \\ &= E^{T}_{R_{i_1} R_{j_1}} E^{T}_{R_{i_2} R_{j_2}} \cdots E^{T}_{R_{i_k} R_{j_k}} \\ &= P^T \end{aligned}$$

Example

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}^T$$