5.3 Matrix Inverses

Definition 5.6: Inverse of a Square Matrix

Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. We say that A is *invertible* if and only if there exists a matrix $C \in \mathbb{R}^{n\times n}$ such that

$$A \cdot C = C \cdot A = I_n$$

where I_n is the $n \times n$ identify matrix. We call $C = A^{-1}$ the *inverse* of A. An invertible matrix is called *non-singular* while a matrix that is not invertible is called *singular*.

Matrix inverses are the matrix-matrix multiplicative inverse. Two-sided inverses are defined only for square matrices.

EXAMPLE 5.3.1

Let's consider the matrices

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \qquad C = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}.$$

Notice that C is the inverse of A, denoted as $C = A^{-1}$, since

$$A \cdot C = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

EXAMPLE 5.3.2

Let's consider the matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix},$$

We see that $C = A^{-1}$ because

$$A \cdot C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Theorem 23: Inverses of Elementary Matrices

Let $i, k, n \in \mathbb{N}$ such that $1 \leq i, k \leq n$ with $i \neq k$ and let $c \in \mathbb{R}$ with $c \neq 0$. Then, the inverses of each of the elementary matrices are given below:

1. Shear matrices: $\Big(S_{ik}(c)\Big)^{-1} = S_{ik}(-c)$
2. Transposition matrices: $\Big(P_{ik}\Big)^{-1} = P_{ik}^T$
3. Dilation matrices: $\Big(D_i(c)\Big)^{-1} = D_i\left(\frac{1}{c}\right)$

Proof. Suppose $i, k, n \in \mathbb{N}$ such that $1 \leq i, k \leq n$ with $i \neq k$ and suppose $c \in \mathbb{R}$ is a nonzero constant. Let's begin by proving that $\Big(S_{ik}(c)\Big)^{-1} = S_{ik}(-c)$. To do so, we will multiply $S_{ik}(c)$ by the stated inverse and show that we produce I_n. Consider

$$\begin{aligned} S_{ik}(c) \cdot S_{ik}(-c) &= \left(I_n + c\,\mathbf{e}_i\,\mathbf{e}_k^T\right) \cdot \left(I_n - c\,\mathbf{e}_i\,\mathbf{e}_k^T\right) \\ &= I_n \cdot \left(I_n - c\,\mathbf{e}_i\,\mathbf{e}_k^T\right) + c\,\mathbf{e}_i\,\mathbf{e}_k^T \cdot \left(I_n - c\,\mathbf{e}_i\,\mathbf{e}_k^T\right) \\ &= I_n - c\,\mathbf{e}_i\,\mathbf{e}_k^T + c\,\mathbf{e}_i\,\mathbf{e}_k^T \cdot I_n - c^2\left(\mathbf{e}_i\,\mathbf{e}_k^T\right) \cdot \left(\mathbf{e}_i\,\mathbf{e}_k^T\right) \\ &= I_n - c\,\mathbf{e}_i\,\mathbf{e}_k^T + c\,\mathbf{e}_i\,\mathbf{e}_k^T - c^2\,\mathbf{e}_i\left(\mathbf{e}_k^T\,\mathbf{e}_i\right)\mathbf{e}_k^T \end{aligned}$$

Notice that the matrix-matrix product $\mathbf{e}_k^T\,\mathbf{e}_i$ results in a scalar output equivalent to the inner product $\mathbf{e}_k \cdot \mathbf{e}_i$. If $j \in \mathbb{N}$ with $1 \leq j \leq n$ we know $\mathbf{e}_j \in \mathbb{R}^n$ is the jth elementary basis vector with all zero entries except the jth coefficient, which has value equal to one. Because of this structure and since $i \neq k$ by assumption, we see that $\mathbf{e}_k^T\,\mathbf{e}_i = \mathbf{e}_k \cdot \mathbf{e}_i = 0$. With this we have,

$$S_{ik}(c) \cdot S_{ik}(-c) = I_n - c\,\mathbf{e}_i\,\mathbf{e}_k^T + c\,\mathbf{e}_i\,\mathbf{e}_k^T = I_n$$

We conclude that $\Big(S_{ik}(c)\Big)^{-1} = S_{ik}(-c)$.

Next, let's prove $\Big(P_{ik}\Big)^{-1} = P_{ik}^T$. We remark that $P_{ik}^T = P_{ik}$ be definition. To this end, let $C = P_{ik} \cdot P_{ik}$. Then for $j \in \mathbb{N}$ with $1 \leq j \leq n$, we can find the jth column of C using the column-partition version of matrix-vector multiplication. Consider

$$C(:, j) = P_{ik} \cdot P_{ik}(:, j), \qquad \text{with} \qquad P_{ik}(:, j) = \begin{cases} \mathbf{e}_j & \text{if } j \neq i \text{ and } j \neq k \\ \mathbf{e}_k & \text{if } j = i \\ \mathbf{e}_i & \text{if } j = k \end{cases}$$

Thus, using these definitions we see that

$$C(:, j) = \mathbf{e}_j$$

for all $1 \leq j \leq n$ and we conclude that $C = I_n$ and we've established $\Big(P_{ik}\Big)^{-1} = P_{ik}^T$.

Finally, let's establish that $\Big(D_i(c)\Big)^{-1} = D_i\left(\frac{1}{c}\right)$. To this end, consider

$$\begin{aligned}
D_i(c) \cdot D_i\,(1/c) &= \left(I_n + (c-1)\,\mathbf{e}_i\,\mathbf{e}_i^T\right) \cdot \left(I_n + (1/c-1)\ \mathbf{e}_i\,\mathbf{e}_i^T\right) \\
&= I_n \cdot \left(I_n + (1/c-1)\ \mathbf{e}_i\,\mathbf{e}_i^T\right) + (c-1)\,\mathbf{e}_i\,\mathbf{e}_i^T \cdot \left(I_n + (1/c-1)\ \mathbf{e}_i\,\mathbf{e}_i^T\right) \\
&= I_n + (1/c-1)\ \mathbf{e}_i\,\mathbf{e}_i^T + (c-1)\,\mathbf{e}_i\,\mathbf{e}_i^T + (c-1)\ (1/c-1)\ \mathbf{e}_i\,\mathbf{e}_i^T\,\mathbf{e}_i\,\mathbf{e}_i^T \\
&= I_n + (1/c+c-2)\ \mathbf{e}_i\,\mathbf{e}_i^T + (c-1)\ (1/c-1)\ \mathbf{e}_i\,\mathbf{e}_i^T
\end{aligned}$$

Using distributivity of scalar multiplication over addition, we see

$$(c-1)\ (1/c-1) = 2 - c - 1/c.$$

Thus, we have

$$D_i(c) \cdot D_i\,(1/c) = I_n + (1/c+c-2)\ \mathbf{e}_i\,\mathbf{e}_i^T + (2-c-1/c)\,\mathbf{e}_i\,\mathbf{e}_i^T = I_n$$

By definition of the matrix inverse, we have $\Big(D_i(c)\Big)^{-1} = D_i\left(\frac{1}{c}\right)$ □

EXAMPLE 5.3.3

The theorem above gives us a very concrete mechanism to invert any elementary matrix E. Choose any elementary matrix and you can immediately produce the inverse.

As we will see, elementary matrices are very special precisely because they are invertible. One of the most powerful realizations in the solution to linear systems is to translate any system $A\mathbf{x} = \mathbf{b}$ into an equivalent system $U\mathbf{x} = \mathbf{y}$ using elementary matrices. This process is a gold mine for producing solution sets.

Theorem 24: Cramer's Rule for Inverse of a 2×2 System

Consider the 2×2 matrix with real entries given by

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Then, the inverse of matrix A is given by

$$A^{-1} = \frac{1}{\delta} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

where $\delta = a_{11}a_{22} - a_{21}a_{12}$.

Proof. Suppose $A \in \mathbb{R}^{2\times 2}$ is nonsingular. Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

In order to produce the inverse of A, let's transform A into I_2 using elementary matrices. To this end, let $\delta = a_{11}a_{22} - a_{21}a_{12}$. As we will see in Chapter 7, we call δ the determinant of A and $\delta \neq 0$ if and only if A is invertible. As we will see in the next section, since A is invertible, we know that either $a_{11} \neq 0$ or $a_{21} \neq 0$. Thus, let's assume that $a_{11} \neq 0$. Now, we reduce A to I_2. Consider the step-by-step calculation beginning with the product

$$S_{21}\left(-a_{21}/a_{11}\right) \cdot A = \begin{bmatrix} 1 & 0 \\ -a_{21}/a_{11} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ 0 & \delta/a_{11} \end{bmatrix}$$

Now we consider

$$D_2\left(a_{11}/\delta\right) \cdot D_1\left(1/a_{11}\right) = \begin{bmatrix} 1/a_{11} & 0 \\ 0 & a_{11}/\delta \end{bmatrix}$$

We can multiply our product by this diagonal matrix to find

$$\begin{bmatrix} 1/a_{11} & 0 \\ 0 & a_{11}/\delta \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ 0 & \delta/a_{11} \end{bmatrix} = \begin{bmatrix} 1 & a_{12}/a_{11} \\ 0 & 1 \end{bmatrix}$$

Finally, multiplying this entire product on the left-hand side by we see

$$\begin{bmatrix} 1 & -a_{12}/a_{11} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & a_{12}/a_{11} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus, we see that $A^{-1} \cdot A = I_2$ with

$$A^{-1} = S_{12}\left(-a_{12}/a_{11}\right) \cdot D_2\left(a_{11}/\delta\right) \cdot D_1\left(1/a_{11}\right) \cdot S_{21}\left(-a_{21}/a_{11}\right).$$

The reduction of matrix A into I_2 above gives a constructive mechanism to explicitly calculate A^{-1}. We begin by finding the product

$$\Big(D_2\left(a_{11}/\delta\right) \cdot D_1\left(1/a_{11}\right)\Big) \cdot S_{21}\left(-a_{21}/a_{11}\right)$$

given as

$$\begin{bmatrix} 1/a_{11} & 0 \\ 0 & a_{11}/\delta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -a_{21}/a_{11} & 1 \end{bmatrix} = \begin{bmatrix} 1/a_{11} & 0 \\ -a_{21}/\delta & a_{11}/\delta \end{bmatrix}$$

We can calculate A^{-1} using the product

$$\begin{bmatrix} 1 & -\dfrac{a_{12}}{a_{11}} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dfrac{1}{a_{11}} & 0 \\ -\dfrac{a_{21}}{\delta} & \dfrac{a_{11}}{\delta} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{a_{11}} + \dfrac{a_{21}a_{12}}{a_{11}\delta} & -\dfrac{a_{12}}{a_{11}} \\ -\dfrac{a_{21}}{\delta} & \dfrac{a_{11}}{\delta} \end{bmatrix} = \frac{1}{\delta} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

This is exactly what we wanted to show. □

EXAMPLE 5.3.4

Show how to use this formula to solve mass-spring chain displacement problem with known input weights.

Theorem 25: Properties of Matrix Inverses

Let $A, B \in \mathbb{R}^{n\times n}$ be square, nonsingular matrices. Then

1. The matrix A^{-1} is unique
2. The linear-system problem $A\mathbf{x} = \mathbf{b}$ has a unique solution.
3. $A \cdot B$ is invertible and $\left(A \cdot B\right)^{-1} = B^{-1} \cdot A^{-1}$
4. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T = A^{-T}$
5. A^{-1} can be written as a product of elementary matrices

For more fun properties see Piziak and Odell p. 18 -19 Theorem 1.2 & 1.3

Theorem 26: RREF Solves the Linear-System Problem

Let $A\mathbf{x} = \mathbf{b}$ be a given linear system problem with $U\mathbf{x} = \mathbf{y}$ the equivalent system with $U = \text{RREF}(A)$. Then, the solution sets to these two linear systems problems are identical.