7.2 Invertible Matrices

DO ANY MATRICES HAVE A MULTIPLICATIVE INVERSE? The definition of matrix multiplication implies that, to have a two-sided inverse, a matrix must have the same number of rows as columns.

7.2.0 Definition. A matrix $\mathbf{A}$ is ***invertible*** if there exists a matrix $\mathbf{B}$ such that $\mathbf{A}\,\mathbf{B} = \mathbf{I}$ and $\mathbf{B}\,\mathbf{A} = \mathbf{I}$. When it exists, the matrix $\mathbf{B}$ is called the ***inverse*** of $\mathbf{A}$ and denoted by $\mathbf{A}^{-1} := \mathbf{B}$.

For example, we have

$$\begin{bmatrix}2 & 1\\ 5 & 3\end{bmatrix}\begin{bmatrix}3 & -1\\ -5 & 2\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \quad\text{and}\quad \begin{bmatrix}3 & -1\\ -5 & 2\end{bmatrix}\begin{bmatrix}2 & 1\\ 5 & 3\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},$$

so both of these matrices are invertible.

7.2.1 Problem. Show that $\begin{bmatrix}2 & 6\\ 1 & 3\end{bmatrix}$ is not invertible.

Solution. If $\begin{bmatrix}w & y\\ x & z\end{bmatrix}$ were the inverse, then we would have

$$\mathbf{I} = \begin{bmatrix}2 & 6\\ 1 & 3\end{bmatrix}\begin{bmatrix}w & y\\ x & z\end{bmatrix} = \begin{bmatrix}2w+6x & 2y+6z\\ w+3x & y+3z\end{bmatrix} \Leftrightarrow \left\{\begin{aligned} 2w+6x &&&= 1\\ w+3x &&&= 0\\ && 2y+6z &= 0\\ && y+2z &= 1 \end{aligned}\right\}.$$

The row reduction algorithm [4.2.0] gives

$$\begin{bmatrix}2 & 6 & 0 & 0 & 1\\ \underline{1} & 3 & 0 & 0 & 0\\ 0 & 0 & 2 & 6 & 0\\ 0 & 0 & \underline{1} & 3 & 1\end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_1 - 2\vec{r}_2 \mapsto \vec{r}_1\\ \vec{r}_3 - 2\vec{r}_4 \mapsto \vec{r}_3}} \begin{bmatrix}0 & 0 & 0 & 0 & 1\\ 1 & 3 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & -2\\ 0 & 0 & 1 & 3 & 1\end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_1 \mapsto \vec{r}_2\\ \vec{r}_2 \mapsto \vec{r}_1\\ \vec{r}_3 \mapsto \vec{r}_4\\ \vec{r}_4 \mapsto \vec{r}_3}} \begin{bmatrix}1 & 3 & 0 & 0 & 0\\ 0 & 0 & 1 & 3 & 0\\ 0 & 0 & 0 & 0 & \underline{1}\\ 0 & 0 & 0 & 0 & -2\end{bmatrix} \xrightarrow[\sim]{\vec{r}_4 + 2\vec{r}_3 \mapsto \vec{r}_4} \begin{bmatrix}\underline{1} & 3 & 0 & 0 & 0\\ 0 & 0 & \underline{1} & 3 & 0\\ 0 & 0 & 0 & 0 & \underline{1}\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}.$$

Since the reduced row echelon form of the augmented matrix has a leading one in its rightmost column, the linear system is inconsistent and no inverse exists. □

Although not explicitly part of the definition, it is straightforward to see that a matrix has at most one inverse.

7.2.2 Lemma. *For an invertible matrix, the inverse matrix is unique.*

Proof. Suppose that $\mathbf{B}$ and $\mathbf{C}$ are both inverses of the matrix $\mathbf{A}$. The definition of an invertible matrix and the properties of matrix multiplication [7.0.4] give $\mathbf{C} = \mathbf{C}\,\mathbf{I} = \mathbf{C}(\mathbf{A}\,\mathbf{B}) = (\mathbf{C}\,\mathbf{A})\mathbf{B} = \mathbf{I}\,\mathbf{B} = \mathbf{B}$. □

Asserting that the coefficient matrix of a linear system is invertible determines the solution set.

7.2.3 Proposition. *For any invertible matrix $\mathbf{A}$, the linear system $\mathbf{A}\,\vec{x} = \vec{b}$ has a unique solution given by $\vec{x} = \mathbf{A}^{-1}\,\vec{b}$.*

Proof. For existence, observe that $\mathbf{A}(\mathbf{A}^{-1}\vec{b}) = (\mathbf{A}\,\mathbf{A}^{-1})\vec{b} = \mathbf{I}\,\vec{b} = \vec{b}$, so $\mathbf{A}^{-1}\vec{b}$ is a solution. For uniqueness, observe that, for any solution $\vec{v}$, we have $\vec{v} = \mathbf{I}\,\vec{v} = (\mathbf{A}^{-1}\mathbf{A})\,\vec{v} = \mathbf{A}^{-1}(\mathbf{A}\,\vec{v}) = \mathbf{A}^{-1}\vec{b}$. □

For (2×2)-matrices, we can easily identify the invertible ones.

7.2.4 Problem. Consider the matrix $\mathbf{A} := \left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]$ where $a, b, c, d \in \mathbb{K}$. When $ad - bc \neq 0$, show that $\mathbf{A}$ is invertible and $\mathbf{A}^{-1} = \frac{1}{ad-bc}\left[\begin{smallmatrix} d & -c \\ -b & a \end{smallmatrix}\right]$. When $ad - bc = 0$, show that $\mathbf{A}$ is not invertible.

Solution. When $ad - bc \neq 0$, we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad-bc & -ab+ba \\ cd-dc & -cb+da \end{bmatrix} = (ad-bc)\,\mathbf{I} = \begin{bmatrix} ad-bc & -ab+ba \\ cd-dc & -cb+da \end{bmatrix} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

so $\mathbf{A}^{-1} = \frac{1}{ad-bc}\left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right]$. Next, suppose that $ad - bc = 0$. If $a \neq 0$, then we would have $d = bc/a$ and

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ ac/a & bc/a \end{bmatrix} = \begin{bmatrix} a & b \\ ka & kb \end{bmatrix},$$

where $k = \frac{c}{a}$. If $\left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]$ were an inverse of A, then we would obtain

$$\mathbf{I} = \begin{bmatrix} a & b \\ ka & kb \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} aw+by & ax+bz \\ k(aw+by) & k(ax+bz) \end{bmatrix}.$$

Applying elementary row operations to the associated augmented matrix yields

$$\begin{bmatrix} \underline{a} & 0 & b & 0 & 1 \\ 0 & \underline{a} & 0 & b & 0 \\ ka & 0 & kb & 0 & 0 \\ 0 & ka & 0 & kb & 1 \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_3 - k\vec{r}_1 \mapsto \vec{r}_3 \\ \vec{r}_4 - k\vec{r}_2 \mapsto \vec{r}_4}} \begin{bmatrix} a & 0 & b & 0 & 1 \\ 0 & a & 0 & b & 0 \\ 0 & 0 & 0 & 0 & -k \\ 0 & 0 & 0 & 0 & \underline{1} \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_1 - \vec{r}_4 \mapsto \vec{r}_1 \\ \vec{r}_3 + k\vec{r}_4 \mapsto \vec{r}_3}} \begin{bmatrix} a & 0 & b & 0 & 0 \\ 0 & a & 0 & b & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow[\sim]{\substack{a^{-1}\vec{r}_1 \mapsto \vec{r}_1 \\ a^{-1}\vec{r}_2 \mapsto \vec{r}_2 \\ \vec{r}_3 \mapsto \vec{r}_4 \\ \vec{r}_4 \mapsto \vec{r}_3}} \begin{bmatrix} 1 & 0 & a^{-1}b & 0 & 0 \\ 0 & 1 & 0 & a^{-1}b & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the linear system is inconsistent, we deduce that the matrix $\mathbf{A}$ does not have an inverse in this case. When $a = 0$, we have $bc = 0$ which implies that either $b = 0$ or $c = 0$. Thus, $\mathbf{A}$ has the form $\left[\begin{smallmatrix} 0 & 0 \\ c & d \end{smallmatrix}\right]$ or $\left[\begin{smallmatrix} 0 & b \\ 0 & d \end{smallmatrix}\right]$ neither of which has an inverse, because no matrix with a zero row or zero column has a inverse. □

For future convenience, we summarize how taking the inverse of a matrix interacts with a few other matrix operations.

7.2.5 Proposition (Properties of invertible matrices)**.** *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be invertible matrices of the same size. For all* $0 \neq c \in \mathbb{K}$ *and all* $k \in \mathbb{N}$*, we have the following:*

(involution)	$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
(compatibility with matrix multiplication)	$(\mathbf{A}\,\mathbf{B})^{-1} = \mathbf{B}^{-1}\,\mathbf{A}^{-1}$
(compatibility with scalar multiplication)	$(c\,\mathbf{A})^{-1} = \frac{1}{c}\,\mathbf{A}^{-1}$
(compatibility with transpose)	$(\mathbf{A}^{\mathsf{T}})^{-1} = (\mathbf{A}^{-1})^{\mathsf{T}}$
(compatibility with powers)	$(\mathbf{A}^{k})^{-1} = (\mathbf{A}^{-1})^{k}$

Proof. Since $\mathbf{A}\,\mathbf{A}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\,\mathbf{A}$, the uniqueness of the inverse for the matrix $\mathbf{A}^{-1}$ implies that $\mathbf{A} = (\mathbf{A}^{-1})^{-1}$. The properties of matrix multiplication [7.0.4] give

$$\begin{aligned}(\mathbf{B}^{-1}\,\mathbf{A}^{-1})(\mathbf{A}\,\mathbf{B}) &= \mathbf{B}^{-1}(\mathbf{A}^{-1}\,\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\,\mathbf{I}\,\mathbf{B} = \mathbf{B}^{-1}\,\mathbf{B} = \mathbf{I}\\ (\mathbf{A}\,\mathbf{B})(\mathbf{B}^{-1}\,\mathbf{A}^{-1}) &= \mathbf{A}^{-1}(\mathbf{B}^{-1}\,\mathbf{B})\mathbf{A} = \mathbf{A}^{-1}\,\mathbf{A} = \mathbf{I}\,,\end{aligned}$$

so the uniqueness of the inverse establishes that $(\mathbf{A}\,\mathbf{B})^{-1} = \mathbf{B}^{-1}\,\mathbf{A}^{-1}$. As $\big(\frac{1}{c}\,\mathbf{A}^{-1}\big)(c\,\mathbf{A}) = \frac{c}{c}(\mathbf{A}^{-1}\,\mathbf{A}) = \mathbf{I}$ and $(c\,\mathbf{A})\big(\frac{1}{c}\,\mathbf{A}^{-1}\big) = \frac{c}{c}(\mathbf{A}\,\mathbf{A}^{-1}) = \mathbf{I}$, the uniqueness of the inverse also establish that $(c\,\mathbf{A})^{-1} = \frac{1}{c}\,\mathbf{A}^{-1}$. The properties of the transpose [5.2.7] give $\mathbf{I} = (\mathbf{A}\,\mathbf{A}^{-1})^{\mathsf{T}} = (\mathbf{A}^{-1})^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}$ and $\mathbf{I} = (\mathbf{A}^{-1}\,\mathbf{A})^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}}(\mathbf{A}^{-1})^{\mathsf{T}}$, so $(\mathbf{A}^{\mathsf{T}})^{-1} = (\mathbf{A}^{-1})^{\mathsf{T}}$. We prove the final property by induction on k. When $k = 0$, we have $(\mathbf{A}^0)^{-1} = \mathbf{I}^{-1} = \mathbf{I} = (\mathbf{A}^{-1})^0$ which prove the base case. The compatibility with matrix multiplication and the induction hypothesis give

$$(\mathbf{A}^k)^{-1} = (\mathbf{A}\,\mathbf{A}^{k-1})^{-1} = (\mathbf{A}^{k-1})^{-1}\mathbf{A}^{-1} = (\mathbf{A}^{-1})^{k-1}\mathbf{A}^{-1} = (\mathbf{A}^{-1})^k\,. \quad \square$$

Exercises

7.2.6 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. Every matrix has a multiplicative inverse.
ii. Every square matrix has a multiplicative inverse.
iii. The identity matrix is equal to its own inverse.
iv. The inverse of a sum of invertible matrices equals the sum of there inverses.

7.2.7 Problem. Consider the matrix

$$\mathbf{P} := \begin{bmatrix} -1 & -1 & -3 \\ 2 & -3 & 2 \\ 1 & -1 & 2 \end{bmatrix}.$$

(i) Demonstrate that $\mathbf{P}^3 + 2\,\mathbf{P}^2 + 2\,\mathbf{P} - 3\,\mathbf{I} = \mathbf{0}$.
(ii) If $\mathbf{Q} := \frac{1}{3}(\mathbf{P}^2 + 2\,\mathbf{P} + 2\,\mathbf{I})$, then verify that $\mathbf{Q} = \mathbf{P}^{-1}$.
(iii) Explain how $\mathbf{Q}$ in part (ii) can be obtained from the equation in part (i).

8 Matrix Factorizations

Expressing a matrix as the product is a general method for solving problems in linear algebra. This chapter uses matrix factorizations to characterize invertible matrices, to solve linear systems, and to better understand the elementary row operations.

8.0 Invertibility

How do we characterize invertible matrices? We first introduce some notation for the simplest nonzero matrices.

8.0.0 Definition. For any positive integers j and k, the ***matrix unit*** $\mathbf{E}_{j,k}$ is the square matrix whose (j,k)-entry is 1 and all other entries are 0.

Among the (3×3)-matrices, there are 9 matrix units:

$$\mathbf{E}_{1,1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{E}_{1,2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\mathbf{E}_{1,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{E}_{2,1} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\mathbf{E}_{2,2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{E}_{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\mathbf{E}_{3,1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \mathbf{E}_{3,2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$
$$\mathbf{E}_{3,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

8.0.1 Definition. An ***elementary matrix*** is any matrix obtained by performing one elementary row operation on an identity matrix $\mathbf{I}$.

For any two scalars $c, d \in \mathbb{K}$ such that $d \neq 0$ and any distinct row and column indices j, k, the three types of elementary matrices are

$$\mathbf{I} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 \cdots 0 & & \\ & & \vdots \ddots \vdots & & \\ & & 0 \cdots 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \xrightarrow[\sim]{\vec{r}_j + c\,\vec{r}_k \,\mapsto\, \vec{r}_j} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 \cdots c & & \\ & & \vdots \ddots \vdots & & \\ & & 0 \cdots 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} = \mathbf{I} + c\,\mathbf{E}_{j,k}$$

$$\mathbf{I} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 \cdots 0 & & \\ & & \vdots \ddots \vdots & & \\ & & 0 \cdots 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_k \,\mapsto\, \vec{r}_j \\ \vec{r}_j \,\mapsto\, \vec{r}_k}} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 \cdots 1 & & \\ & & \vdots \ddots \vdots & & \\ & & 1 \cdots 0 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} = \mathbf{I} - \mathbf{E}_{j,j} - \mathbf{E}_{k,k} + \mathbf{E}_{j,k} + \mathbf{E}_{k,j}$$

$$\mathbf{I} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 \cdots 0 & & \\ & & \vdots \ddots \vdots & & \\ & & 0 \cdots 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \xrightarrow[\sim]{d\,\vec{r}_j \,\mapsto\, \vec{r}_j} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & d \cdots 0 & & \\ & & \vdots \ddots \vdots & & \\ & & 0 \cdots 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} = \mathbf{I} + (d-1)\,\mathbf{E}_{j,j} .$$

Elementary matrices generate the equivalence relation on linear systems.

8.0.2 Lemma (Properties of elementary matrices)**.**

i. *Left multiplication by an elementary matrix is equivalent to performing the corresponding elementary row operation.*

ii. *Elementary matrices are invertible and the inverse of an elementary matrix is an elementary matrix of the same type.*

Proof. Since the j-th row of the matrix unit $\mathbf{E}_{j,k}$ is the standard basis vector $\vec{e}_k$ and it is the only nonzero row, the j-th row in the matrix product $\mathbf{E}_{j,k}\,\mathbf{A}$ equals the same row in $\mathbf{A}$ and all other rows are zero.

i. The properties of matrix multiplication [7.0.4] establishes that $(\mathbf{I} + c\,\mathbf{E}_{j,k})\mathbf{A} \;=\; \mathbf{A} + c\,\mathbf{E}_{j,k}\,\mathbf{A}$ which is the matrix obtained by replacing the j-th row of $\mathbf{A}$ with the sum of c times the k-th row of $\mathbf{A}$ and the j-th row of $\mathbf{A}$. Similarly, we have

$$(\mathbf{I} - \mathbf{E}_{j,j} - \mathbf{E}_{k,k} + \mathbf{E}_{j,k} + \mathbf{E}_{k,j})\mathbf{A} = \mathbf{A} - \mathbf{E}_{j,j}\mathbf{A} - \mathbf{E}_{k,k}\,\mathbf{A} + \mathbf{E}_{j,k}\,\mathbf{A} + \mathbf{E}_{k,j}\,\mathbf{A}\,,$$

which is the matrix obtained by interchanging the j-th and k-th rows in $\mathbf{A}$, and $\big(\mathbf{I} + (d-1)\,\mathbf{E}_{j,j}\big)\mathbf{A} \;=\; \mathbf{A} + d\,\mathbf{E}_{j,j}\,\mathbf{A}$ is the matrix obtained by multiplying the j-th row by d.

ii. Since $\mathbf{E}_{i,j}\,\mathbf{E}_{k,\ell} = \delta_{j,k}\,\mathbf{E}_{i,\ell}$, it follows that $(\mathbf{I} + c\,\mathbf{E}_{j,k})(\mathbf{I} - c\,\mathbf{E}_{j,k}) = \mathbf{I}$, $(\mathbf{I} + \mathbf{E}_{j,k} + \mathbf{E}_{k,j} - \mathbf{E}_{j,j} - \mathbf{E}_{k,k})^2 = \mathbf{I}$, and

$$(\mathbf{I} + (d-1)\,\mathbf{E}_{j,j})(\mathbf{I} + (d^{-1}-1)\,\mathbf{E}_{j,j}) = \mathbf{I}\,. \qquad \square$$

We now enumerate fourteen mathematical synonyms for the phrase “invertible matrix”.

8.0.3 Theorem (Characterizations of invertible matrices)**.** *Let n be a positive integer. For any $(n \times n)$-matrix $\mathbf{A}$, the following are equivalent.*

a. *The matrix $\mathbf{A}$ is invertible.*
b. *There is an $(n \times n)$-matrix $\mathbf{B}$ such that $\mathbf{B}\,\mathbf{A} = \mathbf{I}$.*
c. *There is an $(n \times n)$-matrix $\mathbf{C}$ such that $\mathbf{A}\,\mathbf{C} = \mathbf{I}$.*
d. *The matrix $\mathbf{A}^{\mathsf{T}}$ is invertible.*
e. *For all $\vec{b} \in \mathbb{K}^n$, the linear system $\mathbf{A}\,\vec{x} = \vec{b}$ has a unique solution.*
f. *For all $\vec{b} \in \mathbb{K}^n$, the linear system $\mathbf{A}\,\vec{x} = \vec{b}$ is consistent.*
g. *The homogeneous linear system $\mathbf{A}\,\vec{x} = \vec{0}$ has only the zero solution.*
h. *The reduced row echelon form of $\mathbf{A}$ is the identity matrix $\mathbf{I}_n$.*
i. *The matrix $\mathbf{A}$ is a product of elementary matrices.*
j. *The rank of the matrix $\mathbf{A}$ is n.*
k. *The columns of the matrix $\mathbf{A}$ are linearly independent.*
l. *The rows of the matrix $\mathbf{A}$ are linearly independent.*
m. *The columns of the matrix $\mathbf{A}$ span $\mathbb{K}^n$.*
n. *The rows of the matrix $\mathbf{A}$ span $\mathbb{K}^n$.*

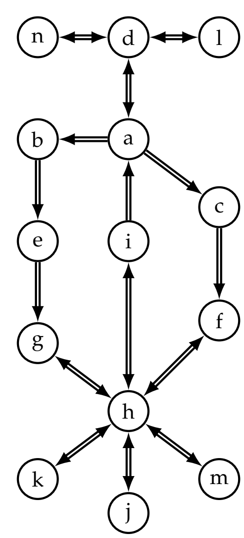

Figure 8.0: Structure of the proof

Proof.
$a \Rightarrow b$: When $\mathbf{A}$ is invertible, set $\mathbf{B} := \mathbf{A}^{-1}$.
$a \Rightarrow c$: When $\mathbf{A}$ is invertible, set $\mathbf{C} := \mathbf{A}^{-1}$.

a ⇔ *d*: Because the properties of invertible matrices [7.2.5] include $(\mathbf{A}^{-1})^{\mathsf{T}} = (\mathbf{A}^{\mathsf{T}})^{-1}$, the matrix $\mathbf{A}$ is invertible if and only if the matrix $\mathbf{A}^{\mathsf{T}}$ is invertible.

b ⇒ *e*: Suppose that $\mathbf{B}\,\mathbf{A} = \mathbf{I}$. Given a solution $\vec{v} \in \mathbb{K}^n$ to $\mathbf{A}\,\vec{x} = \vec{b}$, it follows that $\vec{v} = \mathbf{I}\,\vec{v} = (\mathbf{B}\,\mathbf{A})\,\vec{v} = \mathbf{B}\,(\mathbf{A}\,\vec{v}) = \mathbf{B}\,\vec{b}$.

c ⇒ *f*: Suppose that $\mathbf{A}\,\mathbf{C} = \mathbf{I}$. Since $\mathbf{A}(\mathbf{C}\,\vec{b}) = (\mathbf{A}\,\mathbf{C})\vec{b} = \mathbf{I}\,\vec{b} = \vec{b}$, the vector $\mathbf{C}\,\vec{b} \in \mathbb{K}^n$ is a solution to the linear system $\mathbf{A}\,\vec{x} = \vec{b}$.

f ⇔ *h* ⇔ *j* ⇔ *m*: Since $\mathbf{A}$ is a square matrix, the characterizations of universal consistency [6.1.0] establish these equivalences.

e ⇒ *g*: There exists a unique solution in the special case $\vec{b} = \vec{0}$.

g ⇔ *h* ⇔ *j* ⇔ *k*: Since $\mathbf{A}$ is a square matrix, the characterizations of a unique solution [6.2.4] prove these equivalences.

h ⇔ *i*: This equivalence follows from the interpretation of elementary operations as left multiplication by elementary matrices.

i ⇒ *a*: Since elementary matrices are invertible, the properties of invertible matrices [7.2.5] show that $\mathbf{A}$ is invertible.

d ⇔ *l*: Since we have already established that *a* is equivalent to *k*, the transposed version also holds.

d ⇔ *n*: Since we have already established that *a* is equivalent to *m*, the transposed version also holds. □

8.0.4 Remark. The definition of an invertible matrix requires one matrix be a two-sided inverse. However, parts *b* and *c* in the characterization of invertible matrices demonstrate that it suffices to have a one-sided inverse.

As an immediate corollary, we also obtain an effective method for calculating the inverse of a square matrix.

8.0.5 Algorithm (Finding inverse matrices)**.**

input: an $(n \times n)$-matrix $\mathbf{A}$.
output: the $(n \times n)$-matrix $\mathbf{A}^{-1}$ if it exists.

Set $[\mathbf{B}\ \ \mathbf{C}]$ to be the reduced row echelon form of of the $(n \times 2n)$-matrix $[\mathbf{A}\ \ \mathbf{I}]$. — *apply the row reduction algorithm* [4.2.0]
If $\mathbf{B} \neq \mathbf{I}$, then error "the matrix $\mathbf{A}$ is not invertible". — *decide if* $\mathbf{A}$ *is invertible*
If $\mathbf{B} = \mathbf{I}$, then return $\mathbf{C}$. — *return inverse*

Correctness of the algorithm. The characterization of invertible matrices shows that $\mathbf{A}$ is invertible if and only if its reduced row echelon form equals the identity matrix $\mathbf{I}$. If $\mathbf{E}$ is a product of elementary matrices such that $\mathbf{E}\,\mathbf{A} = \mathbf{I}$, then we have $\mathbf{E} = \mathbf{A}^{-1}$, $\mathbf{B} = \mathbf{E}\,\mathbf{A}$, and $\mathbf{C} = \mathbf{E}\,\mathbf{I}$. □

8.0.6 Problem. Find the inverse of the matrix $\mathbf{A} := \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$.

Proof. Applying the algorithm, we have

$$[\mathbf{A}\ \ \mathbf{I}] = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ \underline{1} & 0 & 3 & 0 & 1 & 0 \\ \underline{4} & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \xrightarrow[\sim]{\vec{r}_3 - 4\vec{r}_2 \mapsto \vec{r}_3} \begin{bmatrix} 0 & \underline{1} & 2 & 1 & 0 & 0 \\ 1 & \underline{0} & 3 & 0 & 1 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{bmatrix}$$

$$\xrightarrow[\sim]{\vec{r}_3 + 3\vec{r}_1 \mapsto \vec{r}_3} \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 0 & \underline{2} & 3 & -4 & 1 \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_1 - \quad \vec{r}_2 \mapsto \vec{r}_1 \\ \vec{r}_2 - (3/2)\vec{r}_2 \mapsto \vec{r}_2}} \begin{bmatrix} 0 & 1 & 0 & -2.0 & 4 & -1.0 \\ 1 & 0 & 0 & -4.5 & 1 & -1.5 \\ 0 & 0 & 2 & 3.0 & -4 & 1.0 \end{bmatrix}$$

$$\xrightarrow[\sim]{\substack{\vec{r}_2 \mapsto \vec{r}_1 \\ \vec{r}_1 \mapsto \vec{r}_2}} \begin{bmatrix} 1 & 0 & 0 & -4.5 & 1 & -1.5 \\ 0 & 1 & 0 & -2.0 & 4 & -1.0 \\ 0 & 0 & 2 & 3.0 & -4 & 1.0 \end{bmatrix} \xrightarrow[\sim]{(1/2)\,\vec{r}_3 \mapsto \vec{r}_3} \begin{bmatrix} 1 & 0 & 0 & -4.5 & 1 & -1.5 \\ 0 & 1 & 0 & -2.0 & 4 & -1.0 \\ 0 & 0 & 1 & 1.5 & -2 & 0.5 \end{bmatrix}.$$

so $\mathbf{A}^{-1} = \frac{1}{2}\begin{bmatrix} -9 & 2 & -3 \\ -4 & 8 & -2 \\ 3 & -4 & 1 \end{bmatrix}$. □

Verification. We have

$$\begin{aligned} &\frac{1}{2}\begin{bmatrix} -9 & 2 & -3 \\ -4 & 8 & -2 \\ 3 & -4 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} (-9)(0)+(2)(1)+(-3)(4) & (-9)(1)+(2)(0)+(-3)(3) & (-9)(2)+(2)(3)+(-3)(8) \\ (-4)(0)+(8)(1)+(-2)(4) & (-4)(1)+(8)(0)+(-2)(3) & (-4)(2)+(8)(3)+(-2)(8) \\ (3)(0)+(-4)(1)+(1)(4) & (3)(1)+(-4)(0)+(1)(3) & (3)(2)+(-4)(3)+(1)(8) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

□

Exercises

8.0.7 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. Every elementary matrix is square.
ii. The identity matrix is an elementary matrix.
iii. The product of two elementary matrices is also an elementary matrix.
iv. If the linear system $\mathbf{A}\vec{x} = \vec{b}$ has a unique solution, then the coefficient matrix $\mathbf{A}$ is invertible.

8.0.8 Problem. Let $\mathbf{M}$ be an invertible $(m \times m)$-matrix, let $\mathbf{N}$ be an invertible $(n \times n)$-matrix, let $\mathbf{P}$ be an $(m \times n)$-matrix, and let $\mathbf{Q}$ be an $(n \times m)$-matrix. Verify the ***Woodbury matrix identity***

$$(\mathbf{M} + \mathbf{P}\,\mathbf{N}\,\mathbf{Q})^{-1} = \mathbf{M}^{-1} - \mathbf{M}^{-1}\,\mathbf{P}(\mathbf{N}^{-1} + \mathbf{Q}\,\mathbf{M}^{-1}\,\mathbf{P})^{-1}\,\mathbf{Q}\,\mathbf{M}^{-1}.$$

8.0.9 Problem. Fix two positive integers m and n. Let $\mathbf{A}$ be an invertible $(m \times m)$-matrix, let $\mathbf{B}$ be an $(n \times m)$-matrix, let $\mathbf{C}$ be an $(m \times n)$-matrix, and let $\mathbf{D}$ be an $(n \times n)$-matrix. When the ***Schur complement*** $\mathbf{S} := \mathbf{D} - \mathbf{B}\mathbf{A}^{-1}\mathbf{C}$ is invertible, establish the blockwise inversion formula

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\,\mathbf{C}\,\mathbf{S}^{-1}\mathbf{B}\,\mathbf{A}^{-1} & -\mathbf{A}^{-1}\,\mathbf{C}\,\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\,\mathbf{B}\,\mathbf{A}^{-1} & \mathbf{S}^{-1} \end{bmatrix}.$$