5. Finding an Inverse using Elementary Row Operations

The system $\mathbf{Ax}=\mathbf{b}$ where $\mathbf{A}$ is non-singular has the solution $\mathbf{x}=\mathbf{A}^{-1}\mathbf{b}$. Earlier we used the Gauss-Jordan elimination method to find the solutions, where they exist, to systems of linear equations.

<u>Example</u> Find the solution or solutions, if any, to the following system

$$\begin{aligned}
&[1] \quad w+x+2y+2z=1\\
&[2] \quad w+x+2y+z=2\\
&[3] \quad 2w+3x+3y+z=1\\
&[4] \quad w+3x+y+2z=1
\end{aligned}$$

Solution:

$$\left[\begin{array}{cccc|c} 1&1&2&2&1\\1&1&2&1&2\\2&3&3&1&1\\1&3&1&2&1\end{array}\right]
\xrightarrow[\substack{[2]=[2]-[1]\\ [3]=[3]-2[1]\\ [4]=[4]-[1]}]{}
\left[\begin{array}{cccc|c} 1&1&2&2&1\\0&0&0&-1&1\\0&1&-1&-3&-1\\0&2&-1&0&0\end{array}\right]
\xrightarrow[{[2]\leftrightarrow[3]}]{}
\left[\begin{array}{cccc|c} 1&1&2&2&1\\0&1&-1&-3&-1\\0&0&0&-1&1\\0&2&-1&0&0\end{array}\right]$$

$$\xrightarrow[{[4]=[4]-2[2]}]{}
\left[\begin{array}{cccc|c} 1&1&2&2&1\\0&1&-1&-3&-1\\0&0&0&-1&1\\0&0&1&6&2\end{array}\right]
\xrightarrow[{[3]\leftrightarrow[4]}]{}
\left[\begin{array}{cccc|c} 1&1&2&2&1\\0&1&-1&-3&-1\\0&0&1&6&2\\0&0&0&-1&1\end{array}\right]_{\text{REF}}$$

$$\xrightarrow[\substack{[3]=[3]+6[4]\\ [2]=[2]-3[4]\\ [1]=[1]+2[4]}]{}
\left[\begin{array}{cccc|c} 1&1&2&0&3\\0&1&-1&0&-4\\0&0&1&0&8\\0&0&0&-1&1\end{array}\right]
\xrightarrow[\substack{[2]=[2]+[3]\\ [1]=[1]-2[3]}]{}
\left[\begin{array}{cccc|c} 1&1&0&0&-13\\0&1&0&0&4\\0&0&1&0&8\\0&0&0&-1&1\end{array}\right]$$

$$\xrightarrow[{[1]=[1]-[2]}]{}
\left[\begin{array}{cccc|c} 1&0&0&0&-17\\0&1&0&0&4\\0&0&1&0&8\\0&0&0&-1&1\end{array}\right]
\xrightarrow[{[4]=-[4]}]{}
\left[\begin{array}{cccc|c} 1&0&0&0&-17\\0&1&0&0&4\\0&0&1&0&8\\0&0&0&1&-1\end{array}\right]_{\text{RREF}}$$

The solution is $(w,x,y,z)=(-17,4,8,-1)$.

The system in matrix form is

$$\mathbf{Ax}=\mathbf{b} \text{ where } \mathbf{A}=\begin{bmatrix}1&1&2&2\\1&1&2&1\\2&3&3&1\\1&3&1&2\end{bmatrix},\ \mathbf{x}=\begin{bmatrix}w\\x\\y\\z\end{bmatrix},\ \mathbf{b}=\begin{bmatrix}1\\2\\1\\1\end{bmatrix}.$$

Each of the elimination steps can be expressed as a matrix multiplication to both sides of $\mathbf{Ax}=\mathbf{b}$. For instance, the first elimination step can be expressed as

$$\begin{bmatrix}1&0&0&0\\-1&1&0&0\\-2&0&1&0\\-1&0&0&1\end{bmatrix}\begin{bmatrix}1&1&2&2\\1&1&2&1\\2&3&3&1\\1&3&1&2\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix}=\begin{bmatrix}1&0&0&0\\-1&1&0&0\\-2&0&1&0\\-1&0&0&1\end{bmatrix}\begin{bmatrix}1\\2\\1\\1\end{bmatrix} \Rightarrow \begin{bmatrix}1&1&2&2\\0&0&0&-1\\0&1&-1&-3\\0&2&-1&0\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix}=\begin{bmatrix}1\\1\\-1\\0\end{bmatrix}$$

Note the location and values of the non-zero off-diagonals.

The next elimination step, which involves switching rows 2 and 3, can be written as

$$\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&1&2&2\\0&0&0&-1\\0&1&-1&-3\\0&2&-1&0\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}1\\1\\-1\\0\end{bmatrix}\Rightarrow\begin{bmatrix}1&1&2&2\\0&1&-1&-3\\0&0&0&-1\\0&2&-1&0\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix}=\begin{bmatrix}1\\-1\\1\\0\end{bmatrix}$$

Notice that the necessary matrix multiplier is obtained from switching rows 2 and 3 of the identity matrix.

The third elimination step is as follows:

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&-2&0&1\end{bmatrix}\begin{bmatrix}1&1&2&2\\0&1&-1&-3\\0&0&0&-1\\0&2&-1&0\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&-2&0&1\end{bmatrix}\begin{bmatrix}1\\-1\\1\\0\end{bmatrix}\Rightarrow\begin{bmatrix}1&1&2&2\\0&1&-1&-3\\0&0&0&-1\\0&0&1&6\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix}=\begin{bmatrix}1\\-1\\1\\2\end{bmatrix}$$

We proceed in similar fashion, until the last step which is

$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix}\begin{bmatrix}-17\\4\\8\\1\end{bmatrix}\Rightarrow\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix}=\begin{bmatrix}-17\\4\\8\\-1\end{bmatrix}$$

The pattern should be clear: to find the appropriate matrix for executing any elimination step, take the identity matrix and apply the same elimination step to it.

Writing the matrix multiplier in each step as $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$, etc. we have

$$\begin{aligned}\mathbf{Ax}=\mathbf{b}\ &\Rightarrow\ \mathbf{E}_n\mathbf{E}_{n-1}\cdots\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{Ax}=\mathbf{E}_n\mathbf{E}_{n-1}\cdots\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{b}\\ &\Rightarrow\ \mathbf{Ix}=\mathbf{E}_n\mathbf{E}_{n-1}\cdots\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{b}\end{aligned}$$

Since

$$\mathbf{E}_n\mathbf{E}_{n-1}\cdots\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}=\mathbf{I}$$

It follows that $\mathbf{A}^{-1}=\mathbf{E}_n\mathbf{E}_{n-1}\cdots\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1$.

If the objective is to find the inverse of **A**, we can do the following: write $\mathbf{A}$ and $\mathbf{I}$ side-by-side. Then apply the same row operators to both $\mathbf{A}$ and $\mathbf{I}$ until $\mathbf{A}$ is reduced to the identity matrix. At the same time, the identity matrix will be "reduced" to the inverse matrix.

$$\begin{array}{r|l}\mathbf{A} & \mathbf{I}\\ \mathbf{E}_1\mathbf{A} & \mathbf{E}_1\mathbf{I}\\ \mathbf{E}_2\mathbf{E}_1\mathbf{A} & \mathbf{E}_2\mathbf{E}_1\mathbf{I}\\ & \vdots \\ \underbrace{\mathbf{E}_n\ldots\mathbf{E}_2\mathbf{E}_1\mathbf{A}}_{\text{reduced to }\mathbf{I}} & \underbrace{\mathbf{E}_n\ldots\mathbf{E}_2\mathbf{E}_1\mathbf{I}}_{=\mathbf{A}^{-1}}\end{array}$$

Each of the elimination step (either swapping row positions, subtracting/adding a multiple of one row from/to another row, multiplying a row by a constant) is called an elementary row operation. These are the only steps you need to solve systems of equations or find inverses. Here is the fully worked out example for the matrix

$$\mathbf{A}=\begin{bmatrix}0&2&4\\3&1&2\\6&2&1\end{bmatrix}$$

Writing $[\mathbf{A} \mid \mathbf{I}]$, and applying the necessary elementary row operations, we have the following, where the

$$\left[\begin{array}{ccc|ccc} 0 & 2 & 4 & 1 & 0 & 0 \\ 3 & 1 & 2 & 0 & 1 & 0 \\ 6 & 2 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[{[1]\leftrightarrow[2]}]{} \left[\begin{array}{ccc|ccc} \boxed{3} & 1 & 2 & 0 & 1 & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 6 & 2 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[{[3]=[3]-2[1]}]{} \left[\begin{array}{ccc|ccc} \boxed{3} & 1 & 2 & 0 & 1 & 0 \\ 0 & \boxed{2} & 4 & 1 & 0 & 0 \\ 0 & 0 & \boxed{-3} & 0 & -2 & 1 \end{array}\right]$$

$$\xrightarrow[{\substack{[2]=[2]-\frac{4}{3}[3] \\ [1]=[1]-\frac{2}{3}[3]}}]{} \left[\begin{array}{ccc|ccc} \boxed{3} & 1 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} \\ 0 & \boxed{2} & 0 & 1 & -\frac{8}{3} & \frac{4}{3} \\ 0 & 0 & \boxed{-3} & 0 & -2 & 1 \end{array}\right]$$

$$\xrightarrow[{[1]=[1]-\frac{1}{2}[2]}]{} \left[\begin{array}{ccc|ccc} \boxed{3} & 0 & 0 & -\frac{1}{2} & 1 & 0 \\ 0 & \boxed{2} & 0 & 1 & -\frac{8}{3} & \frac{4}{3} \\ 0 & 0 & \boxed{-3} & 0 & -2 & 1 \end{array}\right]$$

$$\xrightarrow[{\substack{[1]=(1/3)[1] \\ [2]=(1/2)[2] \\ [3]=-(1/3)[3]}}]{} \left[\begin{array}{ccc|ccc} \boxed{1} & 0 & 0 & -\frac{1}{6} & \frac{1}{3} & 0 \\ 0 & \boxed{1} & 0 & \frac{1}{2} & -\frac{4}{3} & \frac{2}{3} \\ 0 & 0 & \boxed{1} & 0 & \frac{2}{3} & -\frac{1}{3} \end{array}\right]$$

The inverse of $\mathbf{A} = \begin{bmatrix} 0 & 2 & 4 \\ 3 & 1 & 2 \\ 6 & 2 & 1 \end{bmatrix}$ is therefore

$$\begin{bmatrix} 0 & 2 & 4 \\ 3 & 1 & 2 \\ 6 & 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{1}{6} & \frac{1}{3} & 0 \\ \frac{1}{2} & -\frac{4}{3} & \frac{2}{3} \\ 0 & \frac{2}{3} & -\frac{1}{3} \end{bmatrix}.$$

You can verify this by computing

$$\begin{bmatrix} -\frac{1}{6} & \frac{1}{3} & 0 \\ \frac{1}{2} & -\frac{4}{3} & \frac{2}{3} \\ 0 & \frac{2}{3} & -\frac{1}{3} \end{bmatrix}\begin{bmatrix} 0 & 2 & 4 \\ 3 & 1 & 2 \\ 6 & 2 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 2 & 4 \\ 3 & 1 & 2 \\ 6 & 2 & 1 \end{bmatrix}\begin{bmatrix} -\frac{1}{6} & \frac{1}{3} & 0 \\ \frac{1}{2} & -\frac{4}{3} & \frac{2}{3} \\ 0 & \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

to show that the product gives the identity matrix.

Obviously this technique also works to solve the equation $\mathbf{Ax} = \mathbf{b}$: write down $\mathbf{A}$ and $\mathbf{b}$ side-by-side and apply to both sides the elementary row operations needed to reduce $\mathbf{A}$ into the identity matrix. The rationale for this is the same as previously:

$$\begin{array}{r|l} \mathbf{A} & \mathbf{b} \\ \mathbf{E}_1\mathbf{A} & \mathbf{E}_1\mathbf{b} \\ \mathbf{E}_2\mathbf{E}_1\mathbf{A} & \mathbf{E}_2\mathbf{E}_1\mathbf{b} \\ & \vdots \\ \underbrace{\mathbf{E}_n \ldots \mathbf{E}_2\mathbf{E}_1\mathbf{A}}_{\text{reduced to } \mathbf{I}} & \underbrace{\mathbf{E}_n \ldots \mathbf{E}_2\mathbf{E}_1\mathbf{b}}_{=\mathbf{A}^{-1}\mathbf{b}} \end{array}$$

If an $(n \times n)$ matrix is not invertible, there will be fewer than n pivots. Consider

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 4 \\ 2 & 2 & 5 \\ 2 & 2 & 7 \end{bmatrix}$$

Applying elimination, we have

$$\left[\begin{array}{ccc|ccc} 2 & 2 & 4 & 1 & 0 & 0 \\ 2 & 2 & 5 & 0 & 1 & 0 \\ 2 & 2 & 7 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[{[3]=[3]-[1]}]{[2]=[2]-[1]} \left[\begin{array}{ccc|ccc} \boxed{2} & 2 & 4 & 1 & 0 & 0 \\ 0 & 0 & \boxed{1} & -1 & 1 & 0 \\ 0 & 0 & 3 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{[3]=[3]-3[2]} \left[\begin{array}{ccc|ccc} \boxed{2} & 2 & 4 & 1 & 0 & 0 \\ 0 & 0 & \boxed{1} & -1 & 1 & 0 \\ 0 & 0 & 0 & 2 & -3 & 1 \end{array}\right] \qquad (*)$$

$$\xrightarrow{[1]=[1]-4[2]} \left[\begin{array}{ccc|ccc} \boxed{2} & 2 & 0 & 5 & -4 & 0 \\ 0 & 0 & \boxed{1} & -1 & 1 & 0 \\ 0 & 0 & 0 & 2 & -3 & 1 \end{array}\right]$$

$$\xrightarrow{[1]=\frac{1}{2}[1]} \left[\begin{array}{ccc|ccc} \boxed{1} & 1 & 0 & \frac{5}{2} & -2 & 0 \\ 0 & 0 & \boxed{1} & -1 & 1 & 0 \\ 0 & 0 & 0 & 2 & -3 & 1 \end{array}\right] \qquad (**)$$

The matrix on the left in (*) is called the row echelon form of the original matrix. The matrix on the left in (**) is the reduced row echelon form of the original matrix. The definitions are:

- The **row echelon form** of a matrix is the matrix obtained by applying elementary row operations such that (i) all zero rows are at the bottom of the matrix; (ii) all first non-zero entry (the 'leading entry', or 'pivot') of each row is strictly to the right of the first pivots in the rows above.
- The **reduced row echelon form** of a matrix is the matrix obtained by applying elementary row operations so that (i) it is in row echelon form, (ii) the pivots are 1, (iii) all other elements in each pivot's column are zero.

The row echelon forms can be obtained for non-square matrices. Many properties of a matrix can be deduced from its row-echelon form.

Properties of the Inverse

(1) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ when the inverses exist.

Premultiplying $\mathbf{B}^{-1}\mathbf{A}^{-1}$ by $\mathbf{AB}$ gives the identity matrix:

$$\mathbf{AB}\,\mathbf{B}^{-1}\mathbf{A}^{-1} = \mathbf{A}\underbrace{\mathbf{B}\,\mathbf{B}^{-1}}_{=\mathbf{I}}\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}.$$

But $\mathbf{AB}\left(\mathbf{B}^{-1}\mathbf{A}^{-1}\right) = \mathbf{I}$ says that $\mathbf{B}^{-1}\mathbf{A}^{-1}$ is the inverse of $\mathbf{AB}$.

(2) $(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1}$

This says that the inverse of a transpose is the transpose of the inverse. It doesn't matter whether you transpose before computing the inverse, or transpose the inverse. We can take $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ as the starting point. Taking transpose, and noting that the transpose of the identity matrix is itself, we have

$$\left(\mathbf{A}^{-1}\mathbf{A}\right)^{\mathrm{T}} = \mathbf{I}' \quad \Leftrightarrow \quad \mathbf{A}^{\mathrm{T}}(\mathbf{A}^{-1})^{\mathrm{T}} = \mathbf{I}.$$

Now premultiply both sides by $(\mathbf{A}^{\mathrm{T}})^{-1}$, we get

$$\underbrace{(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{A}^{\mathrm{T}}}_{=\mathbf{I}}(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{I}$$

which gives us $(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1}$. A corollary is that the inverse of a symmetric matrix is symmetric:
if $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$, then $(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1} = \mathbf{A}^{-1}$.

Determinants are discussed in the next two sections.

(3) $|\mathbf{A}^{-1}| = \dfrac{1}{|\mathbf{A}|}$.

This arises from the fact that the determinant of a product is the product of the determinints: $|\mathbf{A}^{-1}||\mathbf{A}| = |\mathbf{I}|$.

Exercises

1. Find the inverse of the matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 & -3 \\ 2 & 1 & -3 \\ 2 & 2 & 1 \end{bmatrix}$.

2. The matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 & -3 \\ 2 & 1 & -3 \\ 1 & 2 & -6 \end{bmatrix}$ has no inverse. Apply elementary row operations to $\mathbf{A}$ as though finding the inverse. What happens?

3. Write down the simultaneous equations

$$\begin{aligned} 3x_1 + 3x_2 + 2x_3 &= 6 \\ 2x_1 + x_2 + 3x_3 &= 3 \\ x_1 + 5x_3 + 2x_3 &= 4 \end{aligned}$$

in the form $\mathbf{Ax} = \mathbf{b}$. Solve this by

(i) finding the inverse of $\mathbf{A}$ and computing $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$;

(ii) writing down $\mathbf{A}$ and $\mathbf{b}$ side-by-side:

$$\left[\begin{array}{ccc|c} 3 & 3 & 2 & 6 \\ 2 & 1 & 3 & 3 \\ 1 & 5 & 2 & 4 \end{array}\right]$$

and applying the necessary row operations to reduce the left side of the matrix to the identity matrix.

4. Find the inverse of the matrix

$$\mathbf{D} = \begin{bmatrix} 2 & 2 & 3 & 4 \\ 0 & -1 & 0 & 11 \\ 1 & -1 & 0 & 3 \\ -2 & 0 & -1 & 3 \end{bmatrix}.$$

5(a) Let $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 5 & 2 \\ 5 & -1 \end{bmatrix}$.

Find $(\mathbf{AB})^{-1}$ and verify that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

(b) Let $\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 4 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 5 & 2 \\ 5 & -1 \\ -2 & 1 \end{bmatrix}$.

Find $(\mathbf{AB})^{-1}$. Why would it inappropriate to write $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$?

Note that an alternative notation for transpose is the 'prime', i.e., $\mathbf{X}^{\mathrm{T}} = \mathbf{X}'$. You should be comfortable with both notations. In the following questions, we will use the 'prime' notation. Thereafter, we will switch between the two, depending on which looks better.

6. Let $\mathbf{X}$ be a matrix of dimension $(n \times k)$ such $(\mathbf{X}'\mathbf{X})^{-1}$ exists. We want to show that $\left((\mathbf{X}'\mathbf{X})^{-1}\right)' = (\mathbf{X}'\mathbf{X})^{-1}$. Which of the following arguments are correct?

A. $\mathbf{X}'\mathbf{X}$ is symmetric $(\mathbf{X}'\mathbf{X})' = \mathbf{X}'\mathbf{X}'' = \mathbf{X}'\mathbf{X}$. We know that the inverse of a symmetric matrix is symmetric, therefore

$$\left((\mathbf{X}'\mathbf{X})^{-1}\right)' = (\mathbf{X}'\mathbf{X})^{-1}.$$

B. $\left((\mathbf{X}'\mathbf{X})^{-1}\right)' = \left(\mathbf{X}^{-1}(\mathbf{X}')^{-1}\right)' = \left((\mathbf{X}')^{-1}\right)'\left(\mathbf{X}^{-1}\right)' = \left(\mathbf{X}^{-1}\right)(\mathbf{X}')^{-1} = (\mathbf{X}'\mathbf{X})^{-1}$

C. $\left((\mathbf{X}'\mathbf{X})^{-1}\right)' = \left((\mathbf{X}'\mathbf{X})'\right)^{-1} = (\mathbf{X}'\mathbf{X}'')^{-1} = (\mathbf{X}'\mathbf{X})^{-1}$

(a) A only (b) B and C only (c) All three (d) A and C only.

Explain your choice.