2.5: Solving Matrix Equations AX=B

Learning Objectives

- T/F: To solve the matrix equation \(AX = B\), put the matrix \([A; X]\) into reduced row echelon form and interpret the result properly.
- T/F: The first column of a matrix product \((AB)\) is \((A)\) times the first column of \((B)\).
- Give two reasons why one might solve for the columns of \((X)\) in the equation \((AX=B)\) separately.

We began last section talking about solving numerical equations like \((ax=b)\) for \((x)\). We mentioned that solving matrix equations of the form \((AX=B)\) is of interest, but we first learned how to solve the related, but simpler, equations \((A\vec{x} = \vec{b})\). In this section we will learn how to solve the general matrix equation \((AX=B)\) for \((X)\).

We will start by considering the best case scenario when solving \((A\vec{x} = \vec{b})\); that is, when \((A)\) is square and we have exactly one solution. For instance, suppose we want to solve \((A\vec{x} = \vec{b})\) where

\[
\begin{bmatrix}
1 & 1 \\
2 & 1 \\
\end{bmatrix}
\begin{bmatrix}
x \\
b \\
\end{bmatrix}
= \begin{bmatrix}
0 \\
1 \\
\end{bmatrix}
\]

We know how to solve this; put the appropriate matrix into reduced row echelon form and interpret the result.

\[
\begin{bmatrix}
1 & 0 & 1 \\
0 & 1 & -1 \\
\end{bmatrix}
\]

We read from this that

\[
\begin{bmatrix}
x \\
b \\
\end{bmatrix}
= \begin{bmatrix}
1 \\
-1 \\
\end{bmatrix}
\]
Written in a more general form, we found our solution by forming the augmented matrix
\[
\left[ \begin{array}{cc} A & \vec{b} \end{array} \right]
\]
and interpreting its reduced row echelon form:
\[
\left[ \begin{array}{cc} A & \vec{b} \end{array} \right] \quad \text{rref} \quad \left[ \begin{array}{cc} I & \vec{x} \end{array} \right]
\]
Notice that when the reduced row echelon form of \((A)\) is the identity matrix \((I)\) we have exactly one solution. This, again, is the best case scenario.

We apply the same general technique to solving the matrix equation \((AX=B)\) for \((X)\). We’ll assume that \((A)\) is a square matrix \((\text{(B)} \text{ need not be})\) and we’ll form the augmented matrix
\[
\left[ \begin{array}{cc} A & B \end{array} \right].
\]
Putting this matrix into reduced row echelon form will give us \((X)\), much like we found \((\text{vec}(x))\) before.
\[
\left[ \begin{array}{cc} A & B \end{array} \right] \quad \text{rref} \quad \left[ \begin{array}{cc} I & X \end{array} \right]
\]
As long as the reduced row echelon form of \((A)\) is the identity matrix, this technique works great. After a few examples, we’ll discuss why this technique works, and we’ll also talk just a little bit about what happens when the reduced row echelon form of \((A)\) is not the identity matrix.

First, some examples.

Example \((PageIndex{1})\)

Solve the matrix equation \((AX=B)\) where
\[
A=\left[ \begin{array}{cc} 1 & -1 \\ 5 & 3 \end{array} \right], \quad B=\left[ \begin{array}{ccc} -8 & -13 & 1 \\ 32 & -17 & 21 \end{array} \right].
\]

Solution

To solve \((AX=B)\) for \((X)\), we form the proper augmented matrix, put it into reduced row echelon form, and interpret the result.
\[
\left[ \begin{array}{cccc} 1 & -1 & -8 & -13 \\ 5 & 3 & 32 & -17 \end{array} \right] \quad \text{rref} \quad \left[ \begin{array}{cccc} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -7 \end{array} \right]
\]
We read from the reduced row echelon form of the matrix that
We can easily check to see if our answer is correct by multiplying \(AX\).

Example \(\PageIndex{2}\)

Solve the matrix equation \(AX=B\) where

\[
A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -2 \\ 2 & -1 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -1 & 2 \\ 2 & -6 \\ 2 & -4 \end{bmatrix}
\]

Solution

To solve, let’s again form the augmented matrix

\[
\left[ \begin{array}{ccc} A & B \end{array} \right]
\]

put it into reduced row echelon form, and interpret the result.

\[
\left[ \begin{array}{cccccc} 1 & 0 & 2 & -1 & 2 \\ 0 & -1 & -2 & 2 & -6 \\ 2 & -1 & 0 & 2 & -4 \end{array} \right] \quad \text{vec{rref}} \quad \left[ \begin{array}{cccccc} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 \end{array} \right]
\]

We see from this that

\[
X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}
\]

Why does this work? To see the answer, let’s define five matrices.

\[
A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \vec{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} 5 \\ 6 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} 1 & -1 & 5 \\ 1 & 1 & 6 \end{bmatrix}
\]

Notice that \(\langle \vec{u} \rangle\), \(\langle \vec{v} \rangle\) and \(\langle \vec{w} \rangle\) are the first, second and third columns of \(\langle X \rangle\), respectively. Now consider this list of matrix products: \(\langle A \vec{u} \rangle\), \(\langle A \vec{v} \rangle\), \(\langle A \vec{w} \rangle\) and \(\langle AX \rangle\).
So again note that the columns of \( \langle X \rangle \) are \( \langle \vec{u} \rangle, \langle \vec{v} \rangle \) and \( \langle \vec{w} \rangle \); that is, we can write

\[
\langle X \rangle = \left[ \begin{array}{ccc} \vec{u} & \vec{v} & \vec{w} \end{array} \right]
\]

Notice also that the columns of \( \langle AX \rangle \) are \( \langle A\vec{u} \rangle, \langle A\vec{v} \rangle \) and \( \langle A\vec{w} \rangle \), respectively. Thus we can write

\[
\begin{align}
AX &= A\left[ \begin{array}{ccc} \vec{u} & \vec{v} & \vec{w} \end{array} \right] \\
&= \left[ \begin{array}{ccc} A\vec{u} & A\vec{v} & A\vec{w} \end{array} \right] \\
&= \left[ \begin{array}{ccc} \left[ \begin{array}{c} 3 \\ 7 \end{array} \right] & \left[ \begin{array}{c} 1 \\ 1 \end{array} \right] & \left[ \begin{array}{c} 17 \\ 39 \end{array} \right] \end{array} \right]
\end{align}
\]

We summarize what we saw above in the following statement:

The columns of a matrix product \( \langle AX \rangle \) are \( \langle A \rangle \) times the columns of \( \langle X \rangle \).

How does this help us solve the matrix equation \( \langle AX = B \rangle \) for \( \langle X \rangle \)? Assume that \( \langle A \rangle \) is a square matrix (that forces \( \langle X \rangle \) and \( \langle B \rangle \) to be the same size). We’ll let \( \langle \vec{x}_1 \rangle, \langle \vec{x}_2 \rangle, \ldots, \langle \vec{x}_n \rangle \) denote the columns of the (unknown) matrix \( \langle X \rangle \), and we’ll let \( \langle \vec{b}_1 \rangle, \langle \vec{b}_2 \rangle, \ldots, \langle \vec{b}_n \rangle \) denote the columns of \( \langle B \rangle \). We want to solve \( \langle AX = B \rangle \) for \( \langle X \rangle \). That is, we want \( \langle X \rangle \) where

\[
\begin{align}
AX &= B \\
A \left[ \begin{array}{cccc} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{array} \right] &= \left[ \begin{array}{cccc} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_n \end{array} \right]
\end{align}
\]

If the matrix on the left hand side is equal to the matrix on the right, then their respective columns must be equal. This means we need to solve \( n \) equations:

\[
\begin{align}
A\vec{x}_1 &= \vec{b}_1 \\
A\vec{x}_2 &= \vec{b}_2 \\
\vdots &= \vdots \\
A\vec{x}_n &= \vec{b}_n
\end{align}
\]

We already know how to do this; this is what we learned in the previous section. Let’s do this in a concrete example. In our above work we defined matrices \( \langle A \rangle \) and \( \langle X \rangle \), and looked at the product \( \langle AX \rangle \). Let’s call the product \( \langle B \rangle \); that is, set \( \langle B = AX \rangle \). Now, let’s pretend that we don’t know what \( \langle X \rangle \) is, and let’s try to find the matrix \( \langle X \rangle \) that satisfies the equation \( \langle AX = B \rangle \). As a refresher, recall that

\[
\begin{align}
A &= \left[ \begin{array}{cc} 1 & 2 \\ 3 & 4 \end{array} \right] \\
B &= \left[ \begin{array}{ccc} 3 & 1 & 17 \\ 4 & 1 & 39 \end{array} \right]
\end{align}
\]

Since \( \langle A \rangle \) is a \( 2 \times 2 \) matrix and \( \langle B \rangle \) is a \( 2 \times 3 \) matrix, what dimensions must \( \langle X \rangle \) be in the equation \( \langle AX = B \rangle \)? The number of rows of \( \langle X \rangle \) must match the number of columns of \( \langle A \rangle \); the number of columns of \( \langle X \rangle \) must
match the number of columns of \( (B) \). Therefore we know that \( (X) \) must be a \( (2 \times 3) \) matrix.

We’ll call the three columns of \( (X) \) \( \{ \vec{x}_1 \}, \{ \vec{x}_2 \}, \{ \vec{x}_3 \} \). Our previous explanation tells us that if \( (AX=B) \), then:

\[
\begin{align*}
AX &= B \\
A \left[\begin{array}{c} \vec{x}_1 \end{array}\right] &\quad \left[\begin{array}{c} 3 \\ 7 \end{array}\right] \\
A \left[\begin{array}{c} \vec{x}_2 \end{array}\right] &\quad \left[\begin{array}{c} 1 \\ 1 \end{array}\right] \\
A \left[\begin{array}{c} \vec{x}_3 \end{array}\right] &\quad \left[\begin{array}{c} 17 \\ 39 \end{array}\right]
\end{align*}
\]

Hence

\[
\begin{align*}
A \vec{x}_1 &= \left[\begin{array}{c} 3 \\ 7 \end{array}\right] \\
A \vec{x}_2 &= \left[\begin{array}{c} 1 \\ 1 \end{array}\right] \\
A \vec{x}_3 &= \left[\begin{array}{c} 17 \\ 39 \end{array}\right]
\end{align*}
\]

To find \( \{ \vec{x}_1 \} \), we form the proper augmented matrix and put it into reduced row echelon form and interpret the results.

\[
\left[\begin{array}{ccc} 1 & 2 & 3 \\ 3 & 4 & 7 \end{array}\right] \quad \text{rref} \quad \left[\begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right]
\]

This shows us that

\[
\{ \vec{x}_1 \} = \left[\begin{array}{c} 1 \\ 1 \end{array}\right]
\]

To find \( \{ \vec{x}_2 \} \), we again form an augmented matrix and interpret its reduced row echelon form.

\[
\left[\begin{array}{ccc} 1 & 2 & 1 \\ 3 & 4 & 1 \end{array}\right] \quad \text{rref} \quad \left[\begin{array}{ccc} 1 & 0 & -1 \\ 0 & 1 & 1 \end{array}\right]
\]

Thus

\[
\{ \vec{x}_2 \} = \left[\begin{array}{c} -1 \\ 1 \end{array}\right]
\]

which matches with what we already knew from above.

Before continuing on in this manner to find \( \{ \vec{x}_3 \} \), we should stop and think. If the matrix vector equation \( (A \vec{x} = \vec{b}) \) is consistent, then the steps involved in putting

\[
\left[\begin{array}{ccc} A & \vec{b} \end{array}\right]
\]

into reduced row echelon form depend only on \( (A) \); it does not matter what \( (\vec{b}) \) is. So when we put the two matrices

\[
\left[\begin{array}{ccc} 1 & 2 & 3 \\ 3 & 4 & 7 \end{array}\right] \quad \text{and} \quad \left[\begin{array}{ccc} A & \vec{b} \end{array}\right]
\]

UC Davis ChemWiki is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License.
from above into reduced row echelon form, we performed exactly the same steps! (In fact, those steps are:
\((-3R_1+R_2 \rightarrow R_2); (-\frac12R_2 \rightarrow R_2); (-2R_2+R_1 \rightarrow R_1).\)

Instead of solving for each column of \(X\) separately, performing the same steps to put the necessary matrices into reduced row echelon form three different times, why don’t we just do it all at once? Instead of individually putting

\[
\begin{bmatrix}
1 & 2 & 3 \\
3 & 4 & 7
\end{bmatrix}, \quad
\begin{bmatrix}
1 & 2 & 1 \\
3 & 4 & 1
\end{bmatrix} \quad \text{and} \quad
\begin{bmatrix}
1 & 2 & 17 \\
3 & 4 & 39
\end{bmatrix}
\]

into reduced row echelon form, let’s just put

\[
\begin{bmatrix}
1 & 2 & 3 & 1 & 17 \\
3 & 4 & 7 & 1 & 39
\end{bmatrix}
\]

into reduced row echelon form.

\[
\begin{bmatrix}
1 & 2 & 3 & 1 & 17 & 1 & 0 & 0 & 1 & -1 & 5 \\
3 & 4 & 7 & 1 & 39 & 0 & 1 & 1 & 1 & 5 & 0
\end{bmatrix}
\]

By looking at the last three columns, we see \(X\):

\[
\begin{bmatrix}
1 & 2 & 3 & 1 & 17 & 1 & 0 & 0 & 1 & -1 & 5 \\
3 & 4 & 7 & 1 & 39 & 0 & 1 & 1 & 1 & 5 & 0
\end{bmatrix}
\]

Now that we’ve justified the technique we’ve been using in this section to solve \(AX=B\) for \(X\), we reinforce its importance by restating it as a Key Idea.

**Key Idea \(\PageIndex{1}\): Solving \(AX=B\)**

Let \(A\) be an \(n\times n\) matrix, where the reduced row echelon form of \(A\) is \(I\). To solve the matrix equation \(AX=B\) for \(X\),

1. Form the augmented matrix \(\begin{bmatrix}A&B\end{bmatrix}\).
2. Put this matrix into reduced row echelon form. It will be of the form \(\begin{bmatrix}I&X\end{bmatrix}\), where \(X\) appears in the columns where \(B\) once was.

These simple steps cause us to ask certain questions. First, we specify above that \(A\) should be a square matrix. What happens if \(A\) isn’t square? Is a solution still possible? Secondly, we only considered cases where the reduced row echelon form of \(A\) was \(I\) (and stated that as a requirement in our Key Idea). What if the reduced row echelon form of \(A\) isn’t \(I\)? Would we still be able to find a solution? (Instead of having exactly one solution, could we have no solution? Infinite solutions? How would we be able to tell?)

These questions are good to ask, and we leave it to the reader to discover their answers. Instead of tackling these questions, we instead tackle the problem of “Why do we care about solving \(AX=B\)”? The simple answer is that, for now, we only care...
about the special case when \((B=I)\). By solving \((AX=I)\) for \((X)\), we find a matrix \((X)\) that, when multiplied by \((A)\), gives the identity \((I)\). That will be very useful.

**Footnotes**

[1] One reason to do it three different times is that we enjoy doing unnecessary work. Another reason could be that we are stupid...