Elimination and Matrices

Learning Goals: Put elimination in its natural matrix formulation, and connect it with matrix multiplication.

When doing elimination, it is wasteful to keep writing down all the variables. Instead of
\[ 2x + y + z = 1 \quad 2x + y + z = 1 \]
\[ 4x + y = -2 \rightarrow -y - 2z = -4 \], we might as well only keep the numbers, and write
\[-2x + 2y + z = 7 \quad -2x + 2y + z = 7 \]
down the matrix system
\[
\begin{bmatrix}
2 & 1 & 1 & 1 \\
4 & 1 & 0 & -2 \\
-2 & 2 & 1 & 7
\end{bmatrix}
\rightarrow
\begin{bmatrix}
2 & 1 & 1 & 1 \\
0 & -1 & -2 & -4 \\
-2 & 2 & 1 & 7
\end{bmatrix}
\]
what we do. There are two matrices associated with the system of equations:

- \[
\begin{bmatrix}
2 & 1 & 1 \\
4 & 1 & 0 \\
-2 & 2 & 1
\end{bmatrix}
\]
is the matrix of coefficients
- \[
\begin{bmatrix}
2 & 1 & 1 & 1 \\
4 & 1 & 0 & -2 \\
-2 & 2 & 1 & 7
\end{bmatrix}
\]
is called the augmented matrix for this system. The dashed line in there is artificial, only there to separate the left-hand side coefficients from the right-hand side.

Now we have already seen that multiplying the matrix times the column
\[
\begin{bmatrix}
2 & 1 & 1 \\
4 & 1 & 0 \\
-2 & 2 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
\]
creates a linear combination of the columns: \[ x + y + z = 2x + y + z \]. In other words, we used the necessity of arriving at the given linear system to define matrix multiplication. We can view this two different ways:

- We take a linear combination of the columns of the matrix on the left, weighted by the entries of the column on the right
- We note that each entry in the product is the dot product of the corresponding row in the matrix on the left with the column

Now it frequently happens that we want to use the same coefficients with different right-hand sides. Perhaps the new right hand side is \((7, 6, 5)\). Then we might write both systems simultaneously as:
\[
\begin{bmatrix}
2 & 1 & 1 \\
4 & 1 & 0 \\
-2 & 2 & 1 \\
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
\end{bmatrix}
= 
\begin{bmatrix} 1 \\
-2 \\
7 \\
\end{bmatrix} \quad \text{or simply} \quad
\begin{bmatrix}
2 & 1 & 1 \\
4 & 1 & 0 \\
-2 & 2 & 1 \\
\end{bmatrix}
\begin{bmatrix}
x & a \\
y & b \\
z & c \\
\end{bmatrix}
= 
\begin{bmatrix} 1 & 7 \\
-2 & 6 \\
7 & 5 \\
\end{bmatrix}.
\]

For this to work, we have to define matrix multiplication a specific way:

- We can multiply with a matrix on the left by a matrix of several columns on the right by doing the matrix-times-column thing for each column, and set the columns next to each other.

But we can also look at the system row by row. The first row of the multiplication must yield \([2x + y + z \ 2a + b + c]\) because the first rows of the corresponding systems are \(2x + y + z = 1\) and \(2a + b + c = 7\) respectively. But note that the given row vector can be split apart as \(2[x a] + 1[y b] + 1[z c]\). In other words, the first row of the product is a linear combination of the rows of the matrix on the right, weighted by the entries in the first row of the matrix on the left. This is just the dual point of view of the matrix-times-column view we had before.

Thus we now know how to multiply matrices, and we have three viewpoints for it. To multiply \(AB\) we have

- Column picture: each column of the product is a linear combination of the columns of \(A\), weighted by the entries in the column of \(B\) corresponding to the column of the output we’re finding
- Row picture: each row of the product is a linear combination of the rows of \(B\), weighted by the entries in the row of \(A\) corresponding to the row of the output we’re finding
- Entry-by-entry: each entry in the product is the dot product of the row of \(A\) corresponding to the row of the product with the column of \(B\) corresponding to the column of the product. That is, if \(AB = C\), then \(c_{ij} = (\text{row } i \text{ of } A) \cdot (\text{column } j \text{ of } B)\). We can write this in fancy Σ-notation as \(c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}\) where \(A\) has \(n\) columns and \(B\) has \(n\) rows.