Computing the rank and nullity of a matrix

Example:

\[
A = \begin{bmatrix}
1 & 2 & 3 & 4 \\
1 & 2 & 3 & 5 \\
1 & 2 & 3 & 6 \\
1 & 2 & 3 & 7 \\
1 & 2 & 3 & 8 \\
\end{bmatrix}
\]

By definition, the rank of \( A \) is the dimension of the image of \( f \) and the nullity is the dimension of the kernel of \( f \), where \( f \) is the linear mapping defined by \( f : \mathbb{R}^4 \to \mathbb{R}^5 \) defined by \( f(v) = Av \), interpreting \( v \) as a \( 4 \times 1 \) matrix. The image of \( f \) is the column space of \( A \), which we know has dimension equal to the dimension of the row space of \( A \). In class on Friday I described an algorithm for finding a basis for the row space: Let the rows be \( R_1, R_2, R_3, R_4, R_5 \). We wish to modify these rows by elementary operations without changing the span of these vectors, but in such a way that we can easily pick out a basis for the row space. The operations are these: interchange two rows, or multiply a row by a constant, or add a multiple of one row to another row. Do these operations in such a way that we obtain either zero vectors or vectors with unique leading terms (positions of first non-zero entry are all different). Here’s what happens to our example: First, subtract the first row from the others. This yields

\[
\begin{bmatrix}
1 & 2 & 3 & 4 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 2 \\
0 & 0 & 0 & 3 \\
0 & 0 & 0 & 4 \\
\end{bmatrix}
\]

Next, subtract 2 copies of row 2 from row 3, 3 copies of row 2 from row 4, and 4 copies of row 2 from row 5. This yields

\[
\begin{bmatrix}
1 & 2 & 3 & 4 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

So now we can see that the row space is spanned by the first two rows, and they are obviously linearly independent. Hence the rank of \( A \) is 2.
The same calculations we made in finding a basis for the row space can be used to find a basis for the kernel of \( f \). By definition, the kernel of \( f \) is

\[
\begin{cases}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
\end{bmatrix} \\
\begin{bmatrix}
1 & 2 & 3 & 4 \\
1 & 2 & 3 & 5 \\
1 & 2 & 3 & 6 \\
1 & 2 & 3 & 7 \\
1 & 2 & 3 & 8 \\
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
\end{bmatrix} = \begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
\end{bmatrix}.
\end{cases}
\]

This is equivalent to solving the system of equations

\[
\begin{align*}
1x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\
1x_1 + 2x_2 + 3x_3 + 5x_4 &= 0 \\
1x_1 + 2x_2 + 3x_3 + 6x_4 &= 0 \\
1x_1 + 2x_2 + 3x_3 + 7x_4 &= 0 \\
1x_1 + 2x_2 + 3x_3 + 8x_4 &= 0.
\end{align*}
\]

Each of the row operations we performed above we can perform on this system of equations without changing the solution set. We will arrive at the system

\[
\begin{align*}
1x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\
0x_1 + 0x_2 + 0x_3 + 1x_4 &= 0 \\
0x_1 + 0x_2 + 0x_3 + 0x_4 &= 0 \\
0x_1 + 0x_2 + 0x_3 + 0x_4 &= 0 \\
0x_1 + 0x_2 + 0x_3 + 0x_4 &= 0.
\end{align*}
\]

We can classify all the variables now into two types: those that appear as a leading term in one of the equations, and those that don’t. In our case, \( x_1 \) and \( x_4 \) appear as leading terms, and \( x_2 \) and \( x_3 \) don’t. We can choose values for \( x_2 \) and \( x_3 \) independently and at random, and these determine the values of the leading variables. Setting \( x_2 = a \) and \( x_3 = b \), the equations now read

\[
\begin{align*}
1x_1 + 2a + 3b + 4x_4 &= 0 \\
1x_4 &= 0.
\end{align*}
\]
Working from the bottom equation to the top one, we see that $x_4 = 0$ and $x_1 = -2a - 3b$. So the typical element in the kernel of $f$ is

$$
\begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3 \\
  x_4 \\
\end{bmatrix} =
\begin{bmatrix}
  -2a - 3b \\
  a \\
  b \\
  0 \\
\end{bmatrix}.
$$

This can be resolved into

$$
\begin{bmatrix}
  -2a \\
  a \\
  0 \\
\end{bmatrix} + \begin{bmatrix}
  -3b \\
  0 \\
  b \\
\end{bmatrix} = a \begin{bmatrix}
  -2 \\
  1 \\
  0 \\
\end{bmatrix} + b \begin{bmatrix}
  -3 \\
  0 \\
  1 \\
\end{bmatrix}.
$$

Therefore $\ker f$ is spanned by the set \{ $\begin{bmatrix}
  -2 \\
  1 \\
  0 \\
\end{bmatrix}$, \begin{bmatrix}
  -3 \\
  0 \\
  1 \\
\end{bmatrix}$ \}. The positions of the 1s in these vectors force them to be linearly independent. Hence this is a basis for the kernel of $f$ and the nullity is 2.