The Rank and the Reduced Row-Echelon Form

Learning Goals: continue the exploration of the reduced row echelon form and introduce the rank of a matrix.

As we have seen, we can reduce any matrix via elementary row operations to a unique reduced row-echelon form. Our two examples have been

\[ A = \begin{bmatrix} 2 & 3 & -1 & 2 & 0 \\ -4 & -6 & -2 & -1 & 4 \end{bmatrix} \]

reduced to

\[ R = \begin{bmatrix} 1 & 3/2 & -1/2 & 0 & -4/3 \\ 0 & 0 & 0 & 1 & 4/3 \end{bmatrix} \]

and

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

reduced to

\[ R = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \]

For the record we note a few things:

- \( R \), as stated, is unique. No matter what order we do our various operations, including swapping the rows around at will, we always end up with the same \( R \).
- Columns never move around. The columns that end up with pivots in \( R \) are the columns in \( A \) that had the pivots in the first place.
- The columns without pivots led to free variables, and from \( R \) we can read off the “special” solutions to \( Ax = 0 \). Note that \( Ax = 0 \) if and only if \( Rx = 0 \).

Since the matrix \( R \) does not depend on the order in which we reduce \( A \), only upon \( A \) itself, its properties must be especially related to \( A \). Such properties are called invariants in mathematics. Invariants are things about a mathematical object that appear the same no matter how the object may be disguised. Most of mathematics is the hunt for invariants that are easy to compute and good at distinguishing one object from another!

Since we do elementary matrix operations to turn \( A \) into \( R \), we can effect these by means of elementary matrices multiplied times \( A \). If we put all of these together into one matrix, we find that \( EA = R \). The matrix \( E \) may not be unique, even though \( R \) is (for instance, in the second example above, we could swap rows 4 and 5 and we’d never know the difference!). We could turn this around and write \( A = E^{-1}R \). This is similar to \( A = LU \) (\( R \) taking the place of \( U \)). Unfortunately, we can’t just put the multipliers into \( E^{-1} \) as we did for \( L \), because now we have to eliminate upward as well and that will ruin the earlier entries. We can find \( E \) directly as we compute \( R \), because \( EA = R \) and \( EI = E \), so \( E[A I] = [RE] \). Thus we augment \( A \) with \( I \) and reduce, just as in Gauss-Jordan elimination, and we end with \( E \). In fact, this is Gauss-Jordan elimination applied to a possibly non-invertible, perhaps even non-square matrix. This also tells us that if \( A \) is invertible, the \( R = I \) and \( E = A^{-1} \). In the general case, though, the best we can do is to reduce just the pivot columns to ones and zeroes.

The rank

Taking something simpler than the entire matrix \( R \), we make the following definition:

**Definition:** the rank, \( r \), of a matrix is the number of pivots.
This is clearly the number of non-zero rows of $R$. It is also clearly the number of columns of $R$ (and hence of $A$) that contain pivots. In our above examples, the ranks are 2 and 3 respectively.

The rank tells a lot about a matrix. First, the rank must be less than the number of rows and the number of columns: $r \leq m$ and $r \leq n$. When it is equal, good things happen:

- If $r = n$, the number of columns, then there are no free variables. There is a unique solution to $Ax = 0$.
- If $r = m$, the number of rows, then there is a pivot in every row, and we can therefore solve for any right hand side.
- If $r = m = n$, then both of the above are true. In this case, $A$ is invertible.

Note how two of the conditions for invertibility are contained in the first two statements above, although for non-square matrices we can’t have inverses (we can have one-sided inverses, and these conditions tell us when they are available, but more on that later). For square matrices, the three conditions are really equivalent. The first is the uniqueness of solutions to $Ax = 0$, the second is solvability for any right hand side, and the third is having a complete set of pivots, which we’ve already shown are all the same thing.

At the other extreme, we have rank 0 and rank 1 matrices. Rank 0 isn’t very interesting. The matrix has no pivots—it must therefore be the zero matrix.

Rank one matrices have every row a multiple of just one row. This is because when we do elimination, everything below this row must vanish, otherwise we’d have another pivot. Since each row is just a multiple of one single row, we can factor the matrix in a special way.

For example:

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

Note how the pivot can be volleyed back and forth between the column and the row. Note that we can also think of this as a multiplication by columns. So in a rank one matrix, each column is also a multiple of one of the columns.

The rank also tells us how many “independent” rows there are, and gives us the dimension of the “row space” which is simply all linear combinations of the rows. We will talk more about this, too, later.

Finally, the rank tells us exactly how many special vectors there are in the nullspace. There is one for each non-pivot column, so there must be $n - r$ of them.

### The pivot columns

The columns with pivots in them are special. First, we notice that no combination of pivot columns can ever be zero (except the trivial combination, of course). Why? We know that the solutions to $Ax = 0$ and $Rx = 0$ are exactly the same. If we are combining only the pivot columns, we set all the free variables to zero. But then back-substituting shows that the rest of the variables are zero also. So the only combination of the pivot columns of $R$ that gives zero is the trivial combination. And thus the same applies to $A$. 
Note that the column spaces of $R$ and $A$ are probably different (note that $R$ has bunches of zeroes in the bottom rows) but their nullspaces are the same.

The exact relation is that since $A = E^{-1}R$, the pivot columns of $A$ are exactly the first $r$ columns of $E^{-1}$. Why? Because a pivot column of $A$ is $E^{-1}$ times the corresponding pivot column of $R$. But that column is $[0 \ 0 \ \ldots \ 0 \ 1 \ \ldots \ 0]^T$, where the 1 is in each of the first $r$ slots for each of the $r$ pivot columns. And this multiplication gives us the particular column of $E^{-1}$.

The second thing to note is that each non-pivot column of $A$ is a linear combination of earlier pivot columns of $A$. For each non-pivot column corresponds to a free variable. Set that free variable to one and the rest to zero to produce our special solution. This is a vector $n$ such that $An = 0$. Hence this is a combination of columns that adds to 0. But note that we get $n$ by back substituting in $Rn = 0$, so we get numbers in the earlier basic variables. So we have a combination of pivot columns plus one times the non-pivot column that adds to zero, so we put all the pivot columns on the other side and we have solved for the non-pivot column.

**Example:** Find a combination of the first two columns of
\[
\begin{bmatrix}
1 & 2 & 3 & 4 \\
2 & 4 & 6 & 8 \\
1 & 1 & 1 & 1 \\
1 & 3 & 5 & 7 \\
2 & 3 & 4 & 6 \\
\end{bmatrix}
\]
that adds to the third column (which is a non-pivot column). Well, we have $R = \begin{bmatrix}
1 & 0 & -1 & 0 \\
0 & 1 & 2 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{bmatrix}$. From this, we look at the “special” solution in the third column, which tells us that the third column of $R$ (and hence also $A$) is equal to twice the second minus the first column.

**Reading:** 3.3

**Problems:** 3.3: 1, 2, 4, 6, 8, 10, 12, 16, 17, 18, 19, 20, 22, 24, 25, 27, 28