Row operations, invertible matrices, row reduced echelon form

Last time: Let A be a matrix and let B be the matrix where we take A and double the first row:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} 2a & 2b \\ c & d \end{bmatrix}.$$

Then A and B have the same kernel. The image of B is the set of vectors $\left[\begin{smallmatrix} 2x \\ y \end{smallmatrix}\right]$ for $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ in the image of A.

Last time: Let A be a matrix and let B be the matrix where we take A and double the first row:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} 2a & 2b \\ c & d \end{bmatrix}.$$

Then A and B have the same kernel. The image of B is the set of vectors $\left[\begin{smallmatrix} 2x \\ y \end{smallmatrix}\right]$ for $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ in the image of B.

We can write this more clearly in terms of the matrix $D = \left[\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. So $B = DA$.

Our statements are

$$\mathrm{Ker}(DA) = \mathrm{Ker}(A) \text{ and } \mathrm{Im}(DA) = D\mathrm{Im}(A).$$

The key property of D is that it is ***invertible***. This means that there is a matrix C, called the ***inverse*** of D, with

$$CD = \mathrm{Id} \text{ and } DC = \mathrm{Id}.$$

We write $C = D^{-1}$.

In this case where $D = \left[\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, the inverse matrix is $\left[\begin{smallmatrix} 1/2 & 0 \\ 0 & 1 \end{smallmatrix}\right]$.

We'll see soon that only square matrices can have inverses.

Can you show?

Theorem If A and B are invertible, then AB is invertible.

Theorem A matrix U can only have one inverse. In other words, if $UV = UW = \mathrm{Id}$ and $VU = WU = \mathrm{Id}$, then $V = W$.

Theorem If A and B are invertible, then AB is invertible.

Proof: We claim that $B^{-1}A^{-1}$ is the inverse of AB. Indeed:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A\mathrm{Id}A^{-1} = AA^{-1} = \mathrm{Id} \text{ and}$$

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}\mathrm{Id}B = B^{-1}B = \mathrm{Id}. \quad \square$$

The inverse of "put on your socks, put on your shoes" is "take off your shoes, take off your socks".

Theorem A matrix U can only have one inverse. In other words, if $UV = UW = \mathrm{Id}$ and $VU = WU = \mathrm{Id}$, then $V = W$.

Proof: Consider VUW. We have $VUW = (VU)W = \mathrm{Id}W = W$ but also $VUW = V(UW) = V\mathrm{Id} = V$ so $W = V$. $\square$

Now, back to our main point:

Theorem Let E be an invertible $m \times m$ matrix and let A be any $m \times n$ matrix. Then $\text{Ker}(EA) = \text{Ker}(A)$ and $\text{Im}(EA) = E\text{Im}(A)$.

In other words, $EA\vec{x} = \vec{0}$ if and only if $A\vec{x} = \vec{0}$. For $\vec{y}$ in $\mathbb{R}^m$, there is an $\vec{x}$ with $A\vec{x} = \vec{y}$ if and only if there is a $\vec{w}$ with $EA\vec{w} = E\vec{y}$.

Theorem Let E be an invertible $m \times m$ matrix and let A be any $m \times n$ matrix. Then $\mathrm{Ker}(EA) = \mathrm{Ker}(A)$ and $\mathrm{Im}(EA) = E\mathrm{Im}(A)$.

In other words, $EA\vec{x} = \vec{0}$ if and only if $A\vec{x} = \vec{0}$. For $\vec{y}$ in $\mathbb{R}^m$, there is an $\vec{x}$ with $A\vec{x} = \vec{y}$ if and only if there is a $\vec{w}$ with $EA\vec{w} = E\vec{y}$.

Proof: If $A\vec{x} = \vec{0}$ then $EA\vec{x} = E\vec{0} = \vec{0}$. On the other hand, if $EA\vec{x} = \vec{0}$, then $E^{-1}EA\vec{x} = \mathrm{Id}A\vec{x} = A\vec{x}$ so $A\vec{x} = E^{-1}\vec{0} = \vec{0}$.

If $A\vec{x} = \vec{y}$ then $EA\vec{x} = E\vec{y}$. Conversely, if $EA\vec{x} = E\vec{y}$ then $E^{-1}EA\vec{x} = E^{-1}E\vec{y}$, meaning that $A\vec{x} = \vec{y}$. $\square$

Theorem Let E be an invertible $m \times m$ matrix and let A be any $m \times n$ matrix. Then $\mathrm{Ker}(EA) = \mathrm{Ker}(A)$ and $\mathrm{Im}(EA) = E\mathrm{Im}(A)$.

Take

$$E = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & c & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & & & & & 1 \end{bmatrix} \qquad c \neq 0.$$

We get the row operation of multiplying a single row by a nonzero scalar.

This is one of three row operations:

1. Multiply a row by a nonzero scalar.
2. Switch two rows.
3. Add a multiple of one row to another row.

All of the row operations are multiplication by invertible matrices:

1. Multiply a row by a nonzero scalar.

$$\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$

2. Switch two rows.

$$\begin{bmatrix} 0 & & & & 1 \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ 1 & & & & 0 \end{bmatrix}$$

3. Add a multiple of one row to another row.

$$\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ a & & & & 1 \end{bmatrix}$$

Thus, if we get from one matrix A to another matrix B by row operations, then $\mathrm{Ker}(A) = \mathrm{Ker}(B)$ and the images of A and B are related in a natural way. There will be an invertible matrix U with $B = UA$.

Thus, if we get from one matrix A to another matrix B by row operations, then $\mathrm{Ker}(A) = \mathrm{Ker}(B)$ and the images of A and B are related in a natural way. There will be an invertible matrix U with $B = UA$.

This raises the natural question: How nice can we make a matrix, using row operations? The answer is row reduced echelon form.

Row reduced echelon form

- Either row is either all 0's, or else its first nonzero entry is a 1. This 1 is called a ***pivot***.
- In a column which contains a pivot, called a ***pivot column***, all the other entries are 0.
- The nonzero rows are at the top of the matrix; they are ordered so that the pivots go from left to right as we go down the rows.

$$\begin{bmatrix} \boxed{1} & 3 & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix} \qquad \begin{bmatrix} \boxed{1} & 0 \\ 0 & \boxed{1} \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} \boxed{1} & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Putting a matrix into row reduced echelon form

To make a matrix into row reduced echelon form (rref), we work from left to right. Look at the leftmost column which is not yet a pivot column, and which has a nonzero in a non-pivot row.

- Rescale that entry to be 1.
- Subtract appropriate multiples of the row with the 1 from other rows to make the other entries of that column be 0.
- Switch rows, if needed, to put that row immediately below the already existing pivot rows.

$$\begin{bmatrix} 3 & 6 & 3 & 12 \\ 1 & 2 & 4 & 13 \\ 2 & 4 & 4 & 14 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 2 & 1 & 4 \\ 1 & 2 & 4 & 13 \\ 2 & 4 & 4 & 14 \end{bmatrix} \rightsquigarrow \begin{bmatrix} \boxed{1} & 2 & 1 & 4 \\ 0 & 0 & 3 & 9 \\ 0 & 0 & 2 & 6 \end{bmatrix}$$

$$\rightsquigarrow \begin{bmatrix} \boxed{1} & 2 & 1 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 2 & 6 \end{bmatrix} \rightsquigarrow \begin{bmatrix} \boxed{1} & 2 & 0 & 1 \\ 0 & 0 & \boxed{1} & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Vocabulary related to rref

- The initial leading 1's are called ***pivots***. The columns that contain them are called ***pivot columns***; the corresponding variables in our system of equations are called ***pivot variables***.
- The columns/variables which are not pivot columns/variables are called ***free columns/variables***.
- The number of pivots is called the ***rank***.