Math 1180

1180-09
notes 6

Read: 127-135

We start by repeating some material from the previous notes, for further emphasis and application. First let's recall what a linear combination is.

Definition 1 *Suppose that* $\mathbf{v}_1, ..., \mathbf{v}_n$ *are vectors in* R^n *and* $c_1, ..., c_n$ *are scalars. Then*

$$c_1\mathbf{v}_1 + \cdots c_n\mathbf{v}_n$$

is a linear combination of the vectors $\mathbf{v}_1, ..., \mathbf{v}_n$*.*

Here is an example: Let $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ and $\mathbf{v}_3 = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$. Also let $c_1 = 1, c_2 = -3, c_3 = 2$. Then

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - 3\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + 2\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}.$$

We then have the following statement, also from last time:

Proposition 2 *If* A *is* $m \times n$ *and* B *is* $n \times q$*, so that* AB *is defined, then each column of* AB *is a linear combination of columns of* A*.*

Here is a 3×3 example.

$$A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 1 & -2 \\ -3 & 1 \\ 2 & -3 \end{pmatrix}.$$

Then

$$AB = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -3 & 1 \\ 2 & -3 \end{pmatrix} = \begin{pmatrix} 3 & -19 \\ 3 & -23 \\ 3 & -27 \end{pmatrix}.$$

Notice that the first column of AB is the linear combination we found before.

Looking at the matrix B, we would expect that the second column is given by $-2\mathbf{v}_1 + \mathbf{v}_2 - 3\mathbf{v}_3$. And this is true:

$$-2\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} - 3\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} -19 \\ -23 \\ -27 \end{pmatrix}. \qquad (1)$$

This can also be looked at from the point of view of rows. For example

$$1\,(1,-2) + 4\,(-3,1) + 7\,(2,-3) = \begin{pmatrix} 3 & -19 \end{pmatrix}.$$

Proposition 3 *The rows of AB are linear combinations of the rows of B.*

We now use this second property to develop a connection between matrix multiplication and Guassian elimination. The elementary row operation on a matrix A turns out to be equivalent to multiplying A by a certain kind of matrix called an "elementary matrix."

Definition: *An elementary matrix is a matrix obtained by doing exactly one row operation on the identity matrix I.*

This includes multiplying a row of the identity by 1, so I is itself an elementary matrix.

Example: Consider the 3×3 identity. Add two times the first row to the second. This gives

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This is an elementary matrix. Call it E.

We now multiply another matrix A by E. Let

$$A = \begin{pmatrix} 1 & -3 & 5 \\ -2 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix}.$$

Then

$$EA = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -3 & 5 \\ -2 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -3 & 5 \\ 0 & -5 & 11 \\ 3 & 2 & 1 \end{pmatrix}.$$

Notice that by this operation we have added twice the first row of A to the second row.

This is a general principle: Multiplying A by an elementary matrix is the same as doing the equivalent row operation on A.

This means that Gaussian elimination is the same as multiplying by a sequence of elementary matrices. We continue on with this example. The sequence of Gaussian elimination steps to get to echelon form is

$$A = \begin{pmatrix} 1 & -3 & 5 \\ -2 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -3 & 5 \\ 0 & -5 & 11 \\ 3 & 2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -3 & 5 \\ 0 & -5 & 11 \\ 0 & 11 & -14 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & -3 & 5 \\ 0 & -5 & 11 \\ 0 & 0 & \frac{51}{5} \end{pmatrix}$$

Expressing this as matrix multiplications, let

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}, E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{11}{5} & 1 \end{pmatrix}.$$

Then

$$E_1 A = \begin{pmatrix} 1 & -3 & 5 \\ 0 & -5 & 11 \\ 3 & 2 & 1 \end{pmatrix}, \quad E_2 E_1 A = \begin{pmatrix} 1 & -3 & 5 \\ 0 & -5 & 11 \\ 0 & 11 & -14 \end{pmatrix}$$
$$E_3 E_2 E_1 A = \begin{pmatrix} 1 & -3 & 5 \\ 0 & -5 & 11 \\ 0 & 0 & \frac{51}{5} \end{pmatrix}.$$

Note the order in which the E's appear. We have put the matrix in row echelon form.

This works whether or not the matrix A is square. The matrices E_j are all square, of size $m \times m$ where A is $m \times n$.

The other two elementary row operations are also represented by elementary matrices. For example,

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

interchanges the first and third rows, and

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

multiplies the second row by three.

We can continue the example above further, in order to get the reduced echelon form of the matrix. However the arithmetic would get pretty complicated, so we'll switch to some simpler examples. First a 2×2. Let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

If

$$E_1 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix},$$

then

$$E_1 A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

which is in echelon form. Let

$$E_2 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

We got this by subtracting the second row of $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ from the first. Then

$$E_2 E_1 A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which is in reduced echelon form. But it also happens to be the 2×2 identity I. Notice what happened:

$$E_2 E_1 A = I.$$

Therefore,

$$E_2 E_1 = A^{-1}.$$

Computing this:

$$E_2 E_1 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

We showed that this was the inverse of A in the previous section of notes.

Let's do a 3×3 example:

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -1 & 0 \\ 1 & -2 & 2 \end{pmatrix}.$$

Then:

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_1 A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 1 & -2 & 2 \end{pmatrix},$$

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad E_2 E_1 A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & -1 & 1 \end{pmatrix},$$

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad E_3 E_2 E_1 A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{pmatrix}, \text{ in row echelon form, but keep going}$$

$$E_4 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_4 E_3 E_2 E_1 A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{pmatrix}, \tag{2}$$

$$E_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ (multiplies last row by } -1), \; E_5 E_4 E_3 E_2 E_1 A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}, \tag{3}$$

$$E_6 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_7 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_7 E_6 E_5 E_4 E_3 E_2 E_1 A = I \tag{4}$$

This is nice theoretically. We have proved that A is nonsingular, and that

$$A^{-1} = E_7 E_6 E_5 E_4 E_3 E_2 E_1.$$

The problem is, now we have to multiply all those matrices together, and this is a lot of work. So in practice, this is not how we find the inverse. What we need to do is keep track of the products of the $E's$ as we go.

0.1 Method for finding A^{-1}

The trick is to form a "super augmented matrix", adding the matrix I to the right of A, and put this matrix in reduced echelon form. Start with

$$(A|I) = \begin{pmatrix} 1 & -1 & 1 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 & 1 & 0 \\ 1 & -2 & 2 & 0 & 0 & 1 \end{pmatrix}.$$

Do the first Gaussian step:

$$\begin{pmatrix} 1 & -1 & 1 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 & 1 & 0 \\ 1 & -2 & 2 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 1 & -2 & 2 & 0 & 0 & 1 \end{pmatrix}.$$

Since this step is determined by A, it is the same as before: subtract twice the first row from the second. So we have multiplied the whole 3×6 matrix by E_1 :

$$E_1 (A|I) = (E_1 A|E_1 I) = (E_1 A|E_1).$$

Now do the second Gaussian step, which is the same as multiplying by E_2. But we don't multiply, we do Gaussian elimination:

$$\to \begin{pmatrix} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & -1 & 1 & -1 & 0 & 1 \end{pmatrix} = (E_2 E_1 A \mid E_2 E_1).$$

Notice that we are keeping track of the products of the $E's$ on the right. So if we keep this up, we will get

$$(E_7 E_6 E_5 E_4 E_3 E_2 E_1 A|E_7 E_6 E_5 E_4 E_3 E_2 E_1) = (I, E_7 E_6 E_5 E_4 E_3 E_2 E_1) = \left(I|A^{-1}\right).$$

Thus, continuing from before,

$$\to \begin{pmatrix} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & -1 & -3 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 & -1 & 1 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & -1 & -3 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 & -1 & 1 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & 3 & -1 & -1 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & 0 & 2 & 0 & -1 \\ 0 & 1 & 0 & 4 & -1 & -2 \\ 0 & 0 & 1 & 3 & -1 & -1 \end{pmatrix} = \left(I, A^{-1}\right),$$

where this combined two steps, since both dealt with the 3rd column and so don't interfere with each other. Checking,

$$\begin{pmatrix} 2 & 0 & -1 \\ 4 & -1 & -2 \\ 3 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 2 & -1 & 0 \\ 1 & -2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$