Math 1180

1180-09
notes 7
Homework at end.

There is an exam next Friday, Oct. 2.

How can we tell if the inverse exists? This is easy from what we just did.

Theorem 1 *If A is an $n \times n$ matrix, then A^{-1} exists if and only if some row echelon form of A has n pivots.*

Proof. If A^{-1} exists, then there are matrices $E_1, ..., E_k$ such that

$$E_k E_{k-1} \cdots E_2 E_1 A = I.$$

This means that I is a row echelon form of A, and clearly, I has n pivots.

On the other hand, if $A \to B$ and B is in row echelon form and has n pivots, then each column has a pivot, which means that further Gaussian elimination steps lead from B to I. Thus, translating these steps into matrix multiplications, we get

$$E_k E_{k-1} \cdots E_1 A = I,$$

where k was the total number of steps needed to get from A to I. But then,

$$A^{-1} = E_k \cdots E_1. \tag{1}$$

■

Earlier we said that a matrix was "nonsingular" if it has an inverse. There is another term used for this as well:

Definition 2 *An $n \times n$ matrix A is called "invertible" if A^{-1} exists.*

There is another consequence of formula above. First we have the following easy fact:

Proposition 3 *If A and B are $n \times n$ matrices and each is invertible, then AB is invertible, and*

$$(AB)^{-1} = B^{-1} A^{-1}.$$

Proof. This follows from a simple calculation and the definition of inverse:

$$(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AIA^{-1} = I.$$

Extending that the the product of several matrices, we see from (1) that $A = E_1^{-1}E_2^{-1}\cdots E_k^{-1}$. ■

Hence, any invertible matrix is the product of elementary matrices.

Now recall that if some row echelon form of A has n pivots, then a system of equations

$$A\mathbf{x} = \mathbf{b}$$

has a unique solution for every $\mathbf{b}$. Furthermore, this is true if and only if $A\mathbf{x} = \mathbf{0}$ has only the zero solution.

We therefore have the following theorem:

Theorem 4 *If A is an $n \times n$ matrix, then the following statements are equivalent. (If one is true, then they all are true.)*

(1) A is invertible
(2) $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.
(3) $A\mathbf{x} = \mathbf{0}$ has only the zero solution.
(4) Some row echelon form of A has n pivots.
(5) The rank of A is n
(6) The reduced row echelon form of A is I.
(7) A can be written as the product of elementary matrices.

We will add to this set of equivalent statements later.

0.1 Upper triangular matrices.

Definition 5 *An $n \times n$ matrix A=(a_{ij}) is called "upper triangular" if $a_{ij} = 0$ whenever $i > j$.*

One example makes this obvious:

$$A = \begin{pmatrix} 2 & 3 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is an upper triangular matrix. For example, $a_{21} = 0$, $a_{32} = 0$, $a_{43} = 0$. In fact, any row echelon form of a square matrix is upper triangular.

Definition 6 *An $n \times n$ matrix A is called "lower triangular" if $a_{ij} = 0$ whenever $j > i$.*

Example:

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

Everything above the diagonal is zero.

These two kinds of matrices are important in designing a computer program to solve systems $A\mathbf{x} = \mathbf{b}$ efficiently. The process is given in section 1.11 of the text , "LU factorization". Here is an example of how it works. It consists of reducing a matrix to row echelon form, and as we go along, recording in the slots below the diagonal the negative of the multiplying factor needed to make that entry zero. But it only works if the reduction can be done with no row interchanges.

We start by doing it a little differently from what I described. Suppose that

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 7 & 4 \\ -1 & 4 & 13 \end{pmatrix}.$$

Then

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; E_1 A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 6 \\ -1 & 4 & 13 \end{pmatrix}$$

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \; E_2 E_1 A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 6 \\ 0 & 6 & 12 \end{pmatrix}$$

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}, E_3 E_2 E_1 A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 6 \\ 0 & 0 & 0 \end{pmatrix} = U.$$

Note that $E_3 E_2 E_1 A$ is in upper triangular form. (The matrix A is not invertible, but that is not relevant to whether it can be put in upper triangular form.) From this we see that

$$A = E_1^{-1} E_2^{-1} E_3^{-1} U.$$

Now let's look at the matrix $E_1^{-1} E_2^{-1} E_3^{-1}$. We saw in class how to find the inverses of each elementary matrix. This is also described on page 126 of the text. For example,

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

because in order to undo the row operation performed by E_1, we have to **add** twice the first row to the second. Similarly,

$$E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \; E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

Now let's look closely at what happens when we multiply these in the order $E_1^{-1}E_2^{-1}E_3^{-1}$:

$$E_2^{-1}E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$$

$$E_1^{-1}E_2^{-1}E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$$

This matrix is lower triangular. Moreover, its entries below the diagonal are the multiples of the rows above that we **subtracted** from lower rows in doing the Gaussian elimination. Most interestingly, what went on in each column is not effected by what happens in another column. We could write down $E_1^{-1}E_2^{-1}E_3^{-1}$ as soon as we have done the reduction to upper triangular form, without actually multiplying any matrices together.

Compare this with what the product $E_3E_2E_1$:

$$E_2E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$E_3E_2E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -2 & 1 \end{pmatrix}.$$

The two -2's are not surprising, because they were used in E_1 and E_3. The 3 is surprising, because there was no 3 or -3 in any of the E_j. There is a mixing up in the columns in some way which makes this product hard to predict from the Gaussian elimination steps, while the previous product was predictable.

With this introduction, here is the so-called LU factorization: In the steps below, there are some bold faced numbers. In ordinary elimination you would expect 0 s in those places. See below for an explanation of why there are not zero's.

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 7 & 4 \\ -1 & 4 & 13 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 \\ \mathbf{2} & 3 & 6 \\ -1 & 4 & 13 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 \\ \mathbf{2} & 3 & 6 \\ \mathbf{1} & 6 & 12 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -1 \\ \mathbf{2} & 3 & 6 \\ -\mathbf{1} & \mathbf{2} & 0 \end{pmatrix}.$$

Where you see bold face numbers, ordinarily we would expect 0's. But we use this space to record the "multiplier" that we used for the step. From this we write:

$$U = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 6 \\ 0 & 0 & 0 \end{pmatrix}, L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix},$$

and lo and behold:

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 6 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 7 & 4 \\ -1 & 4 & 13 \end{pmatrix} = A.$$

This is called the "LU factorization" of A. Notice that L has 1's on the diagonal, while U does not.

This helps us to solve equations. Notice that if

$$A\mathbf{x} = \mathbf{b},$$

then

$$L(U\mathbf{x}) = \mathbf{b}.$$

We can solve the system

$$L\mathbf{y} = \mathbf{b}$$

easily, because it only requires "forward substitution". This is because L is lower triangular with 1's on the diagonal. Then we can solve

$$U\mathbf{x} = \mathbf{y}$$

easily by back substitution.

If we need to solve this for a lot of different $\mathbf{b}$'s – say for millions of different $\mathbf{b}$'s, it is relatively easy, because the two steps involved are much easier than repeatedly doing Gaussian elimination. Look at the example of a 4×4 on page 143.

When a row interchange is needed, this does not work in quite the same way. First we need to discuss special matrices called "permutation matrices. Start with I and interchange two rows, for example:

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

This is one of our elementary matrix, and is called an "elementary permutation matrix". Now do this again:

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

This is also a permutation matrix, but it is not an elementary permutation matrix, because it takes two steps to get there. Permutation matrices are an important category of matrix and much can be said about them. But now we concentrate on their use in solving equations.

Making a row interchange is the same as writing equations in a different order. Certainly it does not change the solutions. Sometimes this step is necessary when reducing to row echelon form. But then we are multiplying by a matrix which is not triangular, so we cannot get an LU factorization.

If we knew ahead of time which permutations needed to be done, we could do them at the beginning. In other words, write the equations in a different order from the beginning. Obviously this does not change the solutions. If P is the permutation matrix, then we find PA, and then do LU factorization. We get the formula

$$PA = LU.$$

An example is on page 143.

Homework

1. Find the inverse of the matrix $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. (You will not be able to avoid fractions, but the largest denominator required is 4.)

2. Express the matrix in problem 1 as the product of elementary matrices.

3. Find the LU factorization of the matrix in problem 1.

4. Answer the four true false questions on page 152, with a reason for each answer. If your answer is false, your reason should consist of an explicit matrix A, with numbers, for which you can show that the statement if false.

5. Suppose that A is invertible and A^{-1} is known.

(a) Suppose that B is obtained from A by switching two columns. How can we find B^{-1} from A^{-1}. (Give a reason for your answer.) Hint: Elementary matrices may help you figure out the answer.

(b). Suppose that B is obtained from A by adding a scalar multiple of one column to another. How can we find B^{-1} from A^{-1}?

(c) Suppose B is obtained from A by replacing one column with a new column. If you know that B is invertible, how do your find B^{-1} from A^{-1}. Hint: Suppose the new column is the last column. What do you know about $A^{-1}B$?