# MATH 210 NOTES MATRICES FEB 9

For our purposes a matrix is a rectangular array of numbers: m rows and n columns.We locate the entries by a pair of subscripts, the row and the column. If there are m rows and n columns we refer to A as an m x n matrix and we might write

$$A = ( a_{ij} )$$

We can multiply a matrix by a "scalar" (a number):

$$a\,A = ( a \; a_{ij} )$$

In other words, we multiply each of the entries of the matrix A. We can add two matrices if they are of the same size:

$$A + B = ( a_{ij} + b_{ij}).$$

We just add the corresponding entries. In order for this to make sense the matrices A and B must be of the same size. Matrix addition and scalar multiplication satisfy familiar rules of arithmetic. In particular,

$$A + B = B + A$$
$$A + (B + C) = (A + B) + C$$
$$a\,(A + B) = a\,A + a\,B.$$

For every m and n there is an m x n matrix all of whose entries are 0:

$$O_{mn} = (0)$$

If A is m x n then $A + O_{mn} = A$ and $A - A = O_{mn}$ , where $-A = (-1)A$. In this sense, -A is the additive inverse of A.

We can also multiply matrices if their sizes are compatible. If A is an m x n matrix and B is an n x p matrix then A B is defined and is an m x p matrix.This means that the number of columns of A has to be the number of rows of B for multiplication to be defined. If this is the case and C = A B, then

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj} \; .$$

We get the entries in A B by multiplying the rows of A by the columns of B. This multiplications involves multiplying each entry in a row of A by the corresponding entry in a column of B and adding the results. In the simplest case, when A is a 1 x n matrix (just a single row) and B is a n x 1 matrix (a single column) we are computing the dot product of A and B- each element of A is multiplied by a corresponding element of B and the results are added. The result is just a single number, a 1 x 1 matrix.This multiplication is somewhat complicated and if A and B are large will involve quite a bit of arithmetic. (For instance, if A and B are both 10 x 10, there are 100 entries in the product and each entry is the result of multiplying 10 pairs of numbers and adding them. In all, there would be 1000 ordinary multiplications to compute all of the entries of A B,and 900 additions. Obviously this is best left to a computer.)

Matrix multiplication satisfies some of the familiar rules of arithmetic, but not all. In particular, matrix multiplication is not in general commutative. For one thing, the product AB may be defined while the product BA might not be, but even if both are defined they may not be equal. For instance,

$$\text{if } A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ and } B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

then

$$AB = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix} \text{ and } BA = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}.$$

(You might notice that 19 + 50 = 23 + 46- the sums of the diagonal elements are equal. This always happens.)

Although matrix multiplication is not commutative, it is associative and distributive:

A (B C) = (A B) C and A(B + C) = A B + A C.

(More precisely, in each case if either side of the equation makes sense then both sides are defined and are equal.) The proofs of these two equalities involve writing out the expressions and using familiar rules of arithmetic. For instance, the ij entry of A (B C ) is the sum

$$a_{ik}\, b_{kl}\, c_{lj}$$

over all k and l, which is the same as the corresponding entry of (A B) C. For every integer m there is an m x m identity matrix $I_m$ which has 1's on the main diagonal and 0's elsewhere. For instance, the 3 x 3 identity is

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

If I is an identity matrix of appropriate size then A I = A and I B = B, so the identity matrix acts like the number 1. This raises a natural question: given a matrix A is there a matrix B such that A B = I (or B A = I)? In particular, does a square matrix (m x m) have a multiplicative inverse? This turns out to be a critical question, but the answer is "sometimes". How do you determine whether a square matrix is invertible, and how do you find its inverse if it has one?

One of the reasons the inverse, if it exists, is so useful is that if we want to solve a matrix equation

(*) A X = B

(here, A and B are given and we want to solve for X), then if A has a multiplicative inverse $A^{-1}$then the solution is

X = $A^{-1}$ B.

In this course we are interested in solving systems of simultaneous linear equations, and such equations can be put in matrix form A X = B, so if A is invertible we can solve the system of equations by finding the inverse of A and then multiplying matrices. Since the number 0 does not have a multiplicative inverse you would not expect that all square matrices have inverses, perhaps trivially, but the situation is more complicated. There is a relatively simple way of determining whether a matrix is invertible, and there is a straightforward but laborious procedure for finding the inverse if it exists. In fact, the procedure involves so much arithmetic that it can get out of hand even for a computer as the size of the matrix gets large, and efficient algorithms are very important. (Of course, we would never try to handle a random 100 x 100 matrix by hand, but such do arise in applications.) For instance, $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is not invertible, even though it is not the zero matrix. $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ i s also not invertible. In both cases, it turns out that the dimension of the columns of the matrix is 1- in order for a 2 x 2 matrix to be invertible that dimension will have to be 2. This is an important topic that we will deal with later. The matrix $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ is invertible, since the columns "span $R^2$." This means that every vector in the plane is a combination of the vectors (1,1) and (2,1).

Suppose we have a system of simultaneous linear equations: suppose the coefficients of the unknowns are $a_{ij}$ and the "right-hand sides" of the equations are $b_i$ , so the equations look like

$a_{11}x_1 + ...+ a_{1n}x_n = b_1$

....

$a_{m1}x_1 + ...+ a_{mn}x_n = b_m$

If we let A be the "coefficient matrix" A = ( $a_{ij}$ ) and X and B the column matrices of unknowns and b's, then the system is equivalent to the single matrix equation

A X = B.

Solving the system of simultaneous equations is equivalent to solving the single matrix equation. How do we solve the system? We manipulate the equations in various ways, trying to convert the original system into a simpler equivalent one, so the original matrix equation becomes

A' X = B'.

The typical "row operations" involve interchanging the order in which we write the equations, multiplying an equation by a non-zero number, and adding a multiple of one equation to another. All of these "elementary row operations" are reversible, so any system we derive from the original by a sequence of such operations has the same solutions as the original. If there are as many equations as unknowns, and if we can convert A to the identity matrix I, so A' = I, then the original system becomes

A' X = I X =X = B',

so if we perform a sequence of row operations and convert the original coefficient matrix A into the identity matrix, then the same sequence of row operations will convert B into the solution of the system. One advantage of arranging the equations in matrix form is that we don't have to write down the unknowns and plus and equal signs repeatedly. We can either keep A and B separate, or combine them by "adjoining"B to A to get

[A B]

and then perform a sequence of elementary row operations on this combined matrix so that the first so many columns become the identity matrix and the B-column will then wind up as the solution to the original system: [A B] ⟶[I B'].

In fact, suppose we have several systems to solve, all with the same coefficient matrix. Then we can adjoin several right-hand sides and solve each of the systems at the same time. In particular, suppose we try to solve the system where the right-hand side is $B_i$, the column matrix with a 1 in the ith position and 0's elsewhere. If there are n equations in n unknowns, when we put these B's together we get the n x n identity matrix I, so the systems of equations combine to give

[A I ] . If we convert A to the identity then the columns of I are converted into solutions of the equations with right-hand sides having a single non-zero entry- a 1 in the ith place. Thus

[A I] ⟶[I C]

and the matrix C has the property that A C = $I_n$. In other words, C is the inverse of A.

This means we have the following important theorem:

THEOREM Suppose A is an n x n matrix and I is the corresponding n x n identity matrix. Suppose A is

converted to the identity matrix by a sequence $\rho$ of elementary row operations. If we perform the same sequence on I, then I is converted to the inverse of A. In other words,
if $A^\rho$= I then $I^\rho = A^{-1}$ .

Here is an example. We want to determine whether the matrix
$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ is invertible and, if it is, find its inverse. Suppose we subtract 7 x first row from the third.

This gives us the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & -6 & -12 \end{pmatrix}$. Now, subtract 4 x first row from the second. This gives us
$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix}$. Subtract twice the second row from the third. We get the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix}$.
Since the third row of this matrix is identically 0, this matrix is not invertible, and this says that the original matrix is also NOT INVERTIBLE. It happens that if you replace the 9 in the third row, third column in the original matrix by any other number then the matrix would be invertible. The problem here is that the third column is twice the second minus the first, so the columns are NOT linearly independent. When this happens the matrix is not going to be invertible.You might also notice that the third row is twice the second row minus the first row. The rows are thus not independent. We will eventually see that the rows are independent exactly when the columns are independent. When the rows or columns are not independent then the matrix is not invertible.

Associated with every square matrix is a number called its determinant. It turns out that the square matrix is invertible exactly when its determinant is NOT ZERO. At first glance the way the determinant of a matrix is defined seems complicated: in low dimensions it is straightforward to compute, but as the dimension increases the computation gets exponentially more complicated. However, the determinant ties in very nicely with elementary row operations. For a triangular square matrix the determinant is just the product of the diagonal entries, so if you can triangularize the matrix then you can evaluate the determinant in the process. Here is an example: Suppose

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}.$$

We wish to evaluate det(A). Subtract 7 times the first row from the third and four times the first row from the second. This will not affect the value of the determinant and leads to the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix}.$$

Subtract twice the second row from the third:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{pmatrix}.$$

This is a triangular matrix whose determinant is (1)(-3)(1) =-3, so the determinant of the original matrix is also -3. In particular, A is invertible. You might note that if the (3,3) entry of A had been 9 instead of 10 we would wind up with 0 in the (3,3) place of the last matrix, so its determinant would have been 0. We have already seen that $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ is not invertible.

A permutation of {1,2,...,n} is an arrangement of those n integers in some order, that is, the integers 1,2,...,n written as a sequence. For instance, there are 6 permutations of {1,2,3}:

1, 2 ,3     1 ,3 ,2     2, 3, 1     2, 1,  3     3, 1, 2     3,2,1

In general, there are n! permutations of 1,2,...,n. One of the problems is that factorials grow very rapidly with n, faster than exponential. For instance, 10!=3628800.  100! is approximately 9 x $10^{157}$. If A is an n x n matrix the determinant of A involves evaluating a certain expression for every possible permutation of 1,2,...,n. For every permutation $\pi$ there is a number, either +1 or -1, called the sign of the permutation: sign($\pi$).  If you arrange 1,2,...,n in increasing order you have the identity permutation. Any other permutation can be obtained from the identity permutation by interchanging elements two at a time. For instance, to get 2,1,3 start with 1,2,3 and interchange the 1 and 2. To get 3,1,2 start with 1,2,3 and interchange the 3 and 1, obtaining 3,2,1, and then interchange the 2 and 1, leading to 3,1,2. This required two interchanges. It turns out that although there are in general many ways of gong from the identity to any given permutation, they will all either require an even number of steps, or they will all require an odd number of steps. The permutation is called “even” if it can be obtained from the identity by an even number of such transpositions, and “odd” if it requires an odd number of transpositions. The point is that if, for instance, you can get to the given permutation in three steps you cannot get to it in four or six or any even number. The sign of the permutation is +1 if it can be obtained in an even number of steps and -1 if it takes an odd number.

If A is an n x n determinant here is how the determinant of A, det(A), is defined. Pick one element from each row and each column of A- writing the columns in the order in which you have chosen them gives a permutation $\pi$ of 1,2,...,n.  Write the sequencer as $\pi(1)$, $\pi(2)$,...,$\pi(n)$. Corresponding to this permutation $\pi$ form the sum

$$\sum \text{sgn}(\pi)\ \ a_{1,\pi(1)} \text{ x .....x } a_{n,\pi(n)} \quad ,$$

where the sum if over all permutations of 1,2,...,n. For instance, if n = 10 there are over 3 million terms in this sum. Each term is the product of one element from each row and each column of A, times the sign of the corresponding permutation. So, not only do we have to work out every product but we also have to work out the sign of the permutation each time. At least, if we didn’t have the sgn factor there would be less work.In fact, without that factor the expression is called the permanent of A, and working it out for a 10 x 10 matrix is quite an arithmetical computation.However,  the additional factor of sgn($\pi$) turns out to make the computation feasible, since with that additional factor the determinant ties in very nicely with elementary row operations.

THEOREM     Let A be a square matrix.

(1) If A is upper triangular then det(A) is the product of the diagonal elements of A.

(2) If A’ is obtained from A by multiplying a row of A by a non-zero number b, then det(A’)

(3) If A’ is obtained from A by interchanging two rows of A then det(A’)=- det(A).

(4) If A’s ib obtained from A by adding a multiple of one row to another, then det(A’) = det(A).

This theorem gives us a way of evaluating the determinant of a given matrix: use elementary row operations to put it in triangular form and keep track of the number of times you interchange two rows and of the times you multiply a row by a non-zero number. The product of the diagonal elements of the triangular matrix with +1 or -1 divided by the row multiples gives you the determinant. For instance, here is an example:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 4 & 2 & 7 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -6 & -5 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 7 \end{pmatrix} A'.$$ In this case, det(A)= det(A’) = 7.

Here is a more complicated example:

$$A= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 13 \\ 15 & 18 & 21 & 25 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & -8 & -16 & -23 \\ 0 & -12 & -24 & -35 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$ so det(A) = 0.

THEOREM If A and B are square matrices of the same sizes then det(A B) = det(A) det(B). In particular, if A is invertible then det(A) ≠ 0.

If A is a triangular matrix with non-zero entries on the diagonal then A is invertible. Elementary row operations do not affect the invertibility of a matrix: if A is invertible so is any matrix obtained from A by a sequence of elementary row operations. If you can convert A to a matrix that is invertible by a sequence of elementary row operations, then you know that A is invertible.

With each square matrix there is an associated number called its determinant. The matrix is invertible exactly when the determinant of the matrix is NOT 0. The determinant of a triangular matrix is the product of the diagonal elements. Elementary row operations do not affect whether the determinant is 0 or not.

$$\det(a) = a$$

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\,d - bc$$

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a\,e\,i - a\,f\,h + b\,f\,g - b\,d\,i + c\,d\,h - c\,e\,g$$

All products of one element from each row and each column, some with minus signs.