MA 1B RECITATION 02/12/15

1. Determinants

1.1. **What do determinants represent?** When introduced at this stage of the course, determinants seem like mysterious objects—just a formula that you apply blindly to matrices to get some number out of them. However, there are some alternative ways to think about them; personally, I like to think of determinants as an *oriented* volume. Indeed, this is primarily how determinants arise in calculus, as you will see in Math 1c.

Here's how you view them from this perspective: think of the columns of the matrix as vectors at the origin forming the edges of a (probably skewed) parallelogram/parallelepiped/parallelotope. The determinant gives the volume of that box, multiplied by a ± 1 depending on the "orientation" of the vectors. Let's see this in the first nontrivial case.

Example 1. The determinant of a 2×2 matrix is given by

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

How should we think about this?

Remember that this matrix represents a linear mapping; in particular, one that takes the standard basis vectors and maps them to the columns of A. Alternatively, since $\det(A) = \det(A^T)$, we could also use a mapping that takes the standard basis vectors and maps them to the rows of A. Just for kicks, let's try the mapping to the rows.

The images of the basis vectors forms a parallelogram that represents the image of the unit square under the mapping. The parallelogram defined by the rows of the above matrix is one with vertices at $(0, 0), (a, b), (a + c, b + d)$ and (c, d).

The absolute value of $ad - bc$ is the area of the parallelogram and represents the scale factor by which areas are transformed by A.

The absolute value of the determinant together with the sign becomes the *oriented* area of the parallelogram. The oriented area is the usual area, except that it is negative when the angle from the first to the second vector defining the parallelogram goes clockwise instead of counterclockwise. (Namely, opposite the direct you would get for the identity matrix.) Like most things in math, it's more natural to go counterclockwise than clockwise.

An analogous picture works in higher dimensions.

1.2. **Important Properties of Determinants.** Let A and B be $n \times n$ square matrices.

- $\det(AB) = \det(A)\det(B)$
- $\det(A^T) = \det(A)$
- $\det(I_n) = 1$

Date: February 12, 2015.

- Determinants remain invariant under basis change.
- Exchanging two rows or two columns multiplies the determinant by -1.
- If A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$.
- $\det(cA) = c^n \det(A)$. Note that you need to take it to the nth power! This is a common error on exams, so please remember this. An easy way to remember is to remember the multiplicative formula and realize that you are actually multiplying by cI_n.
- If the columns of A form a linearly dependent set, then $\det(A) = 0$.
- Since $\det(A^T) = \det(A)$, if the rows of A form a linearly dependent set, then $\det(A) = 0$.
- If A is an upper- or lower-triangular matrix, then its determinant is the product of the diagonal entries. (Note that diagonal matrices themselves are both upper- and lower-triangular.)
- $\det(A) = \prod_{i=1}^{n} \lambda_i$, where λ_i are the eigenvalues of A.

The last statement is what's "really" going on when you're calculating the determinant, and is what we're ultimately building up to with the theory of eigenvectors and eigenvalues and reach its full conclusion in this course with the Grand Theorem of Linear Algebra and the Spectral Theorem.

Also, it's worth noting that analogous properties hold when we replace the entries with blocks of matrices. We'll study this more closely soon.

Finally, the main point here is that **every property about determinants can be interpreted geometrically in terms of oriented volumes**. For instance, if two of the columns are linearly dependent, your box is "missing" a dimension and it has been flatted in some way to have zero n-volume.

Thus, we will eventually have three different ways to look at determinants:

- geometrically (as measuring oriented volumes),
- algebraically (as a function of the entries of a matrix), and
- spectrally (as the product of eigenvalues of the transformation).

2. Cramer's Rule

One of the most useful applications is Cramer's Rule, which gives an explicit formula to express a solution to a system of linear equations in a very specific situation.

Theorem 1. *(Cramer's Rule) Let $A\vec{x} = \vec{b}$ be a system of n linear equations in n variables, so A is a square, and assume that $\det(A) \neq 0$. Then*

$$\vec{x} = \left(\frac{\det(A_1)}{\det(A)}, \frac{\det(A_2)}{\det(A)}, \dots, \frac{\det(A_n)}{\det(A)} \right)^T$$

where A_i is the matrix where we replace the ith column of A by b.

One can also understand this result geometrically. (How?) This works because $\vec{x} = A^{-1}\vec{b}$ and we are essentially calculating the inverse of A.

Example 2. Consider the matrix

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \vec{b}.$$

Note that $\det(A) = 6 \neq 0$ and so we can apply Cramer's rule in this situation.

We calculate

$$\det(A_1) = \det \begin{bmatrix} 0 & 4 & 5 \\ 1 & 2 & 6 \\ 0 & 0 & 3 \end{bmatrix} = -12$$

$$\det(A_2) = \det \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 6 \\ 0 & 0 & 3 \end{bmatrix} = 3$$

$$\det(A_3) = \det \begin{bmatrix} 1 & 4 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} = 0.$$

Therefore,

$$\vec{x} = \begin{bmatrix} -2 \\ \frac{1}{2} \\ 0. \end{bmatrix}$$

We check that

$$A\vec{x} = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ \frac{1}{2} \\ 0. \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

3. Expansion by Minors

Determinants of matrices are important to compute, but it's not always easy. It's also not obvious how to find determinants beyond the 2×2 or 3×3 cases, because we no longer have a simple formula that we can get just by multiplying entries diagonally in some particular way.

Minors are very useful because often it's sometimes easier to evaluate determinants by a technique called "expansion by minors."

Definition 1. (Apostol) Given a square matrix A of order $n \geq 2$, the **(k,j)-th minor** of A, denoted A_{kj}, is the square matrix of order $(n-1)$ obtained by deleting the kth row and the jth column of A.

Remark 2. This is a slightly nonstandard definition of a minor of a matrix. We usually call the *determinant* of such matrices the minors. We will return to this in the next section.

Theorem 3 (Expansion by Minors, Apostol, Theorem 3.9)**.** *For each $n > 1$ and each $1 \leq k \leq n$,*

$$\det(A) = \sum_{j=1}^{m} (-1)^{k+j} a_{k,j} \det(A_{k,j}).$$

Remark 4. Note the changes in sign! It is critical that you remember the signs:

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \end{bmatrix}$$

What is this really saying? It's sort of strange that this statement even holds. From the formula alone, it's a little hard to get an idea of what's really going on. Let's try and illustrate this with an easy example.

Example 5. Compute the determinant of

$$\begin{bmatrix} 10 & 0 & -3 \\ -2 & -4 & 1 \\ 3 & 0 & 2 \end{bmatrix}$$

Solution. We have

$$\begin{aligned}
\det \begin{bmatrix} 10 & 0 & -3 \\ -2 & -4 & 1 \\ 3 & 0 & 2 \end{bmatrix} &= 10 \det \begin{bmatrix} -4 & 1 \\ 0 & 2 \end{bmatrix} - (-2) \det \begin{bmatrix} 0 & -3 \\ 0 & 2 \end{bmatrix} + 3 \det \begin{bmatrix} 0 & -3 \\ -4 & 1 \end{bmatrix} \\
&= 10[(-4)(2) - (0)(1)] + 2[(0)(2) - (0)(-3)] + 3[(0)(1) - (-4)(-3)] \\
&= 10(-8) + 2(0) + 3(-12) \\
&= -80 - 36 \\
&= -116.
\end{aligned}$$

We could also have chosen to expand the determinant differently. The easiest would probably be along the second column, since there are two zeroes in it. Let's do it!

$$\begin{aligned}
\det \begin{bmatrix} 10 & 0 & -3 \\ -2 & -4 & 1 \\ 3 & 0 & 2 \end{bmatrix} &= (-0)M + (-4)M + -(0)M \\
&= -4(20 + 9) \\
&= -116.
\end{aligned}$$

We could also expand along rows, using the fact that $\det(A) = \det(A^T)$.

Just as a sanity check, let's try and compute the 3×3 determinant the usual way:

$$\det \begin{bmatrix} 10 & 0 & -3 \\ -2 & -4 & 1 \\ 3 & 0 & 2 \end{bmatrix} = -80 + 0 + 0 - 36 - 0 - 0 = -116.$$

□

3.1. **Minors and Rank.** Minors can be defined for arbitrary matrices.

Definition 6. A $k \times k$ **minor** (a.k.a. order k minor) is the *determinant* of a $k \times k$ submatrix of a $n \times m$ matrix obtained by deleting rows and columns.

Note that while Apostol's "minor" is a *matrix*, the usual definition of "minor" is as a *number* (or a function, if the entries are undefined).

Observation. Let A be an $n \times m$. Then

$$\operatorname{rank}(A) = \text{maximal order of nonzero minors of } A.$$

Idea behind the proof. If a $k \times k$ minor is nonzero, the corresponding columns of A are linearly independent.

This can often make certain calculations easier. Let's see a couple of examples.

Example 7. What is the rank of

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & 4 & 2 \\ 0 & 2 & 2 & 1 \end{bmatrix}?$$

Of course, one way to find this is to transformation A into reduced row echelon form, but let's try and find it using minors.

Since row rank equals column rank equals rank, and we have a 3×4 matrix, the rank of A is at most 3. Note that

$$\det \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 4 \\ 0 & 2 & 2 \end{bmatrix} = 4 + 0 + 0 - 0 - 9 - 0 = -4 \neq 0$$

so the rank of A is at least 3. Therefore, the rank of A is 3.

Example 8. Are $(1, 1, 0)^T$ and $(0, 1, 1)^T$ linearly independent?

Of course, the answer is yes, since neither vector is a multiple of another, but we can also check this via minors. Note that

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

has a nonzero 2×2 minor

$$\det \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = 1.$$

Therefore, A has rank 2. Therefore, the two vectors are linearly independent.

3.2. **Aside: Computation Complexity of Determinant Computation.** For those of you with interests in computer science, applied math, or engineering, one thing you might think about is, what is the computational complexity of computing the determinant of an $n \times n$ matrix? Roughly speaking, computational complexity represents the number of elementary arithmetic operations (like addition and multiplication) that you must perform in order to compute some operation.

As we have seen, calculating the determinant gets much more complicated once we get to 4×4 matrices and higher. If you don't know any special properties about your matrix, e.g. if it isn't in some special form like block-diagonal or if you don't know the eigenvalues, then calculating the determinant will be long and tedious.

Indeed, even if we restrict to matrices with entries in the integers, naively applying expansion by minors will result in an $O(n!)$ operation, that is, the number of operations involved will have a dominant term of $Cn!$ for some positive constant C. Obviously, such a number will grow quickly. There's a very clever way to do this that that will give you an $O(n^4)$ operation, but even such a result is not feasible in practice. Gaussian elimination is an $O(n^3)$ operation, and as we saw, we can find determinants in this way as well, but having $O(n^3)$ is only remotely plausible if we work over *finite fields*, for instance, if only consider entries in the matrices modulo some prime number. There are whole disciplines of applied and computational mathematics that try to find new ways to compute determinants, since it's such a important quantity of compute.

What often happens in practice is that you have to sacrifice a good deal of precision and just find ways to bound your determinant. For instance, you may have to deal with $n \times n$ matrices where $n = 10000$ or so, and without knowing that it has some special properties (e.g. that it's symmetric, Hermitian, or has a nice matrix decomposition), you often just have to settle for whatever Matlab or something can pop out, and for the computation to even finish, you have to set the precision relatively low and hope that you get something useful.

4. Hadamard Matrices

Recall that I gave a challenge problem some time ago involving determinants. It turns out that given the techniques we have learned today, we can prove this result in a very nice way.

Proposition 1. *Prove that the determinant of an $k \times k$ matrix with entries only ± 1 must be divisible by 2^{k-1}.*

Proof. Let A be a given $k \times k$ matrix with entries only ± 1. Add its first row to the other $(k-1)$ rows to get a new $k \times k$ matrix B. By the properties of the determinant, $\det(B) = \det(A)$. Since the elements of A are ± 1, the lower $(k-1)$ submatrix of B consists of entries $2, -2$ or 0.

Expanding by minors across the first row, we obtain

$$\det(B) = \sum_{j=1}^{k} (-1)^{1+j} b_{1j} \det(B_{1j}).$$

Since all elements of $B_{1,j}$ are ± 2 or 0, it means $\det(B_{1j})$ can only be $2^{k-1}, 0$ or -2^{k-1}, which are all divisible by 2^{k-1}. Therefore, $2^{k-1} \mid \det(B) = \det(A)$. $\square$

Such $\{\pm 1\}$ square matrices appear in many parts of mathematics. Indeed, they lead us to an edge of our current knowledge of mathematics.

For instance, over a century about, Jacques Hadamard, a French mathematician, proved that for such matrices,

$$|\det(A)| \leq k^{k/2}.$$

Matrices that meet this upper bound are called *Hadamard matrices* and are the $\{\pm 1\}$-matrices with orthogonal columns. This bound can only be attained when $k = 1, 2$, or a multiple of 4; one reason that you cannot have orthogonal matrices in odd dimensions, because there is no sum of an odd number of ± 1's that equals zero.

However, for k not in the cases above, the bound $|\det(A)|$ is not known in general! For smaller cases, it is known, since we can just consider all such matrices, but this becomes hard as k gets large. Furthermore, it is also open whether a Hadamard matrix of size $4k$ exists for every $k \geq 1$. I believe the smallest k for which this is open is $k = 668$. Note that if we were to try and do this by exhaustive search, we would have to check $2^{668 \times 668}$ matrices. Recall that the number of atoms in the observable universe is close to 10^{80}, so $2^{446224} = 16^{116556}$ is enormous! Barring some extreme change in the nature of computation, it is unfeasible to attack such problems with brute force methods.