LECTURE 1

I. How big is a matrix?

Let me start with a question that's going to be very important to us, not only this week, but next week, and perhaps for the rest of our mathematical lives. How big is a matrix? We know what it means for a *number* to be big– we know that 1000000 is bigger than 5. But how do we tell if one matrix is bigger than another? It's not so clear.

You might counter my question with another question–why does it matter how big a matrix is? I bring to mind a question from the midterm exam.

Namely: Suppose that a vector $\vec{t}_0$ represents a temperature state of a discretely approximated system at time 0.

Then there is a matrix M and a vector $\vec{b}$ such that the temperature distribution an hour later is represented by

$$\vec{t}_1 = M\vec{t} + \vec{b}.$$

In our example, we had

$$M = \begin{bmatrix} 0 & 1/4 & 0 & 1/4 \\ 1/4 & 0 & 1/4 & 0 \\ 0 & 1/4 & 0 & 1/4 \\ 1/4 & 0 & 1/4 & 0 \end{bmatrix}$$

So if for instance we start with

$$\vec{t} = \begin{bmatrix} 100 \\ 100 \\ 100 \\ 100 \end{bmatrix}$$

and boundary conditions

$$\vec{b} = \begin{bmatrix} 160 \\ 80 \\ 0 \\ 0 \end{bmatrix}$$

we get

$$\vec{t}_1 = \begin{bmatrix} 90 \\ 110 \\ 70 \\ 50 \end{bmatrix}$$

Now the question is: what happens ten hours later?

Well, we get

$$\vec{t_2} = M\vec{t_1} + \vec{b} = M(M\vec{t_0} + \vec{b}) + \vec{b} = (I + M)\vec{b} + M^2\vec{t_0}$$

and

$$\vec{t_3} = M\vec{t_2} + \vec{b} = M((I + M)\vec{b} + M^2\vec{t_0}) + \vec{b} = (I + M + M^2)\vec{b} + M^3\vec{t_0}$$

and perhaps you begin to see the pattern. After ten hours, the temperature will be

$$\vec{t_1}0 = (1 + M + M^2 + \ldots + M^9)\vec{b} + M^10\vec{t_0} = \begin{bmatrix} 70.039 \\ 90.039 \\ 50.039 \\ 30.039 \end{bmatrix}$$

and this conforms with our intuition–that in time, the temperature inside the column will approach very closely the equilibrium solution

$$\begin{bmatrix} 70 \\ 90 \\ 50 \\ 30 \end{bmatrix} = (I - M)^{-1}\vec{b}$$

which we found on our exam.

Question: is this a *coincidence?* Surely not. But how, *mathematically*, can we express what's going on?

Well. When you do these calculations, you find something very interesting. You find that

$$M^10 = \frac{1}{2048}\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

In other words, M^10 is very *small.*

So when we compute $\vec{v_1}0$ as above, we can throw out $M^10\vec{t_0}$, because that is very small, and say

$$\vec{t_1}0 \sim (1 + M + M^2 + \ldots + M^9)\vec{b}.$$

Now moreover you might want to say, by your experience with Taylor series, that

$$(1 + M + M^2 + \ldots + M^9) \sim (I - M)^{-1}$$

And you would be right–check that

$$(I - M)(1 + M + M^2 + \ldots + M^9) = I - M^10 \sim I.$$

So to sum up, we conclude that

> If high powers of M are very small, then the process described above converges to the equilibrium solution $(I - M)^{-1}B$.

Remark 1: Note that, according to our physical intiution, M^n should *always* get small as n gets large! Those of you who did a different temperature network on your exams–if you still have the M from your network on your computers, work out a high power of M and see if it is nearly 0!

Remark 2: You might want to compare the discussion above with your knowledge about Taylor series–the fact that, for instance,

$$1 + x + x^2 + \ldots$$

converges to $(1-x)^{-1}$ if and only if $|x| < 1$. In effect, we are asking ourselves: what is the analogue for matrices of the condition $|x| < 1$?

II. When is a matrix invertible?

This question doesn't seem so related to the other, but we'll see when the smoke clears that the two have something in common.

First of all, comment that we already know lots of answers to this question! For instance: if the columns form a basis for $\mathbb{R}^n$, if the rank is n, etc.

THE 2×2 CASE: We have already read (way back in chapter 1) that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

if $ad - cb \neq 0$. In particular,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ is invertible if and only if } ad - bc \neq 0.$$

So we have a very beautiful formula in terms of the entries of the matrix. Could such a thing exist for larger matrices?

THE 3×3 CASE:

Let A be a 3×3 matrix

$$\begin{bmatrix} | & | & | \\ \vec{v}_1 & \ldots & \vec{v}_n \\ | & | & | \end{bmatrix}.$$

How do we know if A is invertible? This is the same as to say that $\vec{v}_1, \vec{v}_2, \vec{v}_3$ span all of $\mathbb{R}^3$. And how can we test this? Well, one way is to examine

$$\vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3.)$$

This is not so desirable, since the use of cross products is restricted to 3 dimensions. Yet it works. For if $\vec{v}_2$ and $\vec{v}_3$ are not independent, the cross-product is 0. If not, and if $\vec{v}_1$ is in the plane spanned by $\vec{v}_2$ and $\vec{v}_3$, then it is orthogonal to $\vec{v}_2 \times \vec{v}_3$.

Note that

$$\vec{v}_2 \times \vec{v}_3 = \begin{bmatrix} a_{22}a_{33} - a_{32}a_{23} \\ a_{32}a_{13} - a_{12}a_{33} \\ a_{12}a_{23} - a_{22}a_{13} \end{bmatrix}.$$

So the dot product is what it is: write all the pluses together, all the minuses together.

Define: The *determinant* of A is

$$\det(A) = \vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3) = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} + a_{21}a_{32}a_{13} - a_{21}a_{12}a_{33} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}.$$

Then do an example:

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 3 \\ 3 & 5 & 0 \end{bmatrix}$$

and we get $10 - 15 = -5$. So this matrix is invertible. Well, great. So what?

Well, we'd like to have a way to do this beyond $3x3$.

Lecture 2

III. Determinant: general definitions

Note that the three terms in the cross product each look like a determinant! Draw the pictures to show the 2×2 matrices whose determinants are the terms of the cross product. (This should remind you of the definition of cross product it seems most of you learned in Math 203...)

This leads us to suggest the following definition.

Definition. A *reduction* M_{ij} of an $n \times n$ matrix is the $(n-1) \times (n-1)$ matrix obtained by removing the ith row and the jth column of A. A *cofactor* A_{ij} of A is $(-1)^{i+j} \det M_{ij}$.

Show some examples.

Definition. The determinant of A is defined to be the sum

$$\sum_{i=1}^{n} a_{i1}A_{i1}$$

Remark. This definition looks slightly different than, but is equivalent to, Strang 4B (see comment (12) on p.227 for the equivalence.)

But isn't this definition circular? Defining the determinant in terms of the determinant? No! It is *recursive.* The crucial point is that we define the determinant of an $n \times n$ matrix in terms of the determinant of $(n-1)\times(n-1)$ matrices, which are in turn defined in terms of determinants of $(n-2)\times(n-2)$ matrices, etc.

Invertibility Theorem 4: (Strang, p.216) A is invertible if and only if $\det A \neq 0$.

So OK. I just made a definition, which perhaps seems completely unmotivated. Let's first begin to explore the properties of this definition–then I'll discuss the geometric interpretation–then, if time permits, I'll prove the invertibility theorem. But first, an example.

Example: Consider the matrix

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

Let's calculate the determinant, 'kay? We get

$$(1-\lambda)\det M_{11} - \det M_{21} - \det M_{41}$$

So what are these? Well,

$$\det M_{11} = (1-\lambda)((1-\lambda)^2 - 1) - (1-\lambda)$$

so

$$(1-\lambda)\det M_{11} = (1-\lambda)^2((1-\lambda)^2-1) + (1-\lambda)^2 = (1-\lambda)^2(-1-2\lambda+\lambda^2)$$

That's the worst of it. Moving on, one can check that

$$\det M_{21} = \det M_{41} = (1-2\lambda+\lambda^2) = (1-\lambda)^2$$

and conclude

$$\det A = (1-\lambda)^2(\lambda^2-2\lambda-3) = (1-\lambda)^2(3-\lambda)(-1-\lambda)$$

As a group exercise, consider the following true-false statements. Do not assume that the claim we've made is true.

1. The determinant of a diagonal matrix is the product of the diagonal entries.

2. The determinant of an upper triangular matrix is the product of its diagonal entries.

3. The determinant of a matrix is equal to the determinant of its transpose.

4. The determinant of a matrix is equal to the determinant of its inverse.

5. The determinant of kA is equal to k times the determinant of A.

This being done, move on to geometry, especially in $\mathbb{R}^2$. Draw the unit box and talk about its image under various transformations. Get to the idea that the determinant of a 2×2 matrix measures the coefficient of *magnification*. Which makes it a good measure (though not, as we're about to see, good enough) for the "size" of a matrix.

LECTURE 3.

I. Proof of Invertibility Theorem 4

I will only outline it.

MAIN LEMMA: Show that if U is the row-reduced echelon form of A, then $\det A = \pm \det U$.

Example: Try

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, U = 120{-}2.$$

Indeed, both have determinant -2.

The lemma being granted, we argue as follows: the determinant of U is the product of the diagonal entries of U, because U is upper triangular. And the product of the diagonal entries of U is the product of the pivots. That is, the following statements are all equivalent:

- $\det A = 0$;
- $\det U = 0$;
- U has a zero diagonal entry;
- U has a zero pivot;
- U is not invertible;
- A is not invertible.

It is only the first equivalence that requires the main lemma.

The proof of the main Lemma requires us to prove:

- If A' is obtained from A by adding a multiple of one row to another row, then $\det A' = \det A$.
- If A' is obtained from A by switching two rows, then $\det A' = -\det A$.

And because these two operations are the only things we have to do to change A into U, we get the Main Lemma.

That's all I'll say about Invertibility Theorem 4.

II. Geometric interpretation

Theorem: $\det(AB) = \det(A)\det(B)$.

This is not so obvious from our definition (or from any of the standard definitions.) Two proofs given on Strang, p.217.

Let's consider some transformations of $\mathbb{R}^2$ and their determinants. Make a table:

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1/4 & 0 \\ 0 & 1/4 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

We see, at least for the first few examples, that the determinant of A. seems to measure the *magnification of area* induced by A.

Theorem. Let R be a region in $\mathbb{R}^n$, and let AR be its image after applying the transformation A. Then the volume of AR is $\det A$ times the volume of R.

This makes it very plain *physically* that an orthogonal matrix has determinant 1, since rotation doesn't change volume. But that's not correct. Talk a bit about signed volume, using the last matrix above as an example.

Remark 1: This description justifies our notion that the determinant is a crude notion of the "size" of a matrix.

Remark 2: With the emphasis on "crude", as we can see if we consider the matrix

$$A = \begin{bmatrix} 9/8 & 7/8 \\ 7/8 & 9/8 \end{bmatrix}.$$

One checks that $\det A = 1/2$. So A is small, right? But if we compute powers, we find

$$A^2 \sim \begin{bmatrix} 2.03 & 1.97 \\ 1.97 & 2.03 \end{bmatrix}$$

Ulp. And

$$A^10 \sim \begin{bmatrix} 512 & 512 \\ 512 & 512 \end{bmatrix}$$

Of course, A^10 is not *equal* to that matrix, since $\det A^10$ is not 0. But it's real close.

So what's going on? Is our notion of determinant as size completely out of whack? Well, frankly, yes.

Consider powers of $\begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}$ geometrically to get the idea.

So we need to undertake a finer analysis. And we'll do so next week.

If there's more time: talk more about the proof of the invertibility theorem, and give the formula for the inverse.