LECTURE 1

So last time we introduced the matrix $\begin{bmatrix} 9/8 & 7/8 \\ 7/8 & 9/8 \end{bmatrix}$, which had small determinant but whose powers grew very quickly.

Another example of such a matrix is $\begin{bmatrix} 2 & 0 \\ 0 & 1/4 \end{bmatrix}$. Here it's quite visible what goes wrong; the matrix has small determinant, but a high power of it is not going to be close to 0. In fact, it's very easy to say what a high power of M is; it is

$$M^n = \begin{bmatrix} 2^n & 0 \\ 0 & 1/4^n \end{bmatrix}.$$

In fact, these examples are related. Draw a picture showing that $\begin{bmatrix} 9/8 & 7/8 \\ 7/8 & 9/8 \end{bmatrix}$ is a stretch of dilation factor 2 in the direction of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, but a shrink of dilation factor 1/4 in the direction of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. That is,

$$A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 1 \end{bmatrix}, A\begin{bmatrix} 1 \\ -1 \end{bmatrix} = (1/4)\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

So this makes it clear what M does to any vector. For instance, this makes it very easy to calculate $M^{10}\vec{e}_1$. Without multiplying it out. Namely, write

$$\vec{e}_1 = (1/2)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (1/2)\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and now have

$$A^{10}\vec{e}_1 = 2^{10}(1/2)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (1/4^{10})(1/2)\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

which is very close to

$$2^{10}(1/2)\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5\vec{1}2512.$$

Also

$$\vec{e}_2 = (1/2)\begin{bmatrix} 1 \\ 1 \end{bmatrix} - (1/2)\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and so one finds similarly that

$$A^{10}\vec{e}_2 \sim 2^{10}(1/2)\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 512 \\ 512 \end{bmatrix}.$$

Conclude that

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

as I told you yesterday. Of course, this is not *exactly* correct, since M^{10} has determinant $1/2^{10}$ and $\begin{bmatrix} 512 & 512 \\ 512 & 512 \end{bmatrix}$ has determinant 0.

So having found these two special vectors was very useful in our analysis of the matrix! They give us

- a geometric sense of what the matrix does;
- an ability to compute the action of large powers of the matrix.

Definition. Let A be an $n \times n$ matrix. An *eigenvector* for A is a *nonzero* vector $\vec{v} \in \mathbb{R}^n$ such that

$$A\vec{v} = \lambda\vec{v}$$

for some scalar λ. The scalar λ is called the *eigenvalue* of the eigenvector $\vec{v}$.

REMARK: The word "eigen" means "characteristic of" or "belonging to" in German and is cognate to the English word "own."

MOTTO: If we understand the eigenvectors and eigenvalues of a matrix, we understand its essence.

Example:

1. The matrix $\begin{bmatrix} 9/8 & 7/8 \\ 7/8 & 9/8 \end{bmatrix}$ has eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, with eigenvalue 2, and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, with eigenvalue $1/4$.

2. Take the matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Does it have an eigenvector? See if anyone offers one. Observe that $\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \vec{e}_1$ is an eigenvector. Are there any others? Hard to say! Let's see. Maybe there's an eigenvector with eigenvalue 2. That is, maybe there's a nonzero vector $\vec{v}$ satisfying

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\vec{v} = 2\vec{v}.$$

That is,

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ x+y \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix}.$$

No, for in that case $x = 2x$, so $x = 0$, and then $x + y = 0 + y = 2y$, so $y = 0$ as well. Let's put this computation another way: from

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \vec{v} = 2\vec{v} = 2I\vec{v}$$

get

$$(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - 2I)\vec{v} = 0$$

.

But one checks that $(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - 2I) = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$ has nonzero determinant, so is invertible. In particular, it has no nullspace. So the equation above has no nonzero solution.

Now the point is, there is nothing special about 2 in the above argument! So, arguing exactly as above, we find that the following statements are equivalent:

- $\vec{v}$ is an eigenvector of A with eigenvalue λ;
- $\vec{v} \in N(A - \lambda I)$.

But note that $N(A - \lambda I)$ contains a nonzero vector if and only if $A - \lambda I$ is not invertible (Rank Theorem!) So this gives us a very hands-on way to compute the eigenvalues. Namely, the following are equivalent:

- λ is an eigenvalue of A;
- $N(A - \lambda I)$ contains a nonzero vector;
- $A - \lambda I$ is not invertible;
- $\det(A - \lambda I) \neq 0$.

This motivates the following definition.

Definition. Let A be an $n \times n$ matrix. The *characteristic polynomial* $P_A(\lambda)$ of A (which should perhaps be called the eigenpolynomial) is $\det(A - \lambda I)$.

Example. If $A = 9\vec{/}87/87/89/8$, we get

$$p_A(\lambda) = \det \begin{bmatrix} 9/8-\lambda & 7/8 \\ 7/8 & 9/8-\lambda \end{bmatrix} = (9/8-\lambda)^2 - (7/8)^2 = 81/64 - 9/4\lambda + \lambda^2 - 49/64 = \lambda^2 - 9/4$$

So indeed, the *only* eigenvalues are 2 and 1/4!

If $S = \vec{1}101$, compute $p_A(\lambda) = (\lambda - 1)^2$. So $\lambda = 1$ is the *only* eigenvalue.

Remark: We can check from the definition of determinant that $p_A(\lambda)$ is a degree n polynomial. So by the fundamental theorem of algebra, it has at most n different roots. So an $n \times n$ matrix A has at most n different eigenvalues. But it really could be less–witness S above, which is 2×2 but only has one eigenvalue.

Remark: The characteristic polynomial method only tells us the eigenvalues; to find the eigenvectors corresponding to an eigenvalue λ we must still compute the nullspace of $A - \lambda I$.

Let's do a bigger example!

Try

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

How to find the eigenvalues? Well, we have to compute

$$P_A(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 1 & 0 & 1 \\ 1 & 1-\lambda & 1 & 0 \\ 0 & 1 & 1-\lambda & 1 \\ 1 & 0 & 1 & 1-\lambda \end{bmatrix}$$

Good thing we computed that last Wednesday! We get

$$P_A(\lambda) = (1-\lambda)^2(3-\lambda)(-1-\lambda).$$

So λ is an eigenvalue of A exactly when $P_A(\lambda) = 0$. That is, the eigenvalues of A are $1, 3$, and -1.

Can we find the eigenvectors associated to these eigenvalues? Well, like for instance let us find the eigenvector associated to the eigenvalue 3. Then we are computing $N(A - 3I)$; that is, the nullspace of

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

We could certainly do Gaussian elimination as usual to compute the nullspace, but let's just observe that

$$\vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

does the trick.

LECTURE 2

1. Trace and determinant

Let me start with a couple of interesting facts, which we can use to check our computations of eigenvalues. First, a remark on the characteristic polynomial.

$$p_A(x) = (-1)^n \prod_{i=1}^{n} (x - \lambda_i)$$

where the λ_i are the eigenvalues of n (counted with multiplicity.) The fact that the λ_i are the roots of $p_A(x)$ tells us that

$$p_A(x) = c \prod_{i=1}^{n} (x - \lambda_i)$$

To find the right value of c, we can expand $\det(A - xI)$ in cofactors and observe that the coefficient of x^n is $(-1)^n$.

Definition: The *trace* of a matrix A (denoted $\text{Tr}A$) is the sum of its eigenvalues (with multiplicity.)

What does "with multiplicity" mean? It means that if $p_A(\lambda)$ has a factor of $(\lambda - a)^m$, then we count the eigenvalue a n times. So for instance the trace of $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is 2, because the eigenvalues are $1, 1$.

Remark: Every matrix has n eigenvalues (counted with multiplicity, and including complex eigenvalues.)

This follows from the fundamental theorem of algebra, which tells us that the degree n polynomial $p_A(\lambda)$ has exactly n roots (counted with multiplicity, and including complex roots.)

Proposition: $\text{Tr}(A)$ is equal to the sum of the diagonal entries of A. $\det(A)$ is equal to the product of the eigenvalues of A (with multiplicity.)

I'm going to stop saying "with multiplicity" now, but it is always there implicitly.

The determinant fact is easy to prove: we know that

$$det(A) = p_A(0) = (-1)^n \prod_{i=1}^{n}(-\lambda_i) = \prod_{i=1}^{n}(\lambda_i)$$

To prove the trace fact (which I won't do–it's interesting but not necessary) one shows that both Tr(A) and the sum of the diagonal entries of A are equal to the negative of the x^{n-1} coefficient of $P_n(x)$.

2. Complex eigenvalues

Now we must face head-on the fact that eigenvalues are not necessarily real numbers.

Example: What are the eigenvalues and eigenvectors of

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

Observe: it shouldn't be so surprising that there are no real eigenvalues–geometrically speaking, it is hard to envision a rotation as "stretch so much in one direction, so much in another direction!"

Easy to compute that $p_A(x) = x^2 + 1$. So the eigenvalues are $\pm i$. What are the eigenvectors? Well, the eigenvector corresponding to i must be in the nullspace of

$$A - iI = \begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix}$$

whose row-reduced form is

$$\begin{bmatrix} -i & 1 \\ 0 & 0 \end{bmatrix}$$

so the nullspace is seen to be spanned by

$$\vec{v_i} = \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

A vector with *complex* coefficients. If our eyes could see the complex numbers, we'd see that rotation was actually a stretch by a factor of i in the $\begin{bmatrix} 1 \\ i \end{bmatrix}$ direction and by a factor of $-i$ in the $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ direction!

Remark If λ is a complex eigenvalue of A, then so is its complex conjugate $\bar{\lambda}$.

3. Diagonalization

Now I want to talk about **diagonalization.** This is a formalization of some of the ideas we talked about Monday–it captures the usefulness of having a basis consisting of eigenvectors for a matrix A.

Remember the example of $\begin{bmatrix} 9/8 & 7/8 \\ 7/8 & 9/8 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 \\ 0 & 1/4 \end{bmatrix}$; these matrices had the same eigenvalues but different eigenvectors, and we found that their "behavior" was in some sense quite similar.

Proposition: Let A be a matrix with n linearly independent eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$. Let S be the matrix

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

and let D be the diagonal matrix

$$S = \begin{bmatrix} \lambda_1 & & \\ & \ldots & \\ & & \lambda_n \end{bmatrix}.$$

Then $S^{-1}AS = D$.

How do we prove this? The idea is to show that $S^{-1}AS$ and D have the same eigenvalues and eigenvectors, and from there to observe that they must be the same.

Proof. Let $M = S^{-1}AS - D$. We're trying to show M is the zero matrix.

Let $\vec{e}_i$ be a standard basis element. Then $D\vec{e}_i = \lambda_i \vec{e}_i$. Now

$$S^{-1}AS\vec{e}_i = S^{-1}A\vec{v}_i = S^{-1}\lambda_i\vec{v}_i = \lambda_i(S^{-1}\vec{v}_i) = \lambda_i e_i.$$

Conclude that

$$M\vec{e}_i = \lambda_i\vec{e}_i - \lambda_i\vec{e}_i = 0$$

Since M kills each of the standard basis vectors, it must be 0.

Definition: We say A is *diagonalizable* if it has n linearly independent eigenvectors.

Remark: We have seen, to our dismay, that not every matrix A is diagonalizable–for instance, $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not.

Corollary: Suppose A be a matrix with n linearly independent eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$, with associated eigenvalues $\lambda_1, \ldots, \lambda_n$. Define S, D as above. Then $A = SDS^{-1}$.

Problem: "Find a matrix A which stretches by a factor of 3 in the $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ direction, and by a factor of 4 in the $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ direction."

In other words: "Find a matrix A which has eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and corresponding eigenvalues $3, 4$."

Then $S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$, so we find

$$A = SDS^{-1} = \begin{bmatrix} 3 & 1 \\ 0 & 4 \end{bmatrix}.$$

4. An application of diagonalizability

It is easy to calculate large powers of a diagonalizable matrix, for example. Suppose $A = SDS^{-1}$. Then

$$A^2 = AA = (SDS^{-1})(SDS^{-1}) = SD^2S^{-1}$$

Likewise, in general,

$$A^n = SD^nS^{-1}.$$

So, for instance, take $A = \begin{bmatrix} 3 & 1 \\ 0 & 4 \end{bmatrix}$. What is A^9? It is

$$SD^9S^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3^9 & 0 \\ 0 & 4^9 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = mattwo 3^9 4^9 - 3^9 0 4^9$$

Much easier than multiplying it out nine times!

LECTURE 3

No notes for lecture 3; JE is out of town!