Lecture 13

Last lecture we defined the determinant and saw some of its properties. The definition was probably different than what you may have seen before. So now we will relate the definition to the Laplace (cofactor) expansion.

Definition 0.1. *Given $A \in M_{n,n}(\mathbb{F})$, the (i,j)-th minor of A, written $A(i|j)$, is the $n-1 \times n-1$ matrix formed by removing the i-th row and the j-th column from A.*

The Laplace expansion is a recursive formula for the determinant. We can write $\det A$ in terms of the determinant of smaller matrices, the minors of A.

Theorem 0.2. *Let $A \in M_{n,n}(\mathbb{F})$ with entries $(a_{i,j})$. Then*

$$\det A = \sum_{i=1}^{n} (-1)^{i-1} a_{i,1} \det A(i|1) \ .$$

Proof. Write the columns of A as $\vec{a}_1, \ldots, \vec{a}_n$ with $\vec{a}_1 = a_{1,1}e_1 + \cdots + a_{n,1}e_n$ (where $e_1, \ldots, e_n$ are the standard basis vectors) and use n-linearity on the matrix $A = (\vec{a}_1, \ldots, \vec{a}_n)$ to get

$$\det A = \det(a_{1,1}e_1, \vec{a}_2, \ldots, \vec{a}_n) + \cdots + \det(a_{n,1}e_n, \vec{a}_2, \ldots, \vec{a}_n) \ . \quad = \sum_{i=1}^{n} a_{i,1} \det(e_i, \vec{a}_2, \ldots, \vec{a}_n) \ .$$

Now we must only show that $\det(e_i, \vec{a}_2, \ldots, \vec{a}_n) = (-1)^{i-1} \det A(i|1)$.

We will need to use two facts from the homework.

1. For any matrix B, $\det B = \det B^t$. As a consequence of this, det is n-linear and alternating when viewed as a function of the *rows* of B.
2. If B is any *block upper triangular matrix*; that is, of the form

   $$B = \begin{pmatrix} C & D \\ 0 & E \end{pmatrix}$$

   for square matrices C and E, then $\det B = \det C \cdot \det E$.

So now use $i-1$ adjacent row swaps to turn the matrix $(e_i, \vec{a}_2, \ldots, \vec{a}_n)$ into

$$\begin{pmatrix} 1 & a_{i,2} & a_{i,3} & \cdots & a_{i,n} \\ 0 & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ & & & \cdots & \\ 0 & a_{i-1,2} & a_{i-1,3} & \cdots & a_{i-1,n} \\ 0 & a_{i+1,2} & a_{i+1,3} & \cdots & a_{i+1,n} \\ & & & \cdots & \\ 0 & a_{n,2} & a_{n,3} & \cdots & a_{n,n} \end{pmatrix} \ .$$

Since we applied $i-1$ transpositions, the determinant of this matrix equals $(-1)^{i-1}$ times $\det(e_i, \vec{a}_2, \ldots, \vec{a}_n)$. Now we apply the block upper-triangular result, noting that this matrix is of the form

$$\begin{pmatrix} 1 & D \\ 0 & A(i|1) \end{pmatrix} .$$

Therefore $\det(e_i, \vec{a}_2, \ldots, \vec{a}_n) = (-1)^{i-1} \det A(i|1)$ and we are done. $\square$

There is a more general version of this result. The above we call "expanding along the first column." We can expand along the j-th column by first applying $j-1$ adjacent column swaps to get

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det A(i|j) .$$

By taking the transpose initially we can expand along any row too.

Eigenvalues

Our goal for most of the rest of the semester is to classify all linear transformations $T : V \to V$ when $\dim V < \infty$. How can we possibly do this? There are so many transformations. Well, let's start with the simplest matrices.

Definition 0.3. *$A \in M_{n,n}(\mathbb{F})$ is called a diagonal matrix if $a_{i,j} = 0$ when $i \neq j$.*

Notice that if A is a diagonal matrix, then A acts very simply on the standard basis. Precisely, if A is diagonal with entries $\lambda_1, \ldots, \lambda_n$, then

$$Ae_i = \lambda_i e_i .$$

For the next couple of lectures we will try to determine exactly when a linear transformation has a diagonal matrix representation (for some basis). This motivates the following definition.

Definition 0.4. *If $T : V \to V$ is linear then a nonzero vector $v \in V$ is called an eigenvector for T with associated eigenvalue λ if $T(v) = \lambda v$. If there is a basis for V consisting of eigenvectors for T then we say T is diagonalizable.*

Proposition 0.5. *Let $T : V \to V$ be linear. Then T is diagonalizable if and only if there is a basis B of V such that $[T]_B^B$ is diagonal.*

Proof. If $[T]_B^B$ is diagonal with entries $\lambda_1, \ldots, \lambda_n$, then writing $B = \{v_1, \ldots, v_n\}$, we have

$$[T(v_i)]_B = [T]_B^B [v_i]_B = [T]_B^B e_i = \lambda_i e_i = [\lambda_i v_i]_B .$$

Therefore $T(v_i) = \lambda_i v_i$. Since v_i is part of a basis, $v_i \neq \vec{0}$ and therefore each v_i is an eigenvector.

Conversely, suppose that $B = \{v_1, \ldots, v_n\}$ is a basis of V consisting of eigenvectors for T. Then the i-th column of $[T]_B^B$ is the column vector $[Tv_i]_B = \lambda_i [v_i]_B = \lambda_i e_i$, so $[T]_B^B$ is diagonal. $\square$

Imagine that we have a linear transformation $T : V \to V$ and we are trying to build a basis of eigenvectors for T. We find first an eigenvector v_1 with eigenvalue λ_1. Next we find v_2 with value λ_2. How do we know they are linearly independent? Here is a sufficient (but not necessary!) condition.

Theorem 0.6. *Let $v_1, \ldots, v_k$ be eigenvectors with respective eigenvalues $\lambda_1, \ldots, \lambda_k$ such that $\lambda_i \neq \lambda_j$ for $i \neq j$. Then $\{v_1, \ldots, v_k\}$ is linearly independent.*

Proof. As usual, suppose that

$$a_1 v_1 + \cdots + a_k v_k = \vec{0} \ .$$

Let's assume that we have already removed all vectors which have nonzero coefficients, and among all such linear combinations equal to zero, this is one with the least number of coefficients. We may assume that there are at least two coefficients, or else we would have $a_1 v_1 = \vec{0}$ and since $v_1 \neq 0$ we would have $a_1 = 0$, meaning all coefficients are zero and $\{v_1, \ldots, v_k\}$ is linearly independent.

So apply T to both sides:

$$a_1 T(v_1) + \cdots + a_k T(v_k) = \vec{0} \ .$$

Since these are eigenvectors, we can rewrite as

$$a_1 \lambda_1 v_1 + \cdots + a_k \lambda_k v_k = \vec{0} \ .$$

However multiplying the linear combination by λ_1 we get

$$a_1 \lambda_1 v_1 + \cdots + a_k \lambda_1 v_k = \vec{0} \ .$$

Subtracting these two,

$$a_2(\lambda_1 - \lambda_2) v_2 + \cdots + a_k(\lambda_1 - \lambda_k) v_k = \vec{0} \ .$$

All λ_i's were distinct and all a_i's were nonzero, so this is a linear combination of the v_i's equal to zero with fewer nonzero coefficients than in the original one, a contradiction. $\square$

For a matrix $A \in M_{n,n}(\mathbb{F})$, we define its eigenvalues and eigenvectors similarly: λ is an eigenvalue of A if there is a nonzero $v \in \mathbb{F}^n$ such that $A \cdot v = \lambda v$.

To find the eigenvalues, we make the following observation.

- λ is an eigenvalue for A if and only if there exists a nonzero v such that $(\lambda I - A)(v) = \vec{0}$. This is true if and only if $\lambda I - A$ is not invertible. Therefore

$$\lambda \text{ is an eigenvalue of } A \Leftrightarrow (\lambda I - A) \text{ not invertible } \Leftrightarrow \det(\lambda I - A) = 0 \ .$$

This leads us to define

Definition 0.7. *The characteristic polynomial of a matrix* $A \in M_{n,n}(\mathbb{F})$ *is the function* $c_A : \mathbb{F} \to \mathbb{F}$ *given by*

$$c_A(x) = \det(xI - A) \ .$$

The definition is similar for a linear transformation. The characteristic polynomial of $T : V \to V$ is $c_T(x) = \det[xI - T]_B^B$, where B is any finite basis of V. (You will show on homework that this definition does not depend on the choice of basis.)

Facts about the characteristic polynomial.

1. c_A is a monic polynomial of degree n.

   Proof. We simply write out the definition of the determinant, using the notation that $A_{i,j}(x)$ is the (i,j)-th entry of $xI - A$:

   $$c_A(x) = \det(xI - A) = \sum_{\pi \in S_n} \operatorname{sgn} \pi A_{\pi(1),1}(x) \cdots A_{\pi(n),n}(x) \ .$$

   Each term in this sum is a product of n polynomials, each of degree at most 1 (it is either a field element or a polynomial of the form $x - a$ for some a). So each term is a polynomial of degree at most n, implying the same for c_A. The only term of degree n corresponds to choosing all the diagonal elements of the matrix – this is the identity permutation. This term is $(x - a_{1,1}) \cdots (x - a_{n,n})$ and the coefficient of x^n is 1. $\square$

2. If we look for terms of degree $n-1$ in $c_A(x)$, we note that in the determinant expansion, each permutation that is not the identity must have at least two numbers $k \neq j$ from 1 to n such that $\pi(k) \neq k$ and $\pi(j) \neq j$. Therefore all nonidentity permutations give terms with degree at most $n-2$. So the only term contributing to degree $n-1$ is the identity. Thus the coefficient of $c_A(x)$ of degree $n-1$ is the same as that of

   $$(x - a_{1,1}) \cdots (x - a_{n,n}) = x^n - x^{n-1}[a_{1,1} + \cdots + a_{n,n}] + \cdots = x^n - x^{n-1}Tr(A) + \cdots$$

   This means the coefficient of x^{n-1} is $-Tr(A)$.

3. The coefficient of degree zero (the last term of c_A) is $(-1)^n \det A$. This follows by just plugging in $x = 0$. Thus

   $$c_A(x) = x^n - x^{n-1}Tr(A) + \cdots + (-1)^n \det A \ .$$

4. The field $\mathbb{F}$ is called *algebraically closed* if every $p \in \mathbb{F}[x]$ with degree at least 1 has a zero in $\mathbb{F}$. (For example, $\mathbb{C}$ is algebraically closed.) You will show on homework that in this case, each polynomial can be factored into factors of the form:

   $$p(x) = a(x - r_1)^{n_1} \cdots (x - r_k)^{n_k} \ ,$$

   for $a, r_1, \ldots, r_k \in \mathbb{F}$ and natural numbers $n_1, \ldots, n_k$. The r_i's are the roots of p and the n_i's are their multiplicities.

So if $\mathbb{F}$ is algebraically closed, we can factor

$$c_A(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k}$$

with $n_1 + \cdots + n_k = n$. Note the leading coefficient a here is 1 since c_A is monic. Expanding this, we see that if $\mathbb{F}$ is algebraically closed, then

$$Tr(A) = \sum_{i=1}^{k} n_i \lambda_i \text{ and } \det A = \prod_{i=1}^{k} \lambda_i^{n_i} .$$

Thus the trace is the sum of eigenvalues (repeated according to multiplicity) and the determinant is the product of eigenvalues (again repeated according to multiplicity).