16. The characteristic polynomial

We have already seen that given a 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

there is a single number $ad - bc$ such that A is invertible if and only if $ad - bc \neq 0$. It is a somewhat amazing fact that one can generalise this function to any n:

Theorem 16.1. *There is a function*

$$\det \colon M_{n,n}(F) \longrightarrow F,$$

called the ***determinant****, which is uniquely determined by the following properties:*

(1) *the determinant respects multiplication, that is*

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

(2)

$$\det E_1(\lambda) = \lambda.$$

There are three points to realise:

- with some, admittedly abstract, algebraic machinery, (16.1) is straightforward to prove;
- the condition that the determinant respects multiplication is very strong. In particular note that the statement that the determinant is uniquely determined by (1) and (2).
- if n is reasonably large, computing the determinant is computationally very expensive. Formulae which compute determinants or formulae with determinants in them are invariably useless.

Lemma 16.2. *Let n be a positive integer and let F be a field.*

(1) $\det I_n = 1$.
(2) $\det 0_n = 0$.
(3) *If A is invertible then* $\det A \neq 0$ *and in fact* $\det A^{-1} = 1/\det A$.
(4) *If A and B are similar then* $\det A = \det B$.
(5) $\det E_{ij}(\lambda) = 1$.
(6) $\det E_i(\lambda) = \lambda$.
(7) $\det P_{ij} = -1$.

Proof. We first prove (1). By assumption det is not the constant function. So there is at least one matrix A such that $\det A \neq 0$. But then

$$\det A = \det I_n A = \det I_n \det A.$$

As $\det A \neq 0$ this implies $\det I_n = 1$. This is (1).

We now prove (2). By assumption there is a matrix B such that $\det B \neq 1$. But then

$$\det 0_n = \det 0_n B = \det 0_n \det B.$$

As $\det B \neq 1$ the only possibility is that $\det 0_n = 0$. This is (2).

We now prove (3). Suppose that B is the inverse of A. As $AB = I_n$ it follows that

$$1 = \det I_n = \det AB = \det A \det B.$$

But then $\det A \neq 0$ and $\det B = 1/\det A$. This is (3).

We now prove (4). Suppose that A and B are similar. Then there is an invertible matrix P such that $B = PAP^{-1}$. Taking determinants, we see that

$$\begin{aligned}\det B &= \det(PAP^{-1}) \\ &= \det P \det A \det P^{-1} \\ &= \det A.\end{aligned}$$

This is (4).

We now prove (5). Suppose that $\lambda \neq 0$. Note that, considering row operations,

$$E_{ij}(\lambda) = E_i(\lambda^{-1})E_{ij}(1)E_i(\lambda).$$

Thus $E_{ij}(\lambda)$ and $E_{ij}(1)$ are similar. By (4) it follows that their determinants are the same. But

$$I_n = E_{ij}(0) = E_{ij}(1)E_{ij}(-1).$$

Taking determinants, we get

$$\det E_{ij}(1)^2 = 1.$$

But the polynomial equation $x^2 = 1$ only has two roots, ± 1, since the polynomial $x^2 - 1$ factors as $(x-1)(x+1)$.

There are two possibilities. If $2 = 0$ (for example, consider the field $\mathbb{F}_2$) then $1 = -1$ and there is nothing to prove. Otherwise consider the equation

$$E_{ij}(2) = E_{ij}(1)^2.$$

Taking determinants, we see that

$$\det E_{ij}(1) = \det E_{ij}(1)^2 = (\det E_{ij}(1))^2,$$

and the only possibility is that $\det E_{ij}(1) = 1$. This is (5).

We now prove (6). Note that $E_i(\lambda)$ is similar to $E_1(\lambda)$. Using (4) this gives (6).

We now prove (7). Note that P_{ij} is similar to P_{12} which in turn is similar to $E_1(-1)$. This proves (7). $\square$

Since the determinant is an invariant of similar matrices, in fact one can define the determinant of any linear function $\phi\colon V \longrightarrow V$, where V is a finite dimensional vector space. Indeed, pick any basis $v_1, v_2, \dots, v_n$ of V. In the usual way, this gives rise to a linear function $\psi\colon F^n \longrightarrow F^n$ and so to a matrix A, representing ψ. The key point is that if we choose a different basis, say $w_1, w_2, \dots, w_n$ of V, we would get a different linear function τ and a different matrix B, but since A and B are similar, the determinant would be the same.

Now that we know what the determinant of an elementary matrix is, we can now compute any determinant by Gauss-Jordan elimination.

For example, if P is any permutation matrix, then its determinant is ± 1, depending on whether P is a product of an even or an odd number of elementary permutation matrices (it is in fact quite hard to prove that the parity is invariant without using determinants).

Now suppose that D is a diagonal matrix, with diagonal entries $\lambda_1, \lambda_2, \dots, \lambda_n$. Then D is a product of $E_i(\lambda_i)$, and so its determinant is the product $\lambda_1 \lambda_2 \dots \lambda_n$.

Now suppose that U is an upper triangular matrix, with 1's on the main diagonal. Then applying Gauss-Jordan elimination, we can row reduce U to the identity matrix I_n, multiplying by only the elementary matrices $E_{ij}(\lambda)$. It follows that the determinant of U is 1. On the other hand, if U has a row of zeroes, then it is easy to see that $\det U = 0$. It is also not hard to see that $\det A^t = \det A$.

Putting all of this together, gives the most efficient method of computing the determinant of a matrix $A \in M_{n,n}(F)$. Applying Gaussian elimination, we can find matrices P, L and U such that

$$A = PLU.$$

The determinant of P is ± 1. The determinant of U is equal to 1. L is the product of L' and D where L' is a lower triangular matrix with 1 on the main diagonal and D is a diagonal matrix with diagonal entries equal to the diagonal entries of L. So the determinant of A is $\pm d$, where d is the product of the diagonal entries of L. We have also proved:

Theorem 16.3. *Let $A \in M_{n,n}(F)$.*

Then $\det A$ is non-zero if and only if A is invertible.

Proof. Indeed, if we write $A = PLU$, then A is invertible if and only if U is invertible, that is U has no rows of zeroes. But $\det U = 0$ if and only if U has a row of zeroes. $\square$

There is a recursive way to compute the determinant of a matrix, which is in someways only of theoretical interest (and of interest if you want to compute the determinant of a 3×3 matrix without using a computer whilst doing your hwk).

The rule is as follows. Take the matrix A and pick a row or a column. For each entry in this row or column, take the determinant of the submatrix you get by deleting the row and the column to which this entry lives, multiply by this entry and take an alternating sum. The sum starts positive if you are in an odd row or column and otherwise starts negatively. For example, consider the following determinant:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

Suppose we decide to expand about the third column. Then we get

$$3\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} - 6\begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} + 9\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}.$$

If we decided instead to expand about the second row, we would start with -4, progress to 5 and finish with -6. In fact notice that all of this is consistent with the fact that the determinant of the tranpose is the same as the original determinant and that we could put the third row into the position of the first row by switching the first and third rows and then switching the second and third rows. This involves two changes of sign and so in total there is no change of sign.

However in this fantasy world, where supposedly it is a good idea to compute determinants by hand, normally you are supposed to use tricks to compute determinants. For example, always expand about rows or columns with lots of zeroes. Look to see if two rows or columns look similar; if so, subtract one from the other to cheaply create lots more zeroes.

Definition 16.4. *Let $A \in M_{n,n}(F)$ be a matrix.*

*The **characteristic polynomial** of A is the polynomial in λ*

$$\det(A - \lambda I_n).$$

Lemma 16.5. *The roots of the characteristic polynomial are the eigenvalues of A.*

Proof. λ_1 is a root of the characteristic polynomial if and only if $A-\lambda_1 I_n$ is not invertible if and only if the eigenspace $E_{\lambda_1}(A)$ is non-trivial if and only if A has an eigenvector with eigenvalue λ_1 □

Theorem 16.6 (Cayley-Hamilton)**.** *Let $A \in M_{n,n}(\mathbb{C})$.*

Then the minimal polynomial divides the characteristic polynomial. In particular A satisfies its own characteristic polynomial.

Proof. We may assume that A is in Jordan canonical form. It is then not hard to see that we may assume that A is a single Jordan block. In this case the minimal polynomial is $(x-\lambda)^n$ and by direct computation the characteristic polynomial is the same. $\square$