5.2 The Characteristic Equation

Let A be an $n \times n$ matrix. Let U be any echelon form obtained from A by row replacements and row interchanges (without scaling). Let r be the number of such row interchanges. The determinant of A is the product of the diagonal entries of U times $(-1)^r$. If A is invertible, then every diagonal entry is a pivot because $A \sim I_n$. Otherwise, at least one of the diagonal entries in U is zero and hence the product of the diagonal entries in U is zero. Therefore,

$$\det A = \begin{cases} (-1)^r \cdot (\text{product of pivots in } U), & \text{when } A \text{ is invertible} \\ 0, & \text{when } A \text{ is not invertible} \end{cases} \tag{1}$$

Theorem (IMT Continued)**.** *Let A be an $n \times n$ matrix. Then A is invertible if and only if*

19. The number 0 *is* not *an eigenvalue of A.*

20. The determinant of A is not *zero.*

Theorem 3 (Properties of Determinants)**.** *Let A and B be $n \times n$ matrices.*

1. *A is invertible if and only if* $\det A \neq 0$.
2. $\det AB = (\det A)(\det B)$.
3. $\det A^T = \det A$.
4. *If A is triangular, then* $\det A$ *is the product of the entries on the main diagonal of A.*
5. *A row replacement operation on A does not change the determinant. A row interchange changes the sign of the determinant. A row scaling also scales the determinant by the same factor.*

The Properties of Determinants Theorem, part 1, shows how to determine when a matrix of the form $A - \lambda I$ is *not* invertible. The scalar equation $\det(A - \lambda I) = 0$ is called the *characteristic equation* of A.

Remark. *A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if λ satisfies the characteristic equation*

$$\det\,(A - \lambda I) = 0$$

If A is an $n \times n$ matrix, then $\det(\mathrm{A} - \lambda \mathrm{I})$ is a polynomial of degree n, called the *characteristic polynomial* of A. The (algebraic) multiplicity of an eigenvalue λ is its multiplicity as a root of the characteristic equation.

Definition. *If A and B are $n \times n$ matrices, then A* is similar to *B if there is an invertible matrix P such that $P^{-1}AP = B$, or equivalently, $A = PBP^{-1}$. Writing $Q = P^{-1}$, we have $Q^{-1}BQ = A$. So B is also similar to A, and we say simply that A and B* are similar. *Changing A into $P^{-1}AP$ is called a* similarity transformation.

Theorem 4. *If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).*

Proof. If $B = P^{-1}AP$, then

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

Using the multiplicative property (2) in Properties of Determinants Theorem, we compute

$$\det(B - \lambda I) = \det[P^{-1}(A - \lambda A)P] = \det(P^{-1}) \cdot \det(A - \lambda I) \cdot \det(P) \tag{2}$$

Sine $\det(P^{-1}) \cdot \det(P) = \det(P^{-1}P) = \det I = 1$, we see from equation (2) that $\det(B - \lambda I) = \det(A - \lambda I)$. □

Remark. *1. The matrices* $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ *and* $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ *are not similar, even though they have the same eigenvalues.*

2. Similarity is not the same as row equivalence. If A is row equivalent to B, then $B = EA$ for some invertible matrix E. Row operations on a matrix usually change its similarity.

Example 1. *Find the characteristic polynomial and the eigenvalues of the matrix:*

$$\begin{bmatrix} 5 & -3 \\ -4 & 3 \end{bmatrix}$$

Example 2. *Find the characteristic polynomial of the matrix. You may use a cofactor expansion.*

$$\begin{bmatrix} 5 & -2 & 3 \\ 0 & 1 & 0 \\ 6 & 7 & -2 \end{bmatrix}$$

Example 3. *List the eigenvalues, repeated according to their multiplicities.*

$$\begin{bmatrix} 5 & 0 & 0 & 0 \\ 8 & -4 & 0 & 0 \\ 0 & 7 & 1 & 0 \\ 1 & -5 & 2 & 1 \end{bmatrix}$$

Often times, we may gain a better insight of linear algebra topics if we investigate them without resorting to determinants.

Consider the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ where

$$T(w, z) = (-z, w). \tag{3}$$

This operator has a nice geometric interpretation: T is just a counterclockwise rotation by 90° about the origin in $\mathbb{R}^2$. An operator has an eigenvalue if and only if there exists a nonzero vector in its domain that gets sent by the operator to a scalar multiple of itself. The rotation of a nonzero vector in $\mathbb{R}^2$ obviously never equals a scalar multiple of itself. Conclusion: the operator T defined in (3) has no eigenvalues. However, if T is defined on C^2 instead, then the story changes. To find eigenvalues of T, we must find the scalars λ such that

$$T(w, z) = \lambda T(w, z)$$

has some solution other than $w = z = 0$. For T defined by (3), the equation above is equivalent to the simultaneous equations

$$-z = \lambda w, \quad w = \lambda z. \tag{4}$$

Substituting the value for w given by the second equation into the first equation gives

$$-z = \lambda^2 z.$$

Now a cannot equal 0 (otherwise (4) implies that $w = 0$; we are looking for solutions to (4) where (w, z) is not the $\mathbf{0}$ vector), so the equation above leads to the equation

$$-1 = \lambda^2.$$

The solutions to this equation are $\lambda = i$ or $\lambda = -i$. We may easily verify that i and $-i$ are eigenvalues of T Indeed, the eigenvectors corresponding to the eigenvalue i are the vectors of the form $(w, -wi)$, with $w \in \mathbb{C}$, and the eigenvectos corresponding to the eigenvalue $-i$ are the vectors of the form (w, wi), with $w \in \mathbb{C}$.

Suppose T is an operator on V. The *multiplicity* of an eigenvalue λ of T is defined to be the dimension of the subspace of generalized eigenvectors corresponding to lam. In other words, the multiplicity of an eigenvalue λ of T equals dim $\text{Nul}(A - \lambda I)^{\dim V}$, where A is the standard matrix of T. If T has an upper-triangular matrix with respect to some basis of V (as always happens when the scalars are complex numbers), then the multiplicity of λ is simply the number of times λ appears on the diagonal of this matrix.

Example 4. *Suppose*

$$T(z_1, z_2, z_3) = (0, z_1, 5z_3). \tag{5}$$

We may verify that 0 *is an eigenvalue of* T *with multiplicity* 2, *that* 5 *is an eigenvalue of* T *with multiplicity* 1, *and that* T *has no additional eigenvalues.*

Example 5. *Suppose* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *is an operator whose matrix is*

$$\begin{bmatrix} 6 & 7 & 7 \\ 0 & 6 & 7 \\ 0 & 0 & 7 \end{bmatrix}. \tag{6}$$

Then 6 *is an eigenvalue of* T *with multiplicity* 2 *and* 7 *is an eigenvalue of* T *with multiplicity* 1.

In each of the examples above, the sum of the multiplicities of the eigenvalues of T equals 3, which is the dimension of the domain of T.

5.2.1 The Characteristic Polynomial

Let's begin with the trivial case of 1-by-1 matrices. Suppose V is a real vector space with dimension 1 and T is an operator on V. If $[\lambda]$ equals the matrix of T with respect to some basis of V, then the matrix of T is λI. We define the characteristic polynomial of $[\lambda]$ to be $x - \lambda$.

Now let's look at 2-by-2 matrices. We define the characteristic polynomial of a 2-by-2 matrix $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$ to be $(x - a)(x - d) - bc$.

Suppose V is a complex vector space and T is an operator on V. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T. Let d_j denoted the multiplicity of λ_j as an eigenvalue of T. The polynomial

$$(z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}$$

is called the *characteristic polynomial* of T. Note that the degree of the characteristic polynomial of T equals dim V. Obviously the roots of the characteristic polynomial of T equal the eigenvalues of T.

Example 6. *The characteristic polynomial of the operator* T *defined by (5) equals* $z^2(z - 5)$.

Example 7. *If* T *is the operator whose matrix is given by (6), then the characteristic polynomial of* T *equals* $(x - 6)^2(x - 7)$.

Now suppose V is a real vector space and T is an operator on V. With respect to some basis of V, T has a block upper-triangular matrix of the form

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix}, \tag{7}$$

where each A_j is a 1-by1 matrix or a 2-by-2 matrix with no eigenvalues. We define the *characteristic polynomial* of T to be the product of the characteristic polynomials of $A_1, \ldots, A_m$. Explicitly, for each j, we define $q_j \in \mathbb{P}$ by

$$q_j(x) = \begin{cases} x - \lambda & \text{if } A_j = [\lambda] \\ (x - a)(x - d) - bc & \text{if } A_j = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \end{cases} \tag{8}$$

Then the characteristic polynomial of T is

$$q_1(x) \cdots q_m(x).$$

Clearly the characterisic polynomial of T has degree dim V. Furthermore, the characteristic polynomial of T depends only on T and not on the choice of a particular basis.