3.5 The Characteristic Polynomial

In this section we will discuss how to determine the eigenvalues of a given matrix. In practice, we cannot always precisely determine them, but we can write down a polynomial whose coefficients depend on the entries of the matrix, and whose roots are the eigenvalues.

Example 3.5.1. *Find a matrix* P *with* $P^{-1}AP$ *diagonal, where* $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$

To answer this, we need to find two linearly independent eigenvectors of A. These are non-zero solutions of

$$\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} \implies \begin{array}{rcl} 2x+2y & = & \lambda x \\ x+3y & = & \lambda y \end{array} \implies \begin{array}{rcl} 0 & = & (\lambda-2)x-2y \\ 0 & = & -x+(\lambda-3)y \end{array} \implies \begin{bmatrix} \lambda-2 & -2 \\ -1 & \lambda-3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So we are looking for non-zero solutions $\begin{bmatrix} x \\ y \end{bmatrix}$ of the system

$$\begin{bmatrix} \lambda-2 & -2 \\ -1 & \lambda-3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

These can occur only if the coefficient matrix is non-invertible. If it is invertible, the only solution is $x = y = 0$.

A 2×2 matrix is non-invertible if and only if its determinant is 0. The *determinant* of the 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $ad - bc$.

$$\det \begin{bmatrix} \lambda-2 & -2 \\ -1 & \lambda-3 \end{bmatrix} = (\lambda-2)(\lambda-3)-(-2)(-1) = \lambda^2 - 5\lambda + 4.$$

The characteristic polynomial of A is

$$\det(\lambda I - A) = \lambda^2 - 5\lambda + 4 = (\lambda-4)(\lambda-1).$$

The *eigenvalues* of A are the solutions of the *characteristic equation* $\det(\lambda I - A) = 0$, 1 and 4. The *eigenspace* of A corresponding to $\lambda = 1$ is the set of all solutions of the system

$$\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 1 \begin{bmatrix} x \\ y \end{bmatrix} \implies \begin{bmatrix} 1-2 & -2 \\ -1 & 1-3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This is the nullspace of the matrix $1I - A = \begin{bmatrix} -1 & -2 \\ -1 & -2 \end{bmatrix}$, which is

$$\left\{ \begin{bmatrix} -2t \\ t \end{bmatrix}, t \in \mathbb{R} \right\}.$$

An eigenvector of A for $\lambda = 1$ is any non-zero element of this space, for example $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

The *eigenspace* of A corresponding to $\lambda = 4$ is the nullspace of the matrix $4I - A = \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix}$, which is

$$\left\{ \begin{bmatrix} t \\ t \end{bmatrix}, t \in \mathbb{R} \right\}.$$

An eigenvector of A for $\lambda = 4$ is any non-zero element of this space, for example $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Conclusion: If P is a matrix whose columns are eigenvectors of A corresponding respectively to the eigenvalues 1 and 4, for example $P = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}$, then $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$. Multiplying each column of this P by a non-zero scalar gives alternative choices of P, with the same diagonal matrix $P^{-1}AP$. Switching the two columns of P would give a matrix Q with $Q^{-1}AQ = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$.

3.5.1 The Determinant (a digression)

For any 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = (ad - bc)I_2.$$

From this matrix equation we have the following observations:

- If $ad - bc = 0$, then A is not invertible, because the columns $\begin{bmatrix} d \\ -c \end{bmatrix}$ and $\begin{bmatrix} -b \\ a \end{bmatrix}$ are in its nullspace (and these are both zero only if A is the zero matrix).
- If $ad - bc \neq 0$, then the equation shows that $A^{-1} = \dfrac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
- The matrix A has an inverse if and only if $ad - bc \neq 0$. This means that the number $ad - bc$ tells us whether or not the columns of A form a basis of $\mathbb{F}^2$ (or $\mathbb{R}^2$).

The equation above also prompts the following definitions

- The number (or field element) $ad - bc$ is the *determinant* of the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, denoted by $\det(A)$ or sometimes $|A|$.
- The matrix $A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ is the *adjugate* (sometimes called the *adjoint*) of A, denoted by $\mathrm{adj}(A)$.

The version of the above equation for a 3×3 matrix $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ is the following:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} ei - fh & -bi + ch & bf - ce \\ -di + fg & ai - cg & -af + cd \\ dh - eg & -ah + bg & ae - bd \end{bmatrix} = (aei - afh - bdi + bfg + cdh - ceg)I_3.$$

This equation can be checked directly. The expression $(aei - afh - bdi + bfg + cdh - ceg)$ is the *determinant* of A, and the adjugate of A is the matrix on the right. Its entries are the determinants of the nine 2×2 submatrices of A (some with a sign change). To see why this definition of the 3×3 determinant is consistent with the 2×2 version, we can write it as follows:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} \\ \begin{vmatrix} d & e \\ g & h \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{bmatrix} = (\underbrace{aei - afh - bdi + bfg + cdh - ceg}_{\det(A)})I_3.$$

Definition 3.5.2. *The* minor $M_{i,j}$ *of the entry in the* (i, j) *position of a* 3×3 *matrix* A *is the determinant of the* 2×2 *matrix that remains when Row* i *and Column* j *are deleted from* A.

Definition 3.5.3. *The* cofactor $C_{i,j}$ *of the entry in the* (i, j) *position of a* 3×3 *matrix* A *is either equal to* $M_{i,j}$ *or to* $-M_{i,j}$*, according to the following pattern of signs:* $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$

Definition 3.5.4. *The* adjugate *of the* 3×3 *matrix* A *is the matrix that has* $C_{j,i}$ *in the* (i, j)*-position. It is the* transpose *of the matrix of cofactors of* A.

By looking at any of the three entries on the main diagonal of the product $A \times \operatorname{adj}(A)$, we can give the following description of the determinant of a 3×3 matrix.

Definition 3.5.5. *The determinant of a* 3×3 *matrix is* A *can be found by choosing any row or column of* A*, multplying each entry of that row or column by its own cofactor, and adding the results.*

NOTES

1. Each of the definitions above applies to $n \times n$ matrices in general, and gives us a way to recursively define a $n \times n$ determinant, in terms of $(n-1) \times (n-1)$ determinants.
2. The *cofactor expansion* method, described in Definition 3.5.5 above, is not generally the most efficient way to compute a determinant (it is ok in the 3×3 case). But it can be taken as the *definition* of a determinant.
3. In some special cases, the determinant is easier to compute. If A is upper or lower triangular, then $\det(A)$ is the product of the entries on the main diagonal of A. If A has a square $k \times k$ block A_1 in the upper left, a square $(n-k) \times (n-k)$ block in the lower right, and only zeros in the lower left $(n-k) \times k$ region, then $\det(A) = \det(A_1)\det(A_2)$.
4. For a pair of $n \times n$ matrices A and B, $\det(AB) = \det(A)\det(B)$. This is the *multplicative property* of the determinant, or the Cauchy-Binet formula. It is not obvious at all.

3.5.2 Algebraic and Geometric Multiplicity

Example 3.5.6. Using cofactor expansion by the first column, we find that the characteristic polynomial of $B = \begin{bmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{bmatrix}$ is

$$\begin{aligned}
\det(\lambda I_3 - B) = \det \begin{bmatrix} \lambda-5 & -6 & -2 \\ 0 & \lambda+1 & 8 \\ -1 & 0 & \lambda+2 \end{bmatrix} & \\
&= (\lambda-5)\left((\lambda+1)(\lambda+2) - 0(8)\right) + (-1)\left((-6)(8) - (\lambda+1)(-2)\right) \\
&= (\lambda-5)(\lambda^2+3\lambda+2) - (2\lambda-46) \\
&= \lambda^3 - 2\lambda^2 - 15\lambda + 36 \\
&= (\lambda-3)(\lambda^2+\lambda-12) \\
&= (\lambda-3)(\lambda+4)(\lambda-3) \\
&= (\lambda-3)^2(\lambda+4)
\end{aligned}$$

The eigenvalues of B are 3 (occurring twice as a root of the characteristic polynomial), and -4 (occurring once). We say that 3 has *algebraic multplicity* 2 and -4 has *algebraic multiplicity* 1 as an eigenvalue of B. The *geometric multplicity* of each eigenvalue is the dimension of its corresponding eigenspace.

The eigenspace of B corresponding to $\lambda = 3$ is the nullspace of the matrix

$$3I_3 - B = \begin{bmatrix} -2 & -6 & -2 \\ 0 & 4 & 8 \\ -1 & 0 & 5 \end{bmatrix}$$

The RREF of this matrix is

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

and the nullspace consists of all vectors $\begin{bmatrix} 5t \\ -2t \\ t \end{bmatrix}$, where $t \in \mathbb{R}$. This is the eigenspace of B corresponding to $\lambda = 3$. It has dimension 1, so 3 has *geometric multiplicity 1* as an eigenvector of B.

Theorem 3.5.7. *The geometric multplicity of an eigenvector is at most equal to its algebraic multiplicity.*

Corollary 3.5.8. *A matrix is diagonalizable if and only if the geometric multiplicity of each of its eigenvalues is equal to the algebraic multiplicity.*