

Chapter 7

Eigenvalues and Eigenvectors

7.1 Let us recall that...

7.1.1 Trace and determinant. If $A \in \mathbb{K}^{n \times n}$, the *trace* of A is the scalar

$$\mathrm{tr}(A) = \sum_{i=1}^n A(i, i),$$

that is, the sum of all the entries on the main diagonal. The determinant of A is the scalar

$$\det(A) := \det(C_1(A), \dots, C_n(A))$$

where $\det : \mathbb{K}^n \times \dots \times \mathbb{K}^n \rightarrow \mathbb{K}$ is the unique alternating multilinear form that takes the value 1 on the canonical basis $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ (cf. Chapter ??).

7.1.2 A square matrix $A \in \mathbb{K}^{n \times n}$ has maximal rank equal to n if and only if its determinant is nonzero. Recall also that the set of all invertible $n \times n$ square matrices forms a group, denoted by $Gl_n(\mathbb{K})$, with respect to matrix multiplication.

7.1.3 Associated matrices. Let $\mathcal{B} := (\mathbf{b}_1, \dots, \mathbf{b}_n)$ be a basis of a vector space V of dimension n . If $\mathbf{v} \in V$, it can be expressed as a unique linear combination of the elements of $\mathcal{B}$, namely $\mathbf{v} = v_1 \mathbf{b}_1 + \dots + v_n \mathbf{b}_n$. Let $(\mathbf{v})_{\mathcal{B}}$ be the column of the components of $\mathbf{v}$ in the basis $\mathcal{B}$. Then the function $V \rightarrow \mathbb{K}^n$ defined by

$$\mathbf{v} \longmapsto (\mathbf{v})_{\mathcal{B}} \tag{7.1}$$

is an isomorphism of $\mathbb{K}$ -vector spaces. It is the unique isomorphism such that

$$(\mathbf{b}_i)_{\mathcal{B}} = \mathbf{e}_i. \tag{7.2}$$

If $\mathcal{C} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ is another basis of V , then $P^{\mathcal{C}, \mathcal{B}} \in Gl_n(\mathbb{K})$ is the unique matrix such that

$$(\mathbf{v})_{\mathcal{C}} = P^{\mathcal{C}, \mathcal{B}}(\mathbf{v})_{\mathcal{B}}.$$

It is easy to see that $P^{\mathcal{C}, \mathcal{B}} P^{\mathcal{B}, \mathcal{C}} = \mathbf{1}_n$. If $f \in End_{\mathbb{K}}(V)$, then the *matrix associated with f with respect to the basis $\mathcal{B}$* is the matrix

$$M_f^{\mathcal{B}, \mathcal{B}} = (f(\mathbf{b}_1)_{\mathcal{B}}, \dots, f(\mathbf{b}_n)_{\mathcal{B}}).$$

Thus

$$(f(\mathbf{v}))_{\mathcal{B}} = M_f^{\mathcal{B}, \mathcal{B}} \cdot (\mathbf{v})_{\mathcal{B}}.$$

Moreover:

$$M_f^{\mathcal{C}, \mathcal{C}} = P^{\mathcal{C}, \mathcal{B}} M_f^{\mathcal{B}, \mathcal{B}} P^{\mathcal{B}, \mathcal{C}} = (P^{\mathcal{B}, \mathcal{C}})^{-1} M_f^{\mathcal{B}, \mathcal{B}} P^{\mathcal{B}, \mathcal{C}}. \quad (7.3)$$

7.2 Similar and diagonalizable matrices

The relation between two square matrices associated with the same endomorphism (with respect to two bases, possibly the same one) suggests the following:

7.2.1 Definition. *Two matrices $A, B \in \mathbb{K}^{n \times n}$ are said to be similar if and only if there exists $P \in Gl_n(\mathbb{K})$ such that $B = P^{-1}AP$.*

Thus one may say that the matrices associated with the same endomorphism, with respect to two arbitrarily chosen bases of $\mathbb{K}^n$, are similar. Conversely, every matrix C similar to a matrix $B := M_f^{\mathcal{B}, \mathcal{B}}$ associated with an endomorphism f with respect to the basis $\mathcal{B}$ is a matrix associated with f . Indeed, let P be such that $C = P^{-1}BP$ and let

$$\mathcal{C} := (\mathbf{c}_1, \dots, \mathbf{c}_n)$$

be such that $((\mathbf{c}_1)_{\mathcal{B}}, \dots, (\mathbf{c}_n)_{\mathcal{B}}) = P$. Then $\mathcal{C}$ is a basis of V and

$$M_f^{\mathcal{C}, \mathcal{C}} = P^{\mathcal{C}, \mathcal{B}} M_f^{\mathcal{B}, \mathcal{B}} P^{\mathcal{B}, \mathcal{C}} = P^{-1}BP = C.$$

7.2.2 Proposition *The relation*

$$A \sim B \iff A \text{ is similar to } B,$$

called similarity (the similarity relation), is an equivalence relation.

Proof. We must prove that the relation is *reflexive*, *symmetric* and *transitive*. First, $A \sim A$, because $A = P^{-1}AP$ where $P = \mathbf{1}_n$. Hence $\sim$ is reflexive. If $A \sim B$, then there exists $P \in Gl_n(\mathbb{K})$ such that $B = P^{-1}AP$. Then $A = Q^{-1}BQ$ where

$Q = P^{-1} \in Gl_n(\mathbb{K})$, hence $B \sim A$, so the relation $\sim$ is symmetric. Finally suppose that A, B, C are three $n \times n$ matrices such that $A \sim B$ and $B \sim C$. Then there exist $P, Q \in Gl_n(\mathbb{K})$ such that $B = P^{-1}AP$ and $C = Q^{-1}BQ$. Thus:

$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (Q^{-1}P^{-1})A(PQ) = (PQ)^{-1}A(PQ).$$

Therefore $C = R^{-1}AR$, where R is the invertible matrix PQ . ■

7.2.3 Characteristic polynomial. Let $\mathbb{K}[T]$ be the vector space of polynomials with coefficients in $\mathbb{K}$. A polynomial $P \in \mathbb{K}[T]$ is called *monic* if the coefficient of the highest power is 1. With every matrix $A \in \mathbb{K}^{n \times n}$ one may associate a monic polynomial of degree n with coefficients in $\mathbb{K}$, denoted by $P_A(T)$, called the *characteristic polynomial of A* and defined by:

$$P_A(T) := |T\mathbf{1}_n - A| = \begin{vmatrix} T - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & T - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & T - a_{nn} \end{vmatrix} \quad (7.4)$$

where $a_{ij} := A(i, j)$. In other words, the characteristic polynomial is the determinant of the matrix

$$T\mathbf{1}_n - A = (T\mathbf{e}_1 - A \cdot \mathbf{e}_1, \dots, T\mathbf{e}_n - A \cdot \mathbf{e}_n),$$

which is obtained by formally adding the indeterminate T to the main diagonal of the matrix $-A$. Since changing the sign of one column of a square matrix changes the sign of the determinant, one also has:

$$P_A(T) = (-1)^n \det(A - T\mathbf{1}_n).$$

Indeed:

$$\begin{aligned} P_A(T) : &= \det(T\mathbf{e}_1 - A \cdot \mathbf{e}_1, \dots, T\mathbf{e}_n - A \cdot \mathbf{e}_n) = \\ &= \det(-(A\mathbf{e}_1 - T\mathbf{e}_1), \dots, -(A\mathbf{e}_n - T\mathbf{e}_n)) = \\ &= (-1)^n \det(A - T\mathbf{1}_n). \end{aligned}$$

That the characteristic polynomial of a matrix is a monic polynomial of degree n is clear from the explicit expression (7.4). It is also straightforward to check that the coefficient of the monomial T^{n-1} is the negative of the *trace* $\text{tr}(A) = \sum_{i=1}^n a_{ii}$, while the constant term, that is the coefficient of degree 0, is the determinant, as is seen by

substituting $T = 0$ in (7.4). We shall write the characteristic polynomial of a matrix $A \in \mathbb{K}^{n \times n}$ in the following form

$$T^n - e_1 T^{n-1} + \dots + (-1)^n e_n = 0,$$

where, as said, $e_1 = \text{tr}(A)$ (the trace of the matrix) and $e_n = \det(A)$. The other coefficients, $e_2, \dots, e_{n-1}$ of $P_A(T)$, together with e_1 and e_n , are called the *principal invariants* of a matrix. Each e_i is a homogeneous polynomial expression of degree i in the entries of the matrix A .

7.2.4 Exercise. Determine the invariant e_2 of a 3×3 matrix:

$$A := \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

(Hint: write out the characteristic polynomial. The coefficient of T^2 is precisely the invariant e_2 .)

7.2.5 Example (The characteristic polynomial of 2×2 matrices). Let

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{K}^{2 \times 2}. \quad (7.5)$$

Then:

$$P_A(T) = (-1)^2 \det(A - T\mathbf{1}_2) = \det(A - T\mathbf{1}_2) = \begin{vmatrix} a - T & b \\ c & d - T \end{vmatrix}.$$

Expanding the determinant gives:

$$P_A(T) = (a - T)(d - T) - bc = T^2 - (a + d)T + ad - bc.$$

Notice that the coefficient of T is the sum of the elements on the main diagonal of the matrix, called the *trace*, while the constant coefficient, that is the coefficient of T^0 , is the determinant of the matrix A . Therefore, if $A \in \mathbb{K}^{2 \times 2}$, the characteristic polynomial of A is:

$$P_A(T) = T^2 - \text{tr}(A)T + \det(A).$$

7.2.6 Esempio. Let

$$A := \begin{pmatrix} 3 & 6 \\ 5 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Then $P_A(T) = T^2 - 4T - 27$.

7.2.7 Exercise. Prove, by a direct computation, that if A is as in (7.5), then

$$A^2 - \operatorname{tr}(A)A + \det(A)\mathbf{1}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This is a special case of a far more general result:

7.2.8 Theorem (Cayley¹ and Hamilton²). *Every matrix $A \in \mathbb{K}^{n \times n}$ is a root of its own characteristic polynomial, that is:*

$$P_A(A) = \mathbf{0}_{n \times n}$$

where $P_A(A)$ is the matrix obtained by substituting A for T in the expression of the characteristic polynomial of A .

The interested reader may find a mechanical proof of the Cayley-Hamilton Theorem in the formidable text [?], on page 62. For a more conceptual proof, based on the notion of a Hasse-Schmidt derivation on an exterior algebra, see [?]³ Here instead we shall prefer to show an application of the theorem to compute the inverse of a matrix. The degree 0 coefficient of the characteristic polynomial of a matrix A is the determinant, so the matrix is invertible if and only if $P_A(T)$ is not divisible by T (equivalently, it does not vanish at $T = 0$). From the Cayley-Hamilton theorem we then have:

$$A^n - e_1 A^{n-1} + \dots + (-1)^n e_n \mathbf{1}_n = 0$$

and hence

$$A(A^{n-1} - e_1 A^{n-2} + \dots + (-1)^{n-1} e_{n-1} \mathbf{1}_n) = (-1)^{n+1} e_n \mathbf{1}_n$$

that is,

$$A^{-1} = (-1)^{n+1} \frac{A^{n-1} - e_1 A^{n-2} + \dots + (-1)^{n-1} e_{n-1} \mathbf{1}_n}{e_n}.$$

¹Arthur Cayley, 16 August 1821–Cambridge, 26 January 1895, the English mathematician whose name is especially linked to invariant theory.

²Sir William Rowan Hamilton, Dublin, 4 August 1805 – Dublin, 2 September 1865, the Irish mathematician best known to the general public for having introduced the hypercomplex numbers called *quaternions*, which have found many beautiful applications also in the description of rigid body mechanics.

³This is perhaps the easiest proof, provided one is familiar with exterior algebra, disguised and smuggled into this text in Chapter ?? on determinants.

7.2.9 Esemplio. Here is how to determine a formula for the inverse of a 2×2 matrix with nonzero determinant:

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

By the Cayley-Hamilton theorem,

$$A^2 - e_1 A + e_2 \mathbf{1} = 0$$

where $e_1 = \text{tr}(A) = a + d$ and $e_2 = \det(A) = ad - bc$. Therefore

$$A^2 - e_1 A = -e_2 \mathbf{1}_2$$

that is,

$$A(A - e_1 \mathbf{1}_2) = -e_2 \mathbf{1}_2.$$

By the definition of inverse,

$$A^{-1} = \frac{e_1 \mathbf{1}_2 - A}{e_2} = \frac{1}{\det(A)} \left[\begin{pmatrix} e_1 & 0 \\ 0 & e_1 \end{pmatrix} - \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right] = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

7.2.10 Esemplio. To find a general formula for the inverse of a 3×3 matrix, one writes the Cayley-Hamilton theorem:

$$A^3 - e_1 A^2 + e_2 A - e_3 \mathbf{1}_3 = \mathbf{0}_{3 \times 3}$$

where $e_1 = \text{tr}(A)$ and $e_3 = \det(A)$, hence

$$A(A^2 - e_1 A + e_2 \mathbf{1}_3) = e_3 \mathbf{1}_3$$

and therefore

$$A^{-1} = \frac{A^2 - e_1 A + e_2 \mathbf{1}_3}{e_3}.$$

7.2.11 Proposition. *Similar matrices have the same characteristic polynomial.*

Proof. Let $A \sim B$. Then there exists $P \in GL_n(\mathbb{K})$ such that $B = P^{-1}AP$. Hence

$$P_B(T) = |T\mathbf{1} - B| = |TP^{-1}\mathbf{1}_n P - P^{-1}AP| = |P^{-1}(T\mathbf{1}_n - A)P|.$$

Using Binet's theorem for the determinant of the product of square matrices:

$$P_B(T) = |P^{-1}| |T\mathbf{1}_n - A| |P| = |T\mathbf{1}_n - A| = P_A(T),$$

having used the commutativity of multiplication in $\mathbb{K}$ as well as the fact that $|P^{-1}||P| = |P^{-1}P| = |\mathbf{1}_n| = 1$. ■

7.2.12 Definition. A matrix $A \in \mathbb{K}^{n \times n}$ is called diagonalizable if and only if A is similar to a diagonal matrix. That is, there exists a matrix $P \in Gl_n(\mathbb{K})$ such that $P^{-1}AP$ is diagonal.

The converse of Proposition 7.2.11 does not hold. For example, the matrices in $\mathbb{R}^{2 \times 2}$

$$\mathbf{1}_2 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad J := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

have the same characteristic polynomial $P_{\mathbf{1}_2}(T) = (T - 1)^2 = P_J(T)$ but are not similar. Indeed, if $\mathbf{1}_2$ and J were similar, there would exist $P \in Gl_2(\mathbb{R})$ such that $J = P^{-1}\mathbf{1}_2P = J$. But this is impossible, because for every invertible P one has:

$$P^{-1}\mathbf{1}_2P = P^{-1}P = \mathbf{1}_2.$$

7.3 Eigenvalues and eigenvectors of an endomorphism

7.3.1 Invariant subspaces. Let $W \in G(V)$ and $f \in End_{\mathbb{K}}(V)$. The set W is called *invariant* under f , or briefly *f-invariant*, if and only if

$$f(W) \subseteq W. \tag{7.6}$$

Equation (7.6) is equivalent to saying that if $\mathbf{v} \in W$, then $f(\mathbf{v}) \in W$. Clearly the trivial subspaces of V (that is, $\{\mathbf{0}\}$ and V itself) are f -invariant for every $f \in End_{\mathbb{K}}(V)$. Particularly interesting are the f -invariant subspaces of V having the smallest possible positive dimension.

7.3.2 Definition. An eigendirection of f is an f -invariant vector line.

Every eigendirection of f is of the form $[\mathbf{v}]$ for some $\mathbf{v} \neq \mathbf{0}$. The vector $\mathbf{v}$ (or any other generator of the same eigendirection of f) is called an *eigenvector* of f .

7.3.3 Eigenvalues. If $[\mathbf{v}]$ is an eigendirection, then $f(\mathbf{v}) \in [\mathbf{v}]$, that is, there exists $\lambda \in \mathbb{K}$ such that $f(\mathbf{v}) = \lambda \cdot \mathbf{v}$. The scalar λ is called the *eigenvalue* of f associated with the eigendirection $[\mathbf{v}]$. This terminology is justified by the fact that this scalar does not depend on the choice of a generator of the eigendirection. Indeed, if $\mathbf{w} \in [\mathbf{v}]$, then $\mathbf{w}$ is a multiple of $\mathbf{v}$, that is $\mathbf{w} = \mu\mathbf{v}$ for some $\mu \in \mathbb{K}$, and hence:

$$f(\mathbf{w}) = f(\mu\mathbf{v}) = \mu f(\mathbf{v}) = \mu\lambda\mathbf{v} = \lambda(\mu\mathbf{v}) = \lambda\mathbf{w},$$

as was to be verified. Alternatively:

7.3.4 Definition. A scalar $\lambda \in \mathbb{K}$ is called an *eigenvalue* of $f \in \text{End}_{\mathbb{K}}(V)$ if and only if there exists $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ such that $f(\mathbf{v}) = \lambda \mathbf{v}$.

The vector $\mathbf{v}$ appearing in the definition is therefore a generator of an eigendirection of f , namely the *eigendirection associated with the eigenvalue* λ .

7.3.5 Equation (7.3) suggests the idea of determining a basis of V with respect to which the matrix associated with an endomorphism f is as simple as possible. Suppose that the endomorphism f has n distinct eigendirections:

$$[\mathbf{v}_1], \dots, [\mathbf{v}_n],$$

such that $V = [\mathbf{v}_1] \oplus \dots \oplus [\mathbf{v}_n]$. Then $\mathcal{V} := (\mathbf{v}_1, \dots, \mathbf{v}_n)$ is a basis of V such that $f(\mathbf{v}_i) \in [\mathbf{v}_i]$. Hence there exists $\lambda_i \in \mathbb{K}$ such that

$$f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i.$$

Therefore

$$M_f^{\mathcal{V}, \mathcal{V}} = (f(\mathbf{v}_1)_{\mathcal{V}}, \dots, f(\mathbf{v}_n)_{\mathcal{V}}) = \quad (7.7)$$

$$= ((\lambda_1 \mathbf{v}_1)_{\mathcal{V}}, \dots, (\lambda_n \mathbf{v}_n)_{\mathcal{V}}) = \quad (7.8)$$

$$= (\lambda_1 (\mathbf{v}_1)_{\mathcal{V}}, \dots, \lambda_n (\mathbf{v}_n)_{\mathcal{V}}) = (\lambda_1 \mathbf{e}_1, \dots, \lambda_n \mathbf{e}_n) \quad (7.9)$$

where in (7.7) we used the definition of associated matrix, in (7.8) the fact that each $\mathbf{v}_i$ is an eigenvector of f , and in (7.9) the linearity of the isomorphism (7.1) and equation (7.2). Conversely, if $M_f^{\mathcal{V}, \mathcal{V}} = (\lambda_1 \mathbf{e}_1, \dots, \lambda_n \mathbf{e}_n)$, then

$$f(\mathbf{v}_i)_{\mathcal{V}} = M_f^{\mathcal{V}, \mathcal{V}} \cdot (\mathbf{v}_i)_{\mathcal{V}} = M_f^{\mathcal{V}, \mathcal{V}} \cdot \mathbf{e}_i \quad (7.10)$$

where in the first equality we used the definition of associated matrix and in the second equation (7.2), which expresses the components of the vectors of a basis $\mathcal{V}$ with respect to $\mathcal{V}$ itself. Since every $m \times n$ matrix multiplied by the coordinate column $\mathbf{e}_i \in \mathbb{K}^n$ gives the i -th column of the matrix itself, we have:

$$f(\mathbf{v}_i)_{\mathcal{V}} = C_i(M_f^{\mathcal{V}, \mathcal{V}}) = \lambda_i \mathbf{e}_i = \lambda_i (\mathbf{v}_i)_{\mathcal{V}}$$

therefore

$$f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$$

and $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ is thus a basis of V made of eigenvectors of f . We have proved:

7.3.6 Proposition. The matrix associated with $f \in \text{End}_{\mathbb{K}}(V)$ with respect to a basis $\mathcal{V}$ of V is diagonal if and only if $\mathcal{V}$ is a basis of eigenvectors of f .

7.4 Determining eigenvalues and eigenvectors

7.4.1 Let now $\mathcal{B} := (\mathbf{b}_1, \dots, \mathbf{b}_n)$ be an arbitrary basis of V and let $\lambda \in \mathbb{K}$ be an arbitrary scalar. Consider the vector subspace of V :

$$V_\lambda(f) := \ker(f - \lambda \text{id}_V).$$

Then $V_\lambda(f) \neq \{\mathbf{0}\}$ if and only if there exists $\mathbf{v} \neq \mathbf{0}$ such that $f(\mathbf{v}) - \lambda \mathbf{v} = \mathbf{0}$, that is, if and only if λ is an eigenvalue of f . Let

$$\text{Spec}(f) := \{\lambda \in \mathbb{K} \mid \ker(f - \lambda \cdot \text{id}_V) \neq \mathbf{0}\}$$

be the set⁴ of the eigenvalues of f . If $\lambda \in \text{Spec}(f)$, the subspace $V_\lambda(f) := \ker(f - \lambda \text{id}_V)$ is called the *eigenspace* of f associated with $\lambda \in \text{Spec}(f)$.

7.4.2 Remark. If f is a non-invertible endomorphism, that is, if $\ker f \neq \{\mathbf{0}\}$, then the kernel is an eigenspace associated with the eigenvalue 0. Let us say this again. A non-invertible endomorphism has at least one zero eigenvalue, and its kernel is the corresponding eigenspace.

7.4.3 Considering the dimensions of the subspaces above, one easily concludes that $\lambda \in \text{Spec}(f)$ if and only if

$$\dim_{\mathbb{K}} \ker(f - \lambda \text{id}_V) > 0. \quad (7.11)$$

The inequality (7.11) expresses the fact that, by definition, an eigenspace contains at least one nonzero vector. Moreover, by the rank-nullity theorem for an endomorphism, one has:

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \ker(f - \lambda \cdot \text{id}_V) + \dim_{\mathbb{K}} \text{Im}(f - \lambda \cdot \text{id}_V),$$

hence $\lambda \in \text{Spec}(f)$ if and only if $\dim_{\mathbb{K}} \text{Im}(f - \lambda \cdot \text{id}_V) < n$.

7.4.4 Let $\lambda \in \text{Spec}(f)$ and let $[\mathbf{v}] \in \mathbb{P}(V)$ be an eigendirection. The equality

$$f(\mathbf{v}) - \lambda \mathbf{v} = \mathbf{0},$$

implies equality of the $\mathcal{B}$ -components:

$$(f(\mathbf{v}) - \lambda \mathbf{v})_{\mathcal{B}} = (\mathbf{0}_V)_{\mathcal{B}},$$

⁴The *spectrum* of an endomorphism is the set of its eigenvalues; for this reason it is denoted by Spec , an abbreviation of the word *spectrum*, widely used in physics. For example, the *spectrum of the energy levels of the hydrogen atom*, described by Rydberg's formula, is described mathematically by means of the spectrum (that is, the eigenvalues) of the Schroedinger operator. See for instance [?, p. 133].

whence, by the definition of the matrix associated with an endomorphism:

$$M_{f-\lambda \text{id}_V}^{\mathcal{B},\mathcal{B}}(\mathbf{v})_{\mathcal{B}} = \mathbf{0}.$$

Using property (??) of associated matrices, one has:

$$(M_f^{\mathcal{B},\mathcal{B}} - \lambda M_{\text{id}_V}^{\mathcal{B},\mathcal{B}}) \cdot (\mathbf{v})_{\mathcal{B}} = \mathbf{0},$$

which, observing that $M_{\text{id}_V}^{\mathcal{B},\mathcal{B}} = \mathbf{1}_n$, may also be written in the form:

$$(M_f^{\mathcal{B},\mathcal{B}} - \lambda \mathbf{1}_n) \cdot (\mathbf{v})_{\mathcal{B}} = \mathbf{0}. \quad (7.12)$$

It follows that $(\mathbf{v})_{\mathcal{B}}$ is a nonzero solution of the linear system

$$(A - \lambda \cdot \mathbf{1}_n) \vec{X} = \mathbf{0}$$

where $A := M_f^{\mathcal{B},\mathcal{B}}$. The rank of the coefficient matrix, $A - \lambda \cdot \mathbf{1}_n$, must therefore be less than n , hence $\lambda \in \text{Spec}(f)$ if and only if $\lambda \in \mathbb{K}$ is a zero of the *characteristic equation*:

$$|A - T\mathbf{1}_n| = 0. \quad (7.13)$$

7.4.5 The *characteristic polynomial* of $f \in \text{End}_{\mathbb{K}}(V)$ is the characteristic polynomial of the matrix $A := M_f^{\mathcal{B},\mathcal{B}}$, where $\mathcal{B}$ is any basis of V . Indeed, it does not depend on the basis. If $\mathcal{C}$ were another basis of V , the matrix $B := M_f^{\mathcal{C},\mathcal{C}}$ would be similar to A , and similar matrices have the same characteristic polynomial by Proposition 7.2.11. We have therefore proved that

7.4.6 Proposition. *A scalar $\lambda \in \mathbb{K}$ is an eigenvalue of an endomorphism f if and only if $P_f(\lambda) = 0$, that is, if and only if it is a zero of the characteristic polynomial of f .* ■

If $A \in \mathbb{K}^{n \times n}$, an eigenvalue of a matrix A is an eigenvalue of A regarded as an endomorphism of $\mathbb{K}^n$:

$$\left\{ \begin{array}{lcl} A & : & \mathbb{K}^n \longrightarrow \mathbb{K}^n \\ & & \mathbf{u} \longmapsto A\mathbf{u} \end{array} \right.$$

Thus $\lambda \in \mathbb{K}$ is an eigenvalue of A if and only if it is a zero of its characteristic polynomial, that is, if and only if

$$\det(A - \lambda \mathbf{1}_n) = 0.$$

7.4.7 Esemplio. The eigenvalues of the matrix $A := \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}$ are 1 and 6, as is found easily by computing the roots of the characteristic polynomial $P_A(T) = T^2 - 7T + 6$.

7.4.8 Exercise. Find the eigenvalues of the matrix

$$A := \begin{pmatrix} 4 & 2 & 2 & 2 \\ 8 & 4 & 4 & 4 \\ 7 & 0 & 3 & 2 \\ -19 & -6 & -9 & -8 \end{pmatrix}$$

by computing the roots of the characteristic polynomial.

Solution. Everything is on page 37 of [?].

7.4.9 Remark. Suppose that λ is an eigenvalue of a matrix $A \in \mathbb{K}^{n \times n}$. If $\mu \in \mathbb{K}$ is any scalar, then $\lambda + \mu$ is an eigenvalue of the matrix $A + \mu \mathbf{1}_n$. Indeed, it is enough to check that $\lambda + \mu$ is a root of the characteristic polynomial of $A + \mu \mathbf{1}_n$. Now:

$$P_{A+\mu \mathbf{1}_n}(T) = |A + \mu \mathbf{1}_n - T \mathbf{1}_n|,$$

hence

$$P_{A+\mu \mathbf{1}_n}(\lambda + \mu) = |A + \mu \mathbf{1}_n - (\lambda + \mu) \mathbf{1}_n| = |A - \lambda \mathbf{1}_n| = 0$$

where the last equality ($= 0$) follows from the hypothesis that λ was an eigenvalue of A .

7.4.10 Esemplio. Consider the matrices

$$A := \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 6 & 1 \\ 1 & 6 \end{pmatrix}.$$

Since 0 and 2 are eigenvalues of the matrix A , the matrix B has eigenvalues 5 and 7. Indeed:

$$B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$$

and therefore, by Remark 7.4.9, the eigenvalues of B are those of A plus 5.

7.5 Simple endomorphisms and diagonalizable matrices

7.5.1 If λ is an eigenvalue of f , then $P_f(\lambda) = 0$ and, by Ruffini's theorem, the polynomial $P_f(T)$ is divisible by $(T - \lambda)$. Thus one may write

$$P_f(T) = Q(T)(T - \lambda)$$

where $Q(T) \in \mathbb{K}[T]$. Let now $m \geq 1$ be an integer such that

$$P_f(T) = Q(T)(T - \lambda)^m$$

with $Q(\lambda) \neq 0$. This means that $Q(T)$ is not itself divisible by $T - \lambda$ and therefore that $(T - \lambda)^m$ is the highest power of $T - \lambda$ dividing $P_f(T)$. This integer m is called the *algebraic multiplicity* of $\lambda \in \text{Spec}(f)$ and will be denoted by $ma(\lambda)$.

7.5.2 Proposition. *An endomorphism f of a vector space V of dimension $n \geq 1$ has at most n eigenvalues, counted with their algebraic multiplicities.*

Proof. Indeed, the characteristic polynomial $P_f(T)$ has degree n , and a polynomial of degree n has at most n roots. ■

7.5.3 Corollary. *If $\mathbb{K} = \mathbb{C}$, or if $\mathbb{K}$ is an algebraically closed field, an endomorphism f of a $\mathbb{K}$ -vector space V of dimension n has exactly n eigenvalues counted with the proper multiplicities.*

Proof. Indeed, every polynomial $P(T)$ of degree n over an algebraically closed field $\mathbb{K}$ has at least one root $\lambda \in \mathbb{K}$. It can therefore be written as

$$P(T) = Q(T)(T - \lambda)$$

where Q is a polynomial with coefficients in $\mathbb{K}$ of degree $n - 1$. By induction, Q has $(n - 1)$ roots counted with their multiplicities. ■

7.5.4 Definition. *The geometric multiplicity $mg(\lambda)$ of an eigenvalue λ is the dimension of the corresponding eigenspace:*

$$mg(\lambda) = \dim_{\mathbb{K}} \ker(f - \lambda \cdot \text{id}_V).$$

7.5.5 Theorem. *For every $\lambda \in \text{Spec}(f)$ the following inequality holds:*

$$1 \leq mg(\lambda) \leq ma(\lambda).$$

Proof. The proof consists in showing that if $ma(\lambda) = m$, then there exists an f -invariant subspace W of V of dimension m and a basis $\mathcal{B}_W$ such that the matrix associated with the restriction of f to W has the eigenvalue λ on every diagonal entry. For this purpose one would need to introduce the minimal polynomial, which lies outside the scope of this text. The curious reader may consult [?, Chapter 4]. ■

7.5.6 Proposition. *Let $\lambda, \mu \in \mathbb{K}$ be two distinct eigenvalues of an endomorphism f . Then*

$$V_\lambda(f) \cap V_\mu(f) = \{\mathbf{0}\}.$$

Proof. Indeed, if $\mathbf{v} \in V_\lambda(f) \cap V_\mu(f)$, one would have $f(\mathbf{v}) = \lambda\mathbf{v} = \mu\mathbf{v}$. Therefore $\lambda\mathbf{v} - \mu\mathbf{v} = \mathbf{0}_V$, and hence $(\lambda - \mu)\mathbf{v} = \mathbf{0}$. This implies $\mathbf{v} = \mathbf{0}_V$, by the initial hypothesis $\lambda \neq \mu$. ■

Proposition 7.5.6 implies in particular that the sum of the geometric multiplicities of the eigenvalues of an endomorphism is less than or equal to n . Indeed, if $\text{Spec}(f) := \{\lambda_1, \dots, \lambda_h\}$, with $\lambda_i \neq \lambda_j$ when $i \neq j$, then

$$\bigoplus_{\lambda \in \text{Spec}(f)} V_\lambda(f) := V_{\lambda_1}(f) \oplus \dots \oplus V_{\lambda_h}(f)$$

is a subspace of V . Hence

$$\sum_{\lambda \in \text{Spec}(f)} mg(\lambda) = \dim_{\mathbb{K}} \left(\bigoplus_{\lambda \in \text{Spec}(f)} V_\lambda(f) \right) = \sum_{\lambda \in \text{Spec}(f)} \dim_{\mathbb{K}} V_\lambda(f) \leq \dim_{\mathbb{K}} V = n,$$

as desired. ■

7.5.7 Definition. *An endomorphism f of a $\mathbb{K}$ -vector space V will be called (semi)simple if V decomposes as the direct sum of eigenspaces of f , or equivalently if V has a basis made of eigenvectors of f .*

Taking Proposition 7.3.6 into account, we have:

7.5.8 Proposition. *An endomorphism is simple if and only if $\sum_{\lambda \in \text{Spec}(f)} ma(\lambda) = n$ and $mg(\lambda) = ma(\lambda)$ for every $\lambda \in \text{Spec}(f)$.*

Proof. The first condition ensures that the endomorphism has n eigenvalues, counted with the proper multiplicity. The second implies that

$$\sum_{\lambda \in \text{Spec}(f)} mg(\lambda) = \sum_{\lambda \in \text{Spec}(f)} ma(\lambda) = n.$$

Let $\mathcal{B}_i$ be a basis of the eigenspace $V_{\lambda_i}(f)$, for every $1 \leq i \leq h$. Since the sum of eigenspaces corresponding to distinct eigenvalues is direct, the union $\mathcal{B} := \mathcal{B}_1 \cup \dots \cup \mathcal{B}_h$, is a basis of V made of eigenvectors of f , and therefore f is simple by 7.5.7. ■

7.5.9 Proposition. *A matrix $A \in \mathbb{K}^{n \times n}$ is diagonalizable if and only if:*

1. $\sum_{\lambda \in \text{Spec}(f)} ma(\lambda) = n$ and
2. *the algebraic multiplicity of each eigenvalue coincides with its geometric multiplicity.*

Proof. Suppose that A is diagonalizable. Then there exists a matrix $P \in GL_n(\mathbb{K})$ such that

$$P^{-1}AP = (\lambda_1 \mathbf{e}_1, \dots, \lambda_n \mathbf{e}_n), \quad (7.14)$$

with the scalars λ_i not necessarily distinct. Since similar matrices have the same characteristic polynomial and the same eigenvalues, it follows that $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A . Moreover, $P\mathbf{e}_i$ is an eigenvector of A with respect to the eigenvalue λ_i . Indeed

$$\begin{aligned}
 A(P\mathbf{e}_i) &= (PP^{-1})A(P\mathbf{e}_i) = && \text{(invertibility of } P) \\
 &= P(P^{-1}AP)\mathbf{e}_i = && \text{(associativity of} \\
 &&& \text{matrix multiplication)} \\
 &= P(\lambda_1 \mathbf{e}_1, \dots, \lambda_i \mathbf{e}_i, \dots, \lambda_n \mathbf{e}_n)\mathbf{e}_i = && \text{(by (7.14))} \\
 &= P(\lambda_i \mathbf{e}_i) = && \text{(product of a matrix} \\
 &&& \text{by the } i\text{-th column)} \\
 &= \lambda_i \cdot P\mathbf{e}_i. && \text{(linearity of matrix multiplication)}
 \end{aligned}$$

The columns of the matrix P therefore form a basis of $\mathbb{K}^n$ made of eigenvectors of A , and there are as many linearly independent eigenvectors associated with an eigenvalue λ as its algebraic multiplicity. Conversely, if $mg(\lambda) = ma(\lambda)$ for every eigenvalue of A , and $\sum ma(\lambda) = n$, then by construction $\mathbb{K}^n$ has a basis $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ made of eigenvectors of A . Setting $P := (\mathbf{u}_1, \dots, \mathbf{u}_n)$, one has

$$\begin{aligned}
 P^{-1}AP &= P^{-1}A(\mathbf{u}_1 \dots, A\mathbf{u}_n) = && \text{(definition of } P) \\
 &= P^{-1}(\lambda_1 \mathbf{u}_1, \dots, \lambda_n \mathbf{u}_n) = && \text{(hypothesis on the } \mathbf{u}_i, \text{ eigenvectors of } A) \\
 &= (\lambda_1 P^{-1}\mathbf{u}_1, \dots, \lambda_n P^{-1}\mathbf{u}_n) = && \text{(linearity of the product of} \\
 &&& \text{a matrix by a column)} \\
 &= (\lambda_1 \mathbf{e}_1, \dots, \lambda_n \mathbf{e}_n).
 \end{aligned}$$

as desired. ■

7.5.10 Notice that, en passant, Proposition 7.5.9 provides a practical method for finding the matrix $P \in Gl_n(\mathbb{K})$ that diagonalizes a diagonalizable matrix A . It is enough to take a P whose columns form a basis of $\mathbb{K}^n$ made of eigenvectors of A .

7.6 Exercises

7.6.1 Let the endomorphism of $\mathbb{R}^2$ be defined, for every vector $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, by:

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x + 4y \\ 4x + 9y \end{pmatrix}$$

1. Find the matrix M of f with respect to the canonical basis $E = (\mathbf{e}_1, \mathbf{e}_2)$;
2. Find the preimages of the vector $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$;
3. Find the preimages of $\mathbf{e}_1, \mathbf{e}_2$;
4. Determine a basis of the kernel and of the image of the matrix $M - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$;
5. Prove that $\vec{u} := \begin{pmatrix} 4 \\ 8 \end{pmatrix}$ is an eigenvector of M and that $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ is not;
6. Determine the eigenvalue of M associated with $\vec{u}$;
7. Prove that 5 is not an eigenvalue of M ;
8. prove that 1 is an eigenvalue of M ;
9. determine all eigenvalues of M ;
10. Determine all eigenvectors of M ;
11. Prove that there exists a basis of orthogonal vectors made of eigenvectors of M ;

12. Find a matrix $P \in Gl_2(\mathbb{R})$ that diagonalizes M .

Soluzione: [?, 10.7.1]

7.6.2 Diagonalize the matrix $A := \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 1 & 1 & 2 \end{pmatrix}$

Soluzione: [?, 10.7.2]. In full detail, on page 25 of the formidable [?], for anyone still able to find a copy.

7.6.3 Exercise. Prove that the matrix

$$A = \begin{pmatrix} -15 & 27 & -10 & -15 \\ -16 & 26 & -8 & -15 \\ -19 & 33 & -12 & -18 \\ 0 & -4 & 4 & 1 \end{pmatrix} \quad (7.15)$$

is diagonalizable, find a diagonal form for it, and determine the matrix $P \in Gl_n(\mathbb{K})$ that diagonalizes it into the chosen diagonal form.

Soluzione: See page 40 of the fantastic [?], for anyone able to obtain a copy. Otherwise, see *Soluzione:* [?, 10.7.3]



7.6.4 Diagonalize the matrix (7.15) using the commands

$$e = \mathbf{eig}(A) \quad \text{and} \quad [P, D] = \mathbf{eig}(A).$$

in MATLAB[®].

Soluzione: [?, 10.7.4]

7.6.5 Let the matrix be

$$A := \begin{pmatrix} 0 & 1 & 0 \\ 2 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}.$$

Prove that it is invertible and find an inverse using the Cayley-Hamilton theorem.

Soluzione: [?, 10.7.5]

7.6.6 Once upon a time there was an era in which, in linear algebra courses, one used to define the *matrix exponential* of a given matrix. If $A \in \mathbb{C}^{n \times n}$, the exponential of A is defined by the power series:

$$\exp(A) := \sum_{n \geq 0} \frac{A^n}{n!}. \quad (7.16)$$

Naturally it is by no means clear that the right-hand side of (7.16) is actually a matrix, that is, that the sum converges. It is not too difficult to show this, however. See for instance [?]. The purpose of this exercise is the following. Let

$$P_A(T) = T^n - e_1(A)T^{n-1} + \cdots + (-1)^n e_n(A)$$

be the characteristic polynomial of $A \in \mathbb{C}^{n \times n}$, with $e_i(A) \in \mathbb{C}$. Prove that

$$\exp(A \cdot x)$$

is a solution of the ordinary differential equation

$$y^{(n)} - e_1(A)y^{(n-1)} + \cdots + (-1)^n e_n(A)y = 0$$

if and only if A satisfies the Cayley-Hamilton theorem 7.2.8⁵

Soluzione: [?, 10.7.6]

7.6.7 Let the matrix be

$$A = \begin{pmatrix} -1 & 0 & 2 \\ -6 & 5 & 2 \\ 0 & 0 & 5 \end{pmatrix}$$

- i) Verify that $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ is an eigenvector of A and determine the corresponding eigenvalue.

⁵For numerical analytic applications of engineering relevance, see the fine article by Falcidieno and Luvison [?].

- ii) Compute the characteristic polynomial $p_A(t)$ of A ;
- iii) Determine all eigenvalues of A ;
- iv) Determine a basis for each eigenspace of A ;
- v) Determine an invertible matrix $P \in \mathbb{R}^{3 \times 3}$ such that $P^{-1}AP$ is diagonal;
- vi) Decide whether there exists a nonzero matrix $Q \in \mathbb{R}^{3 \times 3}$ such that $AQ = 5Q$.

Soluzione: [?, 10.7.7]

7.6.8 Let the matrices be

$$A := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ -1 & -2 & -3 \end{pmatrix}.$$

Which of the two is diagonalizable? And why? For the non-diagonalizable matrix, determine the singular value decomposition.

Soluzione: [?, 11.8.3]

7.6.9 Consider, for every $n \geq 1$, the endomorphism $\alpha \in \mathbb{C}^{n \times n}$ defined by $\alpha(M) = M + 2M^T$ (where M^T , as usual, denotes the transpose of M). Find a polynomial $P \in \mathbb{C}[t]$ such that $P(\alpha) = 0$, where 0 is the zero matrix. Compute the eigenvalues and eigenvectors of α and say whether α is invertible. Determine the kernel and image of α . Whoever wishes and knows how may of course use MATLAB®.

Soluzione: [?, 10.7.8]

