CHAPTER 6

Linear transformations

6.1 Linear transformations

Mappings between two vector spaces are, in many respects, more interesting than vector spaces themselves. This applies especially to linear transformations.

Let V, W be two vector spaces over the same field K. Then a mapping $\mathcal{T} : V \longrightarrow W$ is called a *linear transformation* from V to W if it satisfies the following conditions:

(1) $\mathcal{T}(x + y) = \mathcal{T}(x) + \mathcal{T}(y),\ \forall x, y \in V$

(2) $\mathcal{T}(kx) = k\mathcal{T}(x),\ \forall k \in K, x \in V.$

An immediate consequence of (2) is that the zero vector of V is mapped by every linear transformation into the zero vector of W, that is $\mathcal{T}(0) = 0$.

Sometimes we shall write $\mathcal{T}x$ instead of $\mathcal{T}(x)$.

Example 6.1.1 1) $\mathcal{T} : \mathbb{R}^2 \longrightarrow \mathbb{R}^3,\ \mathcal{T}(x_1, x_2) = (2x_1, -x_2, x_1 + x_2)$ defines a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$.

2) $\mathcal{D} : K[X] \longrightarrow K[X],\ \mathcal{D}(p) = p'\ \forall p \in K[X]$ is a linear transformation, the well-known derivation of polynomials.

It is not difficult to verify that $\mathcal{T} : V \longrightarrow W$ is linear if and only if for any given integer $r \geq 2$ we have

$$\mathcal{T}(k_1x_1 + \cdots + k_rx_r) = k_1\mathcal{T}(x_1) + \cdots + k_r\mathcal{T}(x_r)\ \forall x_i \in V, \forall k_i \in K$$

An important consequence of this property is expressed in the following theorem:

Theorem 6.1 *A linear transformation $\mathcal{T} : V_n \longrightarrow W$ is uniquely determined by the images $\mathcal{T}(b_1), \ldots, \mathcal{T}(b_n)$ of a basis $\{b_1, \ldots, b_n\}$ of V_n.*

Proof. Each vector $x \in V_n$ can be expressed uniquely in the form $x = k_1b_1 + \cdots + k_nb_n, k_i \in K$. Then $\mathcal{T}x = \mathcal{T}(k_1b_1 + \cdots + k_nb_n) = k_1\mathcal{T}b_1 + \cdots + k\mathcal{T}b_n$, hence $\mathcal{T}x$ is uniquely determined. □

For a linear transformation $\mathcal{T} : V \longrightarrow W$ denote

$$Ker(\mathcal{T}) = \{x \in V \mid \mathcal{T}x = 0\}, \quad Im(\mathcal{T}) = \{\mathcal{T}x \mid x \in V\}.$$

$Ker(\mathcal{T})$ is called the *kernel* of $\mathcal{T}$ and $Im(\mathcal{T})$ the *range* of $\mathcal{T}$.

Theorem 6.2 *$Ker(\mathcal{T})$ is a linear subspace of V. $Im(\mathcal{T})$ is a linear subspace of W.*

The (easy) proof is left to the reader.

Suppose that $dimV = n$; it can be shown that

$$dimKer(\mathcal{T}) + dimIm(\mathcal{T}) = n.$$

Definition 6.3 A linear transformation $\mathcal{T} : V \longrightarrow W$ is called an *isomorphism* if it is both one-to-one and onto W. V and W are called *isomorphic.*

The concept of isomorphism is of importance since any two isomorphic vector spaces have identical structure in the sense that any algebraic statement that is true for one space will necessarily be true for the other.

The next theorem is fundamental:

Theorem 6.4 *Two finite dimensional vector spaces over the same field are isomorphic if and only if they have the same dimension.*

We omit the proof but we mention the following obvious

Corollary 6.5 *Any vector space V over K and of dimension n is isomorphic to K^n.*

The reader may question why, in view of this result, we do not restrict our attention to the vector spaces K^n since these exhibit all the algebraic properties of abstract finite-dimensional vector spaces. The answer is that to do so would lead to unnecessary complications, in exactly the same way as in elementary vector analysis it is simpler to work with vectors as such, rather than to reduce every vector to a set of components.

Finally, denote by $\mathcal{L}(V, W)$ the set of all linear transformations from V to W. In $\mathcal{L}(V, W)$ we define addition and scalar multiplication by

$$\begin{array}{rcl} (\mathcal{T}+\mathcal{S})(x) & = & \mathcal{T}x+\mathcal{S}x \\ (k\mathcal{T})(x) & = & k\mathcal{T}x \end{array} \quad \forall x \in V,\ k \in K.$$

It is very easy to verify that, with these operations, $\mathcal{L}(V, W)$ forms a vector space over K.

Moreover, let V, W, U be three linear spaces over the same field K and let $\mathcal{T} \in \mathcal{L}(V, W)$, $\mathcal{S} \in \mathcal{L}(W, U)$. Consider the *composition* $\mathcal{S} \circ \mathcal{T} : V \longrightarrow U$ (called also the *product* and denoted simply by $\mathcal{S}\mathcal{T}$). Then $\mathcal{S}\mathcal{T} \in \mathcal{L}(V, U)$; we leave the proof to the reader.

The elements of $\mathcal{L}(V, V)$ are called the *endomorphisms* of the linear space V. Instead of $\mathcal{L}(V, V)$ we shall write simply $\mathcal{L}(V)$.

Denote by I the identity transformation of V, that is $Ix = x$, $\forall x \in V$. For $\mathcal{T} \in \mathcal{L}(V)$ let $\mathcal{T}^0 = I$, $\mathcal{T}^1 = \mathcal{T}, \mathcal{T}^2 = \mathcal{T}\mathcal{T}, \ldots$.

6.2 The matrix of a linear transformation

Let U, V be vector spaces over the same field K and let $\{u_1, \ldots, u_m\}$, $\{v_1, \ldots, v_n\}$ be bases of U, V respectively. If $\mathcal{T} \in \mathcal{L}(U, V)$ then $\mathcal{T}v_j \in U$

and so we may write

$$\mathcal{T}v_j = \sum_{i=1}^{m} t_{ij}u_i,\ t_{ij} \in K \tag{6.1}$$

The scalars $(t_{1j}, \ldots, t_{mj})$ are, for each j, the coordinates of $\mathcal{T}v_j$ relative to the given basis of U, and so are uniquely determined by $\mathcal{T}$.

Conversely, if we are given any set $\{t_{ij}|i = 1, \ldots, m; j = 1, \ldots, n\}$ of scalars, and bases $\{u_1, \ldots, u_m\}, \{v_1, \ldots, v_n\}$ of U and V, then equation (6.1) determines a unique linear transformation $\mathcal{T} \in \mathcal{L}(V, U)$ (see Th.6.1, Section 5.1).

Write T for the matrix $(t_{ij})_{i=1,\ldots,m;j=1,\ldots,n}$. Then $T \in \mathcal{M}_{m,n}(K)$ will be called *the matrix of* $\mathcal{T}$, or *the matrix representing* $\mathcal{T}$, relative to the given bases of U and V. The columns of T are formed with the coordinates of $\mathcal{T}v_1, \ldots, \mathcal{T}v_n$ relative to the basis $\{u_1, \ldots, u_m\}$.

Since the definition of the scalars t_{ij} by (5.1) depends upon the arbitrarily chosen bases of U and V, many different matrices represent the same linear transformation.

Let $(x_1, \ldots, x_n)$ be the coordinates of $x \in V$ relative to the basis $\{v_1, \ldots, v_n\}$. Let $(y_1, \ldots, y_m)$ be the coordinates of $\mathcal{T}x \in U$ relative to the basis $\{u_1, \ldots, u_m\}$. Denote

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

Theorem 6.6 *The coordinates of x and the coordinates of $\mathcal{T}x$ are connected by the equation*

$$Y = TX. \tag{6.2}$$

Proof. We have $x = \sum\limits_{j=1}^{n} x_j v_j$ and $\mathcal{T}x = \sum\limits_{i=1}^{m} y_i u_i$. On the other hand, $\mathcal{T}x = \mathcal{T}(\sum\limits_{i=1}^{n} x_j v_j) = \sum\limits_{j=1}^{n} x_j \mathcal{T}v_j = \sum\limits_{j=1}^{n} x_j \sum\limits_{i=1}^{m} t_{ij}u_i = \sum\limits_{i=1}^{m}(\sum\limits_{i=1}^{n} t_{ij}x_j)u_i$.

Since the representation of the vector $\mathcal{T}x$ as a linear combination of the elements of the basis $\{u_1, \ldots, u_m\}$ is unique, we may equate the coefficients of u_i, $i = 1, \dots, m$ and so obtain

$$y_i = \sum_{j=1}^{n} t_{ij} x_j \quad , i = 1, \dots, m.$$

This system is equivalent to (6.2). □

When $U = V$, to obtain a matrix representation of $\mathcal{T} \in \mathcal{L}(V, V)$ it is only necessary to choose one basis $\{v_1, \dots, v_n\}$ of V. In this case the theorem must be modified by writing v_i for u_i throughout the statement and proof.

We now interpret, in the language of matrices, the operations on linear transformations defined in Section 5.1.

Theorem 6.7 *Let U, V, W be three vector spaces over the same field K, of dimensions m, n, p respectively, and let $\{u_1, \dots, u_m\}$, $\{v_1, \dots, v_n\}$, $\{w_1, \dots, w_p\}$ be bases of U, V, W. Then, relative to these bases:*

1) *The zero linear transformation $0 \in \mathcal{L}(V, U)$ is represented by the zero matrix $0 \in \mathcal{M}_{m,n}(K)$.*

2) *The identity transformation $I \in \mathcal{L}(V, V)$ is represented by the unit matrix $I \in \mathcal{M}_{n,n}(K)$.*

3) *If $\mathcal{T} \in \mathcal{L}(V, U)$ is represented by the matrix $T \in \mathcal{M}_{m,n}(K)$, then for all $k \in K$ the transformation $k\mathcal{T}$ is represented by the matrix kT.*

4) *If $\mathcal{T}, \mathcal{S} \in \mathcal{L}(V, U)$ are represented by the matrices $T, S \in \mathcal{M}_{m,n}(K)$ respectively, then $\mathcal{T} + \mathcal{S}$ is represented by the matrix $T + S$.*

5) *If $\mathcal{T} \in \mathcal{L}(V, U)$ and $\mathcal{S} \in \mathcal{L}(U, W)$ are represented by $T \in \mathcal{M}_{m,n}(K)$ and $S \in \mathcal{M}_{p,m}$ respectively, then $\mathcal{S}\mathcal{T}$ is represented by ST.*

6) *If $\mathcal{T} \in \mathcal{L}(V, V)$ is non-singular and is represented by the matrix $T \in \mathcal{M}_{m,n}(K)$, then the inverse transformation $\mathcal{T}^{-1}$ is represented by the inverse matrix T^{-1}.*

Proof. All the statements follow immediately from the definitions, and we omit the details. We need also the following result:

Let $\mathcal{T} \in \mathcal{L}(V)$ be represented by the matrix T relative to the basis $B = \{b_1, \ldots, b_n\}$ of V, and by a matrix T' relative to the basis $B' = \{b'_1, \ldots, b'_n\}$ of V. Let C be the transition matrix from B to B'. Then $T' = C^{-1}TC$. □

6.3 Invariant subspaces. Eigenvalues and eigenvectors

We now begin a more detailed study of linear transformations. Throughout the remainder of this chapter we shall be concerned only with linear transformations of a vector space V into itself, that is, with endomorphisms of V.

Definition 6.8 Let $\mathcal{T} \in \mathcal{L}(V)$ and W be a subspace of V with the property that $\mathcal{T}(W) \subset W$. Then $\mathcal{T}$ is called an *invariant subspace* of V under the endomorphism $\mathcal{T}$, or - more briefly - W is said to be $\mathcal{T}$-invariant.

Example 6.3.1 1) The improper subspaces V and $\{0\}$ are invariant under every endomorphism of V. Every subspace of V is invariant under both the identity and zero transformations.

2) $K_n[X]$ is an invariant subspace of $K[X]$ under the endomorphism $\mathcal{D}$ described in Example 5.1.1.

3) $\mathcal{T}\vec{i} = \vec{j}, \mathcal{T}\vec{j} = -\vec{i}$ define an endomorphism of the space $V = \{a\vec{i} + b\vec{j} \mid a, b \in \mathbb{R}\}$. It can be shown that V has no proper invariant subspaces under $\mathcal{T}$. (Exercise!)

Definition 6.9 Let $\mathcal{T} \in \mathcal{L}(V)$. A scalar $\lambda \in K$ is called an *eigenvalue* (or proper value) of $\mathcal{T}$ if there exists a non-zero vector $x \in V$ such that $\mathcal{T}x = \lambda x$. The vector x is called an *eigenvector* (or proper vector) of $\mathcal{T}$.

Let λ be an eigenvalue of $\mathcal{T}$. Denote $E(\lambda) = \{x \in V \mid \mathcal{T}x = \lambda x\}$. Clearly $E(\lambda)$ consists of all the eigenvectors of $\mathcal{T}$ corresponding to λ, together with the vector zero.

It is easy to verify that $E(\lambda)$ is a linear subspace of V and, moreover, it is $\mathcal{T}$-invariant. (Exercise!) It will be called the proper subspace of $\mathcal{T}$ corresponding to the eigenvalue λ.

Let now V_n be an n-dimensional linear space over K and let $B = \{b_1, \ldots, b_n\}$ be a basis of V_n. Let $\lambda \in K$ be an eigenvalue and let $x = x_1 b_1 + \cdots + x_n b_n$ be an eigenvector of $\mathcal{T}$ corresponding to λ. Hence we have $\mathcal{T}x = \lambda x$ and $x \neq 0$. Denote

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Then $(\mathcal{T} - \lambda I)(x) = 0$, which is equivalent to $(T - \lambda I)X = 0$, where T is the matrix of $\mathcal{T}$ relative to the basis B (see Section 5.2).

The equation $(T - \lambda I)X = 0$ may be written in the form

$$\begin{pmatrix} t_{11} - \lambda & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} - \lambda & \ldots & t_{2n} \\ \ldots & & & \\ t_{n1} & t_{n2} & \ldots & t_{nn} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{6.3}$$

This is a linear homogeneous system. Since $x \neq 0$, it has non-trivial solutions, that is $\det(T - \lambda I) = 0$. Let us denote $P(\lambda) = \det(T - \lambda I)$; remark that $P(\lambda)$ is a polynomial of degree n.

Theorem 6.10 *$P(\lambda)$ does not depend on the choice of the basis B.*

Proof. Let B' be another basis of V_n and let T' be the matrix of $\mathcal{T}$ relative to B'. Let C be the transition matrix from B to B'. Then $T' = C^{-1}TC$; see Section 5.2. We have to prove that $\det(T' - \lambda I) = \det(T - \lambda I)$ for all $\lambda \in K$. Indeed,

$$\begin{aligned} \det(T' - \lambda I) &= \det(C^{-1}TC - C^{-1}(\lambda I)C) = \det(C^{-1}(T - \lambda I)C) = \\ &= \det C^{-1} \cdot \det(T - \lambda I) \cdot \det C = \\ &= (\det C)^{-1} \det(T - \lambda I) \det C = \det(T - \lambda I), \end{aligned}$$

so the theorem is proved. $\square$

Since the polynomial $P(\lambda)$ is independent of the choice of the basis B, it will be called the *characteristic polynomial* of $\mathcal{T}$. If a matrix T represents $\mathcal{T}$ with respect to some basis, $P(\lambda)$ will be also called the caracteristic polynomial of T, and we have simply $P(\lambda) = \det(T - \lambda I)$.

Returning to the eigenvalues of $\mathcal{T}$, we see that they are exactly the roots in K of the characteristic polynomial of $\mathcal{T}$. There exist n roots, real or complex. If $K = \mathbb{C}$, all of them are eigenvalues; if $K = \mathbb{R}$, only the real roots (if there exist real roots!) are eigenvalues of $\mathcal{T}$.

Now suppose that λ is an eigenvalue of $\mathcal{T}$. Then (6.3) has non-trivial solutions. Every such non-trivial solution gives us an eigenvector x by means of the formula $x = x_1 b_1 + \cdots + x_n b_n$.

6.4 The Cayley-Hamilton Theorem

Let $P \in K[X]$ be an arbitrary polynomial, $P(X) = a_m X^m + \cdots + a_1 X + a_0$, $a_i \in K$. For a matrix $A \in \mathcal{M}_{n,n}(K)$ let us denote $P(A) = a_m A^m + \cdots + a_1 A + a_0 I$. The Cayley-Hamilton Theorem asserts that *if $P(\lambda) = \det(T - \lambda I)$ is the characteristic polynomial of a matrix $T \in \mathcal{M}_{n,n}(K)$, then $P(T) = 0$.*

We shall use this result in order to prove

Theorem 6.11 *Let $A \in \mathcal{M}_{n,n}(K)$. Then for each $p \geq n$, A^p can be expressed as a linear combination of $I, A, A^2, \ldots, A^{n-1}$.*

Proof. Let $P(\lambda) = \det(A - \lambda I)$ be the characteristic polynomial of the matrix A. By virtue of the Cayley-Hamilton Theorem, we have $P(A) = 0$.

Clearly $P(\lambda) = (-1)^n \lambda^n + k_{n-1} \lambda^{n-1} + \cdots + k_1 \lambda + k_0$, with $k_i \in K$. Hence

$$(-1)^n A^n + k_{n-1} A^{n-1} + \cdots + k_1 A + k_0 I = 0.$$

It follows that

$$A^n = c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I, \quad c_i \in K. \tag{6.4}$$

Thus A^n is a linear combination of $I, A, A^2, \ldots, A^{n-1}$.

From (6.4)we deduce

$$A^{n+1} = c_{n-1}A^n + c_{n-2}A^{n-1} + \cdots + c_1 A^2 + c_0 A \tag{6.5}$$

If we substitute A^n taking into account (6.4), we obtain A^{n+1} as a linear combination of $I, A, \ldots, A^{n-1}$. By repeating this argument we finish the proof. □

6.5 The diagonal form

Let V be a linear space over K.

Theorem 6.12 *Let $\mathcal{T} \in \mathcal{L}(V)$ and let $x_1, \ldots, x_n$ be eigenvectors of $\mathcal{T}$ associated with mutually distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Then the vectors $x_1, \ldots, x_n$ are linearly independent.*

Proof. Suppose that

(1) $\{x_1, \ldots, x_n\}$ is a linearly dependent set
Then there exist $k_1, \ldots, k_n \in K$, not all zero, such that

(2) $k_1 x_1 + \cdots + k_n x_n = 0$.
Renumbering the variables if necessary, we may suppose that

(3) $k_1 \neq 0$.
From (2) we obtain $k_1 \mathcal{T} x_1 + \cdots + k_n \mathcal{T} x_n = 0$. Since $\mathcal{T} x_i = \lambda_i x_i$, it follows that

(4) $k_1 \lambda_1 x_1 + \cdots + k_n \lambda_n x_n = 0$
Now (2) and (4) imply

(5) $k_2(\lambda_2 - \lambda_1)x_2 + \cdots + k_n(\lambda_n - \lambda_1)x_n = 0.$

We claim that $\{x_2, \ldots, x_n\}$ must be linearly dependent. Indeed, if we suppose that they are linearly independent, then $k_2 = \cdots = k_n = 0$ since $\lambda_i - \lambda_1 \neq 0,\ i = 2, \ldots, n$. But (2) implies $k_1 x_1 = 0$. Since x_1 is an eigenvector, it is non-zero. Hence $k_1 = 0$, which contradicts (3).

Thus (1) implies:

(6) $\{x_2, \ldots, x_n\}$ is a linearly dependent set.

Now we repeat the same arguments and conclude that (6) implies:

(7) $\{x_3, \ldots, x_n\}$ is a linearly dependent set.

In this manner we deduce finally that $\{x_n\}$ is a linearly dependent set. On the other hand,the same set is linearly independent, since $x_n \neq 0$ as an eigenvector. This contradiction shows that (1) is false and the theorem is proved. $\square$

Theorem 6.13 *Let $\mathcal{T}$ be an endomorphism of a linear space V_n of finite dimension $n \geq 1$ over K. Suppose that the characteristic polynomial $P(\lambda)$ of $\mathcal{T}$ has n simple roots $\lambda_1, \ldots, \lambda_n$ in the field K. Then there exists a basis of V_n relative to which the matrix of $\mathcal{T}$ is*

$$\begin{pmatrix} \lambda_1 & 0 & 0 \ldots & 0 \\ 0 & \lambda_2 & 0 \ldots & 0 \\ \ldots & & & \\ 0 & 0 & 0 \ldots & \lambda_n \end{pmatrix}$$

Proof. Since the roots $\lambda_1, \ldots, \lambda_n$ are in K, they are eigenvalues of $\mathcal{T}$. For each i choose an eigenvector x_i of $\mathcal{T}$ corresponding to the eigenvalue λ_i. By hypothesis $\lambda_1, \ldots, \lambda_n$ are mutually distinct. Theorem 6.12 shows that $x_1, \ldots, x_n$ are linearly independent. Since $dimV_n = n$, $\{x_1, \ldots, x_n\}$ is a basis. We have $\mathcal{T}x_i = \lambda_i x_i,\ i = 1, \ldots, n$, hence the matrix of $\mathcal{T}$ with respect to this basis is the diagonal matrix of the theorem. $\square$

Corollary 6.14 *Let $T \in \mathcal{M}_{n,n}(K)$. Suppose that the characteristic polynomial of T has n simple roots in K. Then there exists a matrix $C \in \mathcal{M}_{n,n}(K)$ such that*

$$C^{-1}TC = \begin{pmatrix} \lambda_1 & 0 & 0 \dots & 0 \\ 0 & \lambda_2 & 0 \dots & 0 \\ \dots & & & \\ 0 & 0 & 0 \dots & \lambda_n \end{pmatrix},$$

$\lambda_1, \dots, \lambda_n$ being the roots.

Proof. Let B be the canonical basis of K^n. Let $\mathcal{T} \in \mathcal{L}(K^n)$ be the endomorphism which has the matrix T relative to the basis B. Theorem 6.13 shows that there exists a basis B' of K^n relative to which the matrix of $\mathcal{T}$ is

$$T' = \begin{pmatrix} \lambda_1 & 0 & 0 \dots & 0 \\ 0 & \lambda_2 & 0 \dots & 0 \\ \dots & & & \\ 0 & 0 & 0 \dots & \lambda_n \end{pmatrix}$$

Let C be the transition matrix from B to B'. We know that $T' = C^{-1}TC$ and the proof is complete. $\square$

We shall denote

$$diag(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & 0 \dots & 0 \\ 0 & \lambda_2 & 0 \dots & 0 \\ \dots & & & \\ 0 & 0 & 0 \dots & \lambda_n \end{pmatrix}$$

The algebra of matrices applies especially smoothly to diagonal matrices: to add or multiply any two diagonal matrices, one simply adds or multiplies corresponding diagonal entries.

For instance, let T be as in the above corollary. Then it is easy to compute T^p for any $p \geq 1$. Indeed, let $T' = diag(\lambda_1, \dots, \lambda_n)$. Then $C^{-1}TC = T'$, that is $T = CT'C^{-1}$.

We have

$$T^p = (CT'C^{-1}) \cdot (CT'C^{-1}) \cdot \cdots \cdot (CT'C^{-1}) = C(T')^p C^{-1}$$

But $(T')^p = diag(\lambda_1^p, \dots, \lambda_n^p)$ and hence

$$T^p = C \cdot diag(\lambda_1^p, \dots, \lambda_n^p) \cdot C^{-1}.$$

6.6 Reduction to diagonal form

We want to characterize the endomorphisms that can be "diagonalized", that is, for which there exists a basis relative to which the matrix is a diagonal one.

Let V_n be a linear space of finite dimension $n \geq 1$ over the field K. Let $\mathcal{T} \in \mathcal{L}(V_n)$ and let λ_0 be an eigenvalue of $\mathcal{T}$. We know that λ_0 is a root in K of the characteristic polynomial of $\mathcal{T}$. Denote by $m(\lambda_0)$ the multiplicity of λ_0 as a root of this polynomial.

Consider also the proper subspace corresponding to λ_0:

$$E(\lambda_0) = \{x \in V_n \mid \mathcal{T}x = \lambda_0 x\}.$$

Let $B = \{b_1, \dots, b_n\}$ be an arbitrary basis of V_n and let T be the matrix of $\mathcal{T}$ relative to this basis.

Theorem 6.15 $dim E(\lambda_0) = n - rank(T - \lambda_0 I) \leq m(\lambda_0)$

Proof. Let $x \in V_n,\ x = x_1 b_1 + \cdots + x_n b_n$. As usual, denote

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Then the following statements are equivalent:

(1) $x \in E(\lambda_0)$

(2) $(\mathcal{T} - \lambda_0 I)(x) = 0$

(3) $(T - \lambda_0 I) \cdot X = 0$

We conclude that $E(\lambda_0)$ can be identified with the set of the solutions of the linear homogeneous system

$$\begin{pmatrix} t_{11} - \lambda_0 & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} - \lambda_0 & \dots & t_{2n} \\ \dots & & & \\ t_{n1} & t_{n2} & \dots & t_{nn} - \lambda_0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

But this set is a linear subspace of K^n of dimension $n - rank(T - \lambda_0 I)$. Thus $dim E(\lambda_0) = n - rank(T - \lambda_0 I)$, and the first statement of the theorem is proved.

Now denote $q = dim E(\lambda_0)$ and let $\{v_1, \dots, v_q)$ be a basis of $E(\lambda_0)$. Let us complete it in order to obtain a basis $\{v_1, \dots, v_q, v_{q+1}, \dots, v_n\}$ of V_n.

We have $\mathcal{T} v_j = \lambda_0 v_j,\ j = 1, \dots, q$ and $\mathcal{T} v_j = t_{1j} v_1 + \dots + t_{nj} v_n,\ j = q+1, \dots, n,\ t_{ij} \in K$.

Hence the matrix T' of $\mathcal{T}$ relative to the basis $\{v_1, \dots, v_n\}$ is

$$T' = \begin{pmatrix} \lambda_0 & 0 & \dots & 0 & t_{1,q+1} & \dots & t_{1n} \\ 0 & \lambda_0 & \dots & 0 & t_{2,q+1} & \dots & t_{2n} \\ \dots & \dots & & & & & \\ 0 & 0 & \dots & \lambda_0 & t_{q,q+1} & \dots & t_{qn} \\ 0 & 0 & \dots & 0 & t_{q+1,q+1} & \dots & t_{q+1,n} \\ \dots & \dots & & & & & \\ 0 & 0 & \dots & 0 & t_{n,q+1} & \dots & t_{n,n} \end{pmatrix}$$

The characteristic polynomial of $\mathcal{T}$ is $P(\lambda) = \det(T' - \lambda I)$. If we take account of the form of T' we conclude that $P(\lambda)$ is of the form $P(\lambda) = (\lambda_0 - \lambda)^q \cdot Q(\lambda)$, where $Q(\lambda)$ is a polynomial. Now it is clear that the multiplicity of λ_0 as a root of $P(\lambda)$ is at least q, that is $m(\lambda_0) \geq q$.

Thus $n - \mathrm{rank}(T - \lambda_0 I) \leq m(\lambda_0)$ and the theorem is proved. $\square$

Definition 6.16 Let $\mathcal{T}$ be an endomorphism of a vector space V_n of finite dimension n over K. The endomorphism $\mathcal{T}$ is said to be *diagonalizable* if there exists a basis of V_n consisting of eigenvectors of $\mathcal{T}$, in other words a basis relative to which the matrix of $\mathcal{T}$ is diagonal.

Theorem 6.13, Section 5.5 gives a sufficient condition for this to be the case: namely that the roots of the characteristic polynomial of $\mathcal{T}$ all lie in K and are all distinct. But it is easily seen that this condition is not necessary: a trivial example is the identity endomorphism whose matrix with respect to any basis of V_n is diagonal, but whose characteristic polynomial, namely $(1-\lambda)^n$ has no simple roots (assuming that $n > 1$).

6.7 The Jordan canonical form

Let $\mathcal{T} \in \mathcal{L}(V_n)$; suppose that all the roots of the characteristic polynomial are in $\mathbb{K}$. Let λ be such a root, i.e., an eigenvalue of $\mathcal{T}$. Let m be the algebraic multiplicity of λ, and $q = dim E(\lambda)$. Then $m \geq q \geq 1$.

It is possible to find q eigenvectors in $E(\lambda)$ and $m-q$ principal vectors, all of them linearly independent; an eigenvector v and the principal vectors $u_1, \ldots, u_r (r \geq 0)$ corresponding to it satisfy:

$$\mathcal{T}v = \lambda v;\ \mathcal{T}u_1 = \lambda u_1 + v;\ \mathcal{T}u_2 = \lambda u_2 + u_1; \ldots ; \mathcal{T}u_r = \lambda u_r + u_{r-1}.$$

All these eigenvectors and principal vectors, associated to all the eigenvalues of $\mathcal{T}$, form a basis of V_n, called a *Jordan basis* with respect to $\mathcal{T}$. The matrix of $\mathcal{T}$ relative to a Jordan basis is called a *Jordan matrix* of $\mathcal{T}$. Such a matrix has the form $\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \cdots & \\ & & & J_p \end{pmatrix}$, where $J_1, \ldots, J_p$ are called *Jordan cells*. Each cell represents the contribution of an eigenvector v and the corresponding principal vectors

$$u_1, \ldots, u_r : \begin{pmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & & \\ & & \lambda & 1 & & \\ & & & \vdots & 1 & \\ & & & & \vdots & \\ & & & & & \lambda \end{pmatrix} \in M_{r+1}(\mathbb{K}).$$

We see that: the Jordan matrix is a diagonal matrix $\iff$ there are no principal vectors $\iff$ $m(\lambda) = dimE(\lambda)$ for each eigenvalue λ.

Let T be the matrix of $\mathcal{T}$ with respect to a given basis B, and J the Jordan matrix with respect to a Jordan basis B'. Let C be the transition matrix from B to B'. Then $J = C^{-1}TC$, hence $T = CJC^{-1}$. It follows that $T^n = CJ^nC^{-1}$.

The exponential of the matrix T is defined by

$$e^T = I + \frac{1}{1!}T + \frac{1}{2!}T^2 + \cdots + \frac{1}{n!}T^n + \ldots.$$

Example 6.7.1 1. Let $\mathcal{T} \in \mathcal{L}(\mathbb{R}^3)$ have the matrix

$$T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{pmatrix}$$

with respect to the canonical basis.
We find $\lambda_1 = 2,\ m(\lambda_1) = 1$,

$$E(\lambda_1) = \left\{\alpha \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} \;\middle|\; \alpha \in \mathbb{R}\right\}, \text{ hence } q(\lambda_1) = 1.$$

$\lambda_2 = 1,\ m(\lambda_2) = 2$,

$$E(\lambda_2) = \left\{\alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; \alpha \in \mathbb{R}\right\}, \text{ hence } q(\lambda_2) = 1.$$

So $v_1 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$; the principal vector u_1 associated with v_2 satisfies $Tu_1 = u_1 + v_2$. Let $u_1 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$; then

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & -5 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

We get $x = y - 1,\ z = y + 1,\ y \in \mathbb{R}$. Choosing $y = 1$, we obtain

$$u_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

The Jordan basis is $B = \{v_1, v_2, u_1\}$.
Since

$$\begin{aligned} \mathcal{T}v_1 &= 2v_1 + 0v_2 + 0u_1 \\ \mathcal{T}v_2 &= 0v_1 + v_2 + 0u_1 \\ \mathcal{T}u_1 &= 0v_1 + v_2 + u_1, \end{aligned}$$

the Jordan matrix will be

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

The transition matrix from the canonical basis to the Jordan basis is

$$C = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 4 & 1 & 2 \end{pmatrix}.$$

We have $J = C^{-1}TC,\ T = CJC^{-1}, T^n = CJ^nC^{-1},\ J^n = \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$.

2. $T = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ -2 & -3 & -1 \end{pmatrix}$

In this case $\lambda_1 = -1,\ m(\lambda_1) = q(\lambda_1) = 1,\ \lambda_2 = 0,\ m(\lambda_2) = 2,\ q(\lambda_2) = 1$.

$v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ v_2 = \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix},\ u_1 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$.

The Jordan matrix is

$$J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

3. $T = \begin{pmatrix} 1 & 1 & 0 \\ -4 & -2 & 1 \\ 4 & 1 & -2 \end{pmatrix}$.

We find $\lambda_1 = -1,\ m(\lambda_1) = 3,\ q(\lambda_1) = 1$.

$v_1 = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$. The principal vectors u_1 and u_2 associated with v_1 satisfy

$$Tu_1 = -u_1 + v_1$$
$$Tu_2 = -u_2 + u_1.$$

We obtain $u_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},\ u_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

The Jordan basis is $\{v_1, u_1, u_2\}$ and the Jordan matrix is

$$J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Exercices

6.1 Let $\mathcal{T} : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by:
$\mathcal{T}(0,1,2) = (1,0)$, $\mathcal{T}(-1,1,1) = (-1,1)$, $\mathcal{T}(3,0,-1) = (2,1)$. Determine:

a) the matrix of $\mathcal{T}$ relative to the canonical basis in $\mathbb{R}^3$ and $\mathbb{R}^2$

b) bases in the subspaces Ker$\mathcal{T}$ and Im$\mathcal{T}$

6.2 Let $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation given by:
$\mathcal{T}(-1,2) = (-7,6,3)$, $\mathcal{T}(1,3) = (2,9,7)$. Determine:

a) the image of an arbitrary vector of $\mathbb{R}^2$ through $\mathcal{T}$

b) Ker$\mathcal{T}$ and Im$\mathcal{T}$

6.3 Let $\mathcal{T} \in \mathcal{L}(\mathbb{R}^3)$ be defined by $\mathcal{T}x = (x_1+x_2+x_3, x_1+x_2+x_3, x_1+x_2+x_3)$. Determine bases in Ker$\mathcal{T}$ and Im$\mathcal{T}$.

6.4 Consider the basis $B' = \{(1,1,0),(1,0,1),(0,1,1)\}$ in $\mathbb{R}^3$ and the linear transformation $\mathcal{T} : \mathbb{R}^3 \to \mathbb{R}^3$, $\mathcal{T}(x_1,x_2,x_3) = (x_1+x_2-x_3, x_3, 2x_2+3x_3)$. Determine the matrix of $\mathcal{T}$ with respect to the basis B'.

6.5 Determine the eigenvalues and the eigenvectors for the matrix of order n:

$$\begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 0 \end{pmatrix}$$

6.6 Determine the eigenvalues and the eigenvectors for the matrix of order n:

$$\begin{pmatrix} 0 & 0 & \dots & 0 & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

6.7 Denoting by P_2 the space of polynomial functions of degree at most two, let $\mathcal{T} : P_2 \to P_2$ be the linear transformation given by $\mathcal{T}(1+X) = 1 - X^2$, $\mathcal{T}(1+X^2) = -4X$ and $\mathcal{T}(2X^2) = 4X^2$. Find the eigenvalues and the eigenvectors of $\mathcal{T}$.

6.8 Let $V = C(0,1)$, let $T : V \to V$ be an endomorphism defined by $T(f)(x) = xf(x)$. Determine the eigenvalues and eigenvectors of T.

6.9 Let $V = C^\infty(a,b)$, where $0 \notin (a,b)$, let $T : V \to V$ be an endomorphism defined by $T(f)(x) = \dfrac{1}{x} f'(x)$. Determine the eigenvalues and eigenvectors of T.

6.10 Find the Jordan form and the corresponding Jordan basis for:

a) $A = \begin{pmatrix} 6 & 6 & -15 \\ 1 & 5 & -5 \\ 1 & 2 & -2 \end{pmatrix}$, b) $B = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 3 & 1 \end{pmatrix}$,

c) $C = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

6.11 Find the Jordan form and the transfer matrix for:

a) $A = \begin{pmatrix} 4 & 1 & 1 \\ -2 & 1 & -2 \\ 1 & 1 & 4 \end{pmatrix}$, b) $B = \begin{pmatrix} -2 & -1 & 1 \\ 5 & -1 & 4 \\ 5 & 1 & 2 \end{pmatrix}$.

6.12 Determine A^n, $n \in \mathbf{N}$ for:

a) $A = \begin{pmatrix} -1 & 6 & 2 \\ -2 & 6 & 1 \\ 2 & -4 & 1 \end{pmatrix}$, b) $A = \begin{pmatrix} 0 & 2 & -3 \\ 4 & 7 & -12 \\ 3 & 6 & -10 \end{pmatrix}$,

c) $A = \begin{pmatrix} -2 & -1 & 1 \\ 5 & -1 & 4 \\ 5 & 1 & 2 \end{pmatrix}$, d) $A = \begin{pmatrix} -61 & 36 \\ -105 & 62 \end{pmatrix}$.

6.13 Determine e^A, for:

a) $A = \begin{pmatrix} -2 & -4 \\ 3 & 5 \end{pmatrix}$,

b) $A = \begin{pmatrix} 4 & -2 \\ 6 & -3 \end{pmatrix}$.

6.14 For the matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ determine e^A and $\sin A$.

6.15 A matrix $A \in M_n(\mathbf{C})$ is called *self-adjoint* if $\overline{A}^t = A$. Prove that if A is self-adjoint then all the roots of its characteristic polynomial are real, and the eigenvectors corresponding to distinct values are orthogonal.

6.16 A matrix $T \in M_n(\mathbf{C})$ is called *unitary* if $(\overline{T}^t)T = I$. Prove that if T is unitary then

a) For each eigenvalue λ of T we have $|\lambda| = 1$.

b) The eigenvectors corresponding to distinct values are orthogonal.

Solutions

$\boxed{6.1}$ a) Denoting by e_1, e_2 and e_3 the vectors of the canonical basis in $\mathbb{R}^3$, we get the system $\begin{cases} \mathcal{T}e_2 + 2\mathcal{T}e_3 = & (1,0) \\ -\mathcal{T}e_1 + \mathcal{T}e_2 + \mathcal{T}e_3 = & (-1,1) \\ 3\mathcal{T}e_1 - \mathcal{T}e_3 = & (2,1) \end{cases}$ with the solutions $\mathcal{T}e_1 = (1,0)$, $\mathcal{T}e_2 = (-1,2)$ and $\mathcal{T}e_3 = (1,-1)$. So the matrix of $\mathcal{T}$ relative to the canonical basis is $T = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \end{pmatrix}$.

b) For $(x_1, x_2, x_3) \in \mathbb{R}^3$ we have

$$\mathcal{T}(x_1, x_2, x_3) = x_1(1,0) + x_2(-1,2) + x_3(1,-1) = (x_1 - x_2 + x_3, 2x_2 - x_3).$$

The kernel of $\mathcal{T}$ is $\mathrm{Ker}\mathcal{T} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid (x_1 - x_2 + x_3, 2x_2 - x_3) = (0,0)\} = \{(-\alpha, \alpha, 2\alpha) \mid \alpha \in \mathbb{R}\}$, and $\{(-1,1,2)\}$ is a basis of $\mathrm{Ker}\mathcal{T}$. The image is $\mathrm{Im}\mathcal{T} = \{(x_1 - x_2 + x_3, 2x_2 - x_3) \mid x_1, x_2, x_3 \in \mathbb{R}\} = \mathrm{sp}\{(1,0), (-1,2), (1,-1)\} = \mathrm{sp}\{(1,0), (-1,2)\} = \mathbb{R}^2$.

$\boxed{\mathbf{6.2}}$ We have $\mathcal{T}e_1 = (5,0,1)$, $\mathcal{T}e_2 = (-1,3,2)$ and so $\mathcal{T}(x_1,x_2) = (5x_1 - x_2, 3x_2, x_1 + 2x_2)$. Ker$\mathcal{T} = \{(0,0)\}$,. The image is Im$\mathcal{T} = \{(5x_1 - x_2, 3x_2, x_1 + 2x_2) \mid x_1, x_2 \in \mathbb{R}\}$, with a basis $\{(5,0,1), (-1,3,2)\}$. Denoting $5x_1 - x_2 = x$, $3x_2 = y$, $x_1 + 2x_2 = z$ and eliminating the variables x_1, x_2 we get Im$\mathcal{T} = \{(x,y,z) \in \mathbb{R}^3 \mid 3x + 11y - 15z = 0\}$.

$\boxed{\mathbf{6.3}}$ A basis in Ker$\mathcal{T}$ is $\{(1,0,-1), (0,1,-1)\}$ and in Im$\mathcal{T}$ is $\{(1,1,1)\}$.

$\boxed{\mathbf{6.4}}$ The matrix of $\mathcal{T}$ relative to the canonical basis is $T = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 2 & 3 \end{pmatrix}$.

The matrix relative to the new basis B' is $T' = C^{-1}TC$, where C is the transition matrix from the canonical basis to B'. We have

$$C = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}; \quad C^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \text{ so } T' = \begin{pmatrix} 0 & -1 & -2 \\ 2 & 1 & 2 \\ 0 & 2 & 3 \end{pmatrix}$$

Another method to find the same matrix is to write the images of the vectors from the basis B' through $\mathcal{T}$:

$$\begin{aligned} \mathcal{T}(1,1,0) &= (2,0,2) = 2(1,0,1) \\ \mathcal{T}(1,0,1) &= (0,1,3) = -(1,1,0) + (1,0,1) + 2(0,1,1) \\ \mathcal{T}(0,1,1) &= -2(1,1,0) + 2(1,0,1) + 3(0,1,1). \end{aligned}$$

$\boxed{\mathbf{6.5}}$ The characteristic polynomial is $P(\lambda) = (-\lambda - 1)^{n-1}(-\lambda + n - 1)$, so the eigenvalues are $\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} = -1$ and $\lambda_n = n - 1$. For $\lambda = -1$, the eigenvectors are the solutions of the '"system"

$$x_1 + x_2 + \cdots + x_3 = 0,$$

that is, $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{pmatrix}$. For $\lambda = n-1$, we get the system

$$\begin{cases} (1-n)x_1 + x_2 + \cdots + x_n = 0 \\ x_1 + (1-n)x_2 + \cdots + x_n = 0 \\ \cdots\cdots\cdots \\ x_1 + x_2 + \cdots + (1-n)x_n = 0 \end{cases}$$

Expressing $x_n = (n-1)x_1 - x_2 - \cdots - x_{n-1}$ from the first equation and plugging it into the other equations, we get $x_2 = x_1$, $x_3 = x_1$, ... $x_n = x_1$ and the only linearly independent eigenvector is $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$

$\boxed{\textbf{6.6}}$ It is easier to calculate the characteristic polynomial by transposing the determinant, which is a circular determinant:

$$P(\lambda) = C(-\lambda, 0, \ldots, 0, 0, 1) = C(-\lambda, 1, \ldots, 0, 0, 0) = (-1)^n(\lambda^n - 1).$$

The eigenvalues are $\lambda_k = \varepsilon_k$, $k = 0, \ldots, n-1$ (the n-th roots of 1). For each eigenvalue ε_k, we determine the eigenvectors as the solutions of the system:

$$\begin{cases} -\varepsilon_k x_1 + x_n = 0 \\ x_1 - \varepsilon_k x_2 = 0 \\ x_2 - \varepsilon_k x_3 = 0 \\ \cdots\cdots\cdots \\ x_{n-1} - \varepsilon_k x_n = 0 \end{cases}$$

that is $v_k = \begin{pmatrix} \varepsilon_k^{n-1} \\ \varepsilon_k^{n-2} \\ \vdots \\ \varepsilon_k \\ 1 \end{pmatrix}$

$\boxed{\mathbf{6.7}}$ From the given data we get $\mathcal{T}(1) = -4X - 2X^2$, $\mathcal{T}(X) = 1+4X+X^2$, $\mathcal{T}(X^2) = 2X^2$, so the matrix of the transformation, in the canonical basis $1, X, X^2$ is $T = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$. It has a triple eigenvalue $\lambda = 2$, and the subspace of eigenvectors has dimension 2. Two linearly independent eigenvectors are $v_1 = 1 + 2X$ and $v_2 = X^2$.

$\boxed{\mathbf{6.8}}$ If $\lambda \in \mathbb{R}$ is an eigenvalue, then $T(f)(x) = \lambda f(x)$, for any $x \in (0,1)$. We get $(x - \lambda)f(x) = 0$, for any $x \in (0,1)$. We study two situations. If $\lambda \notin (0,1)$, then $x - \lambda \neq 0$, so $f(x) = 0$, for every $x \in (0,1)$, which is not convenient for an eigenvector. If $\lambda \in (0,1)$, then $f(x) = 0$, for $x \neq \lambda$, and $f(x) = \alpha$, $\alpha \in \mathbb{R}$, an arbitrary value. But f has to be continuous, which yields $\alpha = 0$, and $f = 0$, not convenient for an eigenvector. In conclusion, T does not possess eigenvalues.

$\boxed{\mathbf{6.9}}$ If $\lambda \in \mathbb{R}$ is an eigenvalue, then $T(f)(x) = \lambda f(x)$, for any $x \in (a,b)$. We get $\frac{1}{x} f'(x) = \lambda f(x)$ or $\frac{f'(x)}{f(x)} = \lambda x$. Integrating, follows that $ln|f(x)| = \frac{\lambda x^2}{2} + \ln c$, that is $f(x) = ce^{\frac{\lambda x^2}{2}}$, $c \in \mathbb{R}$. So each $\lambda \in \mathbb{R}$ is an eigenvalue, with an infinity of eigenvectors, $f(x) = ce^{\frac{\lambda x^2}{2}}$, $c \in \mathbb{R}^*$.

$\boxed{\mathbf{6.10}}$ a) $J_A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$, with the basis consisting of the eigenvectors $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$ and the principal vector $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. b) $J_B = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, $v_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. c) $J_C = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $v_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

6.11 a) $J_A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$, $C = \begin{pmatrix} -1 & 1 & -1 \\ 0 & -2 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. a) $J_B = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, $C = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$.

6.12 a) The eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = 1$ and $\lambda_3 = 3$, the matrix is diagonalizable. The transition matrix from the canonical basis to the basis consisting of eigenvectors is $C = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix}$, with the inverse $C^{-1} = \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 1 \\ -1 & 2 & 0 \end{pmatrix}$. Finally we get

$$A^n = \begin{pmatrix} 2-3^n & -4+2^{n+1}+2\cdot 3^n & -2+2^{n+1} \\ 1-3^n & -2+2^n+2\cdot 3^n & -1+2^n \\ -1+3^n & 2-2\cdot 3^n & 1 \end{pmatrix}.$$

b) $J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 & 3 \\ 4 & 0 & 0 \\ 3 & 0 & 1 \end{pmatrix}$ and

$$A^n = \begin{pmatrix} (-1)^{n-1}(n-1) & 2n(-1)^{n-1} & 3n(-1)^n \\ 4n(-1)^{n-1} & (-1)^{n-1}(8n-1) & 12n(-1)^n \\ 3n(-1)^{n-1} & 6n(-1)^{n-1} & (-1)^n(9n+1) \end{pmatrix}.$$

c) $A^n = \begin{pmatrix} (-2)^n & -n(-2)^{n-1} & n(-2)^{n-1} \\ 3^n-(-2)^n & n(-2)^{n-1}+(-2)^n & 3^n-(-2)^{n-1}(n-2) \\ 3^n-(-2)^n & n(-2)^{n-1} & 3^n-n(-2)^{n-1} \end{pmatrix}$.

d) $A^n = \begin{pmatrix} 21(-1)^n - 20\cdot 2^n & 12(2^n-(-1)^n) \\ 35((-1)^n-2^n) & 21\cdot\ 2^n - 20(-1)^n \end{pmatrix}$.

6.13 a) The eigenvalues of A are $\lambda_1 = 2$ and $\lambda_2 = 1$, with two corresponding eigenvectors $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$. The diagonal form of A is $J_A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ and the transition matrix is $C = \begin{pmatrix} 1 & 4 \\ -1 & -3 \end{pmatrix}$. From $A^n = CJ_A^nC^{-1}$ we get $A^n = \begin{pmatrix} -3 \cdot 2^n + 4 & -4 \cdot 2^n + 4 \\ 3 \cdot 2^n + 3 & 4 \cdot 2^n - 3 \end{pmatrix}$ and finally $e^A = \begin{pmatrix} -3e^2 + 4e & -4e^2 + 4e \\ 3e^2 + 3e & 4e^2 - 3e \end{pmatrix}$.

b) $\lambda_1 = 0$, $\lambda_2 = 1$. Using the Cayley-Hamilton Theorem we have that $A^2 - A = 0$, so $A^n = A$, for $n \geq 1$. $e^A = \begin{pmatrix} 4e - 3 & 2 - 2e \\ 6e - 6 & 4 - 3e \end{pmatrix}$.

$\boxed{\mathbf{6.14}}$ $A^n = \frac{1}{2}\begin{pmatrix} 3^n + (-1)^n & 3^n - (-1)^n \\ 3^n - (-1)^n & 3^n + (-1)^n \end{pmatrix}$, $e^A = \frac{1}{2}\begin{pmatrix} e^3 + e^{-1} & e^3 - e^{-1} \\ e^3 - e^{-1} & e^3 + e^{-1} \end{pmatrix}$,

$\sin A = \dfrac{1}{1!}A - \dfrac{1}{3!}A^3 + \dfrac{1}{5!}A^5 - \cdots = \dfrac{1}{2}\begin{pmatrix} \sin 3 + \sin(-1) & \sin 3 - \sin(-1) \\ \sin 3 - \sin(-1) & \sin 3 + \sin(-1) \end{pmatrix}$.