

Chapter 6

Linear Maps

6.1 Linear Maps

In what follows, two arbitrary $\mathbb{K}$ -vector spaces U and V will be fixed (Cf. Chapter ??). In order to distinguish them, the zero vectors of U and V will be denoted, respectively, by $\mathbf{0}_U$ and $\mathbf{0}_V$.

6.1.1 Definition. A linear map is a function $f : U \rightarrow V$ such that

$$f(\lambda \mathbf{u}_1 + \mu \mathbf{u}_2) = \lambda f(\mathbf{u}_1) + \mu f(\mathbf{u}_2), \quad (6.1)$$

for every choice of $\mathbf{u}_1, \mathbf{u}_2 \in U$ and $\lambda, \mu \in \mathbb{K}$.

A *linear map* is therefore a function that “respects” linear combinations, that is, such that the image of a linear combination of vectors is the linear combination, with the same coefficients, of the images. In particular f is *additive*:

$$f(\mathbf{u}_1 + \mathbf{u}_2) = f(\mathbf{u}_1) + f(\mathbf{u}_2),$$

(by putting $\lambda = \mu = 1$ in (6.1)) and *homogeneous of degree 1*:

$$f(\lambda \mathbf{u}) = \lambda f(\mathbf{u}) \quad (6.2)$$

by putting $\mathbf{u}_1 = \mathbf{u}$ and $\mu = 0$ in (6.1). We shall denote by $\text{Hom}_{\mathbb{K}}(U, V)$ the set of all linear maps from U to V :

$$\text{Hom}_{\mathbb{K}}(U, V) := \{f : U \rightarrow V \mid f \text{ is linear}\}.$$

6.1.2 If $f \in \text{Hom}_{\mathbb{K}}(U, V)$, the image of the zero vector $\mathbf{0}_U$ of U is the zero vector of V . Indeed:

$$f(\mathbf{0}_U) = f(0 \cdot \mathbf{u}) = 0 \cdot f(\mathbf{u}) = \mathbf{0}_V,$$

where $\mathbf{u}$ is an arbitrary vector of U . In the second equality the linearity of f has been used (applied to the zero scalar). Similarly,

$$f(-\mathbf{u}) = f(-1 \cdot \mathbf{u}) = -f(\mathbf{u})$$

for every $\mathbf{u} \in U$, that is, the image of the opposite is the opposite of the image.

6.1.3 Esempio. Let $f : \mathbb{K}^2 \rightarrow \mathbb{K}^3$ be defined by:

$$\mathbf{u} \mapsto f(\mathbf{u}) = \begin{pmatrix} 2\mathbf{u}(1) - \mathbf{u}(2) \\ \mathbf{u}(1) + \mathbf{u}(2) \\ \mathbf{u}(1) - 3\mathbf{u}(2) \end{pmatrix}$$

For every choice of $\mathbf{u}, \mathbf{v} \in \mathbb{K}^2$ and $\lambda, \mu \in \mathbb{K}$, one has:

$$\begin{aligned} f(\lambda\mathbf{u} + \mu\mathbf{v}) &= \begin{pmatrix} 2(\lambda\mathbf{u} + \mu\mathbf{v})(1) - (\lambda\mathbf{u} + \mu\mathbf{v})(2) \\ (\lambda\mathbf{u} + \mu\mathbf{v})(1) + (\lambda\mathbf{u} + \mu\mathbf{v})(2) \\ (\lambda\mathbf{u} + \mu\mathbf{v})(1) - 3(\lambda\mathbf{u} + \mu\mathbf{v})(2) \end{pmatrix} = \\ &= \begin{pmatrix} \lambda(2\mathbf{u}(1) - \mathbf{u}(2)) + \mu(2\mathbf{v}(1) - \mathbf{v}(2)) \\ \lambda(\mathbf{u}(1) + \mathbf{u}(2)) + \mu(\mathbf{v}(1) + \mathbf{v}(2)) \\ \lambda(\mathbf{u}(1) - 3\mathbf{u}(2)) + \mu(\mathbf{v}(1) - 3\mathbf{v}(2)) \end{pmatrix} = \\ &= \lambda \begin{pmatrix} 2\mathbf{u}(1) - \mathbf{u}(2) \\ \mathbf{u}(1) + \mathbf{u}(2) \\ \mathbf{u}(1) - 3\mathbf{u}(2) \end{pmatrix} + \mu \begin{pmatrix} 2\mathbf{v}(1) - \mathbf{v}(2) \\ \mathbf{v}(1) + \mathbf{v}(2) \\ \mathbf{v}(1) - 3\mathbf{v}(2) \end{pmatrix} = \\ &= \lambda f(\mathbf{u}) + \mu f(\mathbf{v}) \end{aligned}$$

and therefore f is linear.

6.1.4 Esempio. Let V be a vector space over $\mathbb{K}$. The identity function:

$$\text{id}_V : V \rightarrow V$$

defined by $\text{id}_V(\mathbf{v}) = \mathbf{v}$, for every $\mathbf{v} \in V$, is a [linear map](#). Indeed, for every $\mathbf{v}_1, \mathbf{v}_2 \in V$ and every $\lambda, \mu \in \mathbb{K}$, one has:

$$\text{id}_V(\lambda\mathbf{v}_1 + \mu\mathbf{v}_2) = \lambda\mathbf{v}_1 + \mu\mathbf{v}_2 = \lambda \cdot \text{id}_V(\mathbf{v}_1) + \mu \cdot \text{id}_V(\mathbf{v}_2).$$

6.1.5 Esempio. A function $f : \mathbb{R} \rightarrow \mathbb{R}$, real-valued of one real variable, is linear if and only if it is of the form

$$f(x) = ax$$

for some $a \in \mathbb{R}$. Indeed, if $a \in \mathbb{R}$ and $f(x) = ax$, then

$$f(\lambda x_1 + \mu x_2) = a(\lambda x_1 + \mu x_2) = a\lambda x_1 + a\mu x_2 = \lambda f(x_1) + \mu f(x_2)$$

for every choice of $\lambda, \mu, x_1, x_2 \in \mathbb{R}$, hence the linearity of f . Conversely, assuming f linear:

$$\begin{aligned}
f(x) &= f(x \cdot 1) = \\
&= f(1) \cdot x && \text{(using linearity (6.2))} \\
&= a \cdot x && \text{(having put } a := f(1)\text{).}
\end{aligned}$$

6.1.6 Example 6.1.5 is easily generalized to linear functions $\mathbb{K}^n \rightarrow \mathbb{K}$, since if $f : \mathbb{K}^n \rightarrow \mathbb{K}$ is linear, there exist $a_1, \dots, a_n$ such that:

$$f(\mathbf{u}) = a_1 \mathbf{u}(1) + a_2 \mathbf{u}(2) + \dots + a_n \mathbf{u}(n). \quad (6.3)$$

Indeed, for every $\mathbf{u} \in \mathbb{K}^n$:

$$f(\mathbf{u}) = f\left(\sum_{i=1}^n \mathbf{u}(i) \mathbf{e}_i\right) = \sum_{i=1}^n \mathbf{u}(i) f(\mathbf{e}_i).$$

Putting $a_i = f(\mathbf{e}_i)$, one obtains the expression (6.3). One also says that $f : \mathbb{K}^n \rightarrow \mathbb{K}$ is a *linear form*. The set of linear forms is denoted by $(\mathbb{K}^n)^\vee$ and is a vector space with respect to the notion of linear combination:

$$(\lambda f + \mu g)(\mathbf{u}) = \lambda f(\mathbf{u}) + \mu g(\mathbf{u}).$$

Therefore a linear form is uniquely determined by the row

$$(a_1, \dots, a_n) = (f(\mathbf{e}_1), \dots, f(\mathbf{e}_n)),$$

and can therefore be written as:

$$f(\mathbf{u}) = M_f \cdot \mathbf{u} = (a_1, \dots, a_n) \cdot \begin{pmatrix} \mathbf{u}(1) \\ \vdots \\ \mathbf{u}(n) \end{pmatrix},$$

where $M_f := (a_1, \dots, a_n)$ is called the *matrix of the linear form* $f : \mathbb{K}^n \rightarrow \mathbb{K}$ with respect to the canonical basis E_n , and where the *product of a row by a column* is defined by:

$$(a_1, a_2, \dots, a_n) \cdot \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = a_1 u_1 + a_2 u_2 + \dots + a_n u_n.$$

There is therefore a natural identification between the vector space $(\mathbb{K}^n)^\vee$ of linear forms and that of rows of scalars $(a_1, \dots, a_n)$, with respect to the notion of linear combination defined by:

$$\lambda(a_1, \dots, a_n) + \mu(b_1, \dots, b_n) = (\lambda a_1 + \mu b_1, \dots, \lambda a_n + \mu b_n).$$

6.1.7 Linear maps between column spaces. Let

$$f : \mathbb{K}^n \rightarrow \mathbb{K}^m$$

be a linear map. Then, for every $\mathbf{u} \in \mathbb{K}^n$ one has:

$$f(\mathbf{u}) = f(\mathbf{u}(1)\mathbf{e}_1 + \mathbf{u}(2)\mathbf{e}_2 + \dots + \mathbf{u}(n)\mathbf{e}_n)$$

that is, by linearity:

$$f(\mathbf{u}) = \mathbf{u}(1)f(\mathbf{e}_1) + \mathbf{u}(2)f(\mathbf{e}_2) + \dots + \mathbf{u}(n)f(\mathbf{e}_n).$$

Thus one can write:

$$f(\mathbf{u}) = M_f \cdot \mathbf{u}$$

where

$$M_f := (f(\mathbf{e}_1), f(\mathbf{e}_2), \dots, f(\mathbf{e}_n))$$

is the *matrix* formed by the n columns $f(\mathbf{e}_1), \dots, f(\mathbf{e}_n)$, and the product of M_f by a column $\mathbf{u} \in \mathbb{K}^n$ is defined by:

$$M_f \cdot \mathbf{u} := \mathbf{u}(1)f(\mathbf{e}_1) + \dots + \mathbf{u}(n)f(\mathbf{e}_n).$$

6.1.8 A matrix $A \in \mathbb{K}^{m \times n}$ induces a linear map

$$A \cdot : \mathbb{K}^n \rightarrow \mathbb{K}^m$$

defined by $\mathbf{u} \mapsto A\mathbf{u}$. Indeed, for every pair $\mathbf{u}, \mathbf{v} \in \mathbb{K}^n$ and every $\lambda, \mu \in \mathbb{K}$:

$$\begin{aligned} A(\lambda\mathbf{u} + \mu\mathbf{v}) &= \sum_{j=1}^n (\lambda\mathbf{u}(j) + \mu\mathbf{v}(j))C_j(A) = && \text{(definition of} \\ && \text{linear combination)} \\ &= \lambda \sum_{j=1}^n \mathbf{u}(j)C_j(A) + \mu \sum_{j=1}^n \mathbf{v}(j)C_j(A) = && \text{(distributive property} \\ && \text{and factoring out } \lambda \text{ and } \mu) \\ &= \lambda A\mathbf{u} + \mu A\mathbf{v} && \text{(definition of product} \\ && \text{matrix times vector)} \end{aligned}$$

Thus all linear maps from $\mathbb{K}^n$ to $\mathbb{K}^m$ are obtained by multiplying vectors of $\mathbb{K}^n$ by a matrix $A \in \mathbb{K}^{m \times n}$. This generalizes what was said in Example 6.1.5.

6.1.9 Esemplio. Observe that the linear map f of Example 6.1.3 can be written more explicitly as:

$$\begin{pmatrix} 1 & -2 \\ 1 & 1 \\ 1 & -3 \end{pmatrix} : \mathbb{K}^2 \longrightarrow \mathbb{K}^3$$

defined by

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 1 & -2 \\ 1 & 1 \\ 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - 2y \\ x + y \\ x - 3y \end{pmatrix} \in \mathbb{K}^3,$$

and hence f is linear by 6.1.8.

6.1.10 Esemplio. Consider the function

$$f : \mathbb{K}^4 \longrightarrow \mathbb{K}^3$$

defined by

$$f \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x + y - 2z + 3t \\ 2x + y - z - t \\ y - 4t \end{pmatrix}.$$

Again using 6.1.8, one concludes that it is linear. Indeed it can be written as:

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 1 & -2 & 3 \\ 2 & 1 & -1 & -1 \\ 0 & 1 & 0 & -4 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}.$$

6.2 The vector space of linear maps

6.2.1 As already anticipated, the set of linear maps $U \rightarrow V$ is traditionally denoted by

$$\text{Hom}_{\mathbb{K}}(U, V). \quad (6.4)$$

In other words, $f \in \text{Hom}_{\mathbb{K}}(U, V)$ if and only if f is a function from U to V , linear in the sense of (6.1). Linear maps are also called *homomorphisms*, and this explains the notation (6.4). If $U = V$, one sets:

$$\text{End}_{\mathbb{K}}(V) := \text{Hom}_{\mathbb{K}}(V, V),$$

which is the set of linear maps from V to itself. A linear map from V to itself is called an *endomorphism*. The set $\text{End}_{\mathbb{K}}(V)$ of endomorphisms of V is nonempty, whatever the vector space V may be. Indeed $\text{id}_V \in \text{End}(V)$ (Example 6.1.4).

6.2.2 Linear combinations of homomorphisms. Let U and V be two $\mathbb{K}$ -vector spaces and let V^U be the set of all functions $U \rightarrow V$. It is clearly a vector space: the linear combination of $f, g \in V^U$ with coefficients $\lambda, \mu \in \mathbb{K}$ is the unique function

$$\lambda f + \mu g : U \longrightarrow V,$$

such that:

$$(\lambda f + \mu g)(\mathbf{u}) = \lambda \cdot f(\mathbf{u}) + \mu \cdot g(\mathbf{u}),$$

for every $\mathbf{u} \in U$. This is the natural notion of linear combination that makes the set $V^U := \{f : U \rightarrow V\}$ of all functions from U to V into a $\mathbb{K}$ -vector space. Clearly $\text{Hom}_{\mathbb{K}}(U, V) \subset V^U$.

6.2.3 Proposition. *The set $\text{Hom}_{\mathbb{K}}(U, V)$ is a $\mathbb{K}$ -vector subspace of V^U , that is, if $f, g \in \text{Hom}_{\mathbb{K}}(U, V)$ then $\lambda f + \mu g \in \text{Hom}_{\mathbb{K}}(U, V)$, for every $\lambda, \mu \in \mathbb{K}$.*

Proof. Indeed, for every $a, b \in \mathbb{K}$ and every choice of $\mathbf{u}_1, \mathbf{u}_2 \in U$ one has:

$$\begin{aligned} (\lambda f + \mu g)(a\mathbf{u}_1 + b\mathbf{u}_2) &= \lambda \cdot f(a\mathbf{u}_1 + b\mathbf{u}_2) + \mu \cdot g(a\mathbf{u}_1 + b\mathbf{u}_2) && \text{(definition of } \lambda f + \mu g) \\ &= \lambda \cdot (af(\mathbf{u}_1) + bf(\mathbf{u}_2)) + \mu \cdot (ag(\mathbf{u}_1) + bg(\mathbf{u}_2)) && \text{(linearity of } f \text{ and } g) \\ &= a(\lambda f(\mathbf{u}_1) + \mu g(\mathbf{u}_1)) + b(\lambda f(\mathbf{u}_2) + \mu g(\mathbf{u}_2)) = && \text{(factoring out } a \text{ and } b) \\ &= a(\lambda f + \mu g)(\mathbf{u}_1) + b(\lambda f + \mu g)(\mathbf{u}_2), && \text{(definition of } \lambda f + \mu g) \end{aligned}$$

that is, $\lambda f + \mu g \in \text{Hom}_{\mathbb{K}}(U, V)$. ■

6.2.4 Composition of homomorphisms. Let $f : A \rightarrow B$ and $g : B \rightarrow C$ be two functions. The *composite function* of g with f , as one recalls, is the function defined by

$$(g \circ f)(a) = g(f(a)),$$

for every $a \in A$. It is worth recalling that if $h : C \rightarrow D$ is a third function, composition of functions is associative, in the sense that:

$$h \circ (g \circ f) = (h \circ g) \circ f, \tag{6.5}$$

an equality whose verification is based only on the definition. Let now U , V and W be three $\mathbb{K}$ -vector spaces. The composition of linear maps is linear, according to the following:

6.2.5 Proposition. *If $(f, g) \in \text{Hom}_{\mathbb{K}}(U, V) \times \text{Hom}_{\mathbb{K}}(V, W)$, then $g \circ f \in \text{Hom}_{\mathbb{K}}(U, W)$.*

Proof. For every choice of $\mathbf{u}_1, \mathbf{u}_2 \in U$ and $\lambda, \mu \in \mathbb{K}$ one has:

$$\begin{aligned}
 (g \circ f)(\lambda \mathbf{u}_1 + \mu \mathbf{u}_2) &= g(f(\lambda \mathbf{u}_1 + \mu \mathbf{u}_2)) && \text{(definition of } g \circ f) \\
 &= g(\lambda f(\mathbf{u}_1) + \mu f(\mathbf{u}_2)) && \text{(linearity of } f) \\
 &= \lambda g(f(\mathbf{u}_1)) + \mu g(f(\mathbf{u}_2)) && \text{(linearity of } g) \\
 &= \lambda (g \circ f)(\mathbf{u}_1) + \mu (g \circ f)(\mathbf{u}_2), && \text{(definition of } g \circ f)
 \end{aligned}$$

which proves what was required. ■

6.2.6 Proposition 6.2.5 implies, in particular, that the composition of two endomorphisms is an endomorphism. Therefore the set $\text{End}(V)$ is a vector space endowed with an additional multiplicative structure (composition of endomorphisms), which is associative (as a special case of (6.5)) and has an identity element. The *distributive properties* of addition with respect to composition also hold:

$$(\lambda f + \mu g) \circ h = \lambda f \circ h + \mu g \circ h,$$

and

$$h \circ (\lambda f + \mu g) = \lambda h \circ f + \mu h \circ g.$$

The mentioned properties of composition among elements of the vector space $\text{End}_{\mathbb{K}}(V)$ endow the latter with the structure of a $\mathbb{K}$ -algebra, called the $\mathbb{K}$ -algebra of endomorphisms of V . It is a noncommutative $\mathbb{K}$ -algebra.

6.2.7 Esempio. Let $f, g \in \text{End}_{\mathbb{K}}(\mathbb{K}^2)$ be defined by

$$f(\mathbf{u}) = \begin{pmatrix} \mathbf{u}(1) \\ 0 \end{pmatrix} \quad \text{and} \quad g(\mathbf{u}) = \begin{pmatrix} \mathbf{u}(2) \\ \mathbf{u}(2) \end{pmatrix}.$$

Then

$$g(f(\mathbf{u})) = g \begin{pmatrix} \mathbf{u}(1) \\ 0 \end{pmatrix} = \mathbf{0}$$

whereas

$$f(g(\mathbf{u})) = f \begin{pmatrix} \mathbf{u}(2) \\ \mathbf{u}(2) \end{pmatrix} = \begin{pmatrix} \mathbf{u}(2) \\ 0 \end{pmatrix},$$

from which one sees that $g(f(\mathbf{u})) \neq f(g(\mathbf{u}))$ if $\mathbf{u}(2) \neq 0$, that is, $g \circ f \neq f \circ g$.

6.3 Isomorphisms

We have the following fundamental definition:

6.3.1 Definition A bijective linear map $f : U \rightarrow V$, that is, one-to-one and onto, is called an isomorphism. Two vector spaces U and V are called isomorphic if there exists an isomorphism $f : U \rightarrow V$.

Since every bijective function between sets is invertible, every isomorphism $f \in \text{Hom}_{\mathbb{K}}(U, V)$ has a function $f^{-1} : V \rightarrow U$, called its *inverse*, such that

$$f \circ f^{-1} = id_V \quad \text{and} \quad f^{-1} \circ f = id_U.$$

There is, a priori, no guarantee that the inverse f^{-1} is in turn an isomorphism (that is, that in addition to being an invertible function it is also linear). Nevertheless, the following holds:

6.3.2 Proposition. If $f : U \rightarrow V$ is an isomorphism, then $f^{-1} \in \text{Hom}_{\mathbb{K}}(V, U)$.

Proof. Indeed, for every $\mathbf{v}_1, \mathbf{v}_2 \in V$ and every $\lambda, \mu \in \mathbb{K}$ one has:

$$\begin{aligned} f^{-1}(\lambda \mathbf{v}_1 + \mu \mathbf{v}_2) &= f^{-1}(\lambda(f \circ f^{-1})(\mathbf{v}_1) + \mu(f \circ f^{-1})(\mathbf{v}_2)) = && \text{(definition of } f^{-1}) \\ &= f^{-1}(\lambda f(f^{-1}(\mathbf{v}_1)) + \mu f(f^{-1}(\mathbf{v}_2))) = && \text{(definition of } f \circ f^{-1}) \\ &= f^{-1}(f(\lambda f^{-1}(\mathbf{v}_1) + \mu f^{-1}(\mathbf{v}_2))) = && \text{(linearity of } f) \\ &= (f^{-1} \circ f)(\lambda f^{-1}(\mathbf{v}_1) + \mu f^{-1}(\mathbf{v}_2)) = && \text{(definition of com-} \\ &&& \text{position)} \\ &= \lambda f^{-1}(\mathbf{v}_1) + \mu f^{-1}(\mathbf{v}_2), && (f^{-1} \circ f = id_U). \end{aligned}$$

■

Not all endomorphisms of V are invertible, as the following example shows:

6.3.3 Esempio. Let $f : \mathbb{K}^2 \rightarrow \mathbb{K}^2$ be defined by

$$f(\mathbf{u}) = \begin{pmatrix} \mathbf{u}(1) \\ 0 \end{pmatrix}.$$

Clearly f is a nonzero endomorphism of $\mathbb{K}^2$. It is not invertible, however, since if it were, there would have to exist a linear map g such that

$$f(g(\mathbf{u})) = \mathbf{u} \tag{6.6}$$

for every $\mathbf{u} \in \mathbb{K}^2$. Let then $\mathbf{u}$ be such that $\mathbf{u}(2) \neq 0$. Equation (6.6) is equivalent to:

$$\begin{pmatrix} \mathbf{u}(1) \\ \mathbf{u}(2) \end{pmatrix} = f \begin{pmatrix} g(\mathbf{u})(1) \\ g(\mathbf{u})(2) \end{pmatrix} = \begin{pmatrix} g(\mathbf{u})(1) \\ 0 \end{pmatrix}$$

which is absurd. Therefore f is not invertible.

6.3.4 Isomorphisms from V to itself are called *automorphisms* of V . They form a subset of $\text{End}_{\mathbb{K}}(V)$:

$$\text{Aut}_{\mathbb{K}}(V) = \{f \in \text{End}_{\mathbb{K}}(V) \mid \exists f^{-1} : V \rightarrow V\}.$$

Clearly $\text{id}_V \in \text{Aut}(V)$: indeed the *identity endomorphism* is injective and surjective.

6.3.5 Corollary *The set $\text{Aut}(V)$ of invertible endomorphisms is a group with respect to composition of (linear) functions.*

Proof. The operation “ $\circ$ ” is clearly internal: if $f, g \in \text{Aut}(V)$ then $g \circ f$ is an invertible endomorphism. Indeed $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, since

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \text{id}_V \circ g^{-1} = g \circ g^{-1} = \text{id}_V.$$

Moreover $\circ$ is associative because composition of functions is associative in general. The identity endomorphism id_V acts as the neutral element. Indeed: $f \circ \text{id}_V = \text{id}_V \circ f = f$. Finally, every $f \in \text{Aut}(V)$ is invertible by definition of $\text{Aut}(V)$. Thus $\text{Aut}(V)$ is a group, the *group of automorphisms of V* .

6.3.6 Exercise. Prove that, in general, $\text{Aut}(V)$ is not commutative, that is, explicitly produce a vector space V and two isomorphisms from V to itself such that $f \circ g \neq g \circ f$.

6.4 Kernel and image of a linear map

6.4.1 Let $f \in \text{Hom}_{\mathbb{K}}(U, V)$. The set

$$\text{Im}(f) := \{f(\mathbf{u}) \in V \mid \mathbf{u} \in U\}, \tag{6.7}$$

is a subset of V called the *image* of f . It is the set of all images in V of all vectors of U . The set

$$\ker(f) := \{\mathbf{u} \in U \mid f(\mathbf{u}) = \mathbf{0}_V\} \quad (6.8)$$

is the subset of U of all vectors that annihilate f , that is, whose image is the zero vector of V . It is called the *kernel* of the linear map f .

6.4.2 Theorem. *The kernel $\ker(f)$ of f is a vector subspace of U ; the image of f , $\text{Im}(f)$, is a vector subspace of V .*

Proof. If $\mathbf{u}_1, \mathbf{u}_2 \in \ker(f)$, then, for every $\lambda, \mu \in \mathbb{K}$, one has:

$$f(\lambda\mathbf{u}_1 + \mu\mathbf{u}_2) = \lambda f(\mathbf{u}_1) + \mu f(\mathbf{u}_2) = \lambda \cdot \mathbf{0}_V + \mu \cdot \mathbf{0}_V = \mathbf{0}_V + \mathbf{0}_V = \mathbf{0}_V,$$

and therefore $\lambda\mathbf{u}_1 + \mu\mathbf{u}_2 \in \ker(f)$. Thus the kernel of f is a subspace of U .

Let now $\mathbf{v}_1, \mathbf{v}_2 \in \text{Im}(f)$. By definition there exist $\mathbf{u}_1, \mathbf{u}_2 \in U$ such that

$$\mathbf{v}_1 = f(\mathbf{u}_1) \quad \text{and} \quad \mathbf{v}_2 = f(\mathbf{u}_2),$$

so, for every $\lambda, \mu \in \mathbb{K}$:

$$\lambda\mathbf{v}_1 + \mu\mathbf{v}_2 = \lambda f(\mathbf{u}_1) + \mu f(\mathbf{u}_2) = f(\lambda\mathbf{u}_1 + \mu\mathbf{u}_2),$$

proving that $\lambda\mathbf{v}_1 + \mu\mathbf{v}_2$ is itself the image of a vector of U . Therefore $\lambda\mathbf{v}_1 + \mu\mathbf{v}_2 \in \text{Im}(f)$ and the image of f is a subspace. ■

6.4.3 Proposition. *Let $f \in \text{Hom}_{\mathbb{K}}(U, V)$. Then i) f is surjective if and only if $\text{Im}(f) = V$; ii) f is injective if and only if $\ker f = \mathbf{0}_U$.*

Proof. We begin by proving i). If $\text{Im}(f) = V$, every $\mathbf{v} \in V$ is of the form $f(\mathbf{u}) = \mathbf{v}$, for some $\mathbf{u} \in U$, by definition of image. Thus f is surjective. Conversely, if f is surjective, for every $\mathbf{v} \in V$ there exists $\mathbf{u} \in U$ such that $f(\mathbf{u}) = \mathbf{v}$, that is, $\mathbf{v} \in \text{Im}(f)$. Therefore, if f is surjective, $V \subseteq \text{Im}(f)$, which together with the inclusion $\text{Im}(f) \subseteq V$ implies $V = \text{Im}(f)$, as required.

As for ii), if f is injective and $\mathbf{u} \in \ker(f)$, then $f(\mathbf{u}) = \mathbf{0}_V = f(\mathbf{0}_U)$, hence $\mathbf{u} = \mathbf{0}_U$ (by the assumed injectivity of f): therefore $\ker(f)$ contains only the zero vector. Conversely, if $\ker(f) := \{\mathbf{0}_U\}$, the equality $f(\mathbf{u}_1) = f(\mathbf{u}_2)$ implies $f(\mathbf{u}_1) - f(\mathbf{u}_2) = \mathbf{0}_V$. By linearity, $f(\mathbf{u}_1 - \mathbf{u}_2) = \mathbf{0}$, so $\mathbf{u}_1 - \mathbf{u}_2 \in \ker(f)$ by definition of kernel. Hence:

$$\mathbf{u}_1 - \mathbf{u}_2 = \mathbf{0}_U,$$

because $\mathbf{0}_U$ is the only element of $\ker(f)$. Thus $\mathbf{u}_1 = \mathbf{u}_2$ and f is injective. ■

6.4.4 Proposition. *A linear map $f : U \rightarrow V$ is injective if and only if the image of any set of linearly independent vectors of U is a set of linearly independent vectors of V .*

Proof. Suppose f is injective and $\mathbf{u}_1, \dots, \mathbf{u}_h$ are linearly independent. Then the linear relation

$$\lambda_1 f(\mathbf{u}_1) + \dots + \lambda_h f(\mathbf{u}_h) = \mathbf{0}_V \quad (6.9)$$

implies, by linearity, the equality:

$$f(\lambda_1 \mathbf{u}_1 + \dots + \lambda_h \mathbf{u}_h) = \mathbf{0}_V.$$

Thus, by the definition (6.8) of kernel:

$$\lambda_1 \mathbf{u}_1 + \dots + \lambda_h \mathbf{u}_h \in \ker(f). \quad (6.10)$$

Since f is injective, $\ker(f) = \mathbf{0}_U$. Therefore (6.10) is equivalent to:

$$\lambda_1 \mathbf{u}_1 + \dots + \lambda_h \mathbf{u}_h = \mathbf{0}_U.$$

Then $\lambda_1 = \dots = \lambda_h = 0$, because $\mathbf{u}_1, \dots, \mathbf{u}_h$ are linearly independent by hypothesis, and therefore $f(\mathbf{u}_1), \dots, f(\mathbf{u}_h)$ are linearly independent (because (6.9) holds only for all scalars zero).

Conversely, suppose that the image of every linearly independent subset is linearly independent. If the kernel $\ker(f)$ contained a nonzero vector, by hypothesis its image would have to be nonzero (a nonzero vector is linearly independent), contrary to the assumption that it lies in the kernel. Hence the kernel contains only the zero vector.

■

6.4.5 Proposition. *Let $f \in \text{Hom}_{\mathbb{K}}(U, V)$ and let $(\mathbf{b}_1, \dots, \mathbf{b}_n)$ be a basis of U . Then*

$$\text{Im}(f) = [f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)]$$

that is, the image subspace of f in V is generated by the images of a basis of U .

Proof. It is enough to prove that $\text{Im}(f) \subseteq [f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)]$. Let $\mathbf{v} \in \text{Im}(f)$. By definition there exists $\mathbf{u} \in U$ such that $f(\mathbf{u}) = \mathbf{v}$. Hence $\mathbf{u} = \lambda_1 \mathbf{b}_1 + \dots + \lambda_n \mathbf{b}_n$, for certain (unique) $\lambda_1, \dots, \lambda_n$. Then

$$\mathbf{v} = f(\mathbf{u}) = f\left(\sum_{i=1}^n \lambda_i \mathbf{b}_i\right) = \sum_{i=1}^n \lambda_i f(\mathbf{b}_i).$$

This proves that $\mathbf{v} \in [f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)]$, as required. ■

6.4.6 Corollary. *Let $f \in \text{Hom}_{\mathbb{K}}(U, V)$ and $\dim_{\mathbb{K}} U = n < \infty$. Then f is injective if and only if $\dim_{\mathbb{K}} \text{Im}(f) = \dim_{\mathbb{K}} U$, and is surjective if and only if $\dim_{\mathbb{K}} \text{Im}(f) = \dim_{\mathbb{K}} V$.*

Proof. If f is injective, the images of the vectors of a basis are linearly independent vectors. Therefore

$$\dim_{\mathbb{K}} \text{Im}(f) = \dim_{\mathbb{K}} [f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)] = n = \dim_{\mathbb{K}} U.$$

Conversely, if the image of a basis of U is a basis of the image, then f is injective. Indeed, if $\mathbf{u} \in \ker f$, then $\mathbf{u} = u_1 \mathbf{b}_1 + \dots + u_n \mathbf{b}_n$, hence

$$f(\mathbf{u}) = \mathbf{0}_V = \lambda_1 f(\mathbf{b}_1) + \dots + \lambda_n f(\mathbf{b}_n)$$

and all the λ_i are then zero, because the image of the basis $\mathcal{B}$ is a basis of the image. Hence $\mathbf{u} = \mathbf{0}_U$. The assertion concerning surjectivity has already been discussed. ■

6.4.7 Exercise. Construct an example of a non-injective linear map to show that, in general, the image of a basis of U is not a basis of the image.

6.4.8 Proposition. *Let U and V be two $\mathbb{K}$ -vector spaces with $\dim_{\mathbb{K}} U = n$. Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be the vectors of a basis of U and let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be arbitrary vectors of V . Then there exists a unique linear map $f : U \rightarrow V$ such that $f(\mathbf{b}_i) = \mathbf{v}_i$.*

Proof. We begin by constructing the linear map. Every $\mathbf{u} \in U$ can be written uniquely as a linear combination of $(\mathbf{b}_1, \dots, \mathbf{b}_n)$, since this is a basis of U :

$$\mathbf{u} = \lambda_1 \mathbf{b}_1 + \dots + \lambda_n \mathbf{b}_n.$$

Define $f : U \rightarrow V$ as follows:

$$f(\mathbf{u}) = \lambda_1 \mathbf{v}_1 + \dots + \lambda_n \mathbf{v}_n.$$

It follows that $f \in \text{Hom}_{\mathbb{K}}(U, V)$. Indeed, if $\mathbf{u}' = \sum_{i=1}^n \lambda'_i \mathbf{b}_i$, one has:

$$\lambda \mathbf{u} + \lambda' \mathbf{u}' = \lambda \left(\sum_{i=1}^n \lambda_i \mathbf{b}_i \right) + \lambda' \left(\sum_{i=1}^n \lambda'_i \mathbf{b}_i \right) = \sum_{i=1}^n (\lambda \lambda_i + \lambda' \lambda'_i) \mathbf{b}_i$$

and hence:

$$\begin{aligned} f(\lambda \mathbf{u} + \lambda' \mathbf{u}') &= f \sum_{i=1}^n (\lambda \lambda_i + \lambda' \lambda'_i) \mathbf{b}_i = \\ &= \lambda \sum_{i=1}^n \lambda_i \mathbf{v}_i + \lambda' \sum_{i=1}^n \lambda'_i \mathbf{v}_i = \lambda f(\mathbf{u}) + \lambda' f(\mathbf{u}'), \end{aligned}$$

thereby proving the existence of f . Let us now prove the uniqueness of the f just constructed. Suppose that there exists another $g : U \rightarrow V$ such that $g(\mathbf{b}_i) = \mathbf{v}_i$. Then, in this case, for every $\mathbf{u} \in U$ one has:

$$\begin{aligned} g(\mathbf{u}) &= g\left(\sum_{i=1}^n \lambda_i \mathbf{b}_i\right) = (\text{by linearity of } g) = \sum_{i=1}^n \lambda_i g(\mathbf{b}_i) = \\ &= \sum_{i=1}^n \lambda_i \mathbf{v}_i = f(\mathbf{u}). \end{aligned}$$

Thus $g(\mathbf{u}) = f(\mathbf{u})$ for every $\mathbf{u} \in U$, and hence $f = g$. ■

6.4.9 Theorem (dimension theorem for Kernel and Image). *Let $f : U \rightarrow V$ be a linear map between finite-dimensional $\mathbb{K}$ -vector spaces. Then the following fundamental equality holds:*

$$\dim_{\mathbb{K}}(U) = \dim_{\mathbb{K}} \ker(f) + \dim_{\mathbb{K}} \operatorname{Im}(f) \quad (6.11)$$

Proof. Let $(\mathbf{b}_1, \dots, \mathbf{b}_n)$ be a basis of U . If f is injective, then $\ker(f) = \mathbf{0}_U$. Hence $\dim \ker(f) = 0$ and, since by Corollary 6.4.6 $(f(\mathbf{b}_1), \dots, f(\mathbf{b}_n))$ is a basis of $\operatorname{Im}(f)$, one has $\dim \operatorname{Im}(f) = n$. Thus (6.11) is trivially verified if f is injective. In the case where f is not injective, consider a basis $\mathbf{b}_1, \dots, \mathbf{b}_h$ of the kernel of f . This can be completed to a basis $\mathcal{B} := (\mathbf{b}_1, \dots, \mathbf{b}_h, \mathbf{b}_{h+1}, \dots, \mathbf{b}_n)$ of the space U ⁽¹⁾. The theorem follows from the fact that $(f(\mathbf{b}_{h+1}), \dots, f(\mathbf{b}_n))$ is a basis of $\operatorname{Im}(f)$. If this is the case, indeed, one obtains $\dim \operatorname{Im}(f) = n - h = \dim(U) - \dim \ker(f)$, as desired. To prove that $(f(\mathbf{b}_{h+1}), \dots, f(\mathbf{b}_n))$ is a basis of $\operatorname{Im}(f)$, let $\mathbf{u} \in \operatorname{Im}(f)$ be arbitrary. Since the image of the basis $(f(\mathbf{b}_1), \dots, f(\mathbf{b}_n))$ generates $\operatorname{Im}(f)$, one has:

$$\mathbf{u} = \lambda_1 f(\mathbf{b}_1) + \dots + \lambda_h f(\mathbf{b}_h) + \lambda_{h+1} f(\mathbf{b}_{h+1}) + \dots + \lambda_n f(\mathbf{b}_n)$$

for some choice (not necessarily unique)² of scalars $(\lambda_1, \dots, \lambda_n)$. But then:

$$\mathbf{u} = \lambda_{h+1} f(\mathbf{b}_{h+1}) + \dots + \lambda_n f(\mathbf{b}_n)$$

because $\mathbf{b}_1, \dots, \mathbf{b}_h$ belong, by hypothesis, to $\ker(f)$. Thus

$$\operatorname{Im}(f) = [f(\mathbf{b}_{h+1}), \dots, f(\mathbf{b}_n)].$$

¹I am rushing a little. Why?

²Indeed, under the hypothesis, necessarily not unique!

It remains only to prove that the $f(\mathbf{b}_i)$, with $h+1 \leq i \leq n$, are linearly independent. Indeed, if

$$\mu_{h+1}f(\mathbf{b}_{h+1}) + \dots + \mu_n f(\mathbf{b}_n) = \mathbf{0}_V,$$

then

$$f(\mu_{h+1}\mathbf{b}_{h+1} + \dots + \mu_n\mathbf{b}_n) = \mathbf{0}_V$$

by the linearity of f . Hence $\mu_{h+1}\mathbf{b}_{h+1} + \dots + \mu_n\mathbf{b}_n \in \ker(f)$. Therefore

$$\mu_{h+1}\mathbf{b}_{h+1} + \dots + \mu_n\mathbf{b}_n = \mu_1\mathbf{b}_1 + \dots + \mu_h\mathbf{b}_h$$

for a (unique!) choice of scalars $\mu_1, \dots, \mu_h$, since $(\mathbf{b}_1, \dots, \mathbf{b}_h)$ is a basis of $\ker(f)$. But then:

$$\mu_1\mathbf{b}_1 + \dots + \mu_h\mathbf{b}_h - \mu_{h+1}\mathbf{b}_{h+1} - \dots - \mu_n\mathbf{b}_n = \mathbf{0}_U$$

is a zero linear combination of a basis of U and hence, necessarily, $\lambda_1 = \dots = \lambda_h = \lambda_{h+1} = \dots = \lambda_n = 0$, whence, in particular, $\lambda_{h+1} = \dots = \lambda_n = 0$, that is, $(f(\mathbf{b}_{h+1}), \dots, f(\mathbf{b}_n))$ are linearly independent and therefore form a basis of $\text{Im}(f)$, as required. ■

6.4.10 Solved exercise. Let $U = \mathbb{K}^n$ and $V = \mathbb{K}^m$ and let $f \in \text{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$. Prove that if $n > m$, then f cannot be injective.

Solution. It is enough to apply formula (6.11) to the described situation. One has

$$\dim_{\mathbb{K}}(\mathbb{K}^n) = \dim_{\mathbb{K}}(\ker(f)) + \dim_{\mathbb{K}}(\text{Im}(f)).$$

Since $\text{Im}(f) \subseteq V$, it follows that $\dim(\text{Im}(f)) \leq \dim_{\mathbb{K}} \mathbb{K}^m = m$. Thus

$$\dim \ker(f) = n - \dim \text{Im}(f) \geq n - m > 0,$$

where the last inequality follows from the fact that $n > m$. Since $\dim \ker(f) > 0$ implies that $\ker(f) \neq \{\mathbf{0}\}$, f cannot be injective.

6.4.11 Solved exercise. Let $f : \mathbb{K}^3 \rightarrow \mathbb{K}^4$ be linear. Prove that f cannot be surjective.

Solution. If it were surjective, then $\text{Im}(f) = \mathbb{K}^4$ and hence $\dim \text{Im}(f) = 4$. Since $\dim \ker(f) \geq 0$, applying (6.11) one would have:

$$\dim \mathbb{K}^3 = \dim \ker(f) + \dim \text{Im}(f) = \dim \ker(f) + 4 \geq 4$$

but this is impossible, since the dimension of $\mathbb{K}^3$ is 3³.

³Indeed, one of its bases is the canonical one: $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

6.4.12 Exercise. Generalize the solved exercise 6.4.11: if $f \in \text{Hom}_{\mathbb{K}}(U, V)$ and $\dim_{\mathbb{K}} U < \dim_{\mathbb{K}} V$, then f cannot be surjective.

6.4.13 Proposition. Let U and V be two $\mathbb{K}$ -vector spaces of the same finite dimension. Then $f \in \text{Hom}_{\mathbb{K}}(U, V)$ is injective if and only if it is surjective.

Proof. If f is injective, then $\dim \ker(f) = 0$ (by Proposition 6.4.3, ii)) and therefore (6.11) gives:

$$\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} \ker(f) + \dim \text{Im}(f) = 0 + \dim \text{Im}(f),$$

so that $\dim_{\mathbb{K}}(\text{Im}(f)) = \dim_{\mathbb{K}} U = \dim_{\mathbb{K}} V$, whence f is surjective (indeed $\dim \text{Im}(f) = \dim V$ implies $\text{Im}(f) = V$ and hence f is surjective by property 6.4.3 i). Conversely, if f is surjective, then $\text{Im}(f) = V$ (6.4.3), hence $\dim \text{Im}(f) = \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} U$, whence, applying (6.11):

$$\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} \ker(f) + \dim_{\mathbb{K}} \text{Im}(f) = \dim_{\mathbb{K}} \ker(f) + \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \ker(f) + \dim_{\mathbb{K}} U$$

and therefore, evidently, $\dim \ker(f) = 0$, that is, $\ker(f) = \mathbf{0}_U$, hence, by 6.4.3, ii), f is injective. ■

6.4.14 Recall from 6.3 that two vector spaces U and V are called isomorphic if and only if there exists $f \in \text{Hom}_{\mathbb{K}}(U, V)$ which is injective and surjective. In this case f is called an *isomorphism* between U and V .

6.4.15 Theorem. Two finite-dimensional $\mathbb{K}$ -vector spaces are isomorphic if and only if $\dim_{\mathbb{K}}(U) = \dim_{\mathbb{K}}(V)$.

Proof. Indeed, if they are isomorphic, let f be an isomorphism, that is, a linear map from U to V that is injective and surjective. Since f is surjective, one has $\dim_{\mathbb{K}} \text{Im}(f) = \dim_{\mathbb{K}} V$. Moreover, since it is injective, one has $\dim_{\mathbb{K}} \ker(f) = 0$. Thus, applying (6.11), one obtains:

$$\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} \ker(f) + \dim_{\mathbb{K}} \text{Im}(f) = 0 + \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} V,$$

proving that U and V have the same dimension if they are isomorphic. Conversely, let U and V have the same dimension. Let $\mathcal{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ be a basis of U and $\mathcal{C} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ a basis of V . Let $f : U \rightarrow V$ be the unique linear map defined by

$$f(\mathbf{b}_i) = \mathbf{c}_i.$$

This map exists and is unique by Proposition 6.4.8. Moreover it is injective, since $f(\mathbf{u}) = \mathbf{0}_V$ implies that $f(\lambda_1 \mathbf{b}_1 + \dots + \lambda_n \mathbf{b}_n) = \lambda_1 f(\mathbf{b}_1) + \dots + \lambda_n f(\mathbf{b}_n) = \mathbf{0}_V$. That is:

$$\lambda_1 \mathbf{c}_1 + \dots + \lambda_n \mathbf{c}_n = \mathbf{0}_V,$$

hence $\lambda_1 = \dots = \lambda_n = 0$. Therefore $\mathbf{u} = \mathbf{0}_U$ and so f is injective. Then by 6.4.13, f is also surjective. Hence f is an isomorphism. ■

6.5 Change-of-basis matrices

6.5.1 Let U be a $\mathbb{K}$ -vector space of dimension n and let $\mathcal{B} := (\mathbf{b}_1, \dots, \mathbf{b}_n)$ be a basis of it. If $\mathbf{u} = u_1\mathbf{b}_1 + u_2\mathbf{b}_2 + \dots + u_n\mathbf{b}_n$, denote by $(\mathbf{u})_{\mathcal{B}}$ the column, with n scalar entries, of the components of $\mathbf{u}$ with respect to the basis $\mathcal{B}$:

$$(\mathbf{u})_{\mathcal{B}} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

In particular

$$(\mathbf{b}_j)_{\mathcal{B}} = \mathbf{e}_j. \quad (6.12)$$

6.5.2 Proposition. *The function*

$$\begin{cases} \phi_{\mathcal{B}} : U \longrightarrow \mathbb{K}^n \\ \mathbf{u} \longmapsto (\mathbf{u})_{\mathcal{B}}, \end{cases} \quad (6.13)$$

which associates to each $\mathbf{u} \in U$ the column of its components with respect to the basis $\mathcal{B}$, is an isomorphism.

Proof. It is indeed the unique isomorphism between U and $\mathbb{K}^n$ such that $\phi_{\mathcal{B}}(\mathbf{b}_i) = \mathbf{e}_i$. ■

In particular, the $\mathcal{B}$ -components of a linear combination are the linear combination of the $\mathcal{B}$ -components, that is:

$$(\lambda\mathbf{u} + \mu\mathbf{v})_{\mathcal{B}} = \lambda(\mathbf{u})_{\mathcal{B}} + \mu(\mathbf{v})_{\mathcal{B}} \quad (6.14)$$

for every pair $\mathbf{u}, \mathbf{v} \in U$ and every $\lambda, \mu \in \mathbb{K}$.

6.5.3 Let now $\mathcal{B} := (\mathbf{b}_1, \dots, \mathbf{b}_n)$ and $\mathcal{C} := (\mathbf{c}_1, \dots, \mathbf{c}_n)$ be two bases of U . If

$$\mathbf{u} = u_1\mathbf{c}_1 + \dots + u_n\mathbf{c}_n,$$

then:

$$(\mathbf{u})_{\mathcal{B}} = (u_1\mathbf{c}_1 + \dots + u_n\mathbf{c}_n)_{\mathcal{B}} = u_1(\mathbf{c}_1)_{\mathcal{B}} + \dots + u_n(\mathbf{c}_n)_{\mathcal{B}}$$

where the last equality follows from linearity (6.14). Hence:

$$(\mathbf{u})_{\mathcal{B}} = ((\mathbf{c}_1)_{\mathcal{B}}, \dots, (\mathbf{c}_n)_{\mathcal{B}}) \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = P^{\mathcal{B}, \mathcal{C}} \cdot (\mathbf{u})_{\mathcal{C}}$$

where we have put:

$$P^{\mathcal{B}, \mathcal{C}} = ((\mathbf{c}_1)_{\mathcal{B}}, \dots, (\mathbf{c}_n)_{\mathcal{B}}) \in \mathbb{K}^{n \times n}, \quad (6.15)$$

the matrix whose j -th column is the column of the components of the j -th vector $\mathbf{c}_j$ of the basis $\mathcal{C}$ with respect to the basis $\mathcal{B}$.

6.5.4 Proposition.

i) For every basis $\mathcal{B} := (\mathbf{b}_i)_{1 \leq i \leq n}$ of U

$$P^{\mathcal{B}, \mathcal{B}} = \mathbf{1}_n := (\mathbf{e}_1, \dots, \mathbf{e}_n) \quad (6.16)$$

ii) If $\mathcal{B} := (\mathbf{b}_i)_{1 \leq i \leq n}$, $\mathcal{C} := (\mathbf{c}_i)_{1 \leq i \leq n}$, $\mathcal{D} := (\mathbf{d}_i)_{1 \leq i \leq n}$ are three arbitrary bases of U , then:

$$P^{\mathcal{B}, \mathcal{C}} P^{\mathcal{C}, \mathcal{D}} = P^{\mathcal{B}, \mathcal{D}}. \quad (6.17)$$

Proof. For i) one easily has:

$$P^{\mathcal{B}, \mathcal{B}} = ((\mathbf{b}_1)_{\mathcal{B}}, \dots, (\mathbf{b}_n)_{\mathcal{B}}) = (\mathbf{e}_1, \dots, \mathbf{e}_n) = \mathbf{1}_n.$$

As for ii):

$$P^{\mathcal{B}, \mathcal{C}} P^{\mathcal{C}, \mathcal{D}} = P^{\mathcal{B}, \mathcal{C}} ((\mathbf{d}_1)_{\mathcal{C}}, \dots, (\mathbf{d}_n)_{\mathcal{C}}) = \quad (6.18)$$

$$= (P^{\mathcal{B}, \mathcal{C}}(\mathbf{d}_1)_{\mathcal{C}}, \dots, P^{\mathcal{B}, \mathcal{C}}(\mathbf{d}_n)_{\mathcal{C}}) = \quad (6.19)$$

$$= ((\mathbf{d}_1)_{\mathcal{B}}, \dots, (\mathbf{d}_n)_{\mathcal{B}}) = P^{\mathcal{B}, \mathcal{D}}. \quad (6.20)$$

where (6.19) comes from rule (??) for multiplying two matrices. ■

6.5.5 Corollary. The matrix $P^{\mathcal{B}, \mathcal{C}}$ is invertible and its inverse is $P^{\mathcal{C}, \mathcal{B}}$.

Proof. Indeed, by Proposition 6.5.4 ii) and then i), one has

$$P^{\mathcal{B}, \mathcal{C}} \cdot P^{\mathcal{C}, \mathcal{B}} = P^{\mathcal{B}, \mathcal{B}} = \mathbf{1}_n \quad \text{and} \quad P^{\mathcal{C}, \mathcal{B}} \cdot P^{\mathcal{B}, \mathcal{C}} = P^{\mathcal{C}, \mathcal{C}} = \mathbf{1}_n.$$

■

6.6 Matrices associated with linear maps.

6.6.1 Let now $f : U \rightarrow V$ be a linear map between two $\mathbb{K}$ -vector spaces of dimensions n and m , respectively, and let $\mathcal{B} := (\mathbf{b}_1, \dots, \mathbf{b}_n)$ and $\mathcal{C} = (\mathbf{c}_1, \dots, \mathbf{c}_m)$ be bases of U and V , respectively. The matrix

$$M_f^{\mathcal{C}, \mathcal{B}} := (f(\mathbf{b}_1)_{\mathcal{C}}, \dots, f(\mathbf{b}_n)_{\mathcal{C}}) \in \mathbb{K}^{m \times n}, \quad (6.21)$$

is called the *matrix associated* with f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$. Then, for every $\mathbf{u} \in U$:

$$M_f^{\mathcal{C}, \mathcal{B}} \cdot (\mathbf{u})_{\mathcal{B}} = f(\mathbf{u})_{\mathcal{C}}. \quad (6.22)$$

Indeed

$$\begin{aligned} f(\mathbf{u})_{\mathcal{C}} &= f(u_1 \mathbf{b}_1 + \dots + u_n \mathbf{b}_n)_{\mathcal{C}} = && (\mathbf{u} \text{ as a linear combination of the } \mathbf{b}_i) \\ &= (u_1 f(\mathbf{b}_1) + \dots + u_n f(\mathbf{b}_n))_{\mathcal{C}} = && (\text{linearity of } f) \\ &= u_1 (f(\mathbf{b}_1))_{\mathcal{C}} + \dots + u_n (f(\mathbf{b}_n))_{\mathcal{C}} = && (\text{applying (6.14)}) \\ &= (f(\mathbf{b}_1)_{\mathcal{C}}, \dots, f(\mathbf{b}_n)_{\mathcal{C}}) \cdot (\mathbf{u})_{\mathcal{B}}. && (\text{by definition (??) of matrix times column}) \end{aligned}$$

If $\mathbf{0} : U \rightarrow V$ is the linear map that sends every vector $\mathbf{u} \in U$ to the zero vector of V , the matrix $M_{\mathbf{0}}^{\mathcal{C}, \mathcal{B}}$ is the zero matrix for every choice of bases $\mathcal{B}$ in U and $\mathcal{C}$ in V . Indeed

$$M_{\mathbf{0}}^{\mathcal{C}, \mathcal{B}} := (\mathbf{0}(\mathbf{b}_1), \dots, \mathbf{0}(\mathbf{b}_n)).$$

6.6.2 Proposition. Let $f, g \in \text{Hom}_{\mathbb{K}}(U, V)$. Then:

$$M_{\lambda f + \mu g}^{\mathcal{C}, \mathcal{B}} = \lambda M_f^{\mathcal{C}, \mathcal{B}} + \mu M_g^{\mathcal{C}, \mathcal{B}}. \quad (6.23)$$

Proof. It is enough to prove that the columns of the matrices in the left- and right-hand sides of (6.23) are equal. First observe that, for every $1 \leq j \leq n$, one has:

$$C_j(M_{\lambda f + \mu g}^{\mathcal{C}, \mathcal{B}}) = \lambda (f(\mathbf{b}_j))_{\mathcal{C}} + \mu (g(\mathbf{b}_j))_{\mathcal{C}} \quad (6.24)$$

Indeed:

$$\begin{aligned}
C_j(M_{\lambda f + \mu g}^{\mathcal{C}, \mathcal{B}}) &= M_{\lambda f + \mu g}^{\mathcal{C}, \mathcal{B}} \cdot \mathbf{e}_j = && (j\text{-th column}) \\
&= M_{\lambda f + \mu g}^{\mathcal{C}, \mathcal{B}} \cdot (\mathbf{b}_j)_{\mathcal{B}} = && (\text{equality (6.12)}) \\
&= ((\lambda f + \mu g)(\mathbf{b}_j))_{\mathcal{C}} = && (\text{formula (6.22) applied to } \lambda f + \mu g) \\
&= (\lambda f(\mathbf{b}_j) + \mu g(\mathbf{b}_j))_{\mathcal{C}} = && (\text{definition of } \lambda f + \mu g) \\
&= \lambda(f(\mathbf{b}_j))_{\mathcal{C}} + \mu(g(\mathbf{b}_j))_{\mathcal{C}} && (\text{linearity of } \phi_{\mathcal{C}}, \text{ Cf. (6.13)})
\end{aligned}$$

But:

$$\begin{aligned}
\lambda(f(\mathbf{b}_j))_{\mathcal{C}} + \mu(g(\mathbf{b}_j))_{\mathcal{C}} &= \lambda M_f^{\mathcal{C}, \mathcal{B}} \cdot (\mathbf{b}_j)_{\mathcal{B}} + \mu M_g^{\mathcal{C}, \mathcal{B}} \cdot (\mathbf{b}_j)_{\mathcal{B}} = && (\text{formula (6.22) applied to } \lambda f + \mu g) \\
&= (\lambda M_f^{\mathcal{C}, \mathcal{B}} + \mu M_g^{\mathcal{C}, \mathcal{B}}) \cdot (\mathbf{b}_j)_{\mathcal{B}} = && (\text{linearity of the product of matrices}) \\
&= (\lambda M_f^{\mathcal{C}, \mathcal{B}} + \mu M_g^{\mathcal{C}, \mathcal{B}}) \cdot \mathbf{e}_j = && (\text{equality (6.12)}) \\
&= C_j((\lambda M_f^{\mathcal{C}, \mathcal{B}} + \mu M_g^{\mathcal{C}, \mathcal{B}})). && (6.25)
\end{aligned}$$

Equations (6.24) and (6.25) imply the asserted equality between the columns, thus proving the proposition. $\blacksquare$

Similarly, one has:

6.6.3 Proposition. *Let $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ be bases of $\mathbb{K}$ -vector spaces U , V and W , respectively, and let $f \in \text{Hom}_{\mathbb{K}}(U, V)$ and $g \in \text{Hom}_{\mathbb{K}}(V, W)$. Then*

$$M_{g \circ f}^{\mathcal{D}, \mathcal{B}} = M_g^{\mathcal{D}, \mathcal{C}} \cdot M_f^{\mathcal{C}, \mathcal{B}}. \quad (6.26)$$

Proof. Indeed

$$M_{g \circ f}^{\mathcal{D}, \mathcal{B}} = (((g \circ f)(\mathbf{b}_1))_{\mathcal{D}}, \dots, ((g \circ f)(\mathbf{b}_n))_{\mathcal{D}}) = \quad (6.27)$$

$$= ((g(f(\mathbf{b}_1)))_{\mathcal{D}}, \dots, (g(f(\mathbf{b}_n)))_{\mathcal{D}}) = \quad (6.28)$$

$$= (M_g^{\mathcal{D}, \mathcal{C}} \cdot (f(\mathbf{b}_1))_{\mathcal{C}}, \dots, M_g^{\mathcal{D}, \mathcal{C}} \cdot (f(\mathbf{b}_n))_{\mathcal{C}}) = \quad (6.29)$$

$$= M_g^{\mathcal{D}, \mathcal{C}} \cdot ((f(\mathbf{b}_1))_{\mathcal{C}}, \dots, (f(\mathbf{b}_n))_{\mathcal{C}}) = \quad (6.30)$$

$$= M_g^{\mathcal{D}, \mathcal{C}} \cdot M_f^{\mathcal{C}, \mathcal{B}}, \quad (6.31)$$

where (6.27) is the definition of the matrix associated with the homomorphism $g \circ f$, (6.28) uses the definition of composition of functions, (6.29) the definition of the matrix associated with g with respect to the bases $\mathcal{C}$ and $\mathcal{D}$, (6.30) is the rule (??) for multiplying a matrix by a column, while (6.31) uses the definition of the matrix associated with f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$. ■

6.6.4 Solved Exercise. Let $\mathbb{K}_{\leq 2}[X]$ be the $\mathbb{K}$ -vector space of polynomials of degree ≤ 2 .

- a) Prove that $\mathcal{C} := (1 + X, 1 - X, X^2)$ is a basis of $\mathbb{K}_2[X]$;
- b) prove that $\mathcal{B} := (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$, where:

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{b}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

is a basis of $\mathbb{R}^3$;

- c) Let the function $\mathbb{K}_2[X] \longrightarrow \mathbb{R}^3$ be given by:

$$f(a_0 + a_1X + a_2X^2) = \begin{pmatrix} a_0 + 2a_1 + a_2 \\ a_1 - a_2 \\ a_0 + a_1 + a_2 \end{pmatrix}.$$

Prove that it is linear.

- d) Find the matrix associated with f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$;

e) Compute the product

$$M_f^{\mathcal{B}, \mathcal{C}} \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

and interpret the result.

Solution.

a) i) First method. Directly. Let:

$$\lambda(1 + X) + \mu(1 - X) + cX^2 = 0$$

be a zero linear combination of the elements of $\mathcal{C}$. Then

$$(\lambda + \mu) + (\lambda - \mu)X + \nu X^2 = 0.$$

Since $\mathcal{S} := (1, X, X^2)$ is a basis of $\mathbb{K}_{\leq 2}[X]$, it follows that $\lambda + \mu = 0$, $\lambda - \mu = 0$ and $\nu = 0$, that is, $\lambda = \mu = \nu$.

ii) Second method. Since $(1, X, X^2)$ is a basis of $\mathbb{K}_{\leq 2}[X]$, let

$$P := ((1 + X)_{\mathcal{S}}, (1 - X)_{\mathcal{S}}, (X^2)_{\mathcal{S}}) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

be the matrix whose columns are the components of the vectors of $\mathcal{C}$ with respect to the standard basis $(1, X, X^2)$. If the matrix is invertible, then $\mathcal{C}$ is a basis. A quick calculation shows that $\det(P) = -2 \neq 0$, that is, the matrix P is invertible, that is, $\mathcal{C}$ is a basis.

b) Since the components of $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ are expressed in the canonical basis, it is enough to evaluate the determinant of the matrix $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$:

$$\det(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} =$$

which is easily seen to be equal to -2 . Therefore, since it is nonzero, $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ is a basis of $\mathbb{R}^3$.

- c) i) First method. It consists of the direct application of the definition. Let $P_a := a_0 + a_1X + a_2X^2$ and $P_b := b_0 + b_1X + b_2X^2$ be two polynomials of degree ≤ 2 . For every $\lambda, \mu \in \mathbb{K}$ one has:

$$\begin{aligned}
 f(\lambda P_a + \mu P_b) &= f(\lambda(a_0 + a_1X + a_2X^2) + \mu(b_0 + b_1X + b_2X^2)) = \\
 &= f((\lambda a_0 + \mu b_0) + (\lambda a_1 + \mu b_1)X + (\lambda a_2 + \mu b_2)X^2) = \\
 &= \begin{pmatrix} (\lambda a_0 + \mu b_0) + 2(\lambda a_1 + \mu b_1) + (\lambda a_2 + \mu b_2) \\ (\lambda a_0 + \mu b_0) - (\lambda a_1 + \mu b_1) \\ (\lambda a_0 + \mu b_0) + (\lambda a_1 + \mu b_1) + (\lambda a_2 + \mu b_2) \end{pmatrix} = \\
 &= \begin{pmatrix} \lambda a_0 + 2\lambda a_1 + \lambda a_2 \\ \lambda a_1 - \lambda a_2 \\ \lambda a_0 + \lambda a_1 + \lambda a_2 \end{pmatrix} + \begin{pmatrix} \mu b_0 + 2\mu b_1 + \mu b_2 \\ \mu b_1 - \mu b_2 \\ \mu b_0 + \mu b_1 + \mu b_2 \end{pmatrix} = \\
 &= \lambda \begin{pmatrix} a_0 + 2a_1 + a_2 \\ a_1 - a_2 \\ a_0 + a_1 + a_2 \end{pmatrix} + \mu \begin{pmatrix} b_0 + 2b_1 + b_2 \\ b_1 - b_2 \\ b_0 + b_1 + b_2 \end{pmatrix} = \\
 &= \lambda P_a + \mu P_b.
 \end{aligned}$$

and therefore f is linear.

- ii) Second method. One can write:

$$f(a_0 + a_1X + a_2X^2) = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & -3 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}.$$

Thus f can be seen as multiplication of a matrix by a column vector, and this multiplication is linear according to Example 6.1.4; 1).

- d) The matrix $M_f^{\mathcal{B}, \mathcal{C}}$ is given by:

$$M_f^{\mathcal{B}, \mathcal{C}} = ((f(\mathbf{c}_1))_{\mathcal{B}}, f((\mathbf{c}_2))_{\mathcal{B}}, f((\mathbf{c}_3))_{\mathcal{B}}),$$

that is, by the matrix whose j -th column, for $1 \leq j \leq 3$, is the column of the components of the vector $f(\mathbf{c}_j)$ with respect to the basis $\mathcal{B}$. Write

$$f(\mathbf{c}_1) = a_{11}\mathbf{b}_1 + a_{21}\mathbf{b}_2 + a_{31}\mathbf{b}_3$$

that is:

$$f(\mathbf{c}_1) = f(1 + X) = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} = a_{11} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + a_{21} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + a_{31} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

which gives, by equating components:

$$\begin{cases} a_{11} + a_{21} &= 3 \\ a_{11} + a_{31} &= 1 \\ a_{21} + a_{31} &= 2 \end{cases}$$

Solving the system with one of the known ad hoc methods (substitution, linear combination, etc.), one finds:

$$a_{11} = (3 + 1 - 2)/2 = 1; \quad a_{21} = (3 - 1 + 2)/2 = 2; \quad a_{31} = (-3 + 2 + 1)/2 = 0;$$

Similarly,

$$f(\mathbf{c}_2) = f(1 - X) = \begin{pmatrix} 1 - 2 \\ -1 \\ 1 - 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = a_{12} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + a_{22} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + a_{32} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

from which

$$\begin{cases} a_{12} + a_{22} &= -1 \\ a_{12} + a_{32} &= -1 \\ a_{22} + a_{32} &= 0 \end{cases}$$

whose solutions are:

$$a_{12} = (-1 - 1 + 0)/2 = -1; \quad a_{22} = (-1 + 1)/2 = 0; \quad a_{32} = (1 - 1)/2 = 0.$$

Finally:

$$f(\mathbf{c}_3) = f(X^2) = \begin{pmatrix} 0 + 2 \cdot 0 + 1 \\ 0 - 1 \\ 0 + 0 + 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = a_{13} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + a_{23} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + a_{33} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

whence the system:

$$\begin{cases} a_{13} + a_{23} &= 1 \\ a_{13} + a_{33} &= -1 \\ a_{23} + a_{33} &= 0 \end{cases}$$

whose solutions are:

$$a_{13} = -1/2; \quad a_{23} = 3/2, \quad a_{33} = -1/2.$$

In conclusion:

$$M_f^{\mathcal{B}, \mathcal{C}} = ((f(\mathbf{c}_1))_{\mathcal{B}}, (f(\mathbf{c}_2))_{\mathcal{B}}, (f(\mathbf{c}_3))_{\mathcal{B}}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 2 & 0 & \frac{3}{2} \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

e) One has:

$$M_f^{\mathcal{B}, \mathcal{C}} \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 2 & 0 & \frac{3}{2} \\ 0 & 0 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{17}{2} \\ -\frac{3}{2} \end{pmatrix}$$

The result is interpreted as follows: if $(2, 1, 3)^T$ are the components of a vector $P \in \mathbb{K}_{\leq 2}[X]$ in the basis $\mathcal{C}$, then $(-1/2, 17/2, -3/2)^T$ are the components of $f(P)$ in the basis $\mathcal{B}$ of $\mathbb{R}^3$. Let us check it. The vector (= polynomial) P would be:

$$P := 2(1 + X) + 1(1 - X) + 3X^2 = 3 + X + 3X^2;$$

whereas

$$f(P) = -\mathbf{b}_1/2 + 17\mathbf{b}_2/2 - 3\mathbf{b}_3/2 = \begin{pmatrix} -1/2 \\ -1/2 \\ 0 \end{pmatrix} + \begin{pmatrix} 17/2 \\ 0 \\ 17/2 \end{pmatrix} - \begin{pmatrix} 0 \\ 3/2 \\ 3/2 \end{pmatrix} = \begin{pmatrix} 8 \\ -2 \\ 7 \end{pmatrix}.$$

Verification. Indeed:

$$f(P) = f(3 + X + 3X^2) = \begin{pmatrix} 3 + 2 \cdot 1 + 3 \\ 1 - 3 \\ 3 + 1 + 3 \end{pmatrix} = \begin{pmatrix} 8 \\ -2 \\ 7 \end{pmatrix},$$

as expected.

6.6.5 Let $A \in \mathbb{K}^{m \times n}$ be a matrix. Once it is regarded as the linear map $\mathbb{K}^n \rightarrow \mathbb{K}^m$ that sends each $\mathbf{u} \in \mathbb{K}^n$ to $A\mathbf{u} \in \mathbb{K}^m$, one has:

$$M_A^{E_n, E_m} := A$$

where E_n and E_m denote the canonical bases of $\mathbb{K}^n$ and $\mathbb{K}^m$, respectively. In other words, the matrix associated with the linear map $\mathbf{u} \mapsto A\mathbf{u}$ with respect to the canonical bases is the matrix A itself. Since composition of functions is associative, this observation provides another simple proof that matrix multiplication is associative.

6.6.6 Proposition. *Let $\mathcal{B}$ and $\mathcal{C}$ be two bases of U and V . The function*

$$\begin{cases} \psi_{\mathcal{C}, \mathcal{B}} : \text{Hom}_{\mathbb{K}}(U, V) & \longrightarrow & \mathbb{K}^{m \times n}, \\ f & \longmapsto & M_f^{\mathcal{C}, \mathcal{B}}. \end{cases}$$

is an isomorphism.

Proof. First observe that if $f, g \in \text{Hom}_{\mathbb{K}}(U, V)$, then $M_f^{\mathcal{C}, \mathcal{B}} = M_g^{\mathcal{C}, \mathcal{B}}$ if and only if $f(\mathbf{b}_j)_{\mathcal{C}} = g(\mathbf{b}_j)_{\mathcal{C}}$ for every $1 \leq j \leq n$, that is, if and only if $f(\mathbf{b}_j) = g(\mathbf{b}_j)$ for every $1 \leq j \leq n$, since the function $\mathbf{v} \mapsto (\mathbf{v})_{\mathcal{C}}$ is an isomorphism. Thus $\psi_{\mathcal{C}, \mathcal{B}}$ is injective. Let now $A \in \mathbb{K}^{m \times n}$. Define $f_A : U \rightarrow V$ by the equality:

$$(f_A(\mathbf{u}))_{\mathcal{C}} = A \cdot (\mathbf{u})_{\mathcal{B}}.$$

Then

$$M_{f_A}^{\mathcal{C}, \mathcal{B}} = A$$

and this proves that $\psi_{\mathcal{C}, \mathcal{B}}$ is surjective. Finally $\psi_{\mathcal{C}, \mathcal{B}}$ is a homomorphism by Proposition 6.6.2. ■

6.6.7 Corollary: *The dimension of $\text{Hom}_{\mathbb{K}}(U, V)$ is equal to the product of the dimensions of U and V :*

$$\dim_{\mathbb{K}} \text{Hom}_{\mathbb{K}}(U, V) = \dim_{\mathbb{K}} U \cdot \dim_{\mathbb{K}} V.$$

Proof. Indeed, if $\dim_{\mathbb{K}} U = n$ and $\dim_{\mathbb{K}} V = m$, the space $\text{Hom}_{\mathbb{K}}(U, V)$ is isomorphic to $\mathbb{K}^{m \times n}$ by 6.6.6, which has dimension mn . ■

Let now $f \in \text{Hom}_{\mathbb{K}}(U, V)$, let $\mathcal{B}_1, \mathcal{B}_2$ be two bases of U , and let $\mathcal{C}_1, \mathcal{C}_2$ be two bases of V .

6.6.8 Proposition. *The following equality holds:*

$$M_f^{\mathcal{C}_2, \mathcal{B}_2} = P^{\mathcal{C}_2, \mathcal{C}_1} \cdot M_f^{\mathcal{C}_1, \mathcal{B}_1} \cdot P^{\mathcal{B}_1, \mathcal{B}_2}. \quad (6.32)$$

Proof. This is pure routine. Let $\mathcal{B}_2 := (\mathbf{b}_1, \dots, \mathbf{b}_n)$. Then:

$$P^{\mathcal{C}_2, \mathcal{C}_1} \cdot M_f^{\mathcal{C}_1, \mathcal{B}_1} \cdot P^{\mathcal{B}_1, \mathcal{B}_2} = P^{\mathcal{C}_2, \mathcal{C}_1} \cdot M_f^{\mathcal{C}_1, \mathcal{B}_1} \cdot ((\mathbf{b}_1)_{\mathcal{B}_1}, \dots, (\mathbf{b}_n)_{\mathcal{B}_1}) = \quad (6.33)$$

$$= P^{\mathcal{C}_2, \mathcal{C}_1} \cdot (M_f^{\mathcal{C}_1, \mathcal{B}_1} \cdot (\mathbf{b}_1)_{\mathcal{B}_1}, \dots, M_f^{\mathcal{C}_1, \mathcal{B}_1} \cdot (\mathbf{b}_n)_{\mathcal{B}_1}) = \quad (6.34)$$

$$= P^{\mathcal{C}_2, \mathcal{C}_1} ((f(\mathbf{b}_1))_{\mathcal{C}_1}, \dots, (f(\mathbf{b}_n))_{\mathcal{C}_1}) = \quad (6.35)$$

$$= (P^{\mathcal{C}_2, \mathcal{C}_1} (f(\mathbf{b}_1))_{\mathcal{C}_1}, \dots, P^{\mathcal{C}_2, \mathcal{C}_1} (f(\mathbf{b}_n))_{\mathcal{C}_1}) = \quad (6.36)$$

$$= ((f(\mathbf{b}_1))_{\mathcal{C}_2}, \dots, (f(\mathbf{b}_n))_{\mathcal{C}_2}) = M_f^{\mathcal{C}_2, \mathcal{B}_2}, \quad (6.37)$$

where in (6.33) we used the definition (6.15) of transition matrix, in (6.34) we wrote the product of the matrix $M_f^{\mathcal{C}_1, \mathcal{B}_1}$ with $P^{\mathcal{C}_2, \mathcal{C}_1}$ using formula (??), in (6.35) the definition (6.21) of the matrix associated with f (with respect to the given bases), in (6.36) again formula (??) for expressing the product of matrices (6.35) and, finally, in (6.36), the definition of the matrix associated with f with respect to the bases $\mathcal{B}_2$ and $\mathcal{C}_2$ of U and V respectively. ■

6.7 Matrices associated with endomorphisms.

6.7.1 When dealing with endomorphisms of a fixed vector space V , it is sufficient and meaningful to use only one basis. In other words, if $\mathcal{B} := (\mathbf{b}_1, \dots, \mathbf{b}_n)$ is a basis of V , we will be interested in the properties of $M_f^{\mathcal{B}, \mathcal{B}} \in \mathbb{K}^{n \times n}$, the matrix associated with an endomorphism f with respect to the same basis $\mathcal{B}$ in the domain and codomain.

6.7.2 Proposition.

i) Let $\text{id}_V : V \rightarrow V$ be the identity endomorphism. Then

$$M_{\text{id}_V}^{\mathcal{B}, \mathcal{B}} = \mathbf{1}_n := (\mathbf{e}_1, \dots, \mathbf{e}_n).$$

ii) If $f \in \text{Aut}(V)$ then $M_f^{\mathcal{B}, \mathcal{B}} \in \text{Gl}_n(\mathbb{K})$ ⁴

iii) If $\mathcal{C}$ is another basis of V , then the matrices $M_f^{\mathcal{B}, \mathcal{B}}$ and $M_f^{\mathcal{C}, \mathcal{C}}$ associated with f are similar, that is, there exists $P \in \text{Gl}_n(\mathbb{K})$ such that

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

⁴That is: the matrix associated with an automorphism is invertible.

Proof. To prove i) it is enough to use the definition of associated matrix based on equality (6.21):

$$\begin{aligned} M_{\text{id}_V}^{\mathcal{B},\mathcal{B}} &= ((\text{id}_V(\mathbf{b}_1))_{\mathcal{B}}, \dots, (\text{id}_V(\mathbf{b}_n))_{\mathcal{B}}) = \\ &= ((\mathbf{b}_1)_{\mathcal{B}}, \dots, (\mathbf{b}_n)_{\mathcal{B}}) = (\mathbf{e}_1, \dots, \mathbf{e}_n). \end{aligned}$$

As for ii), it is a consequence of Proposition 6.6.3 and of the fact that $f \circ f^{-1} = \text{id}_V$:

$$M_{f \circ f^{-1}}^{\mathcal{B},\mathcal{B}} = M_f^{\mathcal{B},\mathcal{B}} \cdot M_{f^{-1}}^{\mathcal{B},\mathcal{B}} = M_{\text{id}_V}^{\mathcal{B},\mathcal{B}} = \mathbf{1}_n$$

having used i) to write the last equality. Finally iii) is equally obvious. Indeed

$$M_f^{\mathcal{C},\mathcal{C}} = P^{\mathcal{C},\mathcal{B}} \cdot M_f^{\mathcal{B},\mathcal{B}} \cdot P^{\mathcal{B},\mathcal{C}}$$

and it is therefore enough to take $P := P^{\mathcal{B},\mathcal{C}}$ (indeed, by Corollary 6.5.5, $P^{-1} = P^{\mathcal{C},\mathcal{B}}$).

■

6.7.3 Esempio. Let $z \in \mathbb{C}$ and consider the function $z \cdot : \mathbb{C} \rightarrow \mathbb{C}$ which assigns to each $w \in \mathbb{C}$ the product $z \cdot w$. Clearly it is a linear map. Its matrix with respect to the canonical basis⁵ is simply (z) . However, as seen again in the example, $\mathbb{C}$ is also a vector space of dimension 2 over the reals, with basis $\mathcal{R} := \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$:

$$\mathbb{C} = \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \oplus \mathbb{R} \begin{pmatrix} 0 \\ i \end{pmatrix}$$

Let us find the matrix $M_z^{\mathcal{R}}$, associated with multiplication by z . Let $z = a + bi$ (thus written as a vector with real coefficients). We must find the components of $z \cdot 1$ and of $z \cdot i$ with respect to the basis $\mathcal{R}$, and one easily has:

$$M_z^{\mathcal{R}} = ((a + bi)_{\mathcal{R}}, (ai - b)_{\mathcal{R}}) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

that is, the matrix recalled in Example ??.

6.8 Linear systems.

6.8.1 Let $A \in \mathbb{K}^{m \times n}$. Then A can be viewed as a linear map

$$A : \mathbb{K}^n \longrightarrow \mathbb{K}^m.$$

⁵which one?

Since $A = (A\mathbf{e}_1, \dots, A\mathbf{e}_n)$, and since $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ is a basis of $\mathbb{K}^n$, it follows that $\text{Im}(A) = [A \cdot \mathbf{e}_1, \dots, A \cdot \mathbf{e}_n]$, that is, the columns of A generate the image. Let now $\mathbf{w} \in \mathbb{K}^m$ and let

$$A^{-1}\mathbf{w} := \{\mathbf{v} \in \mathbb{K}^n \mid A\mathbf{v} = \mathbf{w}\}.$$

The set $A^{-1}(\mathbf{w})$ is called the set of preimages of $\mathbf{w}$ (not to be confused with the inverse of the matrix A , which may not exist, applied to the vector $\mathbf{w}$). Note that this set may be empty, in the case where $\mathbf{w} \notin \text{Im}(A)$. Observe that if $\mathbf{w} \neq \mathbf{0}$, then $A^{-1}(\mathbf{w})$ is not a vector subspace of $\mathbb{K}^n$. Indeed, if it were, $\mathbf{0}_{\mathbb{K}^n}$ should be one of its elements, but this is impossible, since $A\mathbf{0}_{\mathbb{K}^n} = \mathbf{0}_{\mathbb{K}^m} \neq \mathbf{w}$. Conversely, $A^{-1}(\mathbf{0})$ is a vector subspace, traditionally called the *kernel* of the matrix A . More details follow.

6.8.2 The problem we want to study is that of determining the preimages of a given vector $\mathbf{b}$ of $\mathbb{K}^m$ through the linear map $A : \mathbb{K}^n \rightarrow \mathbb{K}^m$ defined by a matrix $A \in \mathbb{K}^{m \times n}$. This is equivalent to solving the equation

$$A \cdot \mathbf{x} = \mathbf{b} \tag{6.38}$$

with matrix coefficients and vector unknowns. If one sets:

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

then finding the solution of equation (6.38) is equivalent to solving the system of m linear equations in n unknowns:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots + \vdots + \ddots + \vdots = \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \tag{6.39}$$

which can be written in *explicit matrix form* as:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{6.40}$$

or, in explicit vector form:

$$x_1 C_1(A) + x_2 C_2(A) + \dots + x_n C_n(A) = \mathbf{b} \quad (6.41)$$

or, more explicitly:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

6.8.3 Proposition. *The linear system $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b} \in \text{Im}(A)$.*

Proof. The proof is obvious. If $\mathbf{b} \in \text{Im}(A)$, by definition of image there exists at least one $\mathbf{x}_0 \in \mathbb{K}^n$ such that $A\mathbf{x}_0 = \mathbf{b}$, that is, there exists at least one solution of the linear system $A\mathbf{x} = \mathbf{b}$. Conversely, if the linear system $A\mathbf{x} = \mathbf{b}$ has a solution, then, by definition of [image](#), $\mathbf{b} \in \text{Im}(A)$. ■

Let $\ker(A) := \{\mathbf{x} \in \mathbb{K}^n \mid A\mathbf{x} = \mathbf{0}_{\mathbb{K}^m}\}$. In other words, $\ker(A)$ is the set of solutions of the *homogeneous linear system* $A\mathbf{x} = \mathbf{0}_{\mathbb{K}^m}$. Suppose that $\mathbf{b} \in \text{Im}(A)$. Then $A^{-1}\mathbf{b} \neq \emptyset$. Thus $A^{-1}(\mathbf{b})$ is the set of solutions of the system $A\mathbf{x} = \mathbf{b}$. If $\mathbf{x}_0 \in A^{-1}(\mathbf{b})$, that is, if $A\mathbf{x}_0 = \mathbf{b}$, then $\mathbf{x}_0$ will also be called a *particular solution* of equation (6.38).

6.8.4 Definition. *If $A^{-1}(\mathbf{b}) \neq \emptyset$, the dimension of the set of solutions is, by definition, equal to:*

$$\dim A^{-1}(\mathbf{b}) = \dim \ker(A).$$

If $\dim A^{-1}(\mathbf{b}) = n - r$, one also says that the system has ∞^{n-r} solutions.

6.8.5 Proposition. *Suppose that $A^{-1}(\mathbf{b}) \neq \emptyset$ and let $\mathbf{x}_0 \in A^{-1}(\mathbf{b})$. Then:*

$$A^{-1}(\mathbf{b}) := \mathbf{x}_0 + \ker(A) := \{\mathbf{x}_0 + \mathbf{u} \mid \mathbf{u} \in \ker(A)\}. \quad (6.42)$$

Proof. The proof proceeds by double inclusion. First, if $\mathbf{x} \in \mathbf{x}_0 + \ker(A)$, then $\mathbf{x} = \mathbf{x}_0 + \mathbf{u}$ for some $\mathbf{u} \in \ker(A)$. Hence

$$A\mathbf{x} = A(\mathbf{x}_0 + \mathbf{u}) = A\mathbf{x}_0 + A\mathbf{u} = \mathbf{b} + \mathbf{0}_{\mathbb{K}^m}$$

because $\mathbf{u} \in \ker(A)$. This proves that $\mathbf{x}_0 + \ker(A) \subseteq A^{-1}(\mathbf{b})$. Conversely, suppose that $\mathbf{x} \in A^{-1}(\mathbf{b})$. Then $\mathbf{b} = A\mathbf{x} = A\mathbf{x}_0$, that is $A\mathbf{x} - A\mathbf{x}_0 = A(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}_{\mathbb{K}^m}$, hence $\mathbf{u} := \mathbf{x} - \mathbf{x}_0 \in \ker(A)$. This proves that $A^{-1}(\mathbf{b}) \subseteq \mathbf{x}_0 + \ker(A)$, that is, equality (6.42). ■

An element of $\ker(A)$ is also called a *solution of the associated homogeneous system* for A . Thus Proposition 6.8.5 says that *the general solution of the system $A\mathbf{x} = \mathbf{b}$* is the sum of a particular solution and the general solution of the associated homogeneous system.

6.8.6 Remark. Let $C^\infty(\mathbb{R})$ be the vector space of infinitely differentiable real functions of a real variable. Let $D := d/dx$ be the derivative with respect to x . From analysis courses one knows that if $f, g \in C^\infty(\mathbb{R})$ and $\lambda, \mu \in \mathbb{R}$, then

$$D(\lambda \cdot f + \mu \cdot g) = \lambda \cdot Df + \mu \cdot Dg.$$

Thus D is an endomorphism of $C^\infty(\mathbb{R})$. Since composition of endomorphisms is an endomorphism and the sum of endomorphisms is an endomorphism, every expression of the form

$$P(D) := D^n + a_1 D^{n-1} + \dots + a_n \mathbf{1}$$

is an endomorphism of $C^\infty(\mathbb{R})$ called a “differential operator of order n ”. Solving the equation

$$P(D)f = g$$

means solving the associated differential equation:

$$f^{(n)} + a_1 f^{(n-1)} + \dots + a_n f = g.$$

From the theory of differential equations with constant coefficients (see also the fantastic [?]) one knows that the general solution of the differential equation is given by a particular solution added to the general solution of the associated homogeneous equation. In other words, if $p = p(x)$ is a particular solution, the set of solutions is:

$$P(D)^{-1}(p) = \ker P(D) + p.$$

6.8.7 Let now $A\mathbf{x} = \mathbf{b}$ be a linear system. The matrix:

$$(A|\mathbf{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right)$$

is called the *augmented matrix of the system* (the vertical bar is inserted purely for convenience, also to distinguish visually the coefficient matrix from the column of known terms).

6.8.8 Theorem (Rouché–Capelli) *The linear system (6.38) has solutions if and only if the rank of the matrix A coincides with the rank of the augmented matrix $(A|\mathbf{b})$. In this case the dimension of the solution space is $\dim A^{-1}(\mathbf{b}) = \dim \ker(A) = n - r$, where r is the rank of A .*

Proof. Indeed, if the system is solvable, then $\mathbf{b} \in \text{Im}(A)$ and therefore $\mathbf{b}$ is a linear combination of the columns of A . Thus, if $[C_1(A), \dots, C_n(A)] = [C_1(A), \dots, C_n(A), \mathbf{b}]$, then

$$\text{rk}(A) = \dim_{\mathbb{K}}[C_1(A), \dots, C_n(A)] = \dim_{\mathbb{K}}[C_1(A), \dots, C_n(A), \mathbf{b}] = \text{rk}(A|\mathbf{b}).$$

Conversely, if $\text{rk}(A) = \text{rk}(A|\mathbf{b})$, then $\mathbf{b}$ is a linear combination of the columns of A , which generate the image, and therefore $\mathbf{b}$ lies in the image, that is, the system has solutions. ■

6.8.9 Remark. Note that for the homogeneous linear system $A\mathbf{x} = \mathbf{0}$, which always has solutions (at least the zero solution), the Rouché–Capelli condition is trivially satisfied.



6.8.10 If $A\mathbf{x} = \mathbf{b}$ is a Cramer system, that is, if $A \in \mathbb{K}^{n \times n}$ (A is a square matrix), the solution certainly exists if A is invertible. Indeed, if A is invertible, the columns of A form a basis of $\mathbb{K}^n$ and hence every vector $\mathbf{b}$ is a linear combination of those vectors. The condition of non-invertibility of the matrix is $\det(A) = 0$. If $\det(A) \neq 0$, then the system has $\infty^{n-n} = \infty^0$ solutions. The solution is unique. Indeed, if $\mathbf{x}$ and $\mathbf{x}'$ are solutions, then $\mathbf{x} - \mathbf{x}' \in \ker(A)$. Since $\ker(A) = \mathbf{0}$, it follows that $\mathbf{x} = \mathbf{x}'$. ■

6.8.11 Solved Exercise. Study, as the parameters $h, k \in \mathbb{R}$ vary, the solutions of the linear system

$$\begin{cases} x_1 - x_2 + x_3 + 3x_4 = 1 \\ x_1 + kx_2 + \quad \quad kx_4 = k \\ \quad \quad x_2 - x_3 + hx_4 = 0 \end{cases}$$

Solution. The augmented matrix of the system is:

$$(A(h, k)|\mathbf{b}(k)) = \left(\begin{array}{cccc|c} 1 & -1 & 1 & 3 & 1 \\ 1 & k & 0 & k & k \\ 0 & 1 & -1 & h & 0 \end{array} \right)$$

The notation is used to recall the explicit dependence of the coefficient matrix and of the constant column on the parameters h and k . We may now begin a row reduction, as learned in ??, by proposing this first elementary transformation:

$R_3 \rightarrow R_3 + R_1$, obtaining:

$$(A_1(h, k)|\mathbf{b}_1(k)) := \left(\begin{array}{cccc|c} 1 & -1 & 1 & 3 & 1 \\ 1 & k & 0 & k & k \\ 1 & 0 & 0 & h+3 & 1 \end{array} \right)$$

One immediately sees that if $k \neq 0$, for every $h \in \mathbb{R}$, the matrix $A_1(h, k)$ is reduced and therefore $\text{rk}(A) = \text{rk}A_1(h, k) = \text{rk}(A_1(h, k)|\mathbf{b}_1(k)) = 3$. By the Rouché–Capelli theorem, therefore, the system is compatible and has $\infty^{4-3} = \infty^1$ solutions, a way of saying that there is a one-parameter family of solutions:

$$\begin{cases} x_1 = 1 - (3 + h)t \\ x_2 = \frac{1}{k}(k - kt - 1 + 3t + ht) \\ x_3 = ht + \frac{1}{k}(k - kt - 1 + 3t + ht) \\ x_4 = t \end{cases}$$

If $k = 0$, the matrix $(A_1(h, 0), \mathbf{b}_1(0))$ becomes:

$$(A_1(h, 0)|\mathbf{b}_1(0)) = \left(\begin{array}{cccc|c} 1 & -1 & 1 & 3 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & h+3 & 1 \end{array} \right)$$

Performing the further elementary row transformation:

$$R_3(A_1(h, 0), \mathbf{b}_1(0)) \rightarrow R_3(A_1(h, 0), \mathbf{b}_1(0)) - R_2(A_1(h, 0), \mathbf{b}_1(0))$$

one obtains:

$$(A_2(h, 0), \mathbf{b}_2(0)) = \left(\begin{array}{cccc|c} 1 & -1 & 1 & 3 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h+3 & 1 \end{array} \right)$$

For $h \neq -3$, the matrix $A_2(h, 0)$ is reduced (up to exchanging the third and second rows); $\text{rk}(A) = 3$ and $\text{rk}(A|B) = 3$.

Therefore, by the Rouché–Capelli theorem, the system is compatible and has $\infty^{4-3} = \infty^1$ solutions depending on one parameter:

$$\begin{cases} x_1 = 0 \\ x_2 = t \\ x_3 = 1 + t - \frac{3}{h+3} \\ x_4 = \frac{1}{h+3} \end{cases}$$

If $k = 0 \wedge h = -3$, the augmented matrix becomes:

$$(A_2(0, -3), \mathbf{b}_2(0)) = \left(\begin{array}{cccc|c} 1 & -1 & 1 & 3 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

The matrix is reduced; $\text{rk}(A_2(0, -3)) = 2$ and $\text{rk}(A_2(0, -3)|\mathbf{b}_2(0)) = 3$. Therefore, by the Rouché–Capelli theorem, the system has no solutions.

6.8.12 To readers who might wish to complain, arguing that a single example is not enough, we reply: these exercises are all the same anyway... And in any case, more follow right below.

6.9 Exercises

6.9.1 The reader, male or female, should decompose each of the solutions of Exercise 6.8.11 as the sum of a particular solution and a generic vector of the kernel of the matrix $A(h, k)$.

Soluzione: [?, 9.5.1]

6.9.2 Prove that, in general, $\text{Aut}(V)$ is not commutative. That is, explicitly produce a vector space V and two isomorphisms from V to itself such that $f \circ g \neq g \circ f$.

Soluzione: [?, 8.5.1]

6.9.3 Consider the *symmetric* matrix $B := \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Compute the characteristic polynomial and prove that its roots are necessarily real; Could you compute the eigenvectors of such a matrix? Experiments with numerical examples are allowed.

Soluzione: [?, 8.5.2]

6.9.4 Let an arbitrary antisymmetric 3×3 matrix be given:

$$A := \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -c & -b & 0 \end{pmatrix}$$

1. Is it invertible?
2. Prove that every 3×3 matrix can be written uniquely as the sum of a symmetric matrix and an antisymmetric matrix.
3. What is, if it exists, the unique matrix that is both symmetric and antisymmetric?

Soluzione: [?, 8.5.3]

6.9.5 Let the matrix be given:

$$A := \begin{pmatrix} 1 & 2 & 5 \\ 2 & -1 & 0 \\ 5 & 10 & -5 \end{pmatrix}$$

1. Compute $A \cdot A^T$ and $A^T \cdot A$: which of the two products gives a diagonal matrix?
2. Let B be the matrix

$$B := \left(\frac{C_1(A)}{|C_1(A)|}, \frac{C_2(A)}{|C_2(A)|}, \frac{C_3(A)}{|C_3(A)|} \right)$$

Compute $B \cdot B^T$ and $B^T \cdot B$ with **MATLAB**[®] (if you wish). Recall that in **MATLAB**[®], the transpose is indicated by the apostrophe $A^T \rightarrow A'$.

Soluzione: [?, 8.5.4]

6.9.6 Let two endomorphisms of $\mathbb{R}^3$ and $(\mathbb{R}^3)^\vee$ be given by

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2x + y + z \\ y + z \\ 8y - z \end{pmatrix} \quad \text{and} \quad g(a, b, c) = (-2a + b + c, b + c, 8b - c).$$

1. Determine the matrix A of f with respect to the canonical basis of $\mathbb{R}^3$;
2. Determine the matrix B of g with respect to the canonical basis of $(\mathbb{R}^3)^\vee$;
3. Verify that f is an automorphism (= invertible endomorphism);
4. Determine all eigenvalues of f and of g ; (hint for g : look for vectors $\alpha \in (\mathbb{R}^3)^\vee$ such that $g(\alpha)$ is a multiple of α)

5. Determine a basis for each eigenspace of f and one for each eigenspace of g ;
6. Determine an invertible matrix $P \in \mathbb{R}^{3 \times 3}$ such that $P^{-1}AP$ is diagonal.

Soluzione: [?, 9.5.5]

6.9.7 Let the linear map $f : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ be defined by

$$f \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x + z - t \\ y + z \\ x - y - t \end{pmatrix}$$

1. Find the matrix M_f of f with respect to the canonical bases of $\mathbb{R}^4$ and $\mathbb{R}^3$, respectively.
2. Determine a basis of the kernel of f ;
3. Prove that f is not surjective and determine a vector of $\mathbb{R}^3$ with no preimage.
4. Given the vector $v = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^3$, determine the affine space $f^{-1}(v)$ of its preimages.

6.9.8 Consider the linear map $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that

$$f \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1 \quad \text{and} \quad \ker(f) = \left[\begin{pmatrix} 1 \\ 2 \end{pmatrix} \right];$$

1. Find the matrix of f with respect to the canonical bases;
2. Say, justifying the answer, whether f is surjective;
3. Determine the preimage $f^{-1}(2)$ of 2;
4. Compute $f(2\mathbf{e}_1 + 2\mathbf{e}_2)$.

Soluzione: [?, 9.5.7]

6.9.9 Let $f \in \text{End}_{\mathbb{R}}(\mathbb{R}^3)$ be given by:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y + 3z \\ x + 2y + 5z \\ -x - 3y - 7z \end{pmatrix}$$

1. Determine f^2 (the composition of f with itself);
2. Find $\ker(f^2)$. What is its dimension? Determine a basis for it;
3. Determine a complementary subspace to $\ker(f^2)$ in $\mathbb{R}^3$, that is, a subspace W such that $\mathbb{R}^3 = \ker(f^2) \oplus W$;
4. Determine, for every value of a , the vectors $\mathbf{u} \in \mathbb{R}^3$ such that

$$f^2(\mathbf{u}) = \begin{pmatrix} 1 \\ 1 \\ a \end{pmatrix}.$$

Soluzione: [?, 8.5.6]

6.9.10 An endomorphism f of $\mathbb{R}^3$ is given such that

$$f \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = f \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

1. Prove that the vector

$$\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

belongs to the kernel and that $(0, 3, 5)^T$ belongs to the image.

2. Prove that f is not surjective;
3. Prove that the image has dimension at most 2;
4. How many linearly independent vectors of the kernel can be found as the dimension of the image varies?

Soluzione: [?, 9.5.2]

6.9.11 Consider the matrix

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

The *trace* of A is the sum of the elements on the diagonal, that is, $\text{tr}(A) := a + d$. The polynomial

$$p_A(t) = \det(t\mathbf{1} - A)$$

is called the *characteristic polynomial*.

1. Prove that

$$p_A(t) = t^2 - \operatorname{tr}(A)t + \det(A);$$

2. Using the previous point, compute $p_A(t)$ for all the matrices listed below:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 3 & -1 \\ 1 & 4 \end{pmatrix} \quad \begin{pmatrix} 9 & -1 \\ 1 & 8 \end{pmatrix} \quad \begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix}$$

Soluzione: [?, 9.5.3]

6.9.12 Let the matrix be given:

$$A := \begin{pmatrix} 1 & 2 & 5 \\ 2 & -1 & 0 \\ 5 & 10 & -5 \end{pmatrix}$$

1. Compute $A \cdot A^T$ and $A^T \cdot A$: which of the two products gives a diagonal matrix?
2. Let B be the matrix

$$B := \left(\frac{C_1(A)}{|C_1(A)|}, \frac{C_2(A)}{|C_2(A)|}, \frac{C_3(A)}{|C_3(A)|} \right)$$

Compute $B \cdot B^T$ and $B^T \cdot B$ with MATLAB © (if you wish).

Soluzione: [?, 9.5.4]

6.9.13 Let two endomorphisms of $\mathbb{R}^3$ and $(\mathbb{R}^3)^\vee$ be given by

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2x + y + c \\ y + z \\ 8y - z \end{pmatrix} \quad \text{and} \quad g(a, b, c) = (-2a + b + c, b + c, 8b - c).$$

1. Determine the matrix A of f with respect to the canonical basis of $\mathbb{R}^3$;
2. Determine the matrix B of g with respect to the canonical basis of $(\mathbb{R}^3)^\vee$;
3. Verify that f is an automorphism (= invertible endomorphism);
4. Determine all eigenvalues of f and of g ; (hint for g : look for vectors $\alpha \in (\mathbb{R}^3)^\vee$ such that $g(\alpha)$ is a multiple of α)

5. Determine a basis for each eigenspace of f and one for each eigenspace of g ;
6. Determine an invertible matrix $P \in \mathbb{R}^{3 \times 3}$ such that $P^{-1}AP$ is diagonal.

Soluzione: [?, 9.5.5]

6.9.14 Let the linear map $f : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ be defined by

$$f \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x + z - t \\ y + z \\ x - y - t \end{pmatrix}$$

1. Find the matrix M_f of f with respect to the canonical bases of $\mathbb{R}^4$ and $\mathbb{R}^3$, respectively.
2. Determine a basis of the kernel of f ;
3. Prove that f is not surjective and determine a vector of $\mathbb{R}^3$ with no preimage.
4. Given the vector $v = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^3$, determine the affine space $f^{-1}(v)$ of its preimages.

Soluzione: [?, 9.5.6]

6.9.15 The following exercise aims to make readers aware of the difference between writing a vector as a row or writing it as a column. Let the map $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ be given by:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y + 3z \\ x - y + 2z \end{pmatrix}$$

- i) Write the function in the form of the product of a matrix by a vector of unknowns;
- ii) Find $\vec{u} \in \mathbb{R}^3$ such that $f(\vec{u}) = \vec{0}$;
- iii) Find the preimages of the vector $(1, 2)^T \in \mathbb{R}^2$ (T denotes transposition).

Soluzione: [?, 4.4.10]

6.9.16 Let the function $f : (\mathbb{R}^3)^\vee \rightarrow (\mathbb{R}^2)^\vee$ be given by:

$$f(x, y, z) := (x - y - z, 2x - 2y - 2z)$$

1. Write the function f as a product of matrices;
2. Find the preimages of $(0, 0)$;
3. Find the preimages, if they exist, of $(2, 4)$ and of $(2, 5)$.

Soluzione: [?, 4.4.11]

6.9.17 Find an endomorphism all of whose eigenvalues are zero. Find a linear map of zero rank.

Soluzione: [?, 9.5.8]

6.9.18 Recall that if $z := a + bi$ is a complex number, with $i = \sqrt{-1}$, then its *conjugate* is defined by $\bar{z} = a - bi$. Consider then the function $f : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ given by:

$$f \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} z_1 - 2\bar{z}_2 \\ 2z_1 + z_2 \end{pmatrix}$$

- i) Considering $\mathbb{C}^2$ as a vector space over the reals, determine whether f is real-linear;
- ii) What is one of its matrix expressions after identifying $\mathbb{C}^2$ with $\mathbb{R}^4$?
- iii) Is the function complex-linear?
- iv) Is it possible to write a 2×2 matrix with complex coefficients representing f with respect to some basis of $\mathbb{C}^2$?

(hint: is the conjugation function $z \mapsto \bar{z}$ complex-linear? And real-linear?)

Soluzione: [?, 9.5.9]

6.9.19 Reduce the matrix

$$A := \begin{pmatrix} 4 & 6 & 12 & 18 & 16 \\ 1 & 3 & 5 & -1 & 3 \\ 4 & 3 & 0 & 0 & 3 \\ -2 & -1 & 5 & 3 & -2 \end{pmatrix}.$$

both by rows and by columns.

Soluzione: [?, 5.10.9]

6.9.20 Let the matrices be given:

$$A := \begin{pmatrix} 1 & 2 & -1 & 3 \\ -2 & 3 & 1 & 2 \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} -1 & 5 & 0 & 5 \\ -1 & 5 & 0 & 6 \end{pmatrix}.$$

1. Prove that $\mathbb{R}^4 = \text{Im}(A) \oplus \text{Im}(B)$;
2. Determine $\ker(A) \cap \ker(B)$;
3. Establish whether $A : \mathbb{R}^4 \rightarrow \mathbb{R}^2$, defined by $\mathbf{u} \mapsto A\mathbf{u}$, is injective;
4. Establish whether $B : \mathbb{R}^4 \rightarrow \mathbb{R}^2$, defined by $\mathbf{u} \mapsto B\mathbf{u}$, is surjective;
5. Compute the kernel of the matrix

$$\begin{pmatrix} A \\ B \end{pmatrix}$$

whose first two rows are the rows of A and whose second two are those of B ;

6. Generalize the exercise. Prove that if A and B are two matrices with the same number of columns, then the sum $\text{Im}(A) \oplus \text{Im}(B)$ is direct if and only if the rank of the matrix

$$\begin{pmatrix} A \\ B \end{pmatrix}$$

coincides with the sum of the ranks of A and B separately.

Soluzione: [?, 7.4.4]

6.9.21 Find the inverse of the square matrix

$$A := \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$$

by the method of elementary row and column transformations.

Soluzione: [?, 5.10.10]

6.9.22 Let the matrix be given:

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

i) Find the preimages of the vectors

$$C_2(A) = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \quad \text{and} \quad C_3(A) = \begin{pmatrix} -2 \\ 2 \\ 0 \end{pmatrix}$$

ii) Identifying A with a linear map from $\mathbb{R}^4 \rightarrow \mathbb{R}^3$, find the preimages of $C :=$

$$\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix};$$

iii) find at least one vector having no preimage;

iv) Reduce the matrix A by rows;

v) Solve the linear system

$$(M|\vec{b}) = \left(\begin{array}{ccc|c} 1 & 3 & -2 & 1 \\ 2 & 1 & 2 & 3 \\ 3 & 4 & 0 & 4 \end{array} \right)$$

and interpret the result geometrically.

Soluzione: [?, 5.10.13 1]