LINEAR OPERATORS

BRANKO ĆURGUS

Throughout this note $\mathcal{V}$ is a vector space over a scalar field $\mathbb{F}$. $\mathbb{N}$ denotes the set of positive integers and $i, j, k, l, m, n, p \in \mathbb{N}$.

1. Linear operators

In this section $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ are vector spaces over a scalar field $\mathbb{F}$.

1.1. **The definition and the vector space of all linear operators.** A function $T : \mathcal{V} \to \mathcal{W}$ is said to be a *linear operator* if it satisfies the following conditions:

$$\forall\, u \in \mathcal{V} \ \ \forall\, v \in \mathcal{V} \qquad T(u+v) = T(u) + f(v), \tag{1.1} \quad \boxed{\texttt{eq-add}}$$

$$\forall\, \alpha \in \mathbb{F} \ \ \forall\, v \in \mathcal{V} \qquad T(\alpha v) = \alpha\, T(v). \tag{1.2} \quad \boxed{\texttt{eq-hom}}$$

The property (1.1) is called *additivity*, while the property (1.2) is called *homogeneity*. Together additivity and homogeneity are called *linearity*.

Denote by $\mathcal{L}(\mathcal{V}, \mathcal{W})$ the set of all linear operators from $\mathcal{V}$ to $\mathcal{W}$. Define the addition and scaling in $\mathcal{L}(\mathcal{V}, \mathcal{W})$. For $S, T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ and $\alpha \in \mathbb{F}$ we define

$$(S+T)(v) = S(v) + T(v), \qquad \forall\, v \in \mathcal{V}, \tag{1.3} \quad \boxed{\texttt{eq-po+}}$$

$$(\alpha T)(v) = \alpha\, T(v), \qquad \forall\, v \in \mathcal{V}. \tag{1.4} \quad \boxed{\texttt{eq-po-s}}$$

Notice that two plus signs which appear in (1.3) have different meanings. The plus sign on the left-hand side stands for the addition of linear operators that is just being defined, while the plus sign on the right-hand side stands for the addition in $\mathcal{W}$. Notice the analogous difference in empty spaces between α and T in (1.4). Define the zero mapping in $\mathcal{L}(\mathcal{V}, \mathcal{W})$ to be

$$0_{\mathcal{L}(\mathcal{V},\mathcal{W})}(v) = 0_{\mathcal{W}}, \qquad \forall\, v \in \mathcal{V}.$$

For $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ we define its opposite operator by

$$(-T)(v) = -T(v), \qquad \forall\, v \in \mathcal{V}.$$

Proposition 1.1. *The set $\mathcal{L}(\mathcal{V}, \mathcal{W})$ with the operations defined in* (1.3), *and* (1.4) *is a vector space over* $\mathbb{F}$.

For $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ and $v \in \mathcal{V}$ it is customary to write Tv instead of $T(v)$.

Date: February 21, 2023 at 22:37.

Example 1.2. Assume that a vector space $\mathcal{V}$ is a direct sum of its subspaces $\mathcal{U}$ and $\mathcal{W}$, that is $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. Define the function $P : \mathcal{V} \to \mathcal{V}$ by

$$Pv = w \qquad \Leftrightarrow \qquad v = u + w, \quad u \in \mathcal{U}, \quad w \in \mathcal{W}.$$

Then P is a linear operator. It is called the *projection* of $\mathcal{V}$ onto $\mathcal{W}$ parallel to $\mathcal{U}$; it is denoted by $P_{\mathcal{W}\|\mathcal{U}}$.

The definition of the linearity of a function between vector spaces is expressed in the standard functional notation. The next proposition states that a function between vector spaces is linear if and only if its graph is a subspace of the direct product of the domain and the codomain of that function.

pr-lfsub **Proposition 1.3.** *Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over a scalar field $\mathbb{F}$. Let $f : \mathcal{V} \to \mathcal{W}$ be a function and denote by F the graph of f; that is let*

$$\mathcal{F} = \bigl\{ \bigl(v, w\bigr) \in \mathcal{V} \times \mathcal{W} \,:\, v \in \mathcal{V} \text{ and } \ w = f(v) \bigr\} \subseteq \mathcal{V} \times \mathcal{W}.$$

The function f is linear if and only if the set $\mathcal{F}$ is a subspace of the vector space $\mathcal{V} \times \mathcal{W}$.

pr-imsub **Proposition 1.4.** *Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over a scalar field $\mathbb{F}$. Let $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, let $\mathcal{G}$ be a subspace of $\mathcal{V}$ and let $\mathcal{H}$ be a subspace of $\mathcal{W}$. Then*

$$T(\mathcal{G}) = \bigl\{ w \in \mathcal{W} \,:\, \exists v \in \mathcal{G} \text{ such that } \ w = Tv \bigr\}$$

is a subspace of $\mathcal{W}$ and

$$T^{-1}(\mathcal{H}) = \bigl\{ v \in \mathcal{V} \,:\, Tv \in \mathcal{H} \bigr\}$$

is a subspace of $\mathcal{V}$.

1.2. **Composition, inverse, isomorphism.** In the next two propositions we prove that the linearity is preserved under composition of linear operators and under taking the inverse of a linear operator.

Proposition 1.5. *Let $S : \mathcal{U} \to \mathcal{V}$ and $T : \mathcal{V} \to \mathcal{W}$ be linear operators. The composition $T \circ S : \mathcal{U} \to \mathcal{W}$ is a linear operator.*

Proof. Prove this as an exercise. □

When composing linear operators it is customary to write simply TS instead of $T \circ S$.

The identity function on $\mathcal{V}$ is denoted by $I_{\mathcal{V}}$. It is defined by $I_{\mathcal{V}}(v) = v$ for all $v \in \mathcal{V}$. It is clearly a linear operator.

pr-inv-l **Proposition 1.6.** *Let $T : \mathcal{V} \to \mathcal{W}$ be a linear operator which is a bijection. Then the inverse $T^{-1} : \mathcal{W} \to \mathcal{V}$ of T is a linear operator.*

Proof. Since T is a bijection, from what we learned about function, there exists a function $S : \mathcal{W} \to \mathcal{V}$ such that $ST = I_{\mathcal{V}}$ and $TS = I_{\mathcal{W}}$. Since T is linear and $TS = I_{\mathcal{W}}$ we have

$$T\bigl(\alpha Sx + \beta Sy\bigr) = \alpha T(Sx) + \beta T(Sy) = \alpha (TS)x + \beta (TS)y = \alpha x + \beta y$$

for all $\alpha, \beta \in \mathbb{F}$ and all $x, y \in \mathcal{W}$. Applying S to both sides of

$$T\big(\alpha Sx + \beta Sy\big) = \alpha x + \beta y$$

we get

$$(ST)\big(\alpha Sx + \beta Sy\big) = S\big(\alpha x + \beta y\big) \qquad \forall\, \alpha, \beta \in \mathbb{F} \quad \forall x, x \in \mathcal{W}.$$

Since $ST = I_{\mathcal{V}}$, we get

$$\alpha Sx + \beta Sy = S\big(\alpha x + \beta y\big) \qquad \forall\, \alpha, \beta \in \mathbb{F} \quad \forall x, y \in \mathcal{W},$$

thus proving the linearity of S. Since by definition $S = T^{-1}$ the proposition is proved. □

A linear operator $T : \mathcal{V} \to \mathcal{W}$ which is a bijection is called an *isomorphism* between vector spaces $\mathcal{V}$ and $\mathcal{W}$.

By Proposition 1.6 each isomorphism is invertible and its inverse is also an isomorphism.

In the next theorem we introduce the most important isomorphism between a finite-dimensional space $\mathcal{V}$ and a space $\mathbb{F}^n$ where $n = \dim \mathcal{V}$.

th-cor **Theorem 1.7.** *Let $\mathcal{V}$ be a finite dimensional vector space over $\mathbb{F}$, let $n = \dim \mathcal{V}$ and let $\mathcal{B} = \{b_1, \ldots, b_n\}$ be a basis for $\mathcal{V}$. The function $C_{\mathcal{B}} : \mathcal{V} \to \mathbb{F}^n$ defined by: for all $v \in \mathcal{V}$*

$$C_{\mathcal{B}}(v) := \mathbf{a} \quad \textit{where} \quad \mathbf{a} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{F}^n \quad \textit{and} \quad v = \alpha_1 b_1 + \cdots + \alpha_n b_n,$$

is an isomorphism between $\mathcal{V}$ and $\mathbb{F}^n$.

It is important to point out that the formula for the inverse $(C_{\mathcal{B}})^{-1} : \mathbb{F}^n \to \mathcal{V}$ of $C_{\mathcal{B}}$ is given by

$$(C_{\mathcal{B}})^{-1} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \sum_{j=1}^{n} \alpha_j v_j, \quad \text{for all} \quad \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{F}^n. \tag{1.5}$$ eq-CB-i

Notice that (1.5) defines a function from $\mathbb{F}^n$ to $\mathcal{V}$ even if $\mathcal{B}$ is not a basis of $\mathcal{V}$.

exa-LBC **Example 1.8.** Inspired by the definition of $C_{\mathcal{B}}$ and (1.5) we define a general operator of this kind. Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over $\mathbb{F}$. Let $\mathcal{V}$ be finite dimensional, $n = \dim \mathcal{V}$ and let $\mathcal{B}$ be a basis for $\mathcal{V}$. Let $\mathcal{C} = (w_1, \ldots, w_n)$ be any n-tuple of vectors in $\mathcal{W}$. The entries of an n-tuple can be repeated, they can all be equal, for example to $0_{\mathcal{V}}$. We define the linear operator $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \to \mathcal{W}$ by

$$L_{\mathcal{C}}^{\mathcal{B}}(v) = \sum_{j=1}^{n} \alpha_j w_j \qquad \text{where} \qquad \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = C_{\mathcal{B}}(v). \tag{1.6}$$ eq-rCA

In fact, $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \to \mathcal{W}$ is a composition of $C_{\mathcal{B}} : \mathcal{V} \to \mathbb{F}^n$ and the operator $\mathbb{F}^n \to \mathcal{W}$ defined by

$$\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \mapsto \sum_{j=1}^{n} \xi_j w_j \qquad \text{for all} \qquad \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{F}^n. \tag{1.7}$$

eq-rC

It is easy to verify that (1.7) defines a linear operator.

Denote by $\mathcal{E}$ the standard basis of $\mathbb{F}^n$, that is the basis which consists of the columns of the identity matrix I_n. Then $C_{\mathcal{B}} = L_{\mathcal{E}}^{\mathcal{B}}$ and $(C_{\mathcal{B}})^{-1} = L_{\mathcal{B}}^{\mathcal{E}}$.

Exercise 1.9. Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over $\mathbb{F}$. Let $\mathcal{V}$ be finite dimensional, $n = \dim \mathcal{V}$ and let $\mathcal{B}$ be a basis for $\mathcal{V}$. Let $\mathcal{C} = (w_1, \ldots, w_n)$ be a list of vectors in $\mathcal{W}$ with n entries.

(a) Characterize the injectivity of $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \to \mathcal{W}$.
(b) Characterize the surjectivity of $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \to \mathcal{W}$.
(c) Characterize the bijectivity of $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \to \mathcal{W}$.
(d) If $L_{\mathcal{C}}^{\mathcal{B}} : \mathcal{V} \to \mathcal{W}$ is an isomorphism, find a simple formula for $(L_{\mathcal{C}}^{\mathcal{B}})^{-1}$.

1.3. **The nullity-rank theorem.** Let $T : \mathcal{V} \to \mathcal{W}$ is be a linear operator. The linearity of T implies that the set

$$\operatorname{nul} T = \bigl\{ v \in \mathcal{V} : Tv = 0_{\mathcal{W}} \bigr\}$$

is a subspace of $\mathcal{V}$. This subspace is called the *null space* of T. Similarly, the linearity of T implies that the range of T is a subspace of $\mathcal{W}$. Recall that

$$\operatorname{ran} T = \bigl\{ w \in \mathcal{W} : \exists\, v \in \mathcal{V} \ \ w = Tv \bigr\}.$$

Proposition 1.10. *A linear operator $T : \mathcal{V} \to \mathcal{W}$ is an injection if and only if* $\operatorname{nul} T = \{0_{\mathcal{V}}\}$.

Proof. We first prove the "if" part of the proposition. Assume that $\operatorname{nul} T = \{0_{\mathcal{V}}\}$. Let $u, v \in \mathcal{V}$ be arbitrary and assume that $Tu = Tv$. Since T is linear, $Tu = Tv$ implies $T(u-v) = 0_{\mathcal{W}}$. Consequently $u-v \in \operatorname{nul} T = \{0_{\mathcal{V}}\}$. Hence, $u - v = 0_{\mathcal{V}}$, that is $u = v$. This proves that T is an injection.

To prove the "only if" part assume that $T : \mathcal{V} \to \mathcal{W}$ is an injection. Let $v \in \operatorname{nul} T$ be arbitrary. Then $Tv = 0_{\mathcal{W}} = T0_{\mathcal{V}}$. Since T is injective, $Tv = T0_{\mathcal{V}}$ implies $v = 0_{\mathcal{V}}$. Thus we have proved that $\operatorname{nul} T \subseteq \{0_{\mathcal{V}}\}$. Since the converse inclusion is trivial, we have $\operatorname{nul} T = \{0_{\mathcal{V}}\}$. □

Theorem 1.11 (Nullity-Rank Theorem)**.** *Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over a scalar field $\mathbb{F}$ and let $T : \mathcal{V} \to \mathcal{W}$ be a linear operator. If $\mathcal{V}$ is finite dimensional, then* $\operatorname{nul} T$ *and* $\operatorname{ran} T$ *are finite dimensional and*

$$\dim(\operatorname{nul} T) + \dim(\operatorname{ran} T) = \dim \mathcal{V}. \tag{1.8}$$

eq-rnt

Proof. Assume that $\mathcal{V}$ is finite dimensional. We proved earlier that for an arbitrary subspace $\mathcal{U}$ of $\mathcal{V}$ there exists a subspace $\mathcal{X}$ of $\mathcal{V}$ such that

$$\mathcal{U} \oplus \mathcal{X} = \mathcal{V} \qquad \text{and} \qquad \dim \mathcal{U} + \dim \mathcal{X} = \dim \mathcal{V}.$$

Thus, there exists a subspace $\mathcal{X}$ of $\mathcal{V}$ such that

$$(\operatorname{nul} T) \oplus \mathcal{X} = \mathcal{V} \qquad \text{and} \qquad \dim(\operatorname{nul} T) + \dim \mathcal{X} = \dim \mathcal{V}. \tag{1.9}$$

eq-st1-rnt

Since $\dim(\operatorname{nul} T) + \dim \mathcal{X} = \dim \mathcal{V}$, to prove the theorem we only need to prove that $\dim \mathcal{X} = \dim(\operatorname{ran} T)$. To this end, we consider the restriction $T|_{\mathcal{X}} : \mathcal{X} \to \operatorname{ran} T$ of T to the subspace $\mathcal{X}$. This operator is defined by

$$T|_{\mathcal{X}}(v) = Tv \quad \forall v \in \mathcal{X}.$$

We will prove that $T|_{\mathcal{X}}$ is an isomorphism. The first step in this direction is to prove that $T|_{\mathcal{X}}$ is a surjection. That is

$$\operatorname{span}\bigl\{Tx_1, \dots, Tx_m\bigr\} = \operatorname{ran} T. \tag{1.10}$$

eq-span-rnt

Clearly $\bigl\{Tx_1, \dots, Tx_m\bigr\} \subseteq \operatorname{ran} T$. Consequently, since $\operatorname{ran} T$ is a subspace of $\mathcal{W}$, we have $\operatorname{span}\bigl\{Tx_1, \dots, Tx_m\bigr\} \subseteq \operatorname{ran} T$. To prove the converse inclusion, let $w \in \operatorname{ran} T$ be arbitrary. Then, there exists $v \in \mathcal{V}$ such that $Tv = w$. Since $\mathcal{V} = (\operatorname{nul} T) + \mathcal{X}$, there exist $u \in \operatorname{nul} T$ and $x \in \mathcal{X}$ such that $v = u + x$. Then $Tv = T(u+x) = Tu + Tx = Tx$. As $x \in \mathcal{X}$, there exist $\xi_1, \dots, \xi_m \in \mathbb{F}$ such that $x = \sum_{j=1}^{m} \xi_j x_j$. Now we use linearity of T to deduce

$$w = Tv = Tx = \sum_{j=1}^{m} \xi_j T x_j.$$

This proves that $w \in \operatorname{span}\bigl\{Tx_1, \dots, Tx_m\bigr\}$. Since w was arbitrary in $\operatorname{ran} T$ this completes a proof of (1.10).

Next we prove that the vectors $Tx_1, \dots, Tx_m$ are linearly independent. Let $\alpha_1, \dots, \alpha_m \in \mathbb{F}$ be arbitrary and assume that

$$\alpha_1 T x_1 + \cdots + \alpha_m T x_m = 0_{\mathcal{W}}. \tag{1.11}$$

eq-li-rnt

Since T is linear (1.11) implies that

$$\alpha_1 x_1 + \cdots + \alpha_m x_m \in \operatorname{nul} T. \tag{1.12}$$

eq-li2-rnt

Recall that $x_1, \dots, x_m \in cX$ and $\mathcal{X}$ is a subspace of $\mathcal{V}$, so

$$\alpha_1 x_1 + \cdots + \alpha_m x_m \in \mathcal{X}. \tag{1.13}$$

eq-li3-rnt

Now (1.12), (1.13) and the fact that $(\operatorname{nul} T) \cap \mathcal{X} = \{0_{\mathcal{V}}\}$ imply

$$\alpha_1 x_1 + \cdots + \alpha_m x_m = 0_{\mathcal{V}}. \tag{1.14}$$

eq-li4-rnt

Since $x_1, \dots, x_m$ are linearly independent (1.14) yields $\alpha_1 = \cdots = \alpha_m = 0$. This completes a proof of the linear independence of $Tx_1, \dots, Tx_m$.

Thus $\bigl\{Tx_1, \dots, Tx_m\bigr\}$ is a basis for $\operatorname{ran} T$. Consequently $\dim(\operatorname{ran} T) = m$. Since $m = \dim \mathcal{X}$, (1.9) implies (1.8). This completes the proof. $\square$

A direct proof of the Nullity-Rank Theorem is as follows:

Proof. Since $\operatorname{nul} T$ is a subspace of $\mathcal{V}$ it is finite dimensional. Set $k = \dim\bigl(\operatorname{nul} T\bigr)$ and let $\mathcal{C} = \bigl\{u_1, \dots, u_k\bigr\}$ be a basis for $\operatorname{nul} T$.

Since $\mathcal{V}$ is finite dimensional there exists a finite set $\mathcal{F} \subset \mathcal{V}$ such that $\operatorname{span}(\mathcal{F}) = \mathcal{V}$. Then the set $T\mathcal{F}$ is a finite subset of $\mathcal{W}$ and $\operatorname{ran} T =$

$\operatorname{span}\bigl(T\mathcal{F}\bigr)$. Thus $\operatorname{ran} T$ is finite dimensional. Let $\dim\bigl(\operatorname{ran} T\bigr) = m$ and let $\mathcal{E} = \bigl\{w_1, \dots, w_m\bigr\}$ be a basis of $\operatorname{ran} T$.

Since clearly for every $j \in \bigl\{1, \dots, m\bigr\}$, $w_j \in \operatorname{ran} T$, we have that for every $j \in \bigl\{1, \dots, m\bigr\}$ there exists $v_j \in \mathcal{V}$ such that $Tv_j = w_j$. Set $\mathcal{D} = \bigl\{v_1, \dots, v_m\bigr\}$.

Further set $\mathcal{B} = \mathcal{C} \cup \mathcal{D}$.

We will prove the following three facts:

(I) $\mathcal{C} \cap \mathcal{D} = \emptyset$,
(II) $\operatorname{span} \mathcal{B} = \mathcal{V}$,
(III) $\mathcal{B}$ is a linearly independent set.

To prove (I), notice that the vectors in $\mathcal{E}$ are nonzero, since $\mathcal{E}$ is linearly independent. Therefore, for every $v \in \mathcal{D}$ we have that $Tv \neq 0_{\mathcal{W}}$. Since for every $u \in \mathcal{C}$ we have $Tu = 0_{\mathcal{W}}$ we conclude that $u \in \mathcal{C}$ implies $u \notin \mathcal{D}$. This proves (I).

To prove (II), first notice that by the definition of $\mathcal{B} \subset \mathcal{V}$. Since $\mathcal{V}$ is a vector space, we have $\operatorname{span} \mathcal{B} \subseteq \mathcal{V}$.

To prove the converse inclusion, let $v \in \mathcal{V}$ be arbitrary. Then $Tv \in \operatorname{ran} T$. Since $\mathcal{E}$ spans $\operatorname{ran} T$, there exist $\beta_1, \dots, \beta_m \in \mathbb{F}$ such that

$$Tv = \sum_{j=1}^{m} \beta_j w_j.$$

Set

$$v' = \sum_{j=1}^{m} \beta_j v_j.$$

Then, by linearity of T we have

$$Tv' = \sum_{j=1}^{m} \beta_j T v_j = \sum_{j=1}^{m} \beta_j w_j = Tv.$$

The last equality yields and the linearity of T yield $T(v - v') = 0_{\mathcal{W}}$. Consequently, $v - v' \in \operatorname{nul} T$. Since $\mathcal{C}$ spans $\operatorname{nul} T$, there exist $\alpha_1, \dots, \alpha_k \in \mathbb{F}$ such that

$$v - v' = \sum_{j=1}^{k} \alpha_i u_i.$$

Consequently,

$$v = v' + \sum_{j=1}^{k} \alpha_i u_i = \sum_{j=1}^{k} \alpha_i u_i + \sum_{j=1}^{m} \beta_j v_j.$$

This proves that for arbitrary $v \in \mathcal{V}$ we have $v \in \operatorname{span} \mathcal{B}$. Thus $\mathcal{V} \subseteq \operatorname{span} \mathcal{B}$ and (II) is proved.

To prove (III) let $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ and $\beta_1, \ldots, \beta_m \in \mathbb{F}$ be arbitrary and assume that

$$\sum_{j=1}^{k} \alpha_i u_i + \sum_{j=1}^{m} \beta_j v_j = 0_{\mathcal{V}}. \tag{1.15}$$

eq-assu-4-l-i

Applying T to both sides of the last equality, and using the fact that $u_i \in \operatorname{nul} T$ and the definition of v_j we get

$$\sum_{j=1}^{m} \beta_j w_j = 0_{\mathcal{W}}.$$

Since $\mathcal{E}$ is a linearly independent set the last equality implies that $\beta_j = 0$ for all $j \in \{1, \ldots, m\}$. Now substitute these equalities in (1.15) to get

$$\sum_{j=1}^{k} \alpha_i u_i = 0_{\mathcal{V}}.$$

Since $\mathcal{C}$ is a linearly independent set the last equality implies that $\alpha_i = 0$ for all $i \in \{1, \ldots, k\}$. This proves the linear independence of $\mathcal{B}$.

It follows from (II) and (III) that $\mathcal{B}$ is a basis for $\mathcal{V}$. By (I) we have that $|\mathcal{B}| = |\mathcal{C}| + |\mathcal{D}| = k + m$. This completes proof of the theorem. $\square$

The nonnegative integer $\dim(\operatorname{nul} T)$ is called the *nullity* of T; the nonnegative integer $\dim(\operatorname{ran} T)$ is called the *rank* of T.

The nullity-rank theorem in English reads: If a linear operator is defined on a finite dimensional vector space, then its nullity and its rank are finite and they add up to the dimension of the domain.

Proposition 1.12. *Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over $\mathbb{F}$. Assume that $\mathcal{V}$ is finite dimensional. The following statements are equivalent*

(a) *There exists a surjection $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$.*
(b) *$\mathcal{W}$ is finite dimensional and* $\dim \mathcal{V} \geq \dim \mathcal{W}$.

Proposition 1.13. *Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over $\mathbb{F}$. Assume that $\mathcal{V}$ is finite dimensional. The following statements are equivalent*

(a) *There exists an injection $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$.*
(b) *Either $\mathcal{W}$ is infinite dimensional or* $\dim \mathcal{V} \leq \dim \mathcal{W}$.

Proposition 1.14. *Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over $\mathbb{F}$. Assume that $\mathcal{V}$ is finite dimensional. The following statements are equivalent*

(a) *There exists an isomorphism $T : \mathcal{V} \to \mathcal{W}$.*
(b) *$\mathcal{W}$ is finite dimensional and* $\dim \mathcal{W} = \dim \mathcal{V}$.

1.4. **Isomorphism between $\mathcal{L}(\mathcal{V}, \mathcal{W})$ and $\mathbb{F}^{n\times m}$.** Let $\mathcal{V}$ and $\mathcal{W}$ be finite dimensional vector spaces over $\mathbb{F}$, $m = \dim \mathcal{V}$, $n = \dim \mathcal{W}$, let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis for $\mathcal{V}$ and let $\mathcal{C} = \{w_1, \ldots, w_n\}$ be a basis for $\mathcal{W}$. The mapping $C_{\mathcal{B}}$ provides an isomorphism between $\mathcal{V}$ and $\mathbb{F}^m$ and $C_{\mathcal{C}}$ provides an isomorphism between $\mathcal{W}$ and $\mathbb{F}^n$.

Recall that the simplest way to define a linear operator from $\mathbb{F}^m$ to $\mathbb{F}^n$ is to use an $n \times m$ matrix B. It is convenient to consider an $n \times m$ matrix to be an m-tuple of its columns, which are vectors in $\mathbb{F}^n$. For example, let $\mathbf{b}_1, \dots, \mathbf{b}_m \in \mathbb{F}^n$ be columns of an $n \times m$ matrix B. Then we write

$$B = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_m \end{bmatrix}.$$

This notation is convenient since it allows us to write a multiplication of a vector $\mathbf{x} \in \mathbb{F}^m$ by a matrix B as

$$B\mathbf{x} = \sum_{j=1}^{m} \xi_j \mathbf{b}_j \qquad \text{where} \qquad \mathbf{x} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}. \tag{1.16}$$ eq-defBx

Notice the similarity of the definition in (1.16) to the definition (1.6) of the operator $L_{\mathcal{C}}^{\mathcal{B}}$ in Example 1.8. Taking $\mathcal{B}$ to be the standard basis of $\mathbb{F}^m$ and taking $\mathcal{C}$ to me the m-tuple given by B, we have $L_{\mathcal{C}}^{\mathcal{B}}(\mathbf{x}) = B\mathbf{x}$.

Let $T : \mathcal{V} \to \mathcal{W}$ be a linear operator. Our next goal is to connect T in a natural way to a certain $n \times m$ matrix B. That "natural way" is suggested by following diagram:

$$\begin{tikzcd} \mathcal{V} \arrow[r, "T"] \arrow[d, "C_{\mathcal{B}}"'] & \mathcal{W} \arrow[d, "C_{\mathcal{C}}"] \\ \mathbb{F}^m \arrow[r, dashed, "B"'] & \mathbb{F}^n \end{tikzcd}$$

We seek an $n \times m$ matrix B such that the action of T between $\mathcal{V}$ and $\mathcal{W}$ is in some sense replicated by the action of B between $\mathbb{F}^m$ and $\mathbb{F}^n$. Precisely, we seek B such that

$$C_{\mathcal{C}}(Tv) = B\big(C_{\mathcal{B}}(v)\big) \qquad \forall\, v \in \mathcal{V}. \tag{1.17}$$ eq-cdB

In English: multiplying the vector of coordinates of v by B we get exactly the coordinates of Tv.

Using the basis vectors $v_1, \dots, v_n \in \mathcal{B}$ in (1.17) we see that the matrix

$$B = \begin{bmatrix} C_{\mathcal{C}}(Tv_1) & \cdots & C_{\mathcal{C}}(Tv_m) \end{bmatrix} \tag{1.18}$$ eq-defB

has the desired property (1.17).

For an arbitrary $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ the formula (1.18) associates the matrix $B \in \mathbb{F}^{n \times m}$ with T. In other words (1.18) defines a function from $\mathcal{L}(\mathcal{V}, \mathcal{W})$ to $\mathbb{F}^{n \times m}$.

th-MatR **Theorem 1.15.** *Let $\mathcal{V}$ and $\mathcal{W}$ be finite dimensional vector spaces over $\mathbb{F}$, $m = \dim \mathcal{V}$, $n = \dim \mathcal{W}$, let $\mathcal{B} = \{v_1, \dots, v_m\}$ be a basis for $\mathcal{V}$ and let $\mathcal{C} = \{w_1, \dots, w_n\}$ be a basis for $\mathcal{W}$. The function*

$$M_{\mathcal{C}}^{\mathcal{B}} : \mathcal{L}(\mathcal{V}, \mathcal{W}) \to \mathbb{F}^{n \times m}$$

defined by

$$M_{\mathcal{C}}^{\mathcal{B}}(T) = \bigl[C_{\mathcal{C}}(Tv_1) \ \cdots \ C_{\mathcal{C}}(Tv_m)\bigr], \qquad T \in \mathcal{L}(\mathcal{V},\mathcal{W}) \tag{1.19}$$

eq-defM

is an isomorphism.

Proof. It is easy to verify that $M_{\mathcal{C}}^{\mathcal{B}}$ is a linear operator.

Since the definition of $M_{\mathcal{C}}^{\mathcal{B}}(T)$ coincides with (1.18), equality (1.17) yields

$$C_{\mathcal{C}}(Tv) = \bigl(M_{\mathcal{C}}^{\mathcal{B}}(T)\bigr)C_{\mathcal{B}}(v). \tag{1.20}$$

eq-cdMBC

The most direct way to prove that $M_{\mathcal{C}}^{\mathcal{B}}$ is an isomorphism is to construct its inverse. The inverse is suggested by the diagram (1.21).

$$\begin{tikzcd} \mathcal{V} \arrow[r, dashed, red, "T"] \arrow[d, "C_{\mathcal{B}}"'] & \mathcal{W} \\ \mathbb{F}^m \arrow[r, "B"'] & \mathbb{F}^n \arrow[u, "(C_{\mathcal{C}})^{-1}"'] \end{tikzcd} \tag{1.21}$$

cd-rT

Define

$$N_{\mathcal{C}}^{\mathcal{B}} : \mathbb{F}^{n\times m} \to \mathcal{L}(\mathcal{V},\mathcal{W})$$

by

$$\bigl(N_{\mathcal{C}}^{\mathcal{B}}(B)\bigr)(v) = (C_{\mathcal{C}})^{-1}\bigl(B(C_{\mathcal{B}}(v))\bigr), \qquad B \in \mathbb{F}^{n\times m}. \tag{1.22}$$

eq-defN

Next we prove that

$$N_{\mathcal{C}}^{\mathcal{B}} \circ M_{\mathcal{C}}^{\mathcal{B}} = I_{\mathcal{L}(\mathcal{V},\mathcal{W})} \quad \text{and} \quad M_{\mathcal{C}}^{\mathcal{B}} \circ N_{\mathcal{C}}^{\mathcal{B}} = I_{\mathbb{F}^{n\times m}}.$$

First for arbitrary $T \in \mathcal{L}(\mathcal{V},\mathcal{W})$ and arbitrary $v \in \mathcal{V}$ we calculate

$$\begin{aligned} \Bigl(\bigl(N_{\mathcal{C}}^{\mathcal{B}} \circ M_{\mathcal{C}}^{\mathcal{B}}\bigr)(T)\Bigr)(v) &= (C_{\mathcal{C}})^{-1}\Bigl(\bigl(M_{\mathcal{C}}^{\mathcal{B}}(T)\bigr)(C_{\mathcal{B}}(v))\Bigr) && \text{by (1.22)} \\ &= (C_{\mathcal{C}})^{-1}\bigl(C_{\mathcal{C}}(Tv)\bigr) && \text{by (1.20)} \\ &= Tv. \end{aligned}$$

Thus $\bigl(N_{\mathcal{C}}^{\mathcal{B}} \circ M_{\mathcal{C}}^{\mathcal{B}}\bigr)(T) = T$ and thus, since $T \in \mathcal{L}(\mathcal{V},\mathcal{W})$ was arbitrary, $N_{\mathcal{C}}^{\mathcal{B}} \circ M_{\mathcal{C}}^{\mathcal{B}} = I_{\mathcal{L}(\mathcal{V},\mathcal{W})}$.

Let now $B \in \mathbb{F}^{n\times m}$ be arbitrary and calculate

$$\begin{aligned} \bigl(M_{\mathcal{C}}^{\mathcal{B}} \circ N_{\mathcal{C}}^{\mathcal{B}}\bigr)(B) &= M_{\mathcal{C}}^{\mathcal{B}}\bigl(N_{\mathcal{C}}^{\mathcal{B}}(B)\bigr) \\ &= \Bigl[C_{\mathcal{C}}\bigl(\bigl(N_{\mathcal{C}}^{\mathcal{B}}(B)\bigr)(v_1)\bigr) \ \cdots \ C_{\mathcal{C}}\bigl(\bigl(N_{\mathcal{C}}^{\mathcal{B}}(B)\bigr)(v_m)\bigr)\Bigr] && \text{by (1.19)} \\ &= \Bigl[B(C_{\mathcal{B}}(v_1)) \ \cdots \ B(C_{\mathcal{B}}(v_m))\Bigr] && \text{by (1.22)} \\ &= B\Bigl[C_{\mathcal{B}}(v_1) \ \cdots \ C_{\mathcal{B}}(v_m)\Bigr] && \text{matrix mult.} \\ &= B\, I_m && \text{def. of } C_{\mathcal{B}} \\ &= B. \end{aligned}$$

Thus $\bigl(M_{\mathcal{C}}^{\mathcal{B}} \circ N_{\mathcal{C}}^{\mathcal{B}}\bigr)(B) = B$ for all $B \in \mathbb{F}^{n\times m}$, proving that $M_{\mathcal{C}}^{\mathcal{B}} \circ N_{\mathcal{C}}^{\mathcal{B}} = I_{\mathbb{F}^{n\times m}}$.

This completes the proof that $M_{\mathcal{C}}^{\mathcal{B}}$ is a bijection. Since it is linear, $M_{\mathcal{C}}^{\mathcal{B}}$ is an isomorphism. $\square$

th-MTS **Theorem 1.16.** *Let $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ be finite dimensional vector spaces over $\mathbb{F}$, $k = \dim \mathcal{U}$, $m = \dim \mathcal{V}$, $n = \dim \mathcal{W}$, let $\mathcal{A}$ be a basis for $\mathcal{U}$, let $\mathcal{B}$ be a basis for $\mathcal{V}$, and let $\mathcal{C}$ be a basis for $\mathcal{W}$. Let $S \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ and $T \in \mathcal{L}(\mathcal{V}, \mathcal{W})$. Let $M_{\mathcal{B}}^{\mathcal{A}}(S) \in \mathbb{F}^{m\times k}$, $M_{\mathcal{C}}^{\mathcal{B}}(T) \in \mathbb{F}^{n\times m}$ and $M_{\mathcal{C}}^{\mathcal{A}}(TS) \in \mathbb{F}^{n\times k}$ be as defined in Theorem* 1.15. *Then*

$$M_{\mathcal{C}}^{\mathcal{A}}(TS) = M_{\mathcal{C}}^{\mathcal{B}}(T) M_{\mathcal{B}}^{\mathcal{A}}(S).$$

Proof. Let $\mathcal{A} = \{u_, \ldots, u_k\}$ and calculate

$$\begin{aligned}
M_{\mathcal{C}}^{\mathcal{A}}(TS) &= \Big[C_{\mathcal{C}}\big(TSu_1\big) \ \cdots \ C_{\mathcal{C}}\big(TSu_k\big) \Big] && \text{by (1.19)}\\
&= \Big[M_{\mathcal{C}}^{\mathcal{B}}(T)\big(C_{\mathcal{B}}(Su_1)\big) \ \cdots \ M_{\mathcal{C}}^{\mathcal{B}}(T)\big(C_{\mathcal{B}}(Su_k)\big) \Big] && \text{by (1.20)}\\
&= M_{\mathcal{C}}^{\mathcal{B}}(T) \Big[C_{\mathcal{B}}(Su_1) \ \cdots \ C_{\mathcal{B}}(Su_k) \Big] && \text{matrix mult.}\\
&= M_{\mathcal{C}}^{\mathcal{B}}(T) M_{\mathcal{B}}^{\mathcal{A}}(S). && \text{by (1.19)}
\end{aligned}$$

$\square$

The following diagram illustrates the content of Theorem 1.16.

$$\begin{tikzcd}
 & \mathcal{V} \arrow[dr, "T"] \arrow[dd, "C_{\mathcal{B}}"'] & \\
\mathcal{U} \arrow[ur, "S"] \arrow[rr, "TS"] \arrow[dd, "C_{\mathcal{A}}"] & & \mathcal{W} \arrow[dd, "C_{\mathcal{C}}"] \\
 & \mathbb{F}^m \arrow[dr, "M_{\mathcal{C}}^{\mathcal{B}}(T)"] & \\
\mathbb{F}^k \arrow[ur, "M_{\mathcal{B}}^{\mathcal{A}}(S)"'] \arrow[rr, "M_{\mathcal{C}}^{\mathcal{A}}(TS) = M_{\mathcal{C}}^{\mathcal{B}}(T) M_{\mathcal{B}}^{\mathcal{A}}(S)"'] & & \mathbb{F}^n
\end{tikzcd}$$

2. Problems

Problem 2.1. Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over a scalar field $\mathbb{F}$. Let $\mathcal{S}$ be a subspace of the direct product vector space $\mathcal{V}\times\mathcal{W}$, let $\mathcal{G}$ be a subspace of $\mathcal{V}$ and let $\mathcal{H}$ be a subspace of $\mathcal{W}$. Then

$$\mathcal{S}(\mathcal{G}) = \big\{ w \in \mathcal{W} \,:\, \exists v \in \mathcal{G} \text{ such that } \ (v, w) \in \mathcal{S} \big\}$$

is a subspace of $\mathcal{W}$ and

$$\mathcal{S}^{-1}(\mathcal{H}) = \big\{ v \in \mathcal{V} \,:\, \exists w \in \mathcal{H} \text{ such that } \ (v, w) \in \mathcal{S} \big\}$$

is a subspace of $\mathcal{V}$.

Problem 2.2. Let $\mathcal{V}$ and $\mathcal{W}$ be finite-dimensional vector spaces over a scalar field $\mathbb{F}$. Let $\mathcal{S}$ be a subspace of the direct product vector space $\mathcal{V}\times\mathcal{W}$. The following four sets are subspaces

$$\begin{aligned}
\operatorname{dom}\mathcal{S} &= \bigl\{v\in\mathcal{V} \,:\, \exists\, w\in\mathcal{W} \text{ such that } (v,w)\in\mathcal{S}\bigr\},\\
\operatorname{ran}\mathcal{S} &= \bigl\{w\in\mathcal{W} \,:\, \exists\, v\in\mathcal{V} \text{ such that } (v,w)\in\mathcal{S}\bigr\},\\
\operatorname{nul}\mathcal{S} &= \bigl\{v\in\mathcal{V} \,:\, \ (v,0_{\mathcal{W}})\in\mathcal{S}\bigr\},\\
\operatorname{mul}\mathcal{S} &= \bigl\{w\in\mathcal{W} \,:\, \ (0_{\mathcal{V}},w)\in\mathcal{S}\bigr\}.
\end{aligned}$$

and the following equality holds:

$$\dim\operatorname{dom}\mathcal{S}+\dim\operatorname{mul}\mathcal{S}=\dim\operatorname{ran}\mathcal{S}+\dim\operatorname{nul}\mathcal{S}.$$

Hint: The following equivalence holds. For all $v\in\mathcal{V}$ and all $w\in\mathcal{W}$ we have:

$$(v,w)\in\mathcal{S} \quad\Leftrightarrow\quad (v+x,w+y)\in\mathcal{S} \quad \forall x\in\operatorname{nul}\mathcal{S} \text{ and } \forall y\in\operatorname{mul}\mathcal{S}.$$

pb-rev

Problem 2.3. Let $\mathcal{V}$ and $\mathcal{W}$ be finite-dimensional vector spaces over a scalar field $\mathbb{F}$ and recall that $\mathcal{V}\times\mathcal{W}$ and $\mathcal{W}\times\mathcal{V}$ are the direct product vector spaces. Prove that the function

$$R:\mathcal{V}\times\mathcal{W}\to\mathcal{W}\times\mathcal{V}$$

defined by

$$R(v,w)=(w,v) \quad \text{for all} \quad (v,w)\in\mathcal{V}\times\mathcal{W}$$

is an isomorphism.

Problem 2.4. Let $\mathcal{V}$ and $\mathcal{W}$ be finite-dimensional vector spaces over a scalar field $\mathbb{F}$ and recall that $\mathcal{V}\times\mathcal{W}$ and $\mathcal{W}\times\mathcal{V}$ are the direct product vector spaces. Let $\mathcal{T}$ be a subset of $\mathcal{V}\times\mathcal{W}$. Then $\mathcal{T}$ is an isomorphism between $\mathcal{V}$ and $\mathcal{W}$ if and only if the set

$$\bigl\{(w,v)\in\mathcal{W}\times\mathcal{V} \,:\, (v,w)\in\mathcal{T}\bigr\}=R\mathcal{T}$$

is an isomorphism between $\mathcal{W}$ and $\mathcal{V}$. (Use Problem 2.3 and Propositions 1.3 and 1.4 to prove this equivalence.)