The matrix of a linear operator in a pair of ordered bases*

DRAGAN JUKIĆ†

Abstract. *In the lecture it is shown how to represent a linear operator by a matrix. This representation allows us to define an operation with matrices.*

Key words: *linear operator, matrix of a linear operator, matrix operations*

Sažetak. Matrica linearnog operatora u paru uređenih baza. *Na predavanju je pokazano kako se linearni operator može reprezentirati matricom. Ta reprezentacija omogućava nam da definiramo operacije s matricama.*

Ključne riječi: *linearan operator, matrica linearnog operatora, operacije s matricama*

1. Defining a linear operator

The function from one vector space to another vector space is called the **operator**. In this lecture we shall deal only with finite dimensional vector spaces.

Definition 1. *Let V and W be any two finite dimensional real vector spaces. We say that the operator $\mathcal{A} : V \to W$ is* linear *if*

$$\mathcal{A}(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\mathcal{A}(\mathbf{x}) + \mu\mathcal{A}(\mathbf{y})$$

for all scalars $\lambda, \mu \in \mathbb{R}$ and for all vectors $\mathbf{x}, \mathbf{y} \in V$. It is easy to check that the operator $\mathcal{A} : V \to W$ is linear if and only if $\mathcal{A}$ is an additive and homogenous operator, where we define:

Definition 2. *An operator $\mathcal{A} : V \to W$ is said to be:*

a) additive *if $\mathcal{A}(\mathbf{x} + \mathbf{y}) = \mathcal{A}(\mathbf{x}) + \mathcal{A}(\mathbf{y})$ for all vectors $\mathbf{x}, \mathbf{y} \in V$;*

b) homogenous *if $\mathcal{A}(\lambda\mathbf{x}) = \lambda\mathcal{A}(\mathbf{x})$ for each scalar $\lambda \in \mathbb{R}$ and for each vector $\mathbf{x} \in V$.*

*The lecture presented at the MATHEMATICAL COLLOQUIUM in Osijek organized by Croatian Mathematical Society - Division Osijek, October 19, 1996.

†Faculty of Food Technology, F. Kuhača 18, HR-31 000 Osijek, e-mail: `jukicd@oliver.efos.hr`

Properties of additivity and homogenity of a linear operator are graphically illustrated in *Figure 1*.

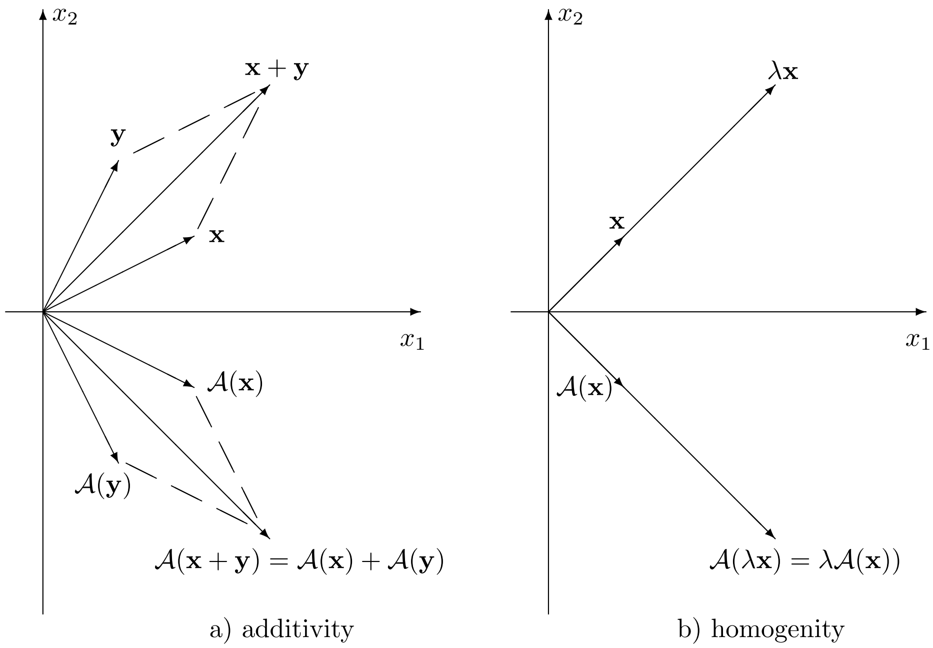

Figure 1.

Example 1. *Let us give some examples of a linear operator $\mathcal{A} : V \to W$:*

a) $V = W = \mathbb{R}^2$, $\mathcal{A}(x_1, x_2) = (x_1, -x_2)$ *(reflection of a plane in the x_1 - axis);*

b) $V = W = \mathbb{R}^2$, $\mathcal{A}(x_1, x_2) = (-x_1, -x_2)$ *(symmetry of a plane about the origin);*

c) $V = W = \mathbb{R}^2$, $\mathcal{A}(x_1, x_2) = (x_1, 0)$ *(orthogonal projection of a plane on x_1 - axis)*

Lemma 1. *Two linear operators $\mathcal{A}, \mathcal{B} : V \to W$ are equal if and only if they attain the same values on the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for V.*

Proof. If two operators are equal, then they attain the same values on basis,

Let us prove the converse. According to the assumption, we have $\mathcal{A}(\mathbf{e}_i) = \mathcal{B}(\mathbf{e}_i)$ for every $i = 1, \ldots, n$. Let us prove that $\mathcal{A}(\mathbf{x}) = \mathcal{B}(\mathbf{x})$ for every $\mathbf{x} \in V$. For that purpose, let $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$ be a linear combination of the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of the basis for V. Then we have

$$\begin{aligned}\mathcal{A}(\mathbf{x}) &= \mathcal{A}(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) = x_1\mathcal{A}(\mathbf{e}_1) + \ldots x_n\mathcal{A}(\mathbf{e}_n) = x_1\mathcal{B}(\mathbf{e}_1) + \ldots x_n\mathcal{B}(\mathbf{e}_n) \\ &= \mathcal{B}(x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n) = \mathcal{B}(\mathbf{x})\end{aligned}$$

from where we conclude that $\mathcal{A} = \mathcal{B}$. □

The next *Theorem* tells us that each linear operator $\mathcal{A} : V \to W$ is completely determined by its values on vectors of the basis for V.

Theorem 1. *Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be any basis for V and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be any n vector in W. Then there is one and only one linear operator $\mathcal{A} : V \to W$ such that*

$$\mathcal{A}(\mathbf{e}_i) = \mathbf{v}_i, \quad i = 1, \ldots, n.$$

Proof. Each vector $\mathbf{x} \in V$ is uniquely expressible as a linear combination of vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$: $\mathbf{x} = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n$. It is easy to check that the operator $\mathcal{A} : V \to W$ defined by the formula $\mathcal{A}(\mathbf{x}) = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n$ is a linear operator and that $\mathcal{A}(\mathbf{e}_i) = \mathbf{v}_i$, $i = 1, \ldots, n$.

If $\mathcal{B} : V \to W$ is a linear operator such that $\mathcal{B}(\mathbf{e}_i) = \mathbf{v}_i$, $i = 1, \ldots, n$, then from *Lemma* 1 we obtain $\mathcal{A} = \mathcal{B}$ □

Because linear operators are functions, they can be added, multiplied by scalars and composed with one another.

- The sum $\mathcal{C} = \mathcal{A} + \mathcal{B}$ of linear operators $\mathcal{A}, \mathcal{B} : V \to W$ is again a linear operator. Namely, for every scalar $\lambda, \mu \in \mathbb{R}$ and for every vector $\mathbf{x}, \mathbf{y} \in V$ there holds:

$$\begin{aligned}\mathcal{C}(\lambda\mathbf{x}+\mu\mathbf{y}) &= \mathcal{A}(\lambda\mathbf{x}+\mu\mathbf{y}) + \mathcal{B}(\lambda\mathbf{x}+\mu\mathbf{y}) = (\lambda\mathcal{A}(\mathbf{x})+\mu\mathcal{A}(\mathbf{y}))+(\lambda\mathcal{B}(\mathbf{x})+\mu\mathcal{B}(\mathbf{y}))\\ &= \lambda\,(\mathcal{A}(\mathbf{x}) + \mathcal{B}(\mathbf{x}))+\mu\,(\mathcal{A}(\mathbf{x}) + \mathcal{B}(\mathbf{x})) = \lambda\mathcal{C}(\mathbf{x}) + \mu\mathcal{C}(\mathbf{y})\end{aligned}$$

- The scalar multiple $\mathcal{C} = \alpha\mathcal{A}$ of the linear operator $\mathcal{A} : V \to W$ by the scalar $\alpha \in \mathbb{R}$ is again a linear operator:

$$\begin{aligned}\mathcal{C}(\lambda\mathbf{x} + \mu\mathbf{y}) &= \alpha\mathcal{A}(\lambda\mathbf{x} + \mu\mathbf{y}) = \alpha\,(\lambda\mathcal{A}(\mathbf{x}) + \mu\mathcal{A}(\mathbf{y})) = \lambda\,(\alpha\mathcal{A}(\mathbf{x})) + \mu\,(\alpha\mathcal{A}(\mathbf{y}))\\ &= \lambda\mathcal{C}(\mathbf{x}) + \mu\mathcal{C}(\mathbf{y}).\end{aligned}$$

- The composition $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$ of linear operators $\mathcal{B} : V \to W$ and $\mathcal{A} : W \to Z$ is the linear operator from V to Z:

$$\begin{aligned}\mathcal{C}(\lambda\mathbf{x}+\mu\mathbf{y}) &= \mathcal{A}\,(\mathcal{B}(\lambda\mathbf{x}+\mu\mathbf{y})) = \mathcal{A}\,(\lambda\mathcal{B}(\mathbf{x})+\mu\mathcal{B}(\mathbf{y})) = \lambda\mathcal{A}\,(\mathcal{B}(\mathbf{x})) + \mu\mathcal{A}\,(\mathcal{B}(\mathbf{y}))\\ &= \lambda\mathcal{C}(\mathbf{x}) + \mu\mathcal{C}(\mathbf{y}).\end{aligned}$$

2. The matrix of a linear operator

In this section we will show how to associate a matrix with each linear operator $\mathcal{A} : V \to W$, where V and W are any two finite dimensional vector spaces.

Suppose $(\mathbf{e}) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ is an ordered basis for the finite dimensional vector space V, and $(\mathbf{f}) = (\mathbf{f}_1, \ldots, \mathbf{f}_m)$ is an ordered basis for the finite dimensional vector space W. According to *Theorem* 1, the operator $\mathcal{A}$ is completly determined by its values $\mathcal{A}(\mathbf{e}_j)$, $j = 1, \ldots, n$, on vectors of the basis. Since $\mathcal{A}(\mathbf{e}_j)$ are vectors in W and $(\mathbf{f}) = (\mathbf{f}_1, \ldots, \mathbf{f}_m)$ is the basis for W, there are unique scalars a_{ij} $(i = 1, \ldots, m, j = 1, \ldots, n)$ such that:

$$\begin{aligned}\mathcal{A}(\mathbf{e}_1) &= \quad a_{11}\mathbf{f}_1 + a_{21}\mathbf{f}_2 + \ldots + a_{m1}\mathbf{f}_m\\ \mathcal{A}(\mathbf{e}_2) &= \quad a_{12}\mathbf{f}_1 + a_{22}\mathbf{f}_2 + \ldots + a_{m2}\mathbf{f}_m\\ &\vdots\\ \mathcal{A}(\mathbf{e}_n) &= \quad a_{1n}\mathbf{f}_1 + a_{2n}\mathbf{f}_2 + \ldots + a_{mn}\mathbf{f}_m\end{aligned} \tag{1}$$

In an ordered pair of the bases (**e**), (**f**) to the operator $\mathcal{A}$ there belong $m \cdot n$ scalars a_{ij} $(i = 1, \ldots, m, j = 1, \ldots, n)$ which can be displayed in a rectangular array

$$\mathbf{A}(\mathbf{f}, \mathbf{e}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{2}$$

called the *matrix of a linear operator* $\mathcal{A}$ in an ordered pair of bases (**e**), (**f**). The matrix (2) has m rows and n columns. Because of this, we say that it has the *order* $m \times n$. The matrix (2) may be written in an abbreviated form as $\mathbf{A} = (a_{ij})$.

Example 1.. *Let $\mathcal{D} : \mathcal{P}^3 \to \mathcal{P}^2$ be a linear operator that assigns to each polynomial its derivative,* $(\mathbf{e}) = (x^3, x^2, x, 1)$ *the basis for* $\mathcal{P}^3$ *and* $(\mathbf{f}) = (x^2, x, 1)$ *the basis for* $\mathcal{P}^2$*. Then*

$$\mathbf{D}(\mathbf{f}, \mathbf{e}) = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

3. The algebra of matrices

Let $M_{m \times n}$ be the set of all $m \times n$ real matrices and $\mathcal{L}(V, W)$ be the set of all linear operators from V to W. By using *Theorem* 1 it is easy to see that the mapping $\mathcal{A} \mapsto \mathbf{A}(\mathbf{e}, \mathbf{f})$ is a bijection from $\mathcal{L}(V, W)$ onto $M_{m \times n}$. This bijection allows us to represent operations by operators and vectors by operations with matrices.

Equality of matrices

Let (2) be a matrix of a linear operator $\mathbf{A} : V \to W$. Furthermore, let $\mathcal{B} : V \to W$ be a linear operator,

$$\begin{aligned} \mathcal{B}(\mathbf{e}_1) = &\quad b_{11}\mathbf{f}_1 + b_{21}\mathbf{f}_2 + \ldots + b_{m1}\mathbf{f}_m \\ \mathcal{B}(\mathbf{e}_2) = &\quad b_{12}\mathbf{f}_1 + b_{22}\mathbf{f}_2 + \ldots + b_{m2}\mathbf{f}_m \\ \vdots \quad & \\ \mathcal{B}(\mathbf{e}_n) = &\quad b_{1n}\mathbf{f}_1 + b_{2n}\mathbf{f}_2 + \ldots + b_{mn}\mathbf{f}_m \end{aligned} \tag{3}$$

its values on vectors of the basis $\mathbf{f}_1, \ldots, \mathbf{f}_m$ and

$$\mathbf{B}(\mathbf{f}, \mathbf{e}) = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \tag{4}$$

its matrix in an ordered pair of bases (**e**), (**f**). According to *Lemma* 1, $\mathcal{A} = \mathcal{B}$ if and only if $\mathcal{A}(\mathbf{e}_i) = \mathcal{B}(\mathbf{e}_i)$ for all $i = 1, \ldots, m$, i.e. if $a_{ij} = b_{ij}$ for all $i = 1, \ldots, m$ and for all $j = 1, \ldots, n$. This gives us a criterion of equality of matrices:

Matrices $\mathbf{A} = (a_{ij})$ *and* $\mathbf{B} = (b_{ij})$ *are equal if and only if they have the same order and* $a_{ij} = b_{ij}$ *for all* $i = 1, \ldots, m$, *and* $j = 1, \ldots, n$.

Scalar multiple of matrices

Multiplying (1) with scalar $\alpha \in \mathbb{R}$ we conclude that in an ordered pair of bases $(\mathbf{e})$, $(\mathbf{f})$ the operator $\alpha\mathcal{A} : V \to W$ has a matrix

$$\begin{bmatrix} \alpha \cdot a_{11} & \alpha \cdot a_{12} & \cdots & \alpha \cdot a_{1n} \\ \alpha \cdot a_{21} & \alpha \cdot a_{22} & \cdots & \alpha \cdot a_{2n} \\ \vdots & & & \\ \alpha \cdot a_{m1} & \alpha \cdot a_{m2} & \cdots & \alpha \cdot a_{mn} \end{bmatrix}$$

Motivated by this, we define:

The scalar multiple $\alpha\mathbf{A}$ *of the matrix* $\mathbf{A}$ *by the scalar* α *is the matrix whose entries are obtained by multiplying all of the entries in* $\mathbf{A}$ *by* α.

Addition of matrices

Adding (1) and (3) we see that to the operator $\mathcal{C} = \mathcal{A} + \mathcal{B}$ in an ordered pair of bases $(\mathbf{e})$, $(\mathbf{f})$, there belongs a matrix

$$\mathbf{C}(\mathbf{f},\mathbf{e}) = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & & & \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & & & \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}.$$

The sum $\mathbf{A} + \mathbf{B}$ *of the matrices* $\mathbf{A} = (a_{ij})$ *and* $\mathbf{B} = (b_{ij})$ *of the order* $m \times n$ *is a matrix* $\mathbf{C} = (c_{ij})$ *of the same order, where*

$$c_{ij} = a_{ij} + b_{ij} \quad (i = 1, \ldots, m; j = 1, \ldots n).$$

Multiplication of matrices

Suppose $\mathcal{B} : V \to W$ and $\mathcal{A} : W \to Z$ are linear operators. Furthermore, suppose $(\mathbf{e}) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$, $(\mathbf{f}) = (\mathbf{f}_1, \ldots, \mathbf{f}_p)$ and $(\mathbf{g}) = (\mathbf{g}_1, \ldots, \mathbf{g}_m)$ are ordered bases for vector spaces V, W and Z, respectively. Let us show how by using matrices $\mathbf{A} = (a_{ij}) := \mathbf{A}(\mathbf{g},\mathbf{f})$ and $\mathbf{B} = (b_{ij}) := \mathbf{B}(\mathbf{f},\mathbf{e})$ one can determine the matrix $\mathbf{C} := \mathbf{C}(\mathbf{g},\mathbf{e})$ of the linear operator $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$. Let $\mathbf{C} = (c_{ij})$. Then

$$\mathcal{C}(\mathbf{e}_j) = \sum_{i=1}^{m} c_{ij}\mathbf{g}_i, \quad j = 1, \ldots, n. \tag{5}$$

On the other hand, we have

$$\begin{aligned}\mathcal{C}(\mathbf{e}_j) = \quad & \mathcal{A}\left(\mathcal{B}(\mathbf{e}_j)\right) = \mathcal{A}\left(\sum_{k=1}^{p} b_{kj}\mathbf{f}_k\right) = \sum_{k=1}^{p} b_{kj}\mathcal{A}(\mathbf{f}_k) = \sum_{k=1}^{p} b_{kj}\sum_{i=1}^{m} a_{ik}\mathbf{g}_i \\ = \quad & \sum_{i=1}^{m}\left(\sum_{k=1}^{p} a_{ik}b_{kj}\right)\mathbf{g}_i.\end{aligned} \tag{6}$$

According to (5) and (6), we find that $\sum_{i=1}^{m} c_{ij}\mathbf{g}_i = \sum_{i=1}^{m}\left(\sum_{k=1}^{p} a_{ik}b_{kj}\right)\mathbf{g}_i$, from where, because of the linear independence of vectors $\mathbf{g}_i$, we obtain:

$$c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}, \quad i = 1,\ldots,m,\ j = 1,\ldots,n. \tag{7}$$

The product $\mathbf{AB}$ *of the matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *is defined only if matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *are conformable for multiplication, i.e. if the number of columns in* $\mathbf{A}$ *is the same as the number of rows in* $\mathbf{B}$*. If* $\mathbf{A}$ *has the order* $m \times p$ *and* $\mathbf{B}$ *has the order* $p \times n$*, then the product* $\mathbf{AB}$ *is an* $m \times n$ *matrix* $\mathbf{C} = (c_{ij})$ *with entries defined by (7).*

We conclude this section by listing the fundamental algebraic properties of matrix addition, scalar multiplication, and matrix multiplication.

PROPERTIES OF MATRIX ADDITION

$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$
$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})$
$\mathbf{O}+\mathbf{A}=\mathbf{A}+\mathbf{O}=\mathbf{A}$, where $\mathbf{O}$ is the matrix with all entries equal to zero
$\mathbf{A}+(-\mathbf{A})=(-\mathbf{A})+\mathbf{A}=\mathbf{O}$, where $-\mathbf{A}=(-1)\mathbf{A}$

PROPERTIES OF SCALAR MULTIPLICATION

$\alpha(\mathbf{A}+\mathbf{B}) = \alpha\mathbf{A}+\alpha\mathbf{B}$
$(\alpha+\beta)\mathbf{A} = \alpha\mathbf{A}+\beta\mathbf{A}$
$\alpha(\beta\mathbf{A}) = (\alpha\beta)\mathbf{A}$
$1\,\mathbf{A} = \mathbf{A}$

PROPERTIES OF MATRIX MULTIPLICATION

$\mathbf{A}(\mathbf{B}+\mathbf{C})=\mathbf{AB}+\mathbf{AC}$
$(\mathbf{A}+\mathbf{B})\mathbf{C}=\mathbf{AC}+\mathbf{BC}$
$\mathbf{A}(\mathbf{BC})=(\mathbf{AB})\mathbf{C}$
$(\alpha\mathbf{A})\mathbf{B}=\alpha(\mathbf{AB})$

These properties hold whenever $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are matrices of appropriate sizes so that indicated operations make sense, and α and β are any scalars.

References

[1] S. KUREPA, *Konačno dimenzionalni prostori i primjene*, Tehnička knjiga, Zagreb, 1967.

[2] S. KUREPA, *Uvod u linearnu algebru*, Školska knjiga, Zagreb, 1978.

[3] J. A. THORPE, P. G. KUMPEL, *Elementary Linear Algebra*, CBS College Publishing, Philadelphia, 1984.