8.A The matrix of a linear map

The matrices provide a simple and clear way for the mathematical calculations of linear maps.

Let V and W be vector spaces over a field K with given bases $\underline{v} = (v_j)_{j\in J}$ and $\underline{w} = (w_i)_{i\in I}$ respectively and let $f : V \longrightarrow W$ be a K-linear map. Then f is uniquely determined (by Theorem 5.D.3) by the images

$$f(v_j) = \sum_{i\in I} a_{ij} w_i \;,\; j \in J.$$

These images are, morover determined by the coefficient system

$$\mathfrak{M}(f) := \mathfrak{M}^{\underline{v}}_{\underline{w}}(f) := (a_{ij}) \in K^{I\times J}.$$

8.A.1 Definition Let K be a field and I, J be (indexed) sets.

(1) The K-vector space $K^{I\times J}$ is called the space of $I\times J$-matrices over K and is denoted by

$$M_{I,J}(K)$$

(2) The matrix $\mathfrak{M}^{\underline{v}}_{\underline{w}}(f) = (a_{ij}) \in M_{I,J}(K)$ is called the matrix of the linear map $f : V \to W$ with respect to the bases $\underline{v} = (v_j)_{j\in J}$ and $\underline{w} = (w_i)_{i\in I}$ of V and W, respectively.

Let $\mathfrak{C} = (c_{ij}) \in M_{I,J}(K)$ be a $I\times J$-matrix over K. For a fixed $i \in I$, the J-tuple $(c_{ij})_{j\in J} \in K^J$ is called the <u>i-th row</u> and for a fixed $j \in J$, the I-tuple $(c_{ij})_{i\in I} \in K^I$ is called the <u>j-th column</u> of $\mathfrak{C}$. Therefore c_{ij} is the coefficient of $\mathfrak{C}$ in the i-th row and j-th column.

<u>Note that</u>: If $f : V \longrightarrow W$ is a linear map and $\underline{v} = (v_j)_{j\in J}$, $\underline{w} = (w_i)_{i\in I}$ are bases of V and W, respectively, then for every $j \in J$, the coefficients of the image $f(v_j) = \sum_{i\in I} a_{ij} w_i$ of v_j, is the j-th column of the matrix $\mathfrak{A} = \mathfrak{M}^{\underline{v}}_{\underline{w}}(f) = (a_{ij})$ of f. In particular, in every column of $\mathfrak{A}$ there are only finitely many non-zero coefficients.

For the finite standard indexed sets of the form $I = \{1, \cdots, m\}$ and $J = \{1, \cdots, n\}$, we denote that a matrix $\mathfrak{C} = (c_{ij}) \in M_{I,J}(K)$ by usual reactangular array

$$\mathfrak{C} = \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mn} \end{pmatrix}.$$

We call these matrices as $m \times n$-matrices and the space of all $m \times n$-matrices over K is denoted by $M_{m,n}(K)$. An I-tuple $(c_i)_{i \in I} \in K^I$, when we consider it as a matrix, in general, is denoted by a column-vector, i.e. as a $I \times \{1\}$ matrix with only one column with coefficients $c_i, i \in I$. Further, if $I = \{1, \cdots, m\}$, then we write

$$(c_i)_{1 \leq i \leq m} = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}.$$

A row-vector is denoted as $\{1\} \times I$ matrix, therefore the vector $(c_i) \in K^m$ is written as

$$(c_i)_{1 \leq i \leq m} = (c_1, \cdots, c_m)$$

As we have always remarked, in the matrix of a linear map in every column there are only finitely many non-zero components. Conversely, every matrix $\mathfrak{A} = (a_{ij})$ with this property is the matrix of a linear map $f : V \longrightarrow W$ with respect to the given bases $\underline{v} = (v_j)_{j \in J}$ of V and $\underline{w} = (w_i)_{i \in I}$ of W, namely the map f defined by $f(v_j) = \sum_{i \in I} a_{ij} w_i$. Since the map $f \longmapsto \mathfrak{M}^{\underline{v}}_{\underline{w}}(f)$ is obviously K-linear, we note the following important result for finite dimensional vector spaces:

8.A.2 Theorem Let V and W be finite dimensional vector spaces over a field K with bases $\underline{v} = (v_j)_{j\in J}$ and $\underline{w} = (w_i)_{i\in I}$, respectively. Then the map

$$\mathrm{Hom}_K(V, W) \longrightarrow M_{I,J}(K),\ f \longmapsto \mathfrak{M}_{\underline{w}}^{\underline{v}}(f),$$

which maps every K-linear map $f: V \longrightarrow W$ to its matrix with respect to the given bases $\underline{v}$ and $\underline{w}$, is a K-isomorphism from $\mathrm{Hom}_K(V, W)$ onto $M_{I,J}(K)$. In particular,

$$\mathrm{Dim}_K\ \mathrm{Hom}_K(V, W) = \mathrm{Dim}_K V \cdot \mathrm{Dim}_K W.$$

We now would like to consider the composition of linear maps and their matrices. Let U, V, and W be vector spaces over a field K with the bases $\underline{u} = (u_r)_{r\in R}$, $\underline{v} = (v_j)_{j\in J}$ and $\underline{w} = (w_i)_{i\in I}$, respectively. Let $g: U \longrightarrow V$ and $f: V \longrightarrow W$ be linear maps with $\mathfrak{B} = \mathfrak{M}_{\underline{v}}^{\underline{u}}(g) = (b_{jr}) \in M_{J,R}(K)$ and $\mathfrak{A} = \mathfrak{M}_{\underline{w}}^{\underline{v}}(f) = (a_{ij}) \in M_{I,J}(K)$, respectively.

Then for $r \in R$ and $j \in J$, we have:

$$g(u_r) = \sum_{j\in J} b_{jr} v_j \quad \text{and} \quad f(v_j) = \sum_{i\in I} a_{ij} w_i$$

Therefore it follows that: for the composition $f\circ g: U \longrightarrow W$,

$$(f\circ g)(u_r) = f\Big(\sum_{j\in J} b_{jr} v_j\Big) = \sum_{j\in J} b_{jr} f(v_j)$$
$$= \sum_{j\in J} b_{jr} \sum_{i\in I} a_{ij} w_i$$
$$= \sum_{i\in I}\Big(\sum_{j\in J} a_{ij} b_{jr}\Big) w_i = \sum_{i\in I} c_{ir} w_i,$$

where $c_{ir} := \sum_{j\in J} a_{ij} b_{jr}$ for $(i,r)\in I\times R$.

(Note here that we have used the commutativity of K.) The matrix $\mathfrak{C} = \mathfrak{M}_{\underline{w}}^{\underline{u}}(f\circ g) = (c_{ir}) \in \mathrm{M}_{I,R}(K)$ of the composition $f\circ g$ with respect to the bases $\underline{u}$ and $\underline{w}$ of V and W, respectively, is therefore the product of the matrices $\mathfrak{A}$ and $\mathfrak{B}$ in the sense of the following definition:

8.A.3 Definition Let I, J, R be indexed sets and $\mathfrak{A} = (a_{ij}) \in \mathrm{M}_{I,J}(K)$, $\mathfrak{B} = (b_{jr}) \in \mathrm{M}_{J,R}(K)$ be matrices over the field K. Suppose that in the matrix $\mathfrak{B}$ every column contains only finitely many non-zero elements. Then the matrix $\mathfrak{C} = (c_{ir}) \in \mathrm{M}_{I,R}(K)$, where

$$c_{ir} := \sum_{j\in J} a_{ij} b_{jr}, \qquad (i,r)\in I\times R,$$

is called the product of $\mathfrak{A}$ and $\mathfrak{B}$ and is denoted by $\mathfrak{A}\cdot\mathfrak{B}$.

The product $\mathfrak{A}\mathfrak{B}$ of two matrices $\mathfrak{A}$ and $\mathfrak{B}$ is defined if the indexed set for the columns of the first (factor) matrix and the indexed set for the rows of the second (factor) matrix are equal. In particular, the product for $\mathfrak{A} \in M_{m,n}(K)$ and $\mathfrak{B} \in M_{p,q}(K)$, $m, n, p, q \in \mathbb{N}$ is defined if and only if $n = p$; and in this case their product $\mathfrak{A}\mathfrak{B} \in M_{m,q}(K)$. At the position (i, r) the product matrix $\mathfrak{A}\mathfrak{B}$ is the product c_{ir} of the i-th row vector $\mathfrak{a}_i = (a_{ij})_{j\in J}$ of $\mathfrak{A}$ with the r-th column vector $\mathfrak{b}_r = (b_{jr})_{j\in J}$ of $\mathfrak{B}$.

If $J = \{1, \cdots, n\}$, then

$$c_{ir} = \mathfrak{a}_i\mathfrak{b}_r = (a_{i1}, \cdots, a_{in}) \begin{pmatrix} b_{1r} \\ \vdots \\ b_{nr} \end{pmatrix} = \sum_{j=1}^{n} a_{ij}b_{jr} =$$

$$a_{i1}b_{1r} + \cdots + a_{in}b_{nr}.$$

If all coefficients of $\mathfrak{A}$ and $\mathfrak{B}$ are in $\mathbb{Z}$ or more generally in a ring A, then all coefficients of the product $\mathfrak{A}\mathfrak{B}$ are also in A. The matrix product is defined ~~so~~ that the following assertion holds:

8.A.4 Theorem Let U, V and W be vector spaces over a field K with bases $\underline{u} = (u_r)_{r\in R}$, $\underline{v} = (v_j)_{j\in J}$ and $\underline{w} = (w_i)_{i\in I}$ respectively. Further, let $g : U \to V$, and $f : V \longrightarrow W$ be linear maps. Then the matrix of the composition $f \circ g : U \longrightarrow W$ with respect to the bases $\underline{u}$ and $\underline{w}$ is equal to the product of the matrix of f with respect to the bases $\underline{v}$ and $\underline{w}$ and the matrix of g with respect to the bases $\underline{u}$ and $\underline{v}$. Therefore:

$$\mathcal{M}_{\underline{w}}^{\underline{u}}(f\circ g) = \mathcal{M}_{\underline{w}}^{\underline{v}}(f)\cdot \mathcal{M}_{\underline{v}}^{\underline{u}}(g).$$

Using the matrix of a linear map one can easily compute the components of the image vectors, for example:

8.A.5 Theorem Let $\mathcal{M}_{\underline{w}}^{\underline{v}}(f) = (a_{ij}) \in M_{I,J}(K)$ be the matrix of a linear map $f: V \longrightarrow W$ with respect to the bases $\underline{v} = (v_j)_{j\in J}$ and $\underline{w} = (w_i)_{i\in I}$ of V and W, respectively. For an arbitrary vector $v = \sum_{j\in J} a_j v_j$, the column vector $\mathfrak{b}$ of the components of the image vector $f(v) = \sum_{i\in I} b_i w_i$ is the matrix product

$$\mathfrak{b} = \mathcal{M}_{\underline{w}}^{\underline{v}}(f)\cdot \mathfrak{a}, \text{ where}$$

$\mathfrak{a} = (a_j)_{j\in J}$ is the column vector of the components, a_j, $j\in J$, of v.

Proof We have $f(v) = \sum_{j\in J} a_j f(v_j) = \sum_{j\in J} a_j \left(\sum_{i\in I} a_{ij} w_i\right) = \sum_{i\in I}\left(\sum_{j\in J} a_{ij} a_j\right) w_i = \sum_{i\in I} b_i w_i.$

8.A.6 Example Let $\mathfrak{A} \in M_{I,J}(K)$ be a matrix, in which the columns have only finitely many non-zero components. Then the matrix of the linear map

$$f_{\mathfrak{A}} : K^{(J)} \longrightarrow K^{(I)}$$

defined by $\mathfrak{a} \longmapsto \mathfrak{A}\cdot\mathfrak{a}$, where $\mathfrak{a} \in K^{(J)}$ is interpreted as a column vector, with respect to the standard bases of $K^{(J)}$ and $K^{(I)}$ is the given matrix $\mathfrak{A}$.

If $f : V \longrightarrow W$ is an arbitrary linear map and the matrix of f with respect to the bases $\underline{v} = (v_j)_{j\in J}$ and $\underline{w} = (w_i)_{i\in I}$ of V and W is the matrix $\mathfrak{A}$, then the diagram

$$\begin{tikzcd} K^{(J)} \arrow[r, "f_{\mathfrak{A}}"] \arrow[d, "\wr"'] & K^{(I)} \arrow[d, "\wr"] \\ V \arrow[r, "f"] & W \end{tikzcd}$$

is commutative, where the maps $K^{(J)} \longrightarrow V$ and $K^{(I)} \longrightarrow W$ are the canonical isomorphisms defined by the bases $\underline{v}$ and $\underline{w}$ of V and W, respectively (see Theorem 5.D.5).

The properties of f and $\mathfrak{A}$ correspond to each other. In concrete case under investigation, therefore one put $\mathfrak{A} = f_{\mathfrak{A}}$ instead of f.

The rules for the composition of linear maps given in 5.C.2 applied to the homomorphisms of type $f_{\mathfrak{A}}$ (see Example 8.A.6) to get the following rules for matrices

8.A.7 Let $\mathfrak{A}, \mathfrak{A}' \in M_{I,J}(K)$, $\mathfrak{B}, \mathfrak{B}' \in M_{J,R}(K)$, and $\mathfrak{C} \in M_{R,S}(K)$. Further, let $a, b \in K$. Suppose that the columns of these matrices have only finitely many non-zero components. Then the following rules hold:

(1) $\mathfrak{A}(\mathfrak{B}+\mathfrak{B}') = \mathfrak{A}\mathfrak{B} + \mathfrak{A}\mathfrak{B}'$

$(\mathfrak{A}+\mathfrak{A}')\mathfrak{B} = \mathfrak{A}\mathfrak{B} + \mathfrak{A}'\mathfrak{B}$

(2) $(a\mathfrak{A})(b\mathfrak{B}) = (ab)(\mathfrak{A}\mathfrak{B})$

(3) $(\mathfrak{A}\mathfrak{B}).\mathfrak{C} = \mathfrak{A}(\mathfrak{B}.\mathfrak{C})$

Even if the matrix product $\mathfrak{A}\mathfrak{B}$ is defined, the matrix product $\mathfrak{B}.\mathfrak{A}$ is in general not defined. Similarly, even if both the products $\mathfrak{A}\mathfrak{B}$ and $\mathfrak{B}\mathfrak{A}$ are defined and have the same format, they need not be equal, since the composition of linear maps is not commutative. For example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1\cdot5+2\cdot7 & 1\cdot6+2\cdot8 \\ 3\cdot5+4\cdot7 & 3\cdot6+4\cdot8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}, \text{ but}$$

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}.$$

Let V be a K-vector space with basis $\underline{v} = (v_i)_{i \in I}$. For simplicity let us assume that V is finite dimensional and hence I is finite. For a linear operator $f : V \longrightarrow V$ the $I \times I$ matrix $\mathfrak{M}_{\underline{v}}^{\underline{v}}(f)$ is called the matrix of f with respect to the basis $\underline{v}$; its rows and columns are indexed by the same indexed set I. Such matrices are called square matrices. The space of square $I \times I$-matrices over K is denoted by $M_I(K)$. In particular, $M_n(K)$ is the space of square $n \times n$ matrices over K. The matrix $\mathfrak{M}_{\underline{v}}^{\underline{v}}(id_V)$ of the identity map $id_V : V \longrightarrow V$ is the well-known unit matrix $\mathfrak{E}_I := (\delta_{ij}) \in M_I(K)$, where $\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$ is the Kronecker's symbol. For $I = \{1, \cdots, n\}$ we denote the unit matrix by $\mathfrak{E}_n$. Therefore

$$\mathfrak{E}_n := \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots \\ & & & 1 \end{pmatrix} \in M_n(K)$$

For arbitrary matrices $\mathfrak{A} \in M_{I,J}(K)$ and $\mathfrak{B} \in M_{J,I}(K)$, we have $\mathfrak{E}_I \mathfrak{A} = \mathfrak{A}$ and $\mathfrak{B} \cdot \mathfrak{E}_I = \mathfrak{B}$. It follows immediately from 8.A.7 that:

8.A.8 Theorem Let K be a field and $n \in \mathbb{N}$ (respectively I be a finite set). Then $M_n(K)$ (respectively $M_I(K)$) is a K-algebra with the unit-matrix $\mathfrak{E}_n$ (respectively $\mathfrak{E}_I$) as unit-element and $\operatorname{Dim}_K M_n(K) = n^2$ (respectively $\operatorname{Dim}_K M_I(K) = \#I^2$).

The map $f \longmapsto \mathfrak{M}_{\underline{v}}^{\underline{v}}(f)$ from $\operatorname{End}_K V$ onto $M_I(K)$ is not only an isomorphism of K-vector spaces, but also (by 8.A.2 and 8.A.4) is compatible with the multiplication of the K-algebras $\operatorname{End}_K V$ and $M_I(K)$, i.e. we have:

8.A.9 Theorem Let V be a finite dimensional K-vector space with the basis $\underline{v} = (v_i)_{i \in I}$. Then the map

$$\operatorname{End}_K V \longrightarrow M_I(K), \; f \longmapsto \mathfrak{M}_{\underline{v}}^{\underline{v}}(f),$$

which maps every K-linear operator $f: V \to V$ to its matrix with respect to the basis $\underline{v}$, is a K-algebra isomorphism of $\operatorname{End}_K V$ onto $M_I(K)$. In particular,

$$\operatorname{Dim}_K \operatorname{End}_K V = \operatorname{Dim}_K M_I(K) = (\#I)^2 = (\operatorname{Dim}_K V)^2.$$

Let A be an arbitrary K-algebra with basis

$\underline{v} = (v_i)_{i\in I}$. For $(a_i), (b_i) \in K^{(I)}$, by the general distributive law, we have:

$$\Big(\sum_{i\in I} a_i v_i\Big)\Big(\sum_{j\in I} b_j v_j\Big) = \sum_{(i,j)\in I\times J} a_i b_j \, v_i \cdot v_j \,.$$

Therefore the multiplication in the K-algebra A is uniquely determined by the products $v_i \cdot v_j$, $i, j \in I$. This can be represented in the given basis and write

$$v_i v_j = \sum_{k\in I} a_{ij}^{(k)} v_k \,,$$

the coefficients $a_{ij}^{(k)} \in K$, $i, j, k \in I$, are called the <u>structure-constants</u> of the K-algebra A with respect to the basis $\underline{v} = (v_i)_{i\in I}$.

For $i, j \in I$, we denote by $\mathcal{E}_{ij} \in M_I(K)$, the matrix whose (i,j)-th entry (the coefficient in the i-th row and j-th column) is 1 and all other entries are 0. Then $\mathcal{E}_{ij}$, $(i,j)\in I\times I$, is the standard basis of $M_I(K) = K^{I\times I}$ and from the equalities

$$\mathcal{E}_{ij}\,\mathcal{E}_{rs} = \delta_{jr}\,\mathcal{E}_{is} \,,\quad (i,j),(r,s) \in I\times I,$$

it follows immediately that the structure-constants of the K-algebra $M_I(K)$ with respect to the basis $\mathcal{E}_{ij}$, $(i,j)\in I\times I$ are 1 and 0.