8.A The matrix of a Linear map

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

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

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.

8A/2

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

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

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

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

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 $f \longmapsto \mathfrak{M}_{\underline{w}}^{\underline{v}}(f)$ is obviously K-linear, we note the following important result for finite dimensional vector spaces:

8A/4

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

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

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 $\mathfrak{B} = \mathcal{M}_{\underline{v}}^{\underline{u}}(g) = (b_{jr}) \in M_{J,R}(K)$ and $\mathfrak{A} = \mathcal{M}_{\underline{w}}^{\underline{v}}(f) = (a_{ij}) \in M_{I,J}(K)$, 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 $\mathfrak{C} = \mathfrak{M}^{\underline{u}}_{\underline{w}}(f\circ g) = (c_{ir}) \in M_{I,R}(K)$ 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 $\mathfrak{A} = (a_{ij}) \in M_{I,J}(K)$, $\mathfrak{B} = (b_{jr}) \in M_{J,R}(K)$ 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 $\mathfrak{C} = (c_{ir}) \in M_{I,R}(K)$, 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 $f(v) = \sum_{j \in J} a_j f(v_j) = \sum_{j \in J} a_j \left(\sum_{i \in I} a_{ij} w_i\right)$
$= \sum_{i \in I} \left(\sum_{j \in J} a_{ij} a_j\right) w_i = \sum_{i \in I} b_i w_i.$

8A/8

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,

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

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

8A|10

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 $\mathfrak{M}^{\underline{v}}_{\underline{v}}(\mathrm{id}_V)$ of the identity map $\mathrm{id}_V : V \longrightarrow V$ 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 E_n (respectively 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 $f : V \longrightarrow V$ 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

8A/12

$\underline{v} = (v_i)_{i\in I}$. 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 structure-constants 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.

For example, in $M_2(K)$

$$\mathcal{E}_{12}\mathcal{E}_{21} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \mathcal{E}_{11} \quad \text{and}$$

$$\mathcal{E}_{12}\mathcal{E}_{11} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0 .$$

More generally, for the unit matrix $\mathcal{E}_I \in M_I(K)$ we have the equation

$$\mathcal{E}_I = \sum_{i\in I} \mathcal{E}_{ii}$$

The K-algebra isomorphism $\mathrm{End}_K V \longrightarrow M_I(K)$ in 8.A.9, maps the unit group of $\mathrm{End}_K V$, i.e. the automorphism group $\mathrm{GL}_K V = \mathrm{Aut}_K V$ of V, onto the unit group of $M_I(K)$. This group is called the <u>group of invertible matrices in $M_I(K)$</u> and is denoted by $\mathrm{GL}_I(K)$; in the case $I = \{1, \dots, n\}$, it is also called the <u>general linear group</u> $\mathrm{GL}_n(K)$. The matrices in $\mathrm{GL}_I(K)$ respectively in $\mathrm{GL}_n(K)$ are called <u>invertible matrices</u>.

We repeat: A square matrix $\mathfrak{A} \in M_I(K)$ (with finite indexed set I) is invertible if and only if there exists a matrix $\mathfrak{A}' \in M_I(K)$ with $\mathfrak{A}\mathfrak{A}' = \mathfrak{A}'\mathfrak{A} = \mathcal{E}_I$. Moreover, matrix $\mathfrak{A}'$ is uniquely determined and is called the <u>inverse matrix</u> of $\mathfrak{A}$ and is denoted by $\mathfrak{A}^{-1}$.

 8A/14

Therefore

(1) $\mathfrak{A}\mathfrak{A}^{-1} = \mathfrak{A}^{-1}\mathfrak{A} = \mathfrak{E}_I$

(2) $(\mathfrak{A}^{-1})^{-1} = \mathfrak{A}$

(3) $\mathfrak{E}_I^{-1} = \mathfrak{E}_I$

Further, the product of invertible matrices $\mathfrak{A}, \mathfrak{B} \in \mathrm{GL}_I(K)$ is again invertible and

(4) $(\mathfrak{A}\mathfrak{B})^{-1} = \mathfrak{B}^{-1}\mathfrak{A}^{-1}$.

The algebra isomorphism $\mathrm{End}_K V \longrightarrow \mathrm{M}_I(K)$ in 8.A.9

8.A.10 Theorem An operator $f \in \mathrm{End}_K V$ is an automorphism of V if and only if the corresponding matrix $\mathfrak{M}_{\underline{v}}^{\underline{v}}(f) \in \mathrm{M}_I(K)$ is invertible. Moreover, in this case

$$\mathfrak{M}_{\underline{v}}^{\underline{v}}(f^{-1}) = \left(\mathfrak{M}_{\underline{v}}^{\underline{v}}(f)\right)^{-1}.$$

We remark here that $\mathrm{GL}_1(K)$ is the multiplicative group $K^\times$ of K. For $\#I \geq 2$, the group $\mathrm{GL}_I(K)$ is not commutative. The matrices $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ are invertible with inverses

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$ Further,

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

If $\#I \geq 2$, then there are non-zero matrices in $\mathrm{M}_I(K)$ which are not invertible, since there are non-zero operators on two dimensional vector space which are not invertible, for example,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = 0, \text{ i.e. } \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq 0$$

is certainly not invertible. The same example, once more shows that the product of two non-zero matrices can be 0.

If a matrix $\mathfrak{A} \in M_I(K)$ is invertible and if $\mathfrak{A}\mathfrak{B} = 0$ for a (not necessary square) matrix $\mathfrak{B} \in M_{I,J}(K)$, then by multiplying by $\mathfrak{A}^{-1}$ on the left, we get $\mathfrak{B} = 0$. Analogously, from $\mathfrak{B}.\mathfrak{A} = 0$ for $\mathfrak{B} \in M_{J,I}(K)$ it follows that $\mathfrak{B} = 0$.

From Theorem 5.D.6 the following important assertions on the invertible matrices are immediate:

8.A.11 Theorem Let I be a finite set and let K be a field. For a square matrix $\mathfrak{A} \in M_I(K)$ the following statements are equivalent:

(1) $\mathfrak{A}$ is invertible

(2) $\mathfrak{A}$ has a left inverse $\mathfrak{B} \in M_I(K)$ in $M_I(K)$, i.e. $\mathfrak{B}.\mathfrak{A} = \mathfrak{E}_I$.

(3) $\mathfrak{A}$ has a right inverse $\mathfrak{C} \in M_I(K)$ in $M_I(K)$, i.e. $\mathfrak{A}.\mathfrak{C} = \mathfrak{E}_I$.

Moreover, if (2) (respectively (3)) holds, then necessarily $\mathfrak{B} = \mathfrak{A}^{-1}$ (respectively $\mathfrak{C} = \mathfrak{A}^{-1}$).

 8A/16

Finally, we shall describe how the matrix of a linear map depends on the choice of bases. For this, we first remark:

8.A.12 Lemma Let $\underline{v}' = (v_i')_{i\in I}$ be a basis of the finite dimensional K-vector space V and $\underline{v} = (v_i)_{i\in I}$ be a family of vectors in V. For every $j \in I$, let $v_j = \sum_{i\in I} b_{ij} v_i'$ with $b_{ij} \in K$. Then $\underline{v}$ is a basis of V if and only if the coefficient matrix $\mathfrak{B} = (b_{ij}) \in M_I(K)$ is invertible.

Proof Note that $\underline{v}$ is a basis of V if and only if the linear map $h : V \longrightarrow V$ defined by $v_i' \longmapsto v_i$, is an automorphism of V. Since $\mathfrak{B} = (b_{ij}) = \mathfrak{M}^{\underline{v}'}_{\underline{v}'}(h)$ is the matrix of h w.r. to the basis $\underline{v}'$, the assertion follows.

In the situation of 8.A.12, now let $\underline{v}$ be a basis of V. Then the matrix $\mathfrak{B} = (b_{ij}) \in GL_I(K)$ is called the <u>transition matrix from the basis $\underline{v}$ to the basis $\underline{v}'$</u>.

Note that one can also interpret $\mathfrak{B}$ as the matrix $\mathfrak{M}^{\underline{v}}_{\underline{v}'}(\mathrm{id}_V)$ of the identity map

w. r. to the (different) bases $\underline{v}$ resp. $\underline{v}'$ of V. The transition matrix from the basis $\underline{v}'$ to the basis $\underline{v}$, is the inverse matrix $\mathfrak{B}^{-1}$ of $\mathfrak{B}$, see 8.A.4.

The coordinates of a vector are transformed under the change of basis as follows:

8.A.13 Lemma Let $\underline{v} = (v_i)_{i\in I}$ and $\underline{v}' = (v_i')_{i\in I}$ be bases of the finite dimensional K-vector space V and $\mathfrak{B} = (b_{ij}) \in GL_I(K)$ with

$$v_j = \sum_{i\in I} b_{ij} v_i', \quad j \in I,$$

be the corresponding transition matrix. For the (column-vectors) coordinate tuples $\underline{a} = (a_i)_{i\in I}$ (resp. $\underline{a}' = (a_i')_{i\in I}$) of a vector

$$x = \sum_{i\in I} a_i v_i \quad \left(\text{resp. } \sum_{i\in I} a_i' v_i'\right),$$

we have $\underline{a}' = \mathfrak{B}\underline{a}$ (resp. $\underline{a} = \mathfrak{B}^{-1}\underline{a}'$.

Proof Since $\sum_{i\in I} a_i' v_i' = x \sum_{j\in I} a_j v_j =$ $\sum_{i\in I}\left(\sum_{j\in I} b_{ij} a_j\right) v_i'$, we have $a_i' = \sum_{j\in I} b_{ij} a_j$ for all $i \in I$.

8A/18

For the change of basis the matrix of a linear map it follows that:

8.A.14 Change of Basis Let V, W be finite dimensional K-vector spaces with bases $\underline{v} = (v_j)_{j\in J}$, $\underline{v}' = (v'_j)_{j \in J}$ resp. $\underline{w} = (w_i)_{i\in I}$, $\underline{w}' = (w'_i)_{i\in I}$. Let $\mathfrak{S} = M^{\underline{v}}_{\underline{v}'}(id_V)$ and $\tau = M^{\underline{w}}_{\underline{w}'}(id_W)$ be the transition matrices from $\underline{v}$ to $\underline{v}'$ and from $\underline{w}$ to $\underline{w}'$, respectively. Then for the matrices $M^{\underline{v}}_{\underline{w}}(f)$ and $M^{\underline{v}'}_{\underline{w}'}(f)$ of a linear map $f: V \longrightarrow W$ with resp. to bases $\underline{v}, \underline{w}$, resp. $\underline{v}', \underline{w}'$, we have:

$$M^{\underline{v}'}_{\underline{w}'}(f) = \tau \cdot M^{\underline{v}}_{\underline{w}}(f) \cdot \mathfrak{S}^{-1}$$

Proof By 8.A.4

$$M^{\underline{v}'}_{\underline{w}'}(f) \cdot \mathfrak{S} = M^{\underline{v}'}_{\underline{w}'}(f)\, M^{\underline{v}}(id_V) = M^{\underline{v}}_{\underline{w}'}(f)$$

and

$$\tau \cdot M^{\underline{v}}_{\underline{w}}(f) = M^{\underline{w}}_{\underline{w}'}(id_W) \cdot M^{\underline{v}}_{\underline{w}}(f) = M^{\underline{v}}_{\underline{w}'}(f).$$

8.A.15 Corollary Let V be a finite dimensional K-vector space with bases $\underline{v} = (v_i)_{i\in I}$ and $\underline{v}' = (v'_i)_{i\in I}$, $\mathfrak{S} = M^{\underline{v}}_{\underline{v}'}(id_V)$ be the transition matrix from $\underline{v}$ to $\underline{v}'$. Then for the matrices $\mathfrak{M}^{\underline{v}}_{\underline{w}}(f)$ and $\mathfrak{M}^{\underline{v}'}_{\underline{w}'}(f)$ of a linear map $f: V \longrightarrow W$ with resp. to bases $\underline{v}, \underline{w}$ and $\underline{v}', \underline{w}'$, respectively, we have:

$$\mathfrak{M}^{\underline{v}'}_{\underline{w}'}(f) = \mathfrak{B}\, \mathfrak{M}^{\underline{v}'}_{\underline{w}'}(f)\, \mathfrak{B}^{-1}$$

Therefore the matrices (and operators on finite dimensional K-vector space V) are similar, in the sense of the following definition:

8.A.16 Definition Let I be a finite set. The matrices $\mathfrak{A}, \mathfrak{A}' \in M_I(K)$ are similar if there exists an invertible matrix $\mathfrak{B} \in GL_I(K)$ such that

$$\mathfrak{A}' = \mathfrak{B} . \mathfrak{A} . \mathfrak{B}^{-1}$$

Note that similar matrices describe the same operator by the choice of appropriate bases.

The "similarity" is an equivalence relation on $M_I(K)$. Its equivalence classes are called similarity classes of matrices.

The main aim of Linear Algebra is to find the simplest possible representatives

 8A/20

for the similarity class of matrices, i.e. classification of operators on finite dimensional K-vector space or characterisation of an operator by using a simplest possible matrix (w.r.to appropriate basis). This will be done in the next section.

Further important operation on matrices is the transpose.

8.A.17 Definition For a matrix $\mathfrak{A} = (a_{ij}) \in M_{I,J}(K)$, the matrix

$$^t\mathfrak{A} = (b_{ji}) \in M_{J,I}(K) \text{ with } b_{ji} := a_{ij},$$

$j \in J$, $i \in I$, is called the transpose (matrix) of $\mathfrak{A}$.

The i-th column of ${}^t\mathfrak{A}$ is the i-th row of $\mathfrak{A}$ and the j-th row of ${}^t\mathfrak{A}$ is the j-th column of $\mathfrak{A}$.

For example: $^t\begin{pmatrix} 1 & -1 & 3 \\ -5 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -5 \\ -1 & 2 \\ 3 & 0 \end{pmatrix}$

The transpose matrix ${}^t\mathfrak{A} \in M_{n,m}(K)$ is flapping around the main diagonal of the matrix $\mathfrak{A}$. For the unit-matrix $\mathfrak{E}_I \in M_I(K)$, ${}^t\mathfrak{E}_I = \mathfrak{E}_I$. Further, we have the following computational-rules (for transpose):

8.A.18 Proposition For matrices $\mathfrak{A}, \mathfrak{A}' \in M_{I,J}(K)$ and $\mathfrak{B} \in M_{J,R}(K)$, and $a \in K$:

(1) ${}^t({}^t\mathfrak{A}) = \mathfrak{A}$ (2) ${}^t(\mathfrak{A}+\mathfrak{A}') = {}^t\mathfrak{A} + {}^t\mathfrak{A}'$
(3) ${}^t(a\mathfrak{A}) = a\,{}^t\mathfrak{A}$ (4) ${}^t(\mathfrak{A}\mathfrak{B}) = {}^t\mathfrak{B}\cdot{}^t\mathfrak{A}$.
(for (4) the index sets I, J, R are finite)

The transpose is compatible with the inverse:

8.A.19 Corollary Let I be a finite set and $\mathfrak{A} \in GL_I(K)$ be an $I\times I$ invertible matrix. Then the transpose ${}^t\mathfrak{A}$ of $\mathfrak{A}$ is also invertible and $({}^t\mathfrak{A})^{-1} = {}^t(\mathfrak{A}^{-1}) =: {}^t\mathfrak{A}^{-1}$.

Proof By 8.A.18(4) ${}^t\mathfrak{A}\,{}^t(\mathfrak{A}^{-1}) = {}^t(\mathfrak{A}^{-1}\mathfrak{A}) = {}^t\mathfrak{E}_I = \mathfrak{E}_I$ and analogously ${}^t(\mathfrak{A}^{-1})\,{}^t\mathfrak{A} = {}^t(\mathfrak{A}\mathfrak{A}^{-1}) = {}^t\mathfrak{E}_I = \mathfrak{E}_I$.

For $\mathfrak{A} \in GL_I(K)$, the matrix ${}^t\mathfrak{A}^{-1} = ({}^t\mathfrak{A})^{-1} = {}^t(\mathfrak{A}^{-1}) \in GL_I(K)$ is called the <u>contra-gradient matrix of $\mathfrak{A}$</u>

The transpose of the matrix of a linear map f is also described by the dual map f^* of f:

8.A.20 Theorem Let V and W be finite dimensional K-vector spaces with bases $\underline{v} = (v_j)_{j\in J}$

and $\underline{w} = (w_i)_{i\in I}$, resp. Further, let $V \xrightarrow{f} W$ be a K-linear map with the matrix $\mathfrak{a} = \mathcal{M}_{\underline{w}}^{\underline{v}}(f)$. Then the matrix of the dual map $f^*: W^* \longrightarrow V^*$ w.r. to the dual bases $w^* = (w_i^*)_{i\in I}$ and $v^* = (v_j^*)_{j\in J}$ is equal to the transpose ${}^t\mathfrak{a}$ of $\mathfrak{a}$, i.e.

$$\mathcal{M}_{\underline{v}^*}^{\underline{w}^*}(f^*) = {}^t\left(\mathcal{M}_{\underline{w}}^{\underline{v}}(f)\right).$$

Proof Let $\mathfrak{a} = (a_{ij}) \in M_{I,J}(K)$, so $f(v_j) = \sum_{i\in I} a_{ij} w_i$, $j \in J$. Further, let $\mathcal{M}_{\underline{v}^*}^{\underline{w}^*}(f^*) = (b_{ji}) \in M_{J,I}(K)$. Therefore $f^*(w_i^*) = \sum_{j\in J} b_{ji} v_j^*$, $i \in I$. Then for all $j \in J$ and $i \in I$, we have

$$b_{ji} = \sum_{k\in J} b_{ki} v_k^*(v_j) = \left(f^*(w_i^*)\right)(v_j) =$$

$$w_i^*(f(v_j)) = w_i^*\left(\sum_{r\in I} a_{rj} w_r\right) = a_{ij}.$$