8.C. Elementary Matrices

In this subsection, we consider number of results regarding the representations of arbitrary matrices as product of particular simple matrices. In the case of invertible matrices with entries in a field K, we use these product decompositions for the relatively simple computations of the inverse.
Also for theoretic considerations, these representations are often very useful.

8.C.1 Definition Let $m, n \in \mathbb{N}$. A matrix $\mathfrak{A} = (a_{ij}) \in M_{m,n}(K)$ is called a diagonal matrix if $a_{ij} = 0$ for all i, j with $1 \leq i \leq m$, $1 \leq j \leq n$ and $i \neq j$.
Such a diagonal matrix $\mathfrak{A}$ is denoted by
$\mathrm{Diag}(a_1, \cdots, a_k)$, where $a_i := a_{ii}$, for $i = 1, \cdots, k := \mathrm{Min}(m, n)$. The entries

of $\mathfrak{A}$ other than on the (main)-diagonal $(1,1), \cdots, (k,k)$-th entries are all 0.

Note that: $\mathrm{Diag}(a_1, \cdots, a_k) =$

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_k & 0 & \cdots & 0 \end{pmatrix} \qquad \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_k \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$k = m \leq n \qquad\qquad k = n \leq m$$

Analogously, a square matrix $\mathfrak{A} = (a_{ij}) \in \mathrm{M}_I(K)$ with arbitrary index-set I is called a <u>diagonal matrix</u> if $a_{ij} = 0$ for all $i, j \in I$, with $i \neq j$. It is also denoted by

$$\mathrm{Diag}(a_i \mid i \in I),$$

where $a_i := a_{ii}$ for all $i \in I$.

<u>8.C.2 Example</u> Let v_i, $i \in I$, be a K-basis of a K-vector space V. The matrix $\mathfrak{M}^{\underline{v}}_{\underline{v}}(f)$ of a K-linear

map $f: V \longrightarrow V$ w.r. to the basis $\underline{v} = (v_i)_{i \in I}$ is the diagonal matrix $\mathrm{Diag}(a_i \mid i \in I)$ if and only if $f(v_i) = a_i v_i$ for all $i \in I$, i.e. every basis vector v_i is mapped to the scalar multiple of itself, i.e. it is an eigenvector of f in the sense of 11.A.2. Such an operator is bijective, i.e. the corresponding diagonal matrix $\mathrm{Diag}(a_i \mid i \in I)$ is invertible if and only if $a_i \neq 0$ for all $i \in I$. In this case

$$\mathrm{Diag}(a_i \mid i \in I)^{-1} = \mathrm{Diag}(a_i^{-1} \mid i \in I).$$

The matrix of the homothecy $a \cdot \mathrm{id}_V : V \longrightarrow V$, $v \longmapsto av$, with resp. to every basis $\underline{v}$ of V is the diagonal matrix with all diagonal entries a.

In Subsection 11.B we will investigate

When an operator is diagonalisable, i.e. if is described by a diagonal matrix w.r.to a suitable basis.

8.C.3 Definition Let $m, n \in \mathbb{N}$. A matrix $\mathfrak{A} = (a_{ij}) \in M_{m,n}(K)$ is called an upper (resp. lower) triangular if $a_{ij} = 0$ for all i, j with $1 \leq i \leq m$, $1 \leq j \leq n$ and $i > j$ (resp. $i < j$), i.e. $\mathfrak{A}$ is of the form:

$$\begin{pmatrix} a_{11} & * & \cdots & * & * & \cdots & * \\ 0 & a_{22} & \cdots & * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{mm} & * & \cdots & * \end{pmatrix} \text{ resp. } \begin{pmatrix} a_{11} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ * & a_{22} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & a_{mm} & 0 & \cdots & 0 \end{pmatrix}$$

$(m \leq n)$

$$\begin{pmatrix} a_{11} & * & \cdots & * \\ 0 & a_{22} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \text{ resp. } \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ * & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & a_{nn} \\ * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{pmatrix}$$

$(n \leq m)$

where at the places $*$ there are arbitrary elements of K. Therefore, in the case of upper (resp. lower) triangular matrix all elements below (resp. above) the main diagonal are 0.

8.C.4 Example Let $\underline{v} = (v_1, \ldots, v_n)$ be a K-basis of an n-dimensional K-vector space V and $V_i := Kv_1 + \cdots + Kv_i$, $i = 1, \ldots, n$. Then a K-linear operator $f : V \longrightarrow V$ on V is described by an upper triangular matrix w.r. to the basis $\underline{v}$ if and only if $f(V_i) \subseteq V_i$ for all $i = 1, \ldots, n$, i.e. if and only if the terms of the flag

$$0 \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_n = V$$

are mapped into themselves by f, i.e. they are f-invariant. Therefore, f is an automorphism if and only if the diagonal entries $a_{11}, \ldots, a_{nn}$ of the corresponding triangular matrix are all $\neq 0$.

The inverse map is also described by an upper triangular matrix with resp. to the given basis $\underline{v} = (v_1, \dots, v_n)$.

The set $UT_n(K)$ of $n \times n$-upper triangular matrices form a K-subalgebra of the matrix algebra $M_n(K)$, i.e. $UT_n(K)$ is a K-subspace of $M_n(K)$, closed w.r. to the matrix multiplication and the unit element $E_n \in UT_n(K)$. This can be seen lucidly if the triangular matrices in $UT_n(K)$ are considered as K-linear operators on K^n which leaves a (fixed) flag in K^n invariant.

8.C.5 Example (Nilpotent operators)

Nilpotent operators on a finite dimensional K-vector space are described by an upper triangular matrices w.r.to appropriate bases.

Recall that an element x in an K-algebra A is called nilpotent if $x^m = 0$ for some natural number $m \in \mathbb{N}$.

If $x \in A$ is nilpotent, then the smallest natural number $m_0 \in \mathbb{N}$ with $x^{m_0}=0$, is called the <u>nilpotent-degree</u> of x.

An operator $f : V \longrightarrow V$ is nilpotent if and only if (as an element of the K-algebra $\mathrm{End}_K V$) its matrix w.r.r. to one (and hence every) basis of V is nilpotent. With this we have:

<u>Let V be n-dimensional K-vector space and $f : V \longrightarrow V$ be a K-linear operator.</u> Then the following statements are equivalent:

(i) f is nilpotent.

(ii) $f^n = 0$, i.e. the nilpotent-degree of $f \le n$.

(iii) There exists a flag $\{0\} = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$ of V with $f(V_i) \subseteq V_{i-1}$ for all $i = 1, \ldots, n$.

(iv) There exists a basis $\underline{v} = (v_1, \ldots, v_n)$ of V such that the matrix $\mathfrak{M}_{\underline{v}}^{\underline{v}}(f)$ of f w.r.to $\underline{v}$ is upper-triangular with all entries on the main diagonal are 0.

Proof We will prove the implications:

$$\begin{array}{ccc} (i) & \Longleftrightarrow & (ii) \\ & \searrow & \Uparrow \\ & & (iii) \Longleftrightarrow (iv) \end{array}$$

Note that (ii) $\Rightarrow$ (i) is trivial.

(i) $\Rightarrow$ (ii) and (iii): Proof by induction on n. If $n>0$ and if f is nilpotent, then f is not an automorphism and hence f is not surjective (see Theorem 5.D.6). Put $V_n = V$ and V_{n-1} be an arbitrary 1-codimensional subspace in V with $\operatorname{Im} f \subseteq V_{n-1}$. Then, clearly $f(V_n) \subseteq V_{n-1}$ and $f|V_{n-1}$ is a nilpotent operator on V_{n-1}. By induction hypothesis $(f|V_{n-1})^{n-1} = 0$ and hence $f^n = 0$. Further, by induction hypothesis there exists a flag $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_{n-1}$ of V_{n-1} with $f(V_i) \subseteq V_{i-1}$ for all $i = 1, \ldots, n-1$.

(iii) $\Rightarrow$ (iv): Let $\underline{v} = (v_1, \ldots, v_n)$ be a K-basis of V with $Kv_1 + \cdots + Kv_i = V_i$ for $i = 0, \ldots, n$. Since $f(V_i) \subseteq V_{i-1}$, $f(v_j) = \sum_{i=1}^{j-1} a_{ij} v_i$ for all $j = 1, \ldots, n$, i.e. the matrix $\mathfrak{M}_{\underline{v}}^{\underline{v}}(f)$ of f w.r.to $\underline{v}$ is of the required form.

(iv) $\Rightarrow$ (iii): Since the flag V_i, $i=0,\dots,n$, with $V_i = Kv_1 + \cdots + Kv_i$ has the property (iii), it follows that $f^i(V_n) \subseteq V_{n-i}$, $i=1,\dots,n$. In particular, $f^n(V)=0$, i.e. $f^n=0$.

Corollary If $\mathfrak{A} \in M_n(K)$ is a nilpotent matrix, then $\mathfrak{A}^n=0$ and $\mathfrak{A}$ is similar to an upper triangular matrix with all main-diagonal entries are 0. The unipotent matrix $\mathfrak{B} = \mathfrak{E}_n + \mathfrak{A}$ is similar to an upper triangular matrix with all entries on the main diagonal are 1.

8.C.6 Example (Elementary Operations)

The elementary row- and column operations on matrices can be realized by matrix multiplications on left and right with appropriate matrices.

Recall that $\mathcal{E}_n \in M_n(K)$ denote the unit-matrix and for $1 \leq r, s \leq n$, $\mathcal{E}_{rs} \in M_n(K)$ whose (r,s)-th entry is 1 and all other entries are 0.

We define:

8.C.7 Definition An elementary matrix in $M_n(K)$ is a matrix of the form:

$$\mathcal{B}_{rs}(a) := \mathcal{E}_n + a\,\mathcal{E}_{rs},$$

where $a \in K$ and $1 \leq r, s \leq n$, $r \neq s$.

Note that:

(1) In the matrices $\mathcal{B}_{rs}(a)$ and the unit matrix $\mathcal{E}_n$, only (r,s)-th entries are different, namely 0 and a, resp.
$\mathcal{B}_{rs}(0) = \mathcal{E}_n$

(2) If $r < s$, then $\mathcal{B}_{rs}(a)$ is upper triangular.

(3) If $r > s$, then $\mathcal{B}_{rs}(a)$ is lower-triangular.

(4) $\mathcal{B}_{rs}(a)$ is invertible with the inverse $\mathcal{B}_{rs}(a)^{-1} = \mathcal{B}_{rs}(-a)$.

More generally, for all $a, b \in K$,

$\mathcal{B}_{rs}(a)\,\mathcal{B}_{rs}(b) = (\mathcal{E}_n + a\mathcal{E}_{rs})(\mathcal{E}_n + b\mathcal{E}_{rs})$

$= \mathcal{E}_n + (a+b)\,\mathcal{E}_{rs} + ab\,\mathcal{E}_{rs}^2 = \mathcal{E}_n + (a+b)\mathcal{E}_{rs}$

$= \mathcal{B}_{rs}(a+b)$, since $r \neq s$, $\mathcal{E}_{rs}\mathcal{E}_{rs} = 0$.

(5) Let $\mathfrak{A} \in M_{m,n}(K)$ be arbitrary.

Then $\mathfrak{A} \cdot \mathcal{B}_{rs}(a)$ is the matrix obtained from $\mathfrak{A}$ by replacing s-th column of $\mathfrak{A}$ by a-times r-th column to the s-th column of $\mathfrak{A}$, i.e. s-th column of $\mathfrak{A} \cdot \mathcal{B}_{rs}(a)$ is

$C_s + aC_r$, where $C_s = (a_{is})_{i \in I}$ and $C_r = (a_{ir})_{i \in I}$ are the s-th and r-th columns of $\mathfrak{A}$, respectively. All other columns of $\mathfrak{A} \cdot \mathcal{B}_{rs}(a)$ are same as those of $\mathfrak{A}$.

Similarly, for $\mathcal{B}_{rs}(a) \in M_m(K)$,

$\mathcal{B}_{rs}(a)\,\mathfrak{A}$ is the matrix obtained from $\mathfrak{A}$ by replacing r-th row by adding a-multiple of s-th row to the r-th row of $\mathfrak{A}$, i.e. r-th row of $\mathcal{B}_{rs}(a) \cdot \mathfrak{A}$ is $R_r + aR_s$, where $R_r = (a_{rj})_{j \in I}$ and $R_s = (a_{sj})_{j \in I}$ are r-th and s-th rows of $\mathfrak{A}$ respectively. All other rows of $B_{rs}(a)\,\mathfrak{A}$ are same as those of $\mathfrak{A}$.

(b) Also, the other elementary operations are represented by similar simple matrices. For $a \in K^\times$ and $1 \leq r \leq n$, let $E_r(a) := E_n + (a-1)E_{rr}$ is the matrix in $M_n(K)$ which differs from the unit-matrix E_n only at the (r,r)-th entry on the main diagonal which is a.

For $\mathfrak{A} \in M_{m,n}(K)$, the matrix $\mathfrak{A}\,E_r(a)$ is obtained by multiplying r-th column by a and all other columns are same as those of $\mathfrak{A}$. Similarly, for $E_r(a) \in M_m(K)$ analogously, the matrix $E_r(a)\cdot\mathfrak{A}$ is obtained from $\mathfrak{A}$ by multiplying r-th row of $\mathfrak{A}$ by a and all other rows are same as those of $\mathfrak{A}$.

The matrices $E_r(a)$ are invertible, with the inverse $E_r(a)^{-1} = E_r(a^{-1})$.

Let $\sigma \in \mathfrak{S}_n$ be a permutation of the indexed set $\mathbb{N}_n^* = \{1, \cdots, n\}$. Then the matrix $\mathfrak{P}_\sigma = (\delta_{i\sigma(j)}) \in M_n(K)$ is called the <u>permutation matrix corresponding to the permutation</u> σ, where δ_{ij} is the Kronecker's symbol.

The i-th row of $\mathfrak{P}_\sigma$ is $(\delta_{i\sigma(j)})_{j\in I}$ and the only non-zero entry in it is in the $\sigma^{-1}(i)$-th column, i.e. i-th row of $\mathfrak{P}_\sigma$ is the $\sigma^{-1}(i)$-th entry of the unit matrix E_n.

Analogously, j-th column of $\mathfrak{P}_\sigma$ is $(\delta_{\sigma^{-1}(i)\,j})_{i\in I}$ and the only non-zero entry in it is in the $\sigma(j)$-th column, i.e. j-th column of $\mathfrak{P}_\sigma$ is the $\sigma(j)$-th column of the unit matrix E_n.

Therefore: the matrix $\mathfrak{P}_\sigma$ is the matrix of the automorphism

$K^n \xrightarrow{\sigma} K^n$, $e_j \longmapsto e_{\sigma(j)}$,

with respect to the standard basis $\underline{e} = (e_1, \dots, e_n)$ of K^n, i.e.

$$\mathcal{M}_{\underline{e}}^{\underline{e}}(\sigma) = \mathcal{P}_\sigma$$

In particular,

$$\mathcal{P}_\sigma \cdot \mathcal{P}_\tau = \mathcal{P}_{\sigma\tau} \text{ for all } \sigma, \tau \in \mathfrak{S}_n.$$

For $\mathfrak{A} = (a_{ij}) \in M_{m,n}(K)$ $(I = \mathbb{N}_m^*, J = \mathbb{N}_n^*)$, the

(i,j)-th entry of $\mathfrak{A} \cdot \mathcal{P}_\sigma$ is $a_{i\sigma(j)}$

i.e. the j-th column of $\mathfrak{A} \cdot \mathcal{P}_\sigma$ is

$(a_{i\sigma(j)})_{i \in I}$ = the $\sigma(j)$-th column of $\mathfrak{A}$.

Similarly, for a permutation $\sigma \in \mathfrak{S}_m$,

the (i,j)-th entry of $\mathcal{P}_\sigma \cdot \mathfrak{A}$ is $a_{\sigma^{-1}(i)j}$,

i.e. the i-th row of $\mathcal{P}_\sigma \cdot \mathfrak{A}$ is

$(a_{\sigma^{-1}(i)j})_{j \in J}$ = the $\sigma^{-1}(i)$-th row of $\mathfrak{A}$.

For $1 \leq r, s \leq n$, let $\mathcal{P}_{rs} \in M_n(K)$ be the permutation matrix corresponding

to the transposition $\langle r,s\rangle$, $r\neq s$.
Then $\mathfrak{P}_{rs}^{-1} = \mathfrak{P}_{rs} = \mathfrak{P}_{sr}$.
Moreover, one can easily check:

$$\mathfrak{P}_{rs} = \mathfrak{B}_{sr}(1)\cdot\mathfrak{B}_{rs}(-1)\cdot\mathfrak{B}_{sr}(1)\,\mathfrak{E}_s(-1),$$

i.e. interchanging r-th and s-th columns (resp. rows) can be realized as multiplication with elementary matrices if s-th column (resp. row) is multiplied by -1.

With these definitions and remarks, we prove the following:

8.C.8 Theorem Let $\mathfrak{A}\in M_{m,n}(K)$ be a $m\times n$ matrix of rank r and $k:=\mathrm{Min}(m,n)$. Then there exist elementary matrices $\mathfrak{L}_1,\dots,\mathfrak{L}_p\in M_m(K)$ and $\mathfrak{T}_1,\dots,\mathfrak{T}_q\in M_n(K)$ such that $\mathfrak{L}_1\cdots\mathfrak{L}_p\,\mathfrak{A}\,\mathfrak{T}_1\cdots\mathfrak{T}_q$ is a diagonal matrix $\mathfrak{D}=\mathrm{Diag}(d_1,\dots,d_k)\in M_{m,n}(K)$, where even one can choose $d_2=\dots=d_r=1$

and $d_{r+1} = \cdots = d_k = 0$. — If $r = n$, then the elementary matrices $\tau_1, \dots, \tau_q$ are not necessary.

Proof In the Example 8.B.4, we have demonstrated that: By interchanging rows and columns, and elementary row operations of type (1), i.e. adding a scalar multiple of one row to another rows, one can transform the matrix $\mathfrak{A}$ to a matrix of the form:

$$\mathfrak{R} := \begin{pmatrix} c_{11} & * & \cdots & * & * & \cdots & * \\ 0 & c_{22} & \cdots & * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_{rr} & * & \cdots & * \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

with $d_j := c_{jj}$, $j = 1, \dots, r$

In the case $r = n$, one do not need to interchange columns; For, suppose the first $j-1$ columns are already in the required form. Then either in in the j-th row or some element in the j-th column below main diagonal must be $\neq 0$; otherwise first j-columns of this matrix are linearly dependent, in particular, its rank $=$ rank $\alpha < n = r =$ column-rank α, a contradiction.

Now by interchanging such a row with $\neq 0$ element (in the j-th column below j-th row) with the j-th row, we get a matrix of the required form.

For $i = r, r-1, \cdots, 2$, adding a suitable scalar multiple of the i-th row to earlier rows, we get a matrix whose entries in the first r columns above the main diagonal are all 0.

Finally, for $j = 1, \cdots, r$, adding a suitable scalar multiple of j-th column to later columns, we get a diagonal matrix $\mathcal{D} = \mathrm{Diag}(d_1, \cdots, d_r, 0, \cdots 0)$. Note that if $r = n$, then this last step is not required, i.e. we do not need to use elementary column operations.

For the last part apply elementary row-operations on the diagonal

8c/18

matrix

$$\mathrm{Diag}(d_1, \cdots, d_{r-1}, d_r, 0, \cdots, 0)$$

to bring it to the form

$$\mathrm{Diag}(d_1, \cdots, d_{r-1} d_r, 1, 0, \cdots, 0)$$

This is done as follows:

$$\begin{pmatrix} d_{r-1} & 0 \\ 0 & d_r \end{pmatrix} \xmapsto[\mathcal{B}_{r-1,r}\left(\frac{1}{d_r}\right)]{R_{r-1} \longmapsto R_{r-1} + \frac{1}{d_r} R_r} \begin{pmatrix} d_{r-1} & 1 \\ 0 & d_r \end{pmatrix}$$

$$\begin{pmatrix} d_{r-1} & 1 \\ 0 & d_r \end{pmatrix} \xmapsto[\mathcal{B}_{r,r-1}(1-d_r)]{R_r \longmapsto R_r + (1-d_r) R_{r-1}} \begin{pmatrix} d_{r-1} & 1 \\ (1-d_r) d_{r-1} & 1 \end{pmatrix}$$

$$\begin{pmatrix} d_{r-1} & 1 \\ (1-d_r) d_{r-1} & 1 \end{pmatrix} \xmapsto[\mathcal{B}_{r-1,r}(-1)]{R_{r-1} \longmapsto R_{r-1} - R_r} \begin{pmatrix} d_{r-1} d_r & 0 \\ (1-d_r) d_{r-1} & 1 \end{pmatrix}$$

$$\begin{pmatrix} d_{r-1} d_r & 0 \\ (1-d_r) d_{r-1} & 1 \end{pmatrix} \xmapsto[\mathcal{B}_{r,r-1}\left(1-\frac{1}{d_r}\right)]{R_r \longmapsto R_r + \left(1-\frac{1}{d_r}\right) R_{r-1}} \begin{pmatrix} d_{r-1} d_r & 0 \\ 0 & 1 \end{pmatrix},$$

i.e. $\mathcal{B}_{r,r-1}\left(1-\frac{1}{d_r}\right) \cdot \mathcal{B}_{r-1,r}(-1) \mathcal{B}_{r,r-1}(-1) \cdot \mathcal{B}_{r-1,r}\left(\frac{1}{d_r}\right) \cdot$

$$\mathrm{Diag}(d_1, \cdots, d_{r-1}, d_r, 0, \cdots, 0) =$$

$$\mathrm{Diag}(d_1, \cdots, d_{r-1} d_r, 0, \cdots, 0).$$

Analogous elementary row-operations for the index pairs $(r-2, r-1), \dots, (1,2)$, will transform the diagonal matrix

$\mathrm{Diag}(d_1, \dots, d_{r-1}d_r, 1, 0, \dots, 0)$

to the diagonal matrix

$\mathrm{Diag}(d, 1, \dots, 1, 0, \dots, 0)$ with $d \in K^\times$. ▪

In the case of a square invertible matrix $\mathfrak{A}$, i.e. $r = m = n$, we get the representation (by 8.C.8)

$$\mathfrak{B}_1 \cdots \mathfrak{B}_p \mathfrak{A} = \mathrm{Diag}(d, 1, \dots, 1).$$

Therefore the inverse $\mathfrak{A}^{-1}$ of $\mathfrak{A}$ is

$$\mathfrak{A}^{-1} = \mathrm{Diag}(d^{-1}, 1, \dots, 1)\mathfrak{B}_1 \cdots \mathfrak{B}_p.$$

This is essentially the method given in Example 8.B.7. For this the following representation is also used.

We say that an upper or lower triangular matrix is <u>normalized</u> if all of its entries on the main-diagonal are 1.

we get a matrix

$$\begin{pmatrix} r & \cdots \\ \vdots & \mathfrak{a}_1 \end{pmatrix} \text{ with } \mathfrak{a}_1 \in M_{m-1,n-1}(\mathbb{Z})$$

and $(k, r) < (k, |a_{11}|)$ in $\mathbb{N} \times \mathbb{N}^*$, we can apply induction.

(2.b) If $j_0 = 1$ then use the same method as in (2.a) replacing column-operations by row-operations.

(2.c) Assume that $i_0 \neq 1$ and $j_0 \neq 1$, and that a_{11} divides both a_{1j_0} and a_{i_01}. We now use the following elementary operations:

$$C_{j_0} \longmapsto C_{j_0} - \frac{a_{1j_0}}{a_{11}} C_1$$

and then

$$R_{i_0} \longmapsto R_{i_0} - \frac{a_{i_01}}{a_{11}} R_1, \quad R_1 \longmapsto R_1 + R_{i_0}$$

then we get a matrix with $(1, j_0)$-th entry $a_{i_0j_0} - \frac{a_{1j_0}}{a_{11}} a_{i_01}$ which is not divisible by a_{11}, since $a_{11} \nmid a_{i_0j_0}$ and $a_{11} \mid a_{i_01}$. With this we obtain a new matrix as in the case (2.a) (see details on page 8C/27)

Analogous elementary row-operations for the index pairs $(r-2, r-1), \ldots, (1,2)$, will transform the diagonal matrix

$\mathrm{Diag}(d_1, \ldots, d_{r-1}d_r, 1, 0, \ldots, 0)$

to the diagonal matrix

$\mathrm{Diag}(d, 1, \ldots, 1, 0, \ldots, 0)$ with $d \in K^\times$. ■

In the case of a square invertible matrix $\mathfrak{A}$, i.e. $r = m = n$, we get the representation (by 8.C.8)

$\mathfrak{B}_1 \cdots \mathfrak{B}_p \mathfrak{A} = \mathrm{Diag}(d, 1, \ldots, 1)$.

Therefore the inverse $\mathfrak{A}^{-1}$ of $\mathfrak{A}$ is

$\mathfrak{A}^{-1} = \mathrm{Diag}(d^{-1}, 1, \ldots, 1)\mathfrak{B}_1 \cdots \mathfrak{B}_p$.

This is essentially the method given in Example 8.B.7. For this the following representation is also used.

We say that an upper or lower triangular matrix is <u>normalized</u> if all of its entries on the main-diagonal are 1.

8.C.9 Theorem Let $\mathfrak{A} \in M_{m,n}(K)$ be a $m \times n$ matrix of rank r. Then there exist a normalized lower-triangular matrix $\mathcal{L} \in M_m(K)$, an upper triangular matrix $\mathcal{R} \in M_{m,n}(K)$ of rank r and permutation matrices $\mathcal{P}_\sigma \in M_m(K)$, $\mathcal{P}_\varphi \in M_n(K)$ such that

$$\mathcal{P}_\sigma \cdot \mathfrak{A} \cdot \mathcal{P}_\varphi = \mathcal{L} \cdot \mathcal{R}$$

If $r = n$, then one can forgo the matrix $\mathcal{P}_\varphi$ and $\mathcal{R}$ can be written as a product of diagonal matrix $\mathcal{D} \in GL_m(K)$ and a normalized upper triangular matrix $\mathcal{R}' \in M_{m,n}(K)$ and obtain:

$$\mathcal{P}_\sigma \mathfrak{A} = \mathcal{L} \cdot \mathcal{D} \mathcal{R}'.$$

Proof We go as in the proof of 8.C.8, where we may need to interchange columns- and rows and hence multiply on right and left by permutation matrices and not with elementary matrices. To arrive at the matrix $\mathfrak{R}$ in the proof of 8.C.8, for $j=1,\dots,r$, we need to make all entries in the j-th column <u>below</u> the main diagonal 0. For doing this, we multiply on the left by appropriate elementary matrices $B_{ij}(a)$, $a\in K$, $i>j$.

These matrices and also their products and inverses are normalized lower triangular.

To prove the first part of the assertion, for the next column, one need to multiply on the left by the permutation matrix $\mathfrak{P}_{st}\in M_m(K)$ with both $s,t>j$ for interchanging s-th and t-th rows. Now, since for $i>j$ and $s,t>j$, we have (in both the cases below)

- $i\notin\{s,t\}$

$$\mathfrak{P}_{st}B_{ij}(a)=\mathfrak{P}_{st}+aE_{ij}=B_{ij}(a)\mathfrak{P}_{st}$$

- $i = s$

$$\mathfrak{P}_{st} \cdot \mathcal{B}_{ij}(a) = \mathfrak{P}_{st} + a\mathcal{E}_{j} = \mathcal{B}_{j}(a)\mathfrak{P}_{st}$$

Therefore, with this process we can bring all permutation matrices together next to the matrix $\mathfrak{A}$. This proves the first assertion.

In the case $r = n$ as in the proof of 8.C.8, one can forgo interchanging columns, i.e. forgo the permutation matrix $\mathfrak{P}_{\varphi}$. Moreover, all entries $d_1, \dots, d_n$ on the main diagonal of $\mathfrak{R}$ are $\neq 0$ and hence $\mathfrak{D} = \mathrm{Diag}(d_1, \dots, d_n, 1, \dots, 1) \in GL_m(K)$ (note $n = r \leq m$). Now, multiplying i-th row of $\mathfrak{R}$ by d_i^{-1} for all $i = 1, \dots, n$, we get the matrix $\mathfrak{R}' \in M_{m,n}(K)$ with $\mathfrak{R} = \mathfrak{D}\mathfrak{R}'$, where $\mathfrak{R}'$ is a normalized upper triangular matrix.

If the matrix $\mathfrak{A}$ in 8.C.9 is invertible and $\mathfrak{D} = \mathrm{Diag}(d_1, \dots, d_n)$. Then the inverse $\mathfrak{A}^{-1}$ of $\mathfrak{A}$ is of the form

$$\mathfrak{A}^{-1} = \mathcal{R}'^{-1}\,\mathrm{Diag}(d_1^{-1},\ldots,d_n^{-1})\,\mathcal{L}^{-1}\mathcal{P}_\sigma$$

where the inverse matrices of the normalized triangular matrices $\mathcal{R}'$ and $\mathcal{L}$ are particularly simple to compute, see Supplement S9.41 also.

8.C.10 Example (Elementary Divisor Theorem) Elementary operations can also be applied on the matrices over arbitrary commutative rings. This often lead to a lucid representations for matrices. For example, we show this, in the case of matrices over the ring of integers $\mathbb{Z}$. This is also very important in the theory of finitely generated abelian groups.

The set $M_{m,n}(\mathbb{Z})$ of $m \times n$ matrices with entries in $\mathbb{Z}$ is a subgroup of $M_{m,n}(\mathbb{Q})$ and, in the case $m=n$, it is a subring of $M_n(\mathbb{Q})$. Every matrix $\mathfrak{A} \in M_{m,n}(\mathbb{Z})$ as an element in $M_{m,n}(\mathbb{Q})$ has a well-defined rank. We now prove the following very importtant result:

8.C.11 Elementary Divisor Theorem.

Let $\mathfrak{a} \in M_{m,n}(\mathbb{Z})$ be a $m \times n$ matrix of rank r with entries in $\mathbb{Z}$ and $k := \mathrm{Min}(m,n)$. Then there exist elementary matrices $\mathfrak{L}_1, \ldots, \mathfrak{L}_p \in M_m(\mathbb{Z})$ and $\tau_1, \ldots, \tau_q \in M_n(\mathbb{Z})$ such that

$$\mathfrak{L}_1 \cdots \mathfrak{L}_p \,\mathfrak{a}\, \tau_1 \cdots \tau_q = \mathfrak{D} = \mathrm{Diag}(e_1, \ldots, e_k)$$

$\in M_{m,n}(\mathbb{Z})$ with $e_i \mid e_{i+1}$ for all $i = 1, \ldots, r-1$ and $e_{r+1} = \cdots = e_k = 0$.

Proof We may assume that $\mathfrak{a} = (a_{ij}) \neq 0$ and by adding a row or column to the first one, we may also assume $a_{11} \neq 0$.

We shall prove this assertion by induction on $(k, |a_{11}|) \in \mathbb{N} \times \mathbb{N}^*$ with respect to the lexicographic order on $\mathbb{N} \times \mathbb{N}^*$.

(1) Suppose that all a_{ij} are divisible by a_{11}. Then substracting successively

a_{1j}/a_{11} - times first column from j-th column for $j=2,\dots,n$ and

a_{i1}/a_{11} - times first row from i-th row for $i=2,\dots,n$, we get a matrix

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathfrak{A}_1 & \\ 0 & & & \end{pmatrix} \quad \text{with } \mathfrak{A}_1 \in M_{m-1,n-1}(\mathbb{Z})$$

and all entries in $\mathfrak{A}_1$ are divisible by a_{11}. Now, apply induction (note that $(k-1,*) < (k,|a_{11}|)$ in $\mathbb{N}\times\mathbb{N}^*$)

(2) Suppose that $a_{i_0 j_0}$ is <u>not</u> divisible by a_{11}.

(2.a) If $i_0 = 1$ and $a_{1j_0} = q\,a_{11} + r$ with $q, r \in \mathbb{Z}$, $0 < r < |a_{11}|$ (Division with remainder in $\mathbb{Z}$), then applying the elementary column operations

$$C_{j_0} \longmapsto C_{j_0} - qC_1 \text{ and then } C_1 \longleftrightarrow C_{j_0}{}^{*}$$

---

* If one would like to avoid interchanging columns, use $C_{j_0} \longmapsto C_{j_0} - (q+1)C_1$

we get a matrix

$$\begin{pmatrix} r & \cdots \\ \vdots & \mathfrak{A}_1 \end{pmatrix} \text{ with } \mathfrak{A}_1 \in M_{m-1,n-1}(\mathbb{Z})$$

and $(k, r) < (k, |a_{11}|)$ in $\mathbb{N} \times \mathbb{N}^*$, we can apply induction.

(2.b) If $j_0 = 1$ then use the same method as in (2.a) replacing column-operations by row-operations.

(2.c) Assume that $i_0 \neq 1$ and $j_0 \neq 1$, and that a_{11} divides both a_{1j_0} and $a_{i_0 1}$. We now use the following elementary operations:

$$C_{j_0} \longmapsto C_{j_0} - \frac{a_{1j_0}}{a_{11}} C_1$$

and then

$$R_{i_0} \longmapsto R_{i_0} - \frac{a_{i_0 1}}{a_{11}} R_1, \quad R_1 \longmapsto R_1 + R_{i_0}$$

then we get a matrix with $(1, j_0)$-th entry $a_{i_0 j_0} - \frac{a_{1j_0}}{a_{11}} a_{i_0 1}$ which is not divisible by a_{11}, since $a_{11} \nmid a_{i_0 j_0}$ and $a_{11} \mid a_{i_0 1}$. With this we obtain a new matrix as in the case (2.a) (see details on page 8C/27)

$$\begin{pmatrix} a_{11} & \cdots & a_{1j_0} & \cdots & a_{1n} \\ a_{i_0 1} & \cdots & a_{i_0 j_0} & \cdots & a_{i_0 n} \\ a_{m1} & & a_{mj_0} & & a_{mn} \end{pmatrix} \xrightarrow{C_{j_0} \mapsto C_{j_0} - \frac{a_{1j_0}}{a_{11}} C_1}$$

$$\begin{pmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & & \vdots & & \\ a_{i_0 1} & & a_{i_0 j_0} - \frac{a_{1j_0}}{a_{11}} a_{i_0 1} & \cdots & a_{i_0 n} \\ \vdots & & \vdots & & \\ a_{m1} & & \vdots & \cdots & a_{mn} \end{pmatrix}$$

$$\xrightarrow{R_{i_0} \mapsto R_{i_0} - \frac{a_{i_0 1}}{a_{11}} R_1}$$

$$\begin{pmatrix} a_{11} & \cdots & 0 & \cdots & \cdots \\ \vdots & & \vdots & & \\ 0 & \cdots & a_{i_0 j_0} - \frac{a_{1j_0}}{a_{11}} a_{i_0 1} & & \cdots \\ \vdots & & \cdots & & \cdots \end{pmatrix}$$

$$\xrightarrow{R_1 \mapsto R_1 + R_{i_0}}$$

$$\begin{pmatrix} a_{11} & \cdots & \boxed{a_{i_0 j_0} - \frac{a_{1j_0}}{a_{11}} a_{i_0 1}} & \cdots \\ \vdots & & & \\ 0 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$