Chapter 5
The Echelon Form and the Rank of Matrices

In this chapter we develop a systematic method for transforming a matrix A with entries from a field into a special form which is called the echelon form of A. The transformation consists of a sequence of multiplications of A from the left by certain "elementary matrices". If A is invertible, then its echelon form is the identity matrix, and the inverse A^{-1} is the product of the inverses of the elementary matrices. For a non-invertible matrix its echelon form is, in some sense, the "closest possible" matrix to the identity matrix. This form motivates the concept of the rank of a matrix, which we introduce in this chapter and will use frequently later on.

5.1 Elementary Matrices

Let R be a commutative ring with unit, $n \in \mathbb{N}$ and $i, j \in \{1, \dots, n\}$. Let $I_n \in R^{n,n}$ be the identity matrix and let e_i be its ith column, i.e., $I_n = [e_1, \dots, e_n]$.

We define

$$E_{ij} := e_i e_j^T = [0, \dots, 0, \underbrace{e_i}_{\text{column } j}, 0, \dots, 0] \in R^{n,n},$$

i.e., the entry (i, j) of E_{ij} is 1, all other entries are 0.

For $n \geq 2$ and $i < j$ we define

$$P_{ij} := [e_1, \dots, e_{i-1}, e_j, e_{i+1}, \dots, e_{j-1}, e_i, e_{j+1}, \dots, e_n] \in R^{n,n}. \tag{5.1}$$

Thus, P_{ij} is a permutation matrix (cp. Definition 4.12) obtained by exchanging the columns i and j of I_n. A multiplication of$A \in R^{n,m}$ from the left with P_{ij} means an exchange ofthe rows i and j of A. For example,

J. Liesen and V. Mehrmann, *Linear Algebra*, Springer Undergraduate
Mathematics Series, DOI 10.1007/978-3-319-24346-7_5

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \quad P_{13} = [e_3, e_2, e_1] = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad P_{13}A = \begin{bmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix}.$$

For $\lambda \in R$ we define

$$M_i(\lambda) := [e_1, \ldots, e_{i-1}, \lambda e_i, e_{i+1}, \ldots, e_n] \in R^{n,n}. \tag{5.2}$$

Thus, $M_i(\lambda)$ is a diagonal matrix obtained by replacing the ith column of I_n by λe_i. A multiplication of $A \in R^{n,m}$ from the left with $M_i(\lambda)$ means a multiplication of the ith row of A by λ. For example,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \quad M_2(-1) = [e_1, -e_2, e_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad M_2(-1)A = \begin{bmatrix} 1 & 2 & 3 \\ -4 & -5 & -6 \\ 7 & 8 & 9 \end{bmatrix}.$$

For $n \geq 2$, $i < j$ and $\lambda \in R$ we define

$$G_{ij}(\lambda) := I_n + \lambda E_{ji} = [e_1, \ldots, e_{i-1}, e_i + \lambda e_j, e_{i+1}, \ldots, e_n] \in R^{n,n}. \tag{5.3}$$

Thus, the lower triangular matrix $G_{ij}(\lambda)$ is obtained by replacing the ith column of I_n by $e_i + \lambda e_j$. A multiplication of $A \in R^{n,m}$ from the left with $G_{ij}(\lambda)$ means that λ times the ith row of A is added to the jth row of A. Similarly, a multiplication of $A \in R^{n,m}$ from the left by the upper triangular matrix $G_{ij}(\lambda)^T$ means that λ times the jth row of A is added to the ith row of A. For example,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \quad G_{23}(-1) = [e_1, e_2 - e_3, e_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix},$$

$$G_{23}(-1)A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 3 & 3 \end{bmatrix}, \quad G_{23}(-1)^T A = \begin{bmatrix} 1 & 2 & 3 \\ -3 & -3 & -3 \\ 7 & 8 & 9 \end{bmatrix}.$$

Lemma 5.1 *The elementary matrices P_{ij}, $M_i(\lambda)$ for invertible $\lambda \in R$, and $G_{ij}(\lambda)$ defined in (5.1), (5.2), and (5.3), respectively, are invertible and have the following inverses:*

(1) $P_{ij}^{-1} = P_{ij}^T = P_{ij}$.
(2) $M_i(\lambda)^{-1} = M_i(\lambda^{-1})$.
(3) $G_{ij}(\lambda)^{-1} = G_{ij}(-\lambda)$.

Proof

(1) The invertibility of P_{ij} with $P_{ij}^{-1} = P_{ij}^T$ was already shown in Theorem 4.16; the symmetry of P_{ij} is easily seen.

(2) Since $\lambda \in R$ is invertible, the matrix $M_i(\lambda^{-1})$ is well defined. A straightforward computation now shows that $M_i(\lambda^{-1})M_i(\lambda) = M_i(\lambda)M_i(\lambda^{-1}) = I_n$.

(3) Since $e_j^T e_i = 0$ for $i < j$, we have $E_{ji}^2 = (e_i e_j^T)(e_i e_j^T) = 0$, and therefore

$$\begin{aligned} G_{ij}(\lambda)G_{ij}(-\lambda) &= (I_n + \lambda E_{ji})(I_n + (-\lambda)E_{ji}) \\ &= I_n + \lambda E_{ji} + (-\lambda)E_{ji} + (-\lambda^2)E_{ji}^2 = I_n. \end{aligned}$$

A similar computation shows that $G_{ij}(-\lambda)G_{ij}(\lambda) = I_n$. □

5.2 The Echelon Form and Gaussian Elimination

The constructive proof of the following theorem relies on the *Gaussian elimination algorithm*.[1] For a given matrix $A \in K^{n,m}$, where K is a field, this algorithm constructs a matrix $S \in GL_n(K)$ such that $SA = C$ is *quasi*-upper triangular. We obtain this special form by left-multiplication of A with elementary matrices P_{ij}, $M_{ij}(\lambda)$ and $G_{ij}(\lambda)$. Each of these left-multiplications corresponds to the application of one of the so-called "elementary row operations" to the matrix A:

- P_{ij}: exchange two rows of A.
- $M_i(\lambda)$: multiply a row of A with an invertible scalar.
- $G_{ij}(\lambda)$: add a multiple of one row of A to another row of A.

We assume that the entries of A are in a field (rather than a ring) because in the proof of the theorem we require that nonzero entries of A are invertible. A generalization of the result which holds over certain rings (e.g. the integers $\mathbb{Z}$) is given by the *Hermite normal form*,[2] which plays an important role in Number Theory.

Theorem 5.2 *Let K be a field and let $A \in K^{n,m}$. Then there exist invertible matrices $S_1, \ldots, S_t \in K^{n,n}$ (these are products of elementary matrices) such that $C := S_t \cdots S_1 A$ is in* echelon form, *i.e., either $C = 0$ or*

[1] Named after Carl Friedrich Gauß (1777–1855). A similar method was already described in Chap. 8, "Rectangular Arrays", of the "Nine Chapters on the Mathematical Art". This text developed in ancient China over several decades BC stated problems of every day life and gave practical mathematical solution methods. A detailed commentary and analysis was written by Liu Hui (approx. 220–280 AD) around 260 AD.

[2] Charles Hermite (1822–1901).

$$C = \left[\begin{array}{c|cc|cc|cc|c|cc}
 & 1 & \star & 0 & \star & 0 & \star & & 0 & \star \\
 & & & 1 & & 0 & & & \vdots & \\
 & & & & & 1 & & & & \\
0 & & 0 & & 0 & & 0 & \ddots & 0 & \\
 & & & & & & & & 1 & \\
 & & & & & & & & & 0
\end{array}\right].$$

Here $\star$ denotes an arbitrary (zero or nonzero) entry of C.
More precisely, $C = [c_{ij}]$ is either the zero matrix, or there exists a sequence of natural numbers $j_1, \ldots, j_r$ (these are called the "steps" of the echelon form), where $1 \leq j_1 < \cdots < j_r \leq m$ and $1 \leq r \leq \min\{n, m\}$, such that

(1) $c_{ij} = 0$ for $1 \leq i \leq r$ and $1 \leq j < j_i$,
(2) $c_{ij} = 0$ for $r < i \leq n$ and $1 \leq j \leq m$,
(3) $c_{i,j_i} = 1$ for $1 \leq i \leq r$ and all other entries in column j_i are zero.

If $n = m$, then $A \in K^{n,n}$ is invertible if and only if $C = I_n$. In this case $A^{-1} = S_t \cdots S_1$.

Proof If $A = 0$, then we set $t = 1$, $S_1 = I_n$, $C = 0$ and we are done.

Now let $A \neq 0$ and let j_1 be the index of the first column of

$$A^{(1)} = \left[a_{ij}^{(1)}\right] := A$$

that does not consist of all zeros. Let $a_{i_1,j_1}^{(1)}$ be the first entry in this column that is nonzero, i.e., $A^{(1)}$ has the form

$$A^{(1)} = \left[\begin{array}{c|c|c}
 & 0 & \\
 & \vdots & \\
 & 0 & \\
0 & a_{i_1,j_1}^{(1)} & \star \\
 & \star & \\
 & \vdots & \\
 & \star &
\end{array}\right].$$
$$\qquad\qquad j_1$$

We then proceed as follows: First we *permute* the rows i_1 and 1 (if $i_1 > 1$). Then we *normalize* the new first row, i.e., we multiply it with $\left(a_{i_1,j_1}^{(1)}\right)^{-1}$. Finally we *eliminate* the nonzero elements below the first entry in column j_1. Permuting and normalizing leads to

$$\widetilde{A}^{(1)} = \left[\tilde{a}_{i,j}^{(1)}\right] := M_1\left(\left(a_{i_1,j_1}^{(1)}\right)^{-1}\right)P_{1,i_1}A^{(1)} = \left[\begin{array}{c|c|c} & 1 & \\ & \widetilde{a}_{2,j_1}^{(1)} & \\ 0 & \vdots & \star \\ & \widetilde{a}_{n,j_1}^{(1)} & \end{array}\right]. \quad (\text{column } j_1)$$

If $i_1 = 1$, then we set $P_{1,1} := I_n$. In order to eliminate below the 1 in column j_1, we multiply $\widetilde{A}^{(1)}$ from the left with the matrices

$$G_{1,n}\left(-\widetilde{a}_{n,j_1}^{(1)}\right), \ldots, G_{1,2}\left(-\widetilde{a}_{2,j_1}^{(1)}\right).$$

Then we have

$$S_1 A^{(1)} = \left[\begin{array}{c|c|c} 0 & 1 & \star \\ \hline & 0 & \\ 0 & \vdots & A^{(2)} \\ & 0 & \end{array}\right], \quad (\text{column } j_1)$$

where

$$S_1 := G_{1,n}\left(-\tilde{a}_{n,j_1}^{(1)}\right)\cdots G_{1,2}\left(-\tilde{a}_{2,j_1}^{(1)}\right)M_1\left(\left(a_{i_1,j_1}^{(1)}\right)^{-1}\right)P_{1,i_1}$$

and $A^{(2)} = [a_{ij}^{(2)}]$ with $i = 2, \ldots, n$, $j = j_1 + 1, \ldots, m$, i.e., we keep the indices of the larger matrix $A^{(1)}$ in the smaller matrix $A^{(2)}$.

If $A^{(2)} = [\ ]$ or $A^{(2)} = 0$, then we are finished, since then $C := S_1 A^{(1)}$ is in echelon form. In this case $r = 1$.

If at least one of the entries of $A^{(2)}$ is nonzero, then we apply the steps described above to the matrix $A^{(2)}$. For $k = 2, 3, \ldots$ we define the matrices S_k recursively as

$$S_k = \begin{bmatrix} I_{k-1} & 0 \\ 0 & \widetilde{S}_k \end{bmatrix}, \quad \text{where} \quad \widetilde{S}_k A^{(k)} = \left[\begin{array}{c|c|c} 0 & 1 & \star \\ \hline & 0 & \\ 0 & \vdots & A^{(k+1)} \\ & 0 & \end{array}\right]. \quad (\text{column } j_k)$$

Each matrix $\tilde{S}_k$ is constructed analogous to S_1: First we identify the first column j_k of $A^{(k)}$ that is not completely zero, as well as the first nonzero entry $a_{i_k,j_k}^{(k)}$ in that column. Then permuting and normalizing yields the matrix

$$\widetilde{A}^{(k)} = [\widetilde{a}_{ij}^{(k)}] := M_k\left(\left(a_{i_k,j_k}^{(k)}\right)^{-1}\right)P_{k,i_k}A^{(k)}.$$

If $k = i_k$, then we set $P_{k,k} := I_{n-k+1}$. Now

$$\widetilde{S}_k = G_{k,n}\left(-\widetilde{a}^{(k)}_{n,j_k}\right)\cdots G_{k,k+1}\left(-\widetilde{a}^{(k)}_{k+1,j_k}\right) M_k\left(\left(a^{(k)}_{i_k,j_k}\right)^{-1}\right) P_{k,i_k},$$

so that S_k is indeed a product of elementary matrices of the form

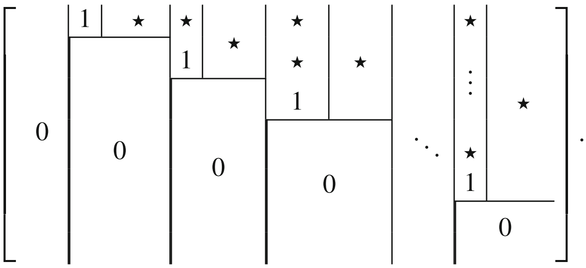
$$\begin{bmatrix} I_{k-1} & 0 \\ 0 & T \end{bmatrix},$$

where T is an elementary matrix of size $(n-k+1)\times(n-k+1)$.

If we continue this procedure inductively, it will end after $r \le \min\{n, m\}$ steps with either $A^{(r+1)} = 0$ or $A^{(r+1)} = [\ ]$.

After r steps we have

$$S_r \cdots S_1 A^{(1)} =$$



(5.4)

By construction, the entries 1 in (5.4) are in the positions

$$(1, j_1),\ (2, j_2),\ \ldots,\ (r, j_r).$$

If $r = 1$, then $S_1 A^{(1)}$ is in echelon form (see the discussion at the beginning of the proof). If $r > 1$, then we still have to eliminate the nonzero entries above the 1 in columns $j_2, \ldots, j_r$. To do this, we denote the matrix in (5.4) by $R^{(1)} = [r^{(1)}_{ij}]$ and form for $k = 2, \ldots, r$ recursively

$$R^{(k)} = [r^{(k)}_{ij}] := S_{r+k-1} R^{(k-1)},$$

where

$$S_{r+k-1} := G_{1,k}\left(-r^{(k-1)}_{1,j_k}\right)^T \cdots G_{k-1,k}\left(-r^{(k-1)}_{k-1,j_k}\right)^T.$$

For $t := 2r - 1$ we have $C := S_t S_{t-1} \cdots S_1 A$ in echelon form.

Suppose now that $n = m$ and that $C = S_t S_{t-1} \cdots S_1 A$ is in echelon form. If A is invertible, then C is a product of invertible matrices and thus invertible. An invertible matrix cannot have a row containing only zeros, so that $r = n$ and hence $C = I_n$. If, on the other hand, $C = I_n$, then the invertibility of the elementary matrices implies that $S_1^{-1} \cdots S_t^{-1} = A$. As a product of invertible matrices, A is invertible and $A^{-1} = S_t \cdots S_1$. □

In the literature, the echelon form is sometimes called *reduced row* echelon form.

Example 5.3 Transformation of a matrix from $\mathbb{Q}^{3,5}$ to echelon form via left multiplication with elementary matrices:

$$\begin{bmatrix} 0 & 2 & 1 & 3 & 3 \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 1 \end{bmatrix}$$

$$\underset{M_1\left(\frac{1}{2}\right)}{\xrightarrow{j_1=2,\, i_1=1}} \begin{bmatrix} 0 & 1 & \frac{1}{2} & \frac{3}{2} & \frac{3}{2} \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 1 \end{bmatrix} \underset{G_{13}(-2)}{\longrightarrow} \begin{bmatrix} 0 & 1 & \frac{1}{2} & \frac{3}{2} & \frac{3}{2} \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 0 & -1 & -2 & -2 \end{bmatrix}$$

$$\underset{G_{12}(-2)}{\longrightarrow} \begin{bmatrix} 0 & 1 & \frac{1}{2} & \frac{3}{2} & \frac{3}{2} \\ 0 & 0 & -1 & -2 & -2 \\ 0 & 0 & -1 & -2 & -2 \end{bmatrix} \underset{M_2(-1)}{\xrightarrow{j_2=3,\, i_2=2}} \begin{bmatrix} 0 & 1 & \frac{1}{2} & \frac{3}{2} & \frac{3}{2} \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & -1 & -2 & -2 \end{bmatrix}$$

$$\underset{G_{23}(1)}{\longrightarrow} \begin{bmatrix} 0 & 1 & \frac{1}{2} & \frac{3}{2} & \frac{3}{2} \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \underset{G_{12}\left(-\frac{1}{2}\right)^T}{\longrightarrow} \begin{bmatrix} 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

MATLAB-Minute.
The echelon form is computed in MATLAB with the command `rref` ("reduced row echelon form"). Apply `rref` to `[A eye(n+1)]` in order to compute the inverse of the matrix `A=full(gallery('tridiag', -ones(n,1),2*ones(n+1,1),-ones(n,1)))` for `n=1,2,3,4,5` (cp. Exercise 5.5).
Formulate a conjecture about the general form of A^{-1}. (Can you prove your conjecture?)

The proof of Theorem 5.2 leads to the so-called *LU-decomposition* of a square matrix.

Theorem 5.4 *For every matrix $A \in K^{n,n}$, there exists a permutation matrix $P \in K^{n,n}$, a lower triangular matrix $L \in GL_n(K)$ with ones on the diagonal and an upper triangular matrix $U \in K^{n,n}$, such that $A = PLU$. The matrix U is invertible if and only if A is invertible.*

Proof For $A \in K^{n,n}$ the Eq. (5.4) has the form $S_n \cdots S_1 A = \widetilde{U}$, where $\widetilde{U}$ is upper triangular. If $r < n$, then we set $S_n = S_{n-1} = \cdots = S_{r+1} = I_n$. Since the matrices $S_1, \ldots, S_n$ are invertible, it follows that $\widetilde{U}$ is invertible if and only if A is invertible. For $i = 1, \ldots, n$ every matrix S_i has the form

$$S_i = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & s_{i,i} & & \\ & & & s_{i+1,i} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & s_{n,i} & & & 1 \end{bmatrix} P_{i,j_i},$$

where $j_i \geq i$ for $i = 1, \ldots, n$ and $P_{i,i} := I_n$ (if $j_i = i$, then no permutation was necessary). Therefore,

$$S_n \cdots S_1 = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & s_{n,n} \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & s_{n-1,n-1} & \\ & & & s_{n,n-1} & 1 \end{bmatrix} P_{n-1,j_{n-1}}$$

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & s_{n-2,n-2} & & \\ & & & s_{n-1,n-2} & 1 & \\ & & & s_{n,n-2} & 0 & 1 \end{bmatrix} P_{n-2,j_{n-2}} \cdots \begin{bmatrix} 1 & & & & \\ & s_{22} & & & \\ & s_{32} & 1 & & \\ & \vdots & & \ddots & \\ & s_{n,2} & & & 1 \end{bmatrix} P_{2,j_2} \begin{bmatrix} s_{11} & & & & \\ s_{21} & 1 & & & \\ s_{31} & & 1 & & \\ \vdots & & & \ddots & \\ s_{n,1} & & & & 1 \end{bmatrix} P_{1,j_1}.$$

The form of the permutation matrices for $k = 2, \ldots, n-1$ and $\ell = 1, \ldots, k-1$ implies that

$$P_{k,j_k} \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & s_{\ell,\ell} & & & \\ & & & s_{\ell+1,\ell} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & s_{n,\ell} & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & s_{\ell,\ell} & & & \\ & & & \widetilde{s}_{\ell+1,\ell} & 1 & & \\ & & & \vdots & & \ddots & \\ & & & \widetilde{s}_{n,\ell} & & & 1 \end{bmatrix} P_{k,j_k}$$

holds for certain $\widetilde{s}_{j,\ell} \in K$, $j = \ell + 1, \ldots, n$. Hence,

$$S_n \cdots S_1 = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & s_{n-1,n-1} & \\ & & & s_{n,n}s_{n,n-1} & s_{n,n} \end{bmatrix} \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & s_{n-2,n-2} & & \\ & & \widetilde{s}_{n-1,n-2} & 1 & \\ & & \widetilde{s}_{n,n-2} & & 1 \end{bmatrix} \cdots$$

$$\begin{bmatrix} 1 & & & \\ & s_{22} & & \\ & \widetilde{s}_{32} & 1 & \\ & \vdots & & \ddots \\ & \widetilde{s}_{n2} & & & 1 \end{bmatrix} \begin{bmatrix} s_{11} & & & & \\ \widetilde{s}_{21} & 1 & & & \\ \widetilde{s}_{31} & & 1 & & \\ \vdots & & & \ddots & \\ \widetilde{s}_{n,1} & & & & 1 \end{bmatrix} P_{n-1,j_{n-1}} \cdots P_{1,j_1}.$$

The invertible lower triangular matrices and the permutation matrices form groups with respect to the matrix multiplication (cp. Theorems 4.13 and 4.16). Thus, $S_n \cdots S_1 = \widetilde{L}\widetilde{P}$, where $\widetilde{L}$ is invertible and lower triangular, and $\widetilde{P}$ is a permutation matrix. Since $\widetilde{L} = [\widetilde{l}_{ij}]$ is invertible, also $D := \mathrm{diag}(\widetilde{l}_{11}, \ldots, \widetilde{l}_{nn})$ is invertible, and we obtain $A = PLU$ with $P := \widetilde{P}^{-1} = \widetilde{P}^T$, $L := \widetilde{L}^{-1}D$ and $U := D^{-1}\widetilde{U}$. By construction, all diagonal entries of L are equal to one. □

Example 5.5 Computation of an LU-decomposition of a matrix from $\mathbb{Q}^{3,3}$:

$$\begin{bmatrix} 2 & 2 & 4 \\ 2 & 2 & 1 \\ 2 & 0 & 1 \end{bmatrix} \xrightarrow[M_1\left(\frac{1}{2}\right)]{j_1 = 2,\, i_1 = 1} \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 1 \\ 2 & 0 & 1 \end{bmatrix} \xrightarrow[G_{13}(-2)]{} \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 1 \\ 0 & -2 & -3 \end{bmatrix}$$

$$\xrightarrow[G_{12}(-2)]{} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & -3 \\ 0 & -2 & -3 \end{bmatrix} \xrightarrow[P_{23}]{} \begin{bmatrix} 1 & 1 & 2 \\ 0 & -2 & -3 \\ 0 & 0 & -3 \end{bmatrix} = \widetilde{U}.$$

Hence, $\widetilde{P} = P_{23}$,

$$\widetilde{L} = G_{12}(-2)G_{13}(-2)M_1\left(\frac{1}{2}\right) = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -2 & 1 & 0 \\ -2 & 1 & 1 \end{bmatrix}, \quad D = \mathrm{diag}\left(\frac{1}{2}, 1, 1\right),$$

and thus, $P = \widetilde{P}^T = P_{23}^T = P_{23}$,

$$L = \widetilde{L}^{-1}D = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad U = D^{-1}\widetilde{U} = \begin{bmatrix} 2 & 2 & 4 \\ 0 & -2 & -3 \\ 0 & 0 & -3 \end{bmatrix}.$$

If $A \in GL_n(K)$, then the LU-decomposition yields $A^{-1} = U^{-1}L^{-1}P^T$. Hence after computing the LU-decomposition, one obtains the inverse of A essentially by inverting the two triangular matrices. Since this can be achieved by the efficient recursive formula (4.4), the LU-decomposition is a popular method in scientific computing applications that require the inversion of matrices or the solution of linear systems of equations (cp. Chap. 6). In this context, however, alternative strategies for the choice of the permutation matrices are used. For example, instead of the first nonzero entry in a column one chooses an entry with large (or largest) absolute value for the row exchange and the subsequent elimination. By this strategy the influence of rounding errors in the computation is reduced.

MATLAB-Minute.
The *Hilbert matrix*[3] $A = [a_{ij}] \in \mathbb{Q}^{n,n}$ has the entries $a_{ij} = 1/(i + j - 1)$ for $i, j = 1, \dots, n$. It can be generated in MATLAB with the command `hilb(n)`. Carry out the command `[L,U,P]=lu(hilb(4))` in order to compute an LU-decomposition of the matrix `hilb(4)`. How do the matrices `P`, `L` and `U` look like?
Compute also the LU-decomposition of the matrix
`full(gallery('tridiag',-ones(3,1),2*ones(4,1),-ones(3,1)))`
and study the corresponding matrices `P`, `L` and `U`.

We will now show that, for a given matrix A, the matrix C in Theorem 5.2 is uniquely determined in a certain sense. For this we need the following definition.

Definition 5.6 If $C \in K^{n,m}$ is in echelon form (as in Theorem 5.2), then the positions of $(1, j_1), \dots, (r, j_r)$ are called the *pivot positions* of C.

We also need the following results.

Lemma 5.7 *If $Z \in GL_n(K)$ and $x \in K^{n,1}$, then $Zx = 0$ if and only if $x = 0$.*

Proof Exercise. □

Theorem 5.8 *Let $A, B \in K^{n,m}$ be in echelon form. If $A = ZB$ for a matrix $Z \in GL_n(K)$, then $A = B$.*

[3] David Hilbert (1862–1943).

Proof If B is the zero matrix, then $A = ZB = 0$, and hence $A = B$.

Let now $B \neq 0$ and let A, B have the respective columns a_i, b_i, $1 \leq i \leq m$. Furthermore, let $(1, j_1), \ldots, (r, j_r)$ be the $r \geq 1$ pivot positions of B. We will show that every matrix $Z \in GL_n(K)$ with $A = ZB$ has the form

$$Z = \left[\begin{array}{c|c} I_r & \star \\ \hline 0 & Z_{n-r} \end{array}\right],$$

where $Z_{n-r} \in GL_{n-r}(K)$. Since B is in echelon form and all entries of B below its row r are zero, it then follows that $B = ZB = A$.

Since $(1, j_1)$ is the first pivot position of B, we have $b_i = 0 \in K^{n,1}$ for $1 \leq i \leq j_1 - 1$ and $b_{j_1} = e_1$ (the first column of I_n). Then $A = ZB$ implies $a_i = 0 \in K^{n,1}$ for $1 \leq i \leq j_1 - 1$ and $a_{j_1} = Zb_{j_1} = Ze_1$. Since Z is invertible, Lemma 5.7 implies that $a_{j_1} \neq 0 \in K^{n,1}$. Since A is in echelon form, $a_{j1} = e_1 = b_{j_1}$. Furthermore,

$$Z = Z_n := \left[\begin{array}{c|c} 1 & \star \\ \hline 0 & Z_{n-1} \end{array}\right],$$

where $Z_{n-1} \in GL_{n-1}(K)$ (cp. Exercise 5.3). If $r = 1$, then we are done.

If $r > 1$, then we proceed with the other pivot positions in an analogous way: Since B is in echelon form, the kth pivot position gives $b_{j_k} = e_k$. From $a_{j_k} = Zb_{j_k}$ and the invertibility of Z_{n-k+1} we obtain $a_{j_k} = b_{j_k}$ and

$$Z = \left[\begin{array}{c|c|c} I_{k-1} & 0 & \star \\ \hline 0 & 1 & \star \\ \hline 0 & 0 & Z_{n-k} \end{array}\right],$$

where $Z_{n-k} \in GL_{n-k}(K)$. □

This result yields the uniqueness of the echelon form of a matrix and its invariance under left-multiplication with invertible matrices.

Corollary 5.9 *For $A \in K^{n,m}$ the following assertions hold:*

(1) There is a unique matrix $C \in K^{n,m}$ in echelon form to which A can be transformed by elementary row operations, i.e., by left-multiplication with elementary matrices. This matrix C is called the echelon form of A.
(2) If $M \in GL_n(K)$, then the matrix C in (1) is also the echelon form of MA, i.e., the echelon form of a matrix is invariant under left-multiplication with invertible matrices.

Proof

(1) If $S_1 A = C_1$ and $S_2 A = C_2$, where C_1, C_2 are in echelon form and S_1, S_2 are invertible, then $C_1 = \left(S_1 S_2^{-1}\right) C_2$. Theorem 5.8 now gives $C_1 = C_2$.
(2) If $M \in GL_n(K)$ and $S_3(MA) = C_3$ is in echelon form, then with $S_1 A = C_1$ from (1) we get $C_3 = \left(S_3 M S_1^{-1}\right) C_1$. Theorem 5.8 now gives $C_3 = C_1$. □

5.3 Rank and Equivalence of Matrices

As we have seen in Corollary 5.9, the echelon form of $A \in K^{n,m}$ is unique. In particular, for every matrix $A \in K^{n,m}$, there exists a unique number of pivot positions (cp. Definition 5.6) in its echelon form. This justifies the following definition.

Definition 5.10 The number r of pivot positions in the echelon form of $A \in K^{n,m}$ is called the *rank*[4] of A and denoted by $\mathrm{rank}(A)$.

We see immediately that for $A \in K^{n,m}$ always $0 \le \mathrm{rank}(A) \le \min\{n, m\}$, where $\mathrm{rank}(A) = 0$ if and only if $A = 0$. Moreover, Theorem 5.2 shows that $A \in K^{n,n}$ is invertible if and only if $\mathrm{rank}(A) = n$. Further properties of the rank are summarized in the following theorem.

Theorem 5.11 *For $A \in K^{n,m}$ the following assertions hold:*

(1) There exist matrices $Q \in GL_n(K)$ and $Z \in GL_m(K)$ with

$$QAZ = \begin{bmatrix} I_r & 0_{r,m-r} \\ 0_{n-r,r} & 0_{n-r,m-r} \end{bmatrix}$$

if and only if $rank(A) = r$.

(2) If $Q \in GL_n(K)$ and $Z \in GL_m(K)$, then $\mathrm{rank}(A) = \mathrm{rank}(QAZ)$.

(3) If $A = BC$ with $B \in K^{n,\ell}$ and $C \in K^{\ell,m}$, then

$$(a) \quad \mathrm{rank}(A) \le \mathrm{rank}(B),$$
$$(b) \quad \mathrm{rank}(A) \le \mathrm{rank}(C).$$

(4) $\mathrm{rank}(A) = \mathrm{rank}(A^T)$.

(5) There exist matrices $B \in K^{n,\ell}$ and $C \in K^{\ell,m}$ with $A = BC$ if and only if $\mathrm{rank}(A) \le \ell$.

Proof

(3a) Let $Q \in GL_n(K)$ be such that QB is in echelon form. Then $QA = QBC$. In the matrix QBC at most the first $\mathrm{rank}(B)$ rows contain nonzero entries. By Corollary 5.9, the echelon form of QA is equal to the echelon form of A. Thus, in the normal echelon form of A also at most the first $\mathrm{rank}(B)$ rows will be nonzero, which implies $\mathrm{rank}(A) \le \mathrm{rank}(B)$.

(1) $\Leftarrow$: If $\mathrm{rank}(A) = r = 0$, i.e., $A = 0$, then $I_r = [\ ]$ and the assertion holds for arbitrary matrices $Q \in GL_n(K)$ and $Z \in GL_m(K)$.
If $r \ge 1$, then there exists a matrix $Q \in GL_n(K)$ such that QA is in echelon form with r pivot positions. Then there exists a permutation matrix $P \in K^{m,m}$, that is a product of elementary permutation matrices P_{ij}, with

[4]The concept of the rank was introduced (in the context of bilinear forms) first in 1879 by Ferdinand Georg Frobenius (1849–1917).

$$PA^TQ^T = \begin{bmatrix} I_r & 0_{r,n-r} \\ V & 0_{m-r,n-r} \end{bmatrix}$$

for some matrix $V \in K^{m-r,r}$. If $r = m$, then $V = [\ ]$. In the following, for simplicity, we omit the sizes of the zero matrices. The matrix

$$Y := \begin{bmatrix} I_r & 0 \\ -V & I_{m-r} \end{bmatrix} \in K^{m,m}$$

is invertible with

$$Y^{-1} = \begin{bmatrix} I_r & 0 \\ V & I_{m-r} \end{bmatrix} \in K^{m,m}.$$

Thus,

$$YPA^TQ^T = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

and with $Z := P^TY^T \in GL_m(K)$ we obtain

$$QAZ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}. \tag{5.5}$$

$\Rightarrow$: Suppose that (5.5) holds for $A \in K^{n,m}$ and matrices $Q \in GL_n(K)$ and $Z \in GL_m(K)$. Then with (3a) we obtain

$$\text{rank}(A) = \text{rank}(AZZ^{-1}) \le \text{rank}(AZ) \le \text{rank}(A),$$

and thus, in particular, $\text{rank}(A) = \text{rank}(AZ)$. Due to the invariance of the echelon form (and hence the rank) under left-multiplication with invertible matrices (cp. Corollary 5.9), we get

$$\text{rank}(A) = \text{rank}(AZ) = \text{rank}(QAZ) = \text{rank}\left(\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\right) = r.$$

(2) If $A \in K^{n\times n}$, $Q \in GL_n(K)$ and $Z \in GL_m(K)$, then the invariance of the rank under left-multiplication with invertible matrices and (3a) can again be used for showing that

$$\text{rank}(A) = \text{rank}(QAZZ^{-1}) \le \text{rank}(QAZ) = \text{rank}(AZ) \le \text{rank}(A),$$

and hence, in particular, $\text{rank}(A) = \text{rank}(QAZ)$.

(4) If $\text{rank}(A) = r$, then by (1) there exist matrices $Q \in GL_n(K)$ and $Z \in GL_m(K)$ with $QAZ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$. Therefore,

$$\operatorname{rank}(A) = \operatorname{rank}(QAZ) = \operatorname{rank}\left(\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}^T\right) = \operatorname{rank}((QAZ)^T)$$
$$= \operatorname{rank}(Z^T A^T Q^T) = \operatorname{rank}(A^T).$$

(3b) Using (3a) and (4), we obtain

$$\operatorname{rank}(A) = \operatorname{rank}(A^T) = \operatorname{rank}(C^T B^T) \le \operatorname{rank}(C^T) = \operatorname{rank}(C).$$

(5) Let $A = BC$ with $B \in K^{n,\ell}$, $C \in K^{\ell,m}$. Then by (3a),

$$\operatorname{rank}(A) = \operatorname{rank}(BC) \le \operatorname{rank}(B) \le \ell.$$

Let, on the other hand, $\operatorname{rank}(A) = r \le \ell$. Then there exist matrices $Q \in GL_n(K)$ and $Z \in GL_m(K)$ with $QAZ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$. Thus, we obtain

$$A = \left(Q^{-1}\begin{bmatrix} I_r & 0_{r,\ell-r} \\ 0_{n-r,r} & 0_{n-r,\ell-r} \end{bmatrix}\right)\left(\begin{bmatrix} I_r & 0_{r,m-r} \\ 0_{\ell-r,r} & 0_{\ell-r,m-r} \end{bmatrix} Z^{-1}\right) =: BC,$$

where $B \in K^{n,\ell}$ and $C \in K^{\ell,m}$. □

Example 5.12 The matrix

$$A = \begin{bmatrix} 0 & 2 & 1 & 3 & 3 \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 1 \end{bmatrix} \in \mathbb{Q}^{3,5}$$

from Example 5.3 has the echelon form

$$\begin{bmatrix} 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since there are two pivot positions, we have $\operatorname{rank}(A) = 2$. Multiplying A from the right by

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 & -1 \end{bmatrix} \in \mathbb{Q}^{5,5},$$

yields $AB = 0 \in \mathbb{Q}^{3,5}$, and hence $\operatorname{rank}(AB) = 0 < \operatorname{rank}(A)$.

Assertion (1) in Theorem 5.11 motivates the following definition.

Definition 5.13 Two matrices $A, B \in K^{n,m}$ are called *equivalent*, if there exist matrices $Q \in GL_n(K)$ and $Z \in GL_m(K)$ with $A = QBZ$.

As the name suggests, this defines an equivalence relation on the set $K^{n,m}$, since the following properties hold:

- Reflexivity: $A = QAZ$ with $Q = I_n$ and $Z = I_m$.
- Symmetry: If $A = QBZ$, then $B = Q^{-1}AZ^{-1}$.
- Transitivity: If $A = Q_1BZ_1$ and $B = Q_2CZ_2$, then $A = (Q_1Q_2)C(Z_2Z_1)$.

The equivalence class of $A \in K^{n,m}$ is given by

$$[A] = \big\{QAZ \mid Q \in GL_n(K) \text{ and } Z \in GL_m(K)\big\}.$$

If $\text{rank}(A) = r$, then by (1) in Theorem 5.11 we have

$$\begin{bmatrix} I_r & 0_{r,m-r} \\ 0_{n-r,r} & 0_{n-r,m-r} \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \in [A]$$

and, therefore,

$$\left[\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\right] = [A].$$

Consequently, the rank of A fully determines the equivalence class $[A]$. The matrix

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \in K^{n,m}$$

is called the *equivalence normal form* of A. We obtain

$$K^{n,m} = \bigcup_{r=0}^{\min\{n,m\}} \left[\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\right], \quad \text{where}$$

$$\left[\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\right] \bigcap \left[\begin{bmatrix} I_\ell & 0 \\ 0 & 0 \end{bmatrix}\right] = \emptyset, \quad \text{if } r \neq \ell.$$

Hence there are $1 + \min\{n, m\}$ pairwise distinct equivalence classes, and

$$\left\{ \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \in K^{n,m} \,\middle|\, r = 0, 1, \ldots, \min\{n, m\} \right\}$$

is a complete set of representatives.

From the proof of Theorem 4.9 we know that $(K^{n,n}, +, *)$ for $n \geq 2$ is a non-commutative ring with unit that contains non-trivial zero divisors. Using the equivalence normal form these can be characterized as follows:

- If $A \in K^{n,n}$ is invertible, then A cannot be a zero divisor, since then $AB = 0$ implies that $B = 0$.

- If $A \in K^{n,n} \setminus \{0\}$ is a zero divisor, then A cannot be invertible, and hence $1 \leq \operatorname{rank}(A) = r < n$, so that the equivalence normal form of A is not the identity matrix I_n. Let $Q, Z \in GL_n(K)$ be given with

$$QAZ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Then for every matrix

$$V := \begin{bmatrix} 0_{r,r} & 0_{r,n-r} \\ V_{21} & V_{22} \end{bmatrix} \in K^{n,n}$$

and $B := ZV$ we have

$$AB = Q^{-1} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0_{r,r} & 0_{r,n-r} \\ V_{21} & V_{22} \end{bmatrix} = 0.$$

If $V \neq 0$, then $B \neq 0$, since Z is invertible.

Exercises

(In the following exercises K is an arbitrary field.)

5.1 Compute the echelon forms of the matrices

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 48 \end{bmatrix} \in \mathbb{Q}^{2,3}, \quad B = \begin{bmatrix} 1 & \mathbf{i} \\ \mathbf{i} & 1 \end{bmatrix} \in \mathbb{C}^{2,2}, \quad C = \begin{bmatrix} 1 & \mathbf{i} & -\mathbf{i} & 0 \\ 0 & 0 & 0 & 1 \\ 5 & 0 & -6\mathbf{i} & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \in \mathbb{C}^{4,4},$$

$$D = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \in (\mathbb{Z}/2\mathbb{Z})^{3,2}, \quad E = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 2 & 0 & 1 & 1 \\ 1 & 2 & 0 & 2 \end{bmatrix} \in (\mathbb{Z}/3\mathbb{Z})^{3,4}.$$

(Here for simplicity the elements of $\mathbb{Z}/n\mathbb{Z}$ are denoted by k instead of $[k]$.) State the elementary matrices that carry out the transformations. If one of the matrices is invertible, then compute its inverse as a product of the elementary matrices.

5.2 Let $A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in K^{2,2}$ with $\alpha\delta \neq \beta\gamma$. Determine the echelon form of A and a formula for A^{-1}.

5.3 Let $A = \begin{bmatrix} 1 & A_{12} \\ 0 & B \end{bmatrix} \in K^{n,n}$ with $A_{12} \in K^{1,n-1}$ and $B \in K^{n-1,n-1}$. Show that $A \in GL_n(K)$ if and only if $B \in GL_{n-1}(K)$.

5.4 Consider the matrix

$$A = \begin{bmatrix} \frac{t+1}{t-1} & \frac{t-1}{t^2} \\ \frac{t^2}{t+1} & \frac{t-1}{t+1} \end{bmatrix} \in (K(t))^{2,2},$$

where $K(t)$ is the field of rational functions (cp. Exercise 3.19). Examine whether A is invertible and determine, if possible, A^{-1}. Verify your result by computing $A^{-1}A$ and AA^{-1}.

5.5 Show that if $A \in GL_n(K)$, then the echelon form of $[A,\ I_n] \in K^{n,2n}$ is given by $[I_n,\ A^{-1}]$.
(The inverse of an invertible matrix A can thus be computed via the transformation of $[A,\ I_n]$ to its echelon form.)

5.6 Two matrices $A, B \in K^{n,m}$ are called *left equivalent*, if there exists a matrix $Q \in GL_n(K)$ with $A = QB$. Show that this defines an equivalence relation on $K^{n,m}$ and determine a most simple representative for each equivalence class.

5.7 Prove Lemma 5.7.

5.8 Determine LU-decompositions (cp. Theorem 5.4) of the matrices

$$A = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 0 & 0 & 1 \\ 5 & 0 & 6 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 & -2 & 0 \\ -4 & 0 & 4 & -1 \\ 0 & -1 & -1 & -2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{4,4}.$$

If one of these matrices is invertible, then determine its inverse using its LU-decomposition.

5.9 Let A be the 4×4 *Hilbert matrix* (cp. the MATLAB-Minute above Definition 5.6). Determine $\operatorname{rank}(A)$. Does A have an LU-decomposition as in Theorem 5.4 with $P = I_4$?

5.10 Determine the rank of the matrix

$$A = \begin{bmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{bmatrix} \in \mathbb{R}^{3,3}$$

in dependence of $\alpha, \beta, \gamma \in \mathbb{R}$.

5.11 Let $A, B \in K^{n,n}$ be given. Show that

$$\operatorname{rank}(A) + \operatorname{rank}(B) \leq \operatorname{rank}\left(\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}\right)$$

for all $C \in K^{n,n}$. Examine when this inequality is strict.

5.12 Let $a, b, c \in \mathbb{R}^{n,1}$.

(a) Determine $\operatorname{rank}(ba^T)$.
(b) Let $M(a, b) := ba^T - ab^T$. Show the following assertions:
   (i) $M(a, b) = -M(b, a)$ and $M(a, b)c + M(b, c)a + M(c, a)b = 0$,
   (ii) $M(\lambda a + \mu b, c) = \lambda M(a, c) + \mu M(b, c)$ for $\lambda, \mu \in \mathbb{R}$,
   (iii) $\operatorname{rank}(M(a, b)) = 0$ if and only if there exist $\lambda, \mu \in \mathbb{R}$ with $\lambda \neq 0$ or $\mu \neq 0$ and $\lambda a + \mu b = 0$,
   (iv) $\operatorname{rank}(M(a, b)) \in \{0, 2\}$.