Chapter 4
Matrices

Matrices are an important tool when dealing with many problems, notably the theory of systems of linear equations and the study of maps (operators) between vector spaces. This chapter is devoted to their basic notions and properties.

4.1 Basic Notions

Definition 4.1.1 A *matrix* M with entries in $\mathbb{R}$ (or a *real matrix*) is a collection of elements $a_{ij} \in \mathbb{R}$, with $i = 1, \dots, m$; $j = 1, \dots, n$ and $m, n \in \mathbb{N}$, displayed as follows

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

The matrix M above is said to be made of m *row* vectors in $\mathbb{R}^n$, that is

$$R_1 = (a_{11}, \dots, a_{1n}) \,, \quad \dots \,, \quad R_m = (a_{m1}, \dots, a_{mn})$$

or by n *column* vectors in $\mathbb{R}^m$, that is

$$C_1 = (a_{11}, \dots, a_{m1}) \,, \quad \dots \,, \quad C_n = (a_{1n}, \dots, a_{mn}).$$

Thus the matrix M above is a $m \times n$-matrix (m rows $R_i \in \mathbb{R}^n$ and n columns $R_i \in \mathbb{R}^n$). As a shorthand, by $M = (a_{ij})$ we shall denote a matrix M with entry a_{ij} at the i-th row and j-th column. We denote by $\mathbb{R}^{m,n}$ the collection of all $m \times n$-matrices whose entries are in $\mathbb{R}$.

G. Landi and A. Zampini, *Linear Algebra and Analytic Geometry for Physical Sciences*, Undergraduate Lecture Notes in Physics,
https://doi.org/10.1007/978-3-319-78361-1_4

Remark 4.1.2 It is sometime useful to consider a matrix $M \in \mathbb{R}^{m,n}$ as the collection of its n columns, or as the collection of its m rows, that is to write

$$M = (C_1 \;\; C_2 \;\; \ldots \;\; C_n) = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix}.$$

An element $M \in \mathbb{R}^{1,n}$ is called a n-dimensional *row matrix*,

$$M = (a_{11} \;\; a_{12} \;\; \ldots \;\; a_{1n})$$

while an element $M \in \mathbb{R}^{m,1}$ is called a m-dimensional *column matrix*,

$$M = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}.$$

A *square* matrices of order n is a $n \times n$ matrix, that is an element in $\mathbb{R}^{n,n}$. An element $M \in \mathbb{R}^{1,1}$ is a scalar, that is a single element in $\mathbb{R}$. If $A = (a_{ij}) \in \mathbb{R}^{n,n}$ is a square matrix, the entries $(a_{11}, a_{22}, \ldots, a_{nn})$ give the (*principal*) *diagonal* of A.

Example 4.1.3 The bold typeset entries in

$$A = \begin{pmatrix} \mathbf{1} & 2 & 2 \\ -1 & \mathbf{0} & 3 \\ 2 & 4 & \mathbf{7} \end{pmatrix}$$

give the diagonal of A.

Proposition 4.1.4 *The set $\mathbb{R}^{m,n}$ is a real vector space whose dimension is mn. With $A = (a_{ij})$, $B = (b_{ij}) \in \mathbb{R}^{m,n}$ and $\lambda \in \mathbb{R}$, the vector space operations are defined by*

$$A + B = (a_{ij} + b_{ij}), \quad \lambda A = (\lambda a_{ij}).$$

Proof We task the reader to show that $\mathbb{R}^{m,n}$ equipped with the above defined operations is a vector space. We only remark that the zero element in $\mathbb{R}^{m,n}$ is given by the matrix A with entries $a_{ij} = 0_{\mathbb{R}}$; such a matrix is also called the *null* matrix and denoted $0_{\mathbb{R}^{m,n}}$.
In order to show that the dimension of $\mathbb{R}^{m,n}$ is mn, consider the *elementary* $m \times n$-matrices

$$E_{rs} = (e^{(rs)}_{jk}), \quad \text{with} \quad e^{(rs)}_{jk} = \begin{cases} 1 \text{ if } (j,k) = (r,s) \\ 0 \text{ if } (j,k) \neq (r,s) \end{cases}.$$

Thus the matrix E_{rs} has entries all zero but for the entry (r, s) which is 1. Clearly there are mn of them and it is immediate to show that they form a basis for $\mathbb{R}^{m,n}$. $\square$

Exercise 4.1.5 Let $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \in \mathbb{R}^{2,3}$. One computes

$$\begin{aligned}
\begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 2\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\
&\quad + 0\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= E_{11} + 2E_{12} - E_{13} - E_{22} + E_{23}.
\end{aligned}$$

In addition to forming a vector space, matrices of 'matching size' can be multiplied.

Definition 4.1.6 If $A = (a_{ij}) \in \mathbb{R}^{m,n}$ and $B = (b_{jk}) \in \mathbb{R}^{n,p}$ the *product* between A and B is the matrix in $\mathbb{R}^{m,p}$ defined by

$$C = (c_{ik}) = AB \in \mathbb{R}^{m,p}, \quad \text{where} \quad c_{ik} = R_i^{(A)} \cdot C_k^{(B)} = \sum_{j=1}^{n} a_{ij} b_{jk},$$

with $i = 1, \dots, m$ and $k = 1, \dots, p$. Here $R_i^{(A)} \cdot C_k^{(B)}$ denotes the scalar product in $\mathbb{R}^n$ between the i-th row vector $R_i^{(A)}$ of A with the k-th column vector $C_k^{(B)}$ of B.

Remark 4.1.7 Notice that the product AB—also called the *row by column* product—is defined only if the number of columns of A equals the number of rows of B.

Exercise 4.1.8 Consider the matrices

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{2,3}, \qquad B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 4 \end{pmatrix} \in \mathbb{R}^{3,2}.$$

One has $AB = C = (c_{ik}) \in \mathbb{R}^{2,2}$ with

$$C = \begin{pmatrix} 2 & 0 \\ 6 & 10 \end{pmatrix}.$$

On the other hand, $BA = C' = (c'_{st}) \in \mathbb{R}^{3,3}$ with

$$C' = \begin{pmatrix} 7 & 2 & 1 \\ 5 & 4 & -1 \\ 15 & 6 & 1 \end{pmatrix}.$$

Clearly, comparing C with C' is meaningless, since they are in different spaces.

Remark 4.1.9 With $A \in \mathbb{R}^{m,n}$ and $B \in \mathbb{R}^{p,q}$, the product AB is defined only if $n = p$, giving a matrix $AB \in \mathbb{R}^{m,q}$. It is clear that the product BA is defined only if $m = q$ and in such a case one has $BA \in \mathbb{R}^{p,n}$. When both the conditions $m = q$ and $n = p$ are satisfied both products are defined. This is the case of the matrices A and B in the previous exercise.
Let us consider the space $\mathbb{R}^{n,n}$ of square matrices of order n. If A, B are in $\mathbb{R}^{n,n}$ then both AB and BA are square matrices of order n. An example shows that in general one has $AB \neq BA$. If

$$A = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix},$$

one computes that

$$AB = \begin{pmatrix} 3 & -1 \\ 0 & -1 \end{pmatrix} \neq BA = \begin{pmatrix} 0 & 3 \\ 1 & 2 \end{pmatrix}.$$

Thus the product of matrices is *non commutative*. We shall say that two matrices $A, B \in \mathbb{R}^{n,n}$ commute if $AB = BA$. On the other hand, the associativity of the product in $\mathbb{R}$ and its distributivity with respect to the sum, allow one to prove analogous properties for the space of matrices.

Proposition 4.1.10 *The following identities hold:*

(i) $A(BC) = (AB)C$, *for any* $A \in \mathbb{R}^{m,n}$, $B \in \mathbb{R}^{n,p}$, $C \in \mathbb{R}^{p,q}$,
(ii) $A(B + C) = AB + AC$, *for any* $A \in \mathbb{R}^{m,n}$, $B, C \in \mathbb{R}^{n,p}$,
(iii) $\lambda(AB) = (\lambda A)B = A(\lambda B)$, *for any* $A \in \mathbb{R}^{m,n}$, $B \in \mathbb{R}^{n,p}$, $\lambda \in \mathbb{R}$.

Proof (i) Consider three matrices $A = (a_{ih}) \in \mathbb{R}^{m,n}$, $B = (b_{hk}) \in \mathbb{R}^{n,p}$ and $C = (c_{kj}) \in \mathbb{R}^{p,q}$. From the definition of row by column product one has $AB = (d_{ik})$ with $d_{ik} = \sum_{h=1}^{n} a_{ih}b_{hk}$, while $BC = (e_{hj})$ with $e_{hj} = \sum_{k=1}^{p} b_{hk}c_{kj}$. The ij-entries of $(AB)C$ and $A(BC)$ are

$$\sum_{k=1}^{p} d_{ik}c_{kj} = \sum_{k=1}^{p} \left(\sum_{h=1}^{n} a_{ih}b_{hk} \right) c_{kj} = \sum_{k=1}^{p} \sum_{h=1}^{n} (a_{ih}b_{hk}c_{kj}),$$
$$\sum_{h=1}^{n} a_{ih}e_{hj} = \sum_{h=1}^{n} a_{ih} \left(\sum_{k=1}^{p} b_{hk}c_{kj} \right) = \sum_{h=1}^{n} \sum_{k=1}^{p} (a_{ih}b_{hk}c_{kj}).$$

These two expressions coincide (the last equality on both lines follows from the distributivity in $\mathbb{R}$ of the product with respect to the sum).
(ii) Take matrices $A = (a_{ih}) \in \mathbb{R}^{m,n}$, $B = (b_{hj}) \in \mathbb{R}^{n,p}$ and $C = (c_{hj}) \in \mathbb{R}^{n,p}$. The equality $A(B + C) = AB + AC$ is proven again by a direct computation of the ij-entry for both sides:

$$
\begin{aligned}
[A(B+C)]_{ij} &= \sum_{h=1}^{n} a_{ih}(b_{hj}+c_{hj}) \\
&= \sum_{h=1}^{n} a_{ih}b_{hj} + \sum_{h=1}^{n} a_{ih}c_{hj} \\
&= [AB]_{ij} + [AC]_{ij} \\
&= [AB+AC]_{ij}.
\end{aligned}
$$

(iii) This is immediate. □

The matrix product in $\mathbb{R}^{n,n}$ is inner and it has a neutral element, a multiplication unit.

Definition 4.1.11 The *unit* matrix of order n, denoted by I_n, is the element in $\mathbb{R}^{n,n}$ given by

$$I_n = (\delta_{ij}) , \quad \text{with} \quad \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} .$$

The diagonal entries of I_n are all 1, while its off-diagonal entries are all zero.

Remark 4.1.12 It is easy to prove that, with $A \in \mathbb{R}^{m,n}$, one has

$$AI_n = A \quad \text{and} \quad I_m A = A.$$

Proposition 4.1.13 *The space $\mathbb{R}^{n,n}$ of square matrices of order n, endowed with the sum and the product as defined above, is a non abelian ring.*

Proof Recall the definition of a ring given in A.1.6. The matrix product is an inner operation in $\mathbb{R}^{n,n}$, so the claim follows from the fact that $(\mathbb{R}^{n,n}, +, 0_{\mathbb{R}^{n,n}})$ is an abelian group and the results of the Proposition 4.1.10. □

Definition 4.1.14 A matrix $A \in \mathbb{R}^{n,n}$ is said to be *invertible* (also *non-singular* or *non-degenerate*) if there exists a matrix $B \in \mathbb{R}^{n,n}$, such that $AB = BA = I_n$. Such a matrix B is denoted by A^{-1} and is called the *inverse* of A.

Definition 4.1.15 If a matrix is non invertible, then it is called *singular* or *degenerate*.

Exercise 4.1.16 An element of the ring $\mathbb{R}^{n,n}$ needs not be invertible. The matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2,2}$$

is invertible, with inverse

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

as it can be easily checked. On the other hand, the matrix

$$A' = \begin{pmatrix} 1 & 0 \\ k & 0 \end{pmatrix} \in \mathbb{R}^{2,2}$$

is non invertible, for any value of the parameter $k \in \mathbb{R}$. It is easy to verify that the matrix equation

$$\begin{pmatrix} 1 & 0 \\ k & 0 \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

has no solutions.

Proposition 4.1.17 *The subset of invertible elements in* $\mathbb{R}^{n,n}$ *is a group with respect to the matrix product. It is called the* general linear group *of order n and is denoted by* $\mathrm{GL}(n, \mathbb{R})$ *or simply by* $\mathrm{GL}(n)$.

Proof Recall the definition of a group in A.2.7. We observe first that if A and B are both invertible then AB is invertible since $AB(B^{-1}A^{-1}) = (B^{-1}A^{-1})AB = I_n$; this means that $(AB)^{-1} = B^{-1}A^{-1}$ so $\mathrm{GL}(n)$ is closed under the matrix product. It is evident that $I_n^{-1} = I_n$ and that if A is invertible, then A is the inverse of A^{-1}, thus the latter is invertible. □

Notice that since AB is in general different from BA the group $\mathrm{GL}(n)$ is non abelian. As an example, the non commuting matrices A and B considered in the Remark 4.1.9 are both invertible.

Definition 4.1.18 Given $A = (a_{ij}) \in \mathbb{R}^{m,n}$ its *transpose*, denoted by ${}^tA$, is the matrix obtained from A when exchanging its rows with its columns, that is ${}^tA = (b_{ij}) \in \mathbb{R}^{n,m}$ with $b_{ij} = a_{ji}$.

Exercise 4.1.19 The matrix

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{2,3}$$

has transpose ${}^tA \in \mathbb{R}^{3,2}$ given by

$${}^tA = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 1 \end{pmatrix}.$$

Proposition 4.1.20 *The following identities hold:*

(i) ${}^t(A + B) = {}^tA + {}^tB$*, for any* $A, B \in \mathbb{R}^{m,n}$,
(ii) ${}^t(AB) = {}^tB\,{}^tA$*, for* $A \in \mathbb{R}^{m,n}$ *and* $B \in \mathbb{R}^{n,p}$,
(iii) if $A \in \mathrm{GL}(n)$ *then* ${}^tA \in \mathrm{GL}(n)$ *that is, if A is invertible its transpose is invertible as well with* $({}^tA)^{-1} = {}^t(A^{-1})$.

Proof (i) It is immediate.

(ii) Given $A = (a_{ij})$ and $B = (b_{ij})$, denote ${}^tA = (a'_{ij})$ and ${}^tB = (b'_{ij})$ with $a'_{ij} = a_{ji}$ and $b'_{ij} = b_{ji}$. If $AB = (c_{ij})$, then $c_{ij} = \sum_{h=1}^{n} a_{ih}b_{hj}$. The ij-element in ${}^t(AB)$ is then $\sum_{h=1}^{n} a_{jh}b_{hi}$; the ij-element in ${}^tB{}^tA$ is $\sum_{h=1}^{n} b'_{ih}a'_{hj}$ and these elements clearly coincide, for any i and j.

(iii) It is enough to show that ${}^t(A^{-1})\,{}^tA = I_n$. From (ii) one has indeed

$${}^t(A^{-1})\,{}^tA = {}^t(AA^{-1}) = {}^tI_n = I_n.$$

This finishes the proof. □

Definition 4.1.21 A square matrix of order n, $A = (a_{ij}) \in \mathbb{R}^{n,n}$, is said to be *symmetric* if ${}^tA = A$ that is, if for any i, j it holds that $a_{ij} = a_{ji}$.

Exercise 4.1.22 The matrix $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 1 \\ -1 & 1 & -1 \end{pmatrix}$ is symmetric.

4.2 The Rank of a Matrix

Definition 4.2.1 Let $A = (a_{ij})$ be a matrix in $\mathbb{R}^{m,n}$. We have seen that the m rows of A,

$$\begin{aligned} R_1 &= (a_{11}, \dots, a_{1n}), \\ &\vdots \\ R_m &= (a_{m1}, \dots, a_{mn}) \end{aligned}$$

are elements (vectors, indeed) in $\mathbb{R}^n$. By $R(A)$ we denote the vector subspace of $\mathbb{R}^n$ generated by the vectors $R_1, \dots, R_m$ that is,

$$R(A) = \mathcal{L}(R_1, \dots, R_m).$$

We call $R(A)$ the *row space* of A. Analogously, given the columns

$$C_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \cdots \quad , C_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

of A, we define the vector subspace $C(A)$ in $\mathbb{R}^m$,

$$C(A) = \mathcal{L}(C_1, \dots, C_n)$$

as the *column space* of A.

Remark 4.2.2 Clearly $C({}^tA) = R(A)$ since the columns of ${}^tA$ are the rows of A.

Theorem 4.2.3 *Given* $A = (a_{ij}) \in \mathbb{R}^{m,n}$ *one has that* $\dim(R(A)) = \dim(C(A))$.

Proof Since A is fixed, to simplify notations we set $R = R(A)$ and $C = C(A)$. The first step is to show that $\dim(C) \le \dim(R)$. Let $\dim R = r$; up to a permutation, we can take the first r rows in A as linearly independent. The remaining rows $R_{r+1}, \dots, R_m$ are elements in $R = \mathcal{L}(R_1, \dots, R_r)$ and we can write

$$A = \begin{pmatrix} R_1 \\ \vdots \\ R_r \\ \sum_{i=1}^r \lambda_i^{r+1} R_i \\ \vdots \\ \sum_{i=1}^r \lambda_i^m R_i \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rn} \\ \sum_{i=1}^r \lambda_i^{r+1} a_{i1} & \cdots & \sum_{i=1}^r \lambda_i^{r+1} a_{in} \\ \vdots & & \vdots \\ \sum_{i=1}^r \lambda_i^m a_{i1} & \cdots & \sum_{i=1}^r \lambda_i^m a_{in} \end{pmatrix}$$

for suitable scalars λ_i^j (with $i \in 1, \dots, r$, and $j \in r+1, \dots, m$). Given $h = 1, \dots, n$, consider the h-th column,

$$C_h = \begin{pmatrix} a_{1h} \\ a_{2h} \\ \vdots \\ a_{rh} \\ \sum_{i=1}^r \lambda_i^{r+1} a_{ih} \\ \vdots \\ \sum_{i=1}^r \lambda_i^m a_{ih} \end{pmatrix} = a_{1h} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \lambda_1^{r+1} \\ \vdots \\ \lambda_1^m \end{pmatrix} + \cdots + a_{rh} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \lambda_r^{r+1} \\ \vdots \\ \lambda_r^m \end{pmatrix}.$$

This means that C is generated by the r columns

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \lambda_1^{r+1} \\ \vdots \\ \lambda_1^m \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \lambda_r^{r+1} \\ \vdots \\ \lambda_r^m \end{pmatrix},$$

so we have $\dim(C) \le r = \dim R$. By exchanging the rows with columns, a similar argument shows also that $\dim(C) \ge \dim(R)$ thus the claim. □

This theorem shows that $\dim(R(A)) = \dim(C(A))$ is an integer number that characterises A.

Definition 4.2.4 Given a matrix $A \in \mathbb{R}^{m,n}$, its *rank* is the number

$$\mathrm{rk}(A) = \dim(C(A)) = \dim(R(A))$$

that is the common dimension of its space of rows, or columns.

Corollary 4.2.5 *For any* $A \in \mathbb{R}^{m,n}$ *one has* $\mathrm{rk}(A) = \mathrm{rk}({}^tA)$.

Proof This follows from Remark 4.2.2 since $C({}^tA) = R(A)$. □

It is clear that $\mathrm{rk}(A) \leq \min(m, n)$.

Definition 4.2.6 A matrix $A \in \mathbb{R}^{m,n}$ has *maximal rank* if $\mathrm{rk}(A) = \min(m, n)$.

Our task next is to give methods to compute the rank of a given matrix. We first identify a class of matrices whose rank is easy to determine.

Remark 4.2.7 It is immediate to convince one-self that the rank of a matrix A does not change by enlarging it with an arbitrary number of zero rows or columns. Moreover, if a matrix B is obtained from a matrix A by a permutation of either its rows or columns, that is, if it is

$$A = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} R_{\sigma(1)} \\ R_{\sigma(2)} \\ \vdots \\ R_{\sigma(m)} \end{pmatrix}$$

(where σ denotes a permutation of m objects) or if

$$A = (C_1, \dots, C_n) \qquad \text{and} \qquad B' = (C_{\sigma'(1)}, \dots, C_{\sigma'(n)})$$

(where σ' denotes a permutation of n objects), then $\mathrm{rk}(A) = \mathrm{rk}(B) = \mathrm{rk}(B')$. These equalities are true since the dimension of a vector space does not depend on the ordering of its basis.

Definition 4.2.8 A square matrix $A = (a_{ij}) \in \mathbb{R}^{n,n}$ is called *diagonal* if $a_{ij} = 0$ for $i \neq j$.

Exercise 4.2.9 The following matrix is diagonal,

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

Its rows and columns are vectors in $\mathbb{R}^4$, with $R_1 = e_1$, $R_2 = 2e_2$, $R_3 = 0$, $R_4 = -3e_4$ with respect to the canonical basis. As a consequence $R(A) = \mathcal{L}(R_1, R_2, R_3, R_4) = \mathcal{L}(e_1, e_2, e_4)$ so that $\mathrm{rk}(A) = 3$.

The rank of a diagonal matrix of order n coincides with the number of its non zero diagonal elements, since, as the previous exercise shows, its non zero rows or columns correspond to multiples of vectors of the canonical basis of $\mathbb{R}^n$. Beside the diagonal ones, a larger class of matrices for which the rank is easy to compute is given in the following definition.

Definition 4.2.10 Let $A = (a_{ij})$ be a square matrix in $\mathbb{R}^{n,n}$. The matrix A is called *upper triangular* if $a_{ij} = 0$ for $i > j$. An upper triangular matrix for which $a_{ii} \neq 0$ for any i, is called a *complete upper triangular* matrix.

Exercise 4.2.11 Given

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & -1 \end{pmatrix},$$

then A is upper triangular and B is complete upper triangular.

Theorem 4.2.12 *Let $A \in \mathbb{R}^{n,n}$ be a complete upper triangular matrix. Then,*

$$\mathrm{rk}(A) = n.$$

Proof Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

To prove the claim we show that the n columns $C_1, \ldots, C_n$ of A are linearly independent. The equation $\lambda_1 C_1 + \cdots + \lambda_n C_n = 0$ can be written in the form

$$\begin{pmatrix} \lambda_1 a_{11} + \cdots + \lambda_{n-1} a_{1n-1} + \lambda_n a_{1n} \\ \vdots \\ \lambda_{n-1} a_{n-1\,n-1} + \lambda_n a_{n-1\,n} \\ \lambda_n a_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

Equating term by term, one has for the n-th component $\lambda_n a_{nn} = 0$, which gives $\lambda_n = 0$ since $a_{nn} \neq 0$. For the $(n-1)$-th component, one has

$$\lambda_{n-1} a_{n-1,n-1} + \lambda_n a_{n-1,n} = 0$$

which gives, from $\lambda_n = 0$ and $a_{n-1,n-1} \neq 0$, that $\lambda_{n-1} = 0$. This can be extended step by step to all components, thus getting $\lambda_n = \lambda_{n-1} = \cdots = \lambda_1 = 0$. □

The notion of upper triangularity can be extended to non square matrices.

Definition 4.2.13 A matrix $A = (a_{ij}) \in \mathbb{R}^{m,n}$ is called *upper triangular* if it satisfies $a_{ij} = 0$ for $i > j$ and *complete upper triangular* if it is upper triangular with $a_{ii} \neq 0$ for any i.

Remark 4.2.14 Given a matrix $A \in \mathbb{R}^{m,n}$ set $p = \min(m, n)$. If A is a complete upper triangular matrix, the submatrix B made by the first p rows of A when $m > n$, or the first p columns of A when $m < n$, is a square complete upper triangular matrix of order p.

Exercise 4.2.15 The following matrices are complete upper triangular:

$$A = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A' = \begin{pmatrix} 1 & 2 & 3 & 9 \\ 0 & 2 & 0 & 7 \\ 0 & 0 & 4 & -3 \end{pmatrix}.$$

The submatrices

$$B = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad B' = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

are (square) complete upper triangular as mentioned in the previous remark.

Corollary 4.2.16 *If $A \in \mathbb{R}^{m,n}$ is a complete upper triangular matrix then* $\mathrm{rk}(A) = \min(m, n)$.

Proof We consider two cases.

- $n \geq m$. One has $\mathrm{rk}(A) \leq \min(m, n) = m$, with

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m-1} & a_{1m} & * & \dots & * \\ 0 & a_{22} & a_{23} & \dots & a_{2m-1} & a_{2m} & * & \dots & * \\ 0 & 0 & a_{33} & \dots & a_{3m-1} & a_{3m} & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 & a_{mm} & * & \dots & * \end{pmatrix}.$$

  Let B be the submatrix of A given by the its first m columns. Since B is (Remark 4.2.14) a complete upper triangular square matrix of order m, the columns $C_1, \dots, C_m$ are linearly independent. This means that $\mathrm{rk}(A) \geq m$ and the claim follows.
- $n < m$. One has $\mathrm{rk}(A) \leq \min(m, n) = n$, with

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

By deleting all zero rows, one gets a matrix of the previous type, thus $\mathrm{rk}(A) = n$. □

The matrices A and A' in the Exercise 4.2.15 are both complete upper triangular. Their rank is 3.

Remark 4.2.17 The notions introduced in the present section can be formulated by considering columns instead of rows. One has:

- A matrix $A \in \mathbb{R}^{m,n}$ is called *lower triangular* if $a_{ij} = 0$ for $i < j$. A lower triangular matrix is called *complete* if $a_{ii} \neq 0$ for any i.
- Given $A \in \mathbb{R}^{m,n}$, one has that A is (complete) upper triangular if and only if ${}^tA$ is (complete) lower triangular.
- If $A \in \mathbb{R}^{m,n}$ is a complete lower triangular matrix then $\mathrm{rk}(A) = \min(m, n)$.

4.3 Reduced Matrices

Definition 4.3.1 A matrix $A \in \mathbb{R}^{m,n}$ is said to be *reduced by rows* if any non zero row has a non zero element such that the entries below it are all zero. Such an element, which is not necessarily unique if $m \leq n$, is called the *pivot* of its row.

Exercise 4.3.2 The matrix

$$A = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

is reduced by row. The pivot element for the first row is 1, the pivot element for the third row is 2, the pivot element for the fourth row is -1. Note that $\mathrm{rk}(A) = 3$ since the three non zero rows are linearly independent.

Exercise 4.3.3 Any complete upper triangular matrix is reduced by rows.

Theorem 4.3.4 *The rank of a matrix A which is reduced by row coincides with the number of its non zero rows. Indeed, the non zero rows of a reduced by rows matrix are linearly independent.*

Proof Let A be a reduced by rows matrix and let A' be the submatrix of A obtained by deleting the zero rows of A. From the Remark 4.2.7, $\mathrm{rk}(A') = \mathrm{rk}(A)$. Let A'' be the matrix obtained by A' by the following permutation of its columns: the first column of A'' is the column of A' containing the pivot element for the first row of A', the second column of A'' is the column of A' containing the pivot element for the second row of A' and so on. By such a permutation A'' turns out to be a complete upper triangular matrix and again from the Remark 4.2.7 it is $\mathrm{rk}(A') = \mathrm{rk}(A'')$. Since A'' is complete upper triangular its rank is given by the number of its rows, the rank of A is given by the number of non zero rows of A. □

Since the proof of such a theorem is constructive, an example clarifies it.

Example 4.3.5 Let us consider the following matrix A which is reduced by rows (its pivot elements are bold typed):

$$A = \begin{pmatrix} 1 & -\mathbf{1} & 1 & 1 \\ 0 & 0 & \mathbf{2} & -1 \\ \mathbf{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{pmatrix}.$$

The first column of A' is the column containing the pivot element for the first row of A, the second column of A' is the column containing the pivot element for the second row of A and so on. The matrix A' is then

$$A' = \begin{pmatrix} -1 & 1 & 1 & 1 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and A' is complete upper triangular; so $\mathrm{rk}(A) = \mathrm{rk}(A') = 4$.

Remark 4.3.6 As we noticed in Remark 4.2.17, the notions introduced above can be formulated by exchanging the role of the columns with that of the rows of a matrix.

- A matrix $A \in \mathbb{R}^{m,n}$ is said to be *reduced by columns* if any non zero column has a non zero element such that the entries at its right are all zero. Such an element, which is not necessarily unique, is called the *pivot* of its column.
- If A is a reduced by columns matrix its rank coincides with the number of its non zero columns. The non zero columns are linearly independent.
- By mimicking the proof of the Theorem 4.3.4 it is clear that a matrix A is reduced by rows if and only if ${}^{t}A$ is reduced by column.

4.4 Reduction of Matrices

In the previous section we have learnt how to compute the rank of a reduced matrix. In this section we outline a procedure that associates to any given matrix a reduced matrix having the same rank.
We shall consider the following set of transformations acting on the rows of a matrix. They are called *elementary transformations of rows* and their action preserves the vector space structure of the space of rows.

- (λ) The transformation $R_i \mapsto \lambda R_i$ that replace the row R_i with its multiple λR_i, with $\mathbb{R} \ni \lambda \neq 0$,
- (e) The transformation $R_i \leftrightarrow R_j$, that exchanges the rows R_i and R_j,
- (D) The transformation $R_i \mapsto R_i + aR_j$ that replace the row R_i with the linear combination $R_i + aR_j$, with $a \in \mathbb{R}$ and $i \neq j$.

Given a matrix $A \in \mathbb{R}^{m,n}$ the matrix $A' \in \mathbb{R}^{m,n}$ is said to be *row-transformed* from A if A' is obtained from A by the action of a finite number of the elementary transformations (λ), (e) and (D) listed above.

Proposition 4.4.1 *Let $A \in \mathbb{R}^{m,n}$ and $A' \in \mathbb{R}^{m,n}$ be row-transformed form A. Then $R(A) = R(A')$ as vector spaces and* $\mathrm{rk}(A) = \mathrm{rk}(A')$.

Proof It is obvious that for an elementary transformation (e) or (λ) the vector spaces $R(A)$ and $R(A')$ coincide. Let us take A' to be row-transformed from A by a transformation (D). Since

$$R(A) = \mathcal{L}(R_1, \dots, R_{i-1}, R_i, R_{i+1}, \dots, R_m)$$

and

$$R(A') = \mathcal{L}(R_1, \dots, R_{i-1}, R_i + aR_j, R_{i+1}, \dots, R_m)$$

it is clear that $R(A') \subseteq R(A)$. To prove the opposite inclusion, $R(A) \subseteq R(A')$, it is enough to show that the row R_i in A is in the linear span of the rows of A'. Indeed $R_i = (R_i + aR_j) - aR_j$, thus the claim. $\square$

Exercise 4.4.2 Let

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

We act on A with the following (D) elementary transformations:

$$A \xrightarrow{R_2 \mapsto R_2 - 2R_1} A' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \\ -1 & 1 & 0 \end{pmatrix}$$

$$A' \xrightarrow{R_3' \mapsto R_3' + R_1'} A'' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \\ 0 & 1 & 1 \end{pmatrix}$$

$$A'' \xrightarrow{R_3'' \mapsto R_3'' - R_2''} A''' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 4 \end{pmatrix}.$$

The matrix A''' is reduced by rows with $\mathrm{rk}(A''') = 3$. From the proposition above, we conclude that $\mathrm{rk}(A) = \mathrm{rk}(A''') = 3$. This exercise shows how the so called *Gauss' algorithm* works.

Proposition 4.4.3 *Given any matrix A it is always possible to find a finite sequence of type* (D) *elementary transformations whose action results in a matrix (say B) which is reduced by rows.*

Proof Let $A = (a_{ij}) \in \mathbb{R}^{m,n}$. We denote by R_i the first non zero row in A and by a_{ij} the first non zero element in R_i. In order to obtain a matrix A' such that the elements under a_{ij} are zero one acts with the following (D) transformation

$$R_k \;\mapsto\; R_k - a_{kj}{a_{ij}}^{-1}R_i, \quad \text{for any } k > i.$$

We denote such a transformed matrix by $A' = (a'_{ij})$. Notice that the first i rows in A' coincide with the first i rows in A, with all the elements in the column j below the element $a'_{ij} = a_{ij}$ being null. Next, let R'_h be the first non zero row in A' with $h > i$ and let a'_{hp} be the first non zero element in R'_h. As before we now act with the following (D) elementary transformation

$$R'_k \longrightarrow R'_k - a'_{kp}{a'_{hp}}^{-1}R'_h, \quad \text{for any } k > h.$$

Let A'' the matrix obtained with this transformation and iterate. It is clear that a finite number of iterations of this procedure yield a matrix B which is—by construction—reduced by row. □

With the expression of *reduction by rows* of a matrix A we mean a finite sequence of elementary transformations on the rows of A whose final image is a matrix A' which is reduced by rows.

Remark 4.4.4 The proof of the Proposition 4.4.3 made use only of type (D) transformations. It is clear that, depending on the specific elements of the matrix one is considering, it can be easier to use also type (e) and (λ) transformations. The claim of the Proposition 4.4.1 does not change.

Exercise 4.4.5 Let us reduce by rows the following matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 9 \\ 1 & 3 & 1 & 5 \end{pmatrix}.$$

This matrix can be reduced as in the proof of the Proposition 4.4.3 by type (D) transformations alone. A look at it shows that it is convenient to swap the first row with the fourth. We have

$$A \xrightarrow{R_1 \leftrightarrow R_4} \begin{pmatrix} \mathbf{1} & 3 & 1 & 5 \\ 0 & 1 & \mathbf{2} & -1 \\ 0 & 0 & 0 & \mathbf{9} \\ 0 & \mathbf{1} & 0 & 0 \end{pmatrix} = B.$$

It is evident that the matrix B is *already* reduced by row so we can write $\mathrm{rk}(A) = \mathrm{rk}(B) = 4$.

Exercise 4.4.6 Let us consider the matrix

$$A = \begin{pmatrix} 2 & 1 & -1 & 1 \\ 3 & 1 & 1 & -1 \\ 0 & 1 & 1 & 9 \end{pmatrix}.$$

To reduce A we start with the type (D) transformation $R_2 \mapsto R_2 - 3/2R_1$, that leads to

$$A' = \begin{pmatrix} 2 & 1 & -1 & 1 \\ 0 & -1/2 & 5/2 & -5/2 \\ 0 & 1 & 1 & 9 \end{pmatrix}.$$

Since we are interested in computing the rank of the matrix A in order to avoid non integers matrix entries (which would give heavier computations) we can instead reduce by rows the matrix A' as

$$A' \xrightarrow{R_2 \mapsto 2R_2} A'' = \begin{pmatrix} 2 & 1 & -1 & 1 \\ 0 & -1 & 5 & -5 \\ 0 & 1 & 1 & 9 \end{pmatrix}$$

$$A'' \xrightarrow{R'_3 \mapsto R'_2 + R'_3} A''' = \begin{pmatrix} 2 & 1 & -1 & 1 \\ 0 & -1 & 5 & -5 \\ 0 & 0 & 6 & 4 \end{pmatrix}.$$

The matrix A''' is upper triangular so we have $\mathrm{rk}(A) = 3$.

The method of reducing by rows a matrix can be used to select a basis for a vector space V given as a linear span of a system of vectors in some $\mathbb{R}^n$, that is

$V = \mathcal{L}(v_1, \ldots, v_r)$. To this end, given the vectors $v_1, \ldots, v_r$ spanning V, one considers the matrix A with rows $v_1, \ldots, v_r$ or alternatively a matrix B with columns $v_1, \ldots, v_r$:

$$A = \begin{pmatrix} v_1 \\ \vdots \\ v_r \end{pmatrix}, \qquad B = (v_1 \cdots v_r)\,.$$

One then has $R(A) = V$ using A, which is reduced by rows to a matrix

$$A' = \begin{pmatrix} w_1 \\ \vdots \\ w_r \end{pmatrix}.$$

Clearly $V = R(A) = R(A')$ and $\dim(V) = \dim(R(A)) = \mathrm{rk}(A) = \mathrm{rk}(A')$. That is $\dim(V)$ is the number of non zero rows in A' and these non zero rows in A' are a basis for V.

Exercise 4.4.7 In $\mathbb{R}^4$ consider the system of vectors $I = \{v_1, v_2, v_3, v_4, v_5\}$ with $v_1 = (1, -1, 2, 1), v_2 = (-2, 2, -4, -2), v_3 = (1, 1, 1, -1), v_4 = (-1, 3, -3, -3)$, $v_5 = (1, 2, 1, 2)$. We would like to

(a) exhibit a basis $\mathcal{B}$ for $V = \mathcal{L}(I) \subset \mathbb{R}^4$, with $\mathcal{B} \subset I$,
(b) complete $\mathcal{B}$ to a basis $\mathcal{C}$ for $\mathbb{R}^4$.

For point (a) we let A be the matrix whose rows are the vectors in I that is,

$$A = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 & 1 \\ -2 & 2 & -4 & -2 \\ 1 & 1 & 1 & -1 \\ -1 & 3 & -3 & -3 \\ 1 & 2 & 1 & 2 \end{pmatrix}.$$

We reduce the matrix A by rows using the following transformations:

$$A \xrightarrow[\substack{R_4 \mapsto R_4 + R_1 \\ R_5 \mapsto R_5 - R_1}]{\substack{R_2 \mapsto R_2 + 2R_1 \\ R_3 \mapsto R_3 - R_1}} A' = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & -1 & -2 \\ 0 & 2 & -1 & -2 \\ 0 & 3 & -1 & 1 \end{pmatrix}$$

$$A' \xrightarrow[R'_5 \mapsto 2R'_5 - 3R'_3]{R'_4 \mapsto R'_4 - R'_3} A'' = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & -1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 8 \end{pmatrix}.$$

As a result we have $\mathrm{rk}(A) = 3$ and then $\dim(V) = 3$. A basis for V is for example given by the three non zero rows in A'' since $R(A) = R(A'')$. The basis $\mathcal{B}$ is made by the vectors in I corresponding to the three non zero rows in A'' that is $\mathcal{B} = (v_1, v_3, v_5)$. Cleary, with the transformations given above one has also that

$$B = \begin{pmatrix} v_1 \\ v_3 \\ v_5 \end{pmatrix} \quad \mapsto \quad B' = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 2 & -1 & -2 \\ 0 & 0 & 1 & 8 \end{pmatrix}.$$

To complete the basis $\mathcal{B}$ to a basis for $\mathbb{R}^4$ one can use the vectors of the canonical basis. From the form of the matrix B' it is clear that it suffices to add the vector e_4 to the three row vectors in B to meet the requirement:

$$\begin{pmatrix} v_1 \\ v_3 \\ v_5 \\ e_4 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 2 & -1 & -2 \\ 0 & 0 & 1 & 8 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We can conclude that $\mathcal{C} = (v_1, v_3, v_5, e_4)$.

Exercise 4.4.8 Let $I = \{v_1, v_2, v_3, v_4\} \subset \mathbb{R}^4$ be given by $v_1 = (0, 1, 2, 1)$, $v_2 = (0, 1, 1, 1)$, $v_3 = (0, 2, 3, 2)$, $v_4 = (1, 2, 2, 1)$. With $V = \mathcal{L}(I) \subset \mathbb{R}^4$:

(a) determine a basis $\mathcal{B}$ for V, with $\mathcal{B} \subset I$,
(b) complete $\mathcal{B}$ to a basis $\mathcal{C}$ for $\mathbb{R}^4$.

Let A be the matrix whose rows are given by the vectors in I that is,

$$A = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 3 & 2 \\ 1 & 2 & 2 & 1 \end{pmatrix}.$$

After swapping $R_1 \leftrightarrow R_4$, the matrix can be reduced following the lines above, leading to

$$\begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 3 & 2 \\ 0 & 1 & 2 & 1 \end{pmatrix} \xrightarrow[R_4 \mapsto R_4 - R_2]{R_3 \mapsto R_3 - 2R_2} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow[R_4 \mapsto R_4 - R_3]{} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We can then take $\mathcal{B} = (v_4, v_2, v_3)$. Analogously to what we did in the Exercise 4.4.7, we have

$$\begin{pmatrix} v_4 \\ v_2 \\ v_3 \\ e_4 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and such a matrix shows that we can tale $\mathcal{C} = (v_4, v_2, v_3, e_4)$ as a basis for $\mathbb{R}^4$.

Exercise 4.4.9 Consider again the set I given in the previous exercise. We now look for a basis $\mathcal{B} \subset I$ via the constructive proof of the Theorem 2.4.2. The reduction by rows procedure can be used in this case as well. Start again with

$$A = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 3 & 2 \\ 1 & 2 & 2 & 1 \end{pmatrix}.$$

The swap operated in Exercise 4.4.8 is not admissible with the procedure in the Theorem 2.4.2 so we use type (D) transformations:

$$A \xrightarrow[\substack{R_3 \mapsto R_3 - 2R_2 \\ R_4 - R_4 - R_1}]{R_2 \mapsto R_2 - R_1} \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \xrightarrow[R_3' \mapsto R_3' - R_2']{} \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

These computations show that $R_3' - R_2' = 0$, $R_3' = R_3 - 2R_1$ and $R_2' = R_2 - R_1$. From these relations we have that $R_3 - R_2 - R_1 = 0$ which is equivalent to $v_3 = v_1 + v_2$: this shows that v_3 is a linear combination of v_1 and v_2, so we recover the set $\{v_1, v_2, v_4\}$ as a basis for $\mathcal{L}(I)$.

The method we just illustrated in order to exhibit the basis of a vector subspace of $\mathbb{R}^n$ can be used with any vector space: the entries of the relevant matrix will be given by the components of a system of vectors with respect to a fixed basis.

Exercise 4.4.10 Let $V = \mathcal{L}(I) \subset \mathbb{R}^{2,3}$ with $I = \{M_1, M_2, M_3, M_4\}$ given by

$$M_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$
$$M_3 = \begin{pmatrix} 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix}, \quad M_4 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix};$$

(a) exhibit a basis $\mathcal{B}$ for V, with $\mathcal{B} \subset I$,
(b) complete $\mathcal{B}$ to a basis $\mathcal{C}$ for $\mathbb{R}^{2,3}$.

In order to use the reduction method we need to represent the matrices M_1, M_2, M_3, M_4 as row vectors. The components of these vectors will be given by the components of the matrices in a basis. This we may take to be the basis $\mathcal{E} = (E_{ij} \mid i = 1, 2; j = 1, 2, 3)$ of $\mathbb{R}^{2,3}$ made of elementary matrices as introduced in the proof of the Proposition 4.1.4. One has, for example,

$$M_1 = E_{11} + E_{12} + E_{13} + E_{22} = (1, 1, 1, 0, 1, 0)_{\mathcal{E}}.$$

Proceeding analogously we write the matrix

$$A = \begin{pmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 1 & 1 \\ 2 & 3 & 2 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 & 1 & -1 \end{pmatrix}.$$

With a suitable reduction we have

$$A \mapsto \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & -1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \end{pmatrix},$$

from which we have $\mathcal{B} = (M_1, M_2, M_4)$.
We complete $\mathcal{B}$ to a basis $\mathcal{C}$ for $\mathbb{R}^{2,3}$ by considering 3 elements in $\mathcal{E}$ and the same reduction:

$$\begin{pmatrix} M_1 \\ M_2 \\ M_4 \\ E_{13} \\ E_{21} \\ E_{22} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Since this matrix is reduced by row, the vectors $\{M_1, M_2, M_4, E_{13}, E_{21}, E_{22}\}$ are 6 linearly independent vectors in $\mathbb{R}^{2,3}$ (whose dimension is 6). This is enough to say that they give a basis $\mathcal{C}$ for $\mathbb{R}^{2,3}$ completing $\mathcal{B}$.

4.5 The Trace of a Matrix

We end this chapter with another useful notion for square matrices.

Definition 4.5.1 The *trace* of a square matrix is the function $\mathrm{tr} : \mathbb{R}^{n,n} \to \mathbb{R}$ defined as follows. If $A = (a_{ij})$ its *trace* is given by

$$\mathrm{tr}(A) \;=\; a_{11} + a_{22} + \cdots + a_{nn} = \sum_{j=1}^{n} a_{jj}.$$

That is, the trace of a matrix is the sum of its diagonal elements.

The following proposition proves an important property of the trace function for a matrix.

Proposition 4.5.2 *With $A = (a_{ij})$ and $B = (b_{ij}) \in \mathbb{R}^{n,n}$ it holds that*

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

Proof The entry (i, j) in AB is $(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$, while the entry (i, j) in BA is $(BA)_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj}$. From the row by column product of square matrices one obtaines

$$\begin{aligned} \mathrm{tr}(AB) = \sum_{j=1}^{n} (AB)_{jj} &= \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} b_{kj} \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} b_{kj} a_{jk} \\ &= \sum_{k=1}^{n} (BA)_{kk} = \mathrm{tr}(BA), \end{aligned}$$

which is the claim. □

Because of the above property one says that the trace is cyclic.