Note 8

Determinants

8.1 Determinant of a square matrix

In this section, we will introduce the *determinant* of a square matrix. Determinants will be useful when investigating if a given matrix is invertible, but will also become very useful in later chapters. We start with a notational convention:

Definition 8.1

Let $\mathbf{A} = (a_{ij})_{1 \le i \le n, 1 \le j \le n} \in \mathbb{F}^{n \times n}$ be a given square matrix. Then we define the matrix $\mathbf{A}(i;j) \in \mathbb{F}^{(n-1)\times(n-1)}$ as:

$$\mathbf{A}(i;j) = \begin{bmatrix} a_{11} & \dots & a_{1\,j-1} & a_{1\,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1\,1} & \dots & a_{i-1\,j-1} & a_{i-1\,j+1} & \dots & a_{i-1\,n} \\ a_{i+1\,1} & \dots & a_{i+1\,j-1} & a_{i+1\,j+1} & \dots & a_{i+1\,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n\,j-1} & a_{n\,j+1} & \dots & a_{nn} \end{bmatrix}.$$

In words: the matrix $\mathbf{A}(i;j)$ is obtained from $\mathbf{A}$ by deleting the i-th row and j-th column of $\mathbf{A}$. With this in place, we can define the determinant of a square matrix recursively as follows:

Definition 8.2

Let $\mathbf{A} \in \mathbb{F}^{n \times n}$ be a square matrix. Then we define

$$\det(\mathbf{A}) = \begin{cases} \mathbf{A} & \text{if } n = 1, \\ \sum_{i=1}^{n} (-1)^{i+1} \cdot a_{i1} \cdot \det(\mathbf{A}(i;1)) & \text{if } n \geq 2. \end{cases}$$

Instead of using the summation symbol, one may also write:

$$\det(\mathbf{A}) = a_{11} \cdot \det(\mathbf{A}(1;1)) - a_{21} \cdot \det(\mathbf{A}(2;1)) + \cdots + (-1)^{n+1} \cdot a_{n1} \cdot \det(\mathbf{A}(n;1)).$$

Example 8.3

Let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

To compute the determinant of this matrix, we will use Definition 8.2. First of all, note that $\mathbf{A}(1;1) = a_{22}$ and $\mathbf{A}(2;1) = a_{12}$. Therefore

$$\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = a_{11} \cdot \det(a_{22}) - a_{21} \cdot \det(a_{12}) = a_{11}a_{22} - a_{21}a_{12}. \tag{8-1}$$

When given the task to compute the determinant of a 2×2 matrix, this equation can be practical.

Example 8.4

As in Example 7.21, let $\mathbb{F} = \mathbb{R}$ and

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}.$$

Compute the determinant of $\mathbf{A}$.

Answer: First of all, note that

$$\mathbf{A}(1;1) = \begin{bmatrix} 5 & 6 \\ 7 & 9 \end{bmatrix}, \mathbf{A}(2;1) = \begin{bmatrix} 2 & 3 \\ 7 & 9 \end{bmatrix}, \text{ and } \mathbf{A}(3;1) = \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}.$$

Hence using Definition 8.2, we obtain that

$$\det(\mathbf{A}) = 1 \cdot \det\left(\begin{bmatrix} 5 & 6 \\ 7 & 9 \end{bmatrix}\right) - 4 \cdot \det\left(\begin{bmatrix} 2 & 3 \\ 7 & 9 \end{bmatrix}\right) + 5 \cdot \det\left(\begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}\right).$$

Using equation (8-1), we can quickly compute the determinants of 2×2 matrices. Then we obtain that

$$\det(\mathbf{A}) = 1 \cdot (45 - 42) - 4 \cdot (18 - 21) + 5 \cdot (12 - 15) = 3 + 12 - 15 = 0.$$

Later, we will have a few more techniques at our disposal for computing determinants of matrices, but for now we consider one particular class of matrices. Given any square matrix $\mathbf{A} = (a_{ij})_{1 \le i \le n; 1 \le j \le n}$, the entries $a_{11}, \ldots, a_{nn}$ are called the diagonal entries of $\mathbf{A}$.

Definition 8.5

A matrix $\mathbf{A} = \mathbb{F}^{n \times n}$ is called a *diagonal matrix*, if there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_{n-1} & 0 \\ 0 & \ldots & 0 & 0 & \lambda_n \end{bmatrix}.$$

In other words: a diagonal matrix is a square matrix all of whose entries are zeroes, except possibly on its diagonal. For example, the identity matrix $\mathbf{I}_n$ mentioned in the beginning of Section 7.3, is a diagonal matrix, with diagonal entries all equal to 1.

Proposition 8.6

Let $\mathbf{A} = \mathbb{F}^{n \times n}$ be a diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. Then

$$\det(\mathbf{A}) = \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n.$$

In particular $\det(\mathbf{I}_n) = 1$.

Proof. We show this using induction on n. Indeed, if $n = 1$, then $\mathbf{A} = \lambda_1$ and Definition 8.2 implies that $\det(\mathbf{A}) = \lambda_1$. Now assume that $n \geq 2$ and that the proposition holds for diagonal matrices in $\mathbb{F}^{(n-1)\times(n-1)}$. Using Definition 8.2, we then see that:

$$\begin{aligned}\det(\mathbf{A}) &= \lambda_1 \cdot \det(\mathbf{A}(1;1)) - 0 \cdot \det(\mathbf{A}(2;1)) + \cdots + (-1)^{n+1} \cdot 0 \cdot \det(\mathbf{A}(n;1)) \\ &= \lambda_1 \cdot \det(\mathbf{A}(1;1)) \\ &= \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n,\end{aligned}$$

where in the last equality we used the induction hypothesis. The induction hypothesis applies, since $\mathbf{A}(1;1)$ is a diagonal matrix with diagonal entries $\lambda_2, \ldots, \lambda_n$. This completes the induction step. Using the induction principle, we conclude that the proposition is true. The particular case of the identity matrix now also follows, since then all diagonal entries are equal to one. $\square$

We can in fact at this point already give a formula for the determinant of a larger class of matrices called upper triangular matrices:

|||| Definition 8.7

A matrix $\mathbf{A} = \mathbb{F}^{n\times n}$ is called an *upper triangular matrix*, if there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and $a_{ij} \in \mathbb{F}$ for $1 \leq i < j \leq n$, such that

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & \lambda_2 & a_{23} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} & a_{n-1\,n} \\ 0 & \cdots & 0 & 0 & \lambda_n \end{bmatrix}.$$

In words: an upper triangular matrix has all its nonzero entries above or on its diagonal. In particular, all entries below the diagonal of an upper triangular matrix are zero.

|||| Theorem 8.8

Let $\mathbf{A} = \mathbb{F}^{n\times n}$ be an upper triangular matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. Then

$$\det(\mathbf{A}) = \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n.$$

Proof. The proof is very similar as the proof of Proposition 8.6 and left to the reader. □

Another type of matrices, in the same spirit as upper triangular matrices, is the following:

Definition 8.9

A matrix $\mathbf{A} = \mathbb{F}^{n\times n}$ is called an *lower triangular matrix*, if there exist $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ and $a_{ij} \in \mathbb{F}$ for $1 \le j < i \le n$, such that

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ a_{21} & \lambda_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-1\,1} & \dots & a_{n-1\,n-2} & \lambda_{n-1} & 0 \\ a_{n\,1} & \dots & a_{n\,n-2} & a_{n\,n-1} & \lambda_n \end{bmatrix}.$$

In words: a lower triangular matrix has all its nonzero entries below or on its diagonal. In particular, all entries above the diagonal of a lower triangular matrix are zero. Also here, we can find a formula for its determinant. Before showing that, we need a lemma that can be useful in its own right.

Lemma 8.10

If a square matrix in $\mathbb{F}^{n\times n}$ contains a zero row, its determinant is zero.

Proof. This can be shown using induction on n. Providing the details, is left to the reader. □

Theorem 8.11

Let $\mathbf{A} = \mathbb{F}^{n\times n}$ be a lower triangular matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. Then

$$\det(\mathbf{A}) = \lambda_1 \cdot \dots \cdot \lambda_n.$$

Proof. We show this using induction on n. Indeed, if $n = 1$, then $\mathbf{A} = \lambda_1$ and Definition 8.2 implies that $\det(\mathbf{A}) = \lambda_1$. Now assume that $n \geq 2$ and that the proposition holds for lower diagonal matrices in $\mathbb{F}^{(n-1)\times(n-1)}$. Now note that $\mathbf{A}(1;1)$ is a lower diagonal matrix with diagonal entries $\lambda_2, \ldots, \lambda_n$. Hence the induction hypothesis implies that $\det(\mathbf{A}(1;1)) = \lambda_2 \cdot \cdots \cdot \lambda_n$. The matrices $\mathbf{A}(2;1), \ldots, \mathbf{A}(n;1)$ all have the zero row as first row. The reason for this is that the first row of $\mathbf{A}$ only has a nonzero entry in its first position, but this position has been removed when constructing the matrices $\mathbf{A}(2;1), \ldots, \mathbf{A}(n;1)$. By Lemma 8.10, we therefore have $\det(\mathbf{A}(2;1)) = 0, \ldots, \det(\mathbf{A}(n;1)) = 0$.

Using Definition 8.2, we then see that:

$$
\begin{aligned}
\det(\mathbf{A}) &= \lambda_1 \cdot \det(\mathbf{A}(1;1)) - a_{21} \cdot \det(\mathbf{A}(2;1)) + \cdots + (-1)^{n+1} \cdot a_{n1} \cdot \det(\mathbf{A}(n;1)) \\
&= \lambda_1 \cdot \det(\mathbf{A}(1;1)) - a_{21} \cdot 0 + \cdots + (-1)^{n+1} \cdot a_{n1} \cdot 0 \\
&= \lambda_1 \cdot \det(\mathbf{A}(1;1)) \\
&= \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n,
\end{aligned}
$$

where in the last equality we used the induction hypothesis. This completes the induction step. Using the induction principle, we conclude that the theorem is true. □

8.2 Determinants and elementary row operations

Using Definition 8.2 is not always the fastest way to compute the determinant of a square matrix. When studying systems of linear equations, three types of elementary row operations could be used to simplify a given system immensely. Motivated by this, we now study the effect of these three types of elementary row operations on the value of a determinant. The easiest to deal with is an elementary row operation of the form $R_i \leftarrow c \cdot R_i$. We start by proving a more general result.

Theorem 8.12

Consider the following three matrices in $\mathbb{F}^{n\times n}$:

$$\mathbf{A} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix}, \mathbf{B} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{b}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix}, \text{ and } \mathbf{C} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & c\cdot\mathbf{a}_i + \mathbf{b}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix},$$

where $c \in \mathbb{F}$. Then $\det(\mathbf{C}) = c \cdot \det(\mathbf{A}) + \det(\mathbf{B})$.

Proof. We use induction on n. If $n = 1$, we have $\mathbf{A} = a$ for some $a \in \mathbb{F}$, $\mathbf{B} = b$ for some $b \in \mathbb{F}$ and $\mathbf{C} = c \cdot a + b$. Then according to Definition 8.2, we see that $\det(\mathbf{C}) = c \cdot a + b = c \cdot \det(\mathbf{A}) + \det(\mathbf{B})$.

Now assume that $n \geq 2$ and that the theorem holds for $(n-1) \times (n-1)$ matrices. We know from Definition 8.2 that

$$\det(\mathbf{C}) = \sum_{k=1}^{n} (-1)^{k+1} \cdot c_{k1} \cdot \det(\mathbf{C}(k;1)).$$

Let us denote by $\sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot c_{k1} \cdot \det(\mathbf{C}(k;1))$ the summation one obtains by letting k range from 1 to n, except that now the value i is skipped. Then we can write

$$\det(\mathbf{C}) = \sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot c_{k1} \cdot \det(\mathbf{C}(k;1)) + (-1)^{i+1} \cdot c_{i1} \cdot \det(\mathbf{C}(i;1)).$$

For all k different from i, the induction hypothesis implies that $\det(\mathbf{C}(k;1)) = c \cdot \det(\mathbf{A}(k;1)) + \det(\mathbf{B}(k;1))$. Further $\mathbf{C}(i;1) = \mathbf{A}(i;1) = \mathbf{B}(i;1)$, since the i-th row is the only row in which the matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ differ. Now using that $c_{i1} = c \cdot a_{i1} + b_{i1}$ and $c_{k1} = a_{k1}$ if

$k \neq i$, we see that

$$\begin{aligned}
\det(\mathbf{C}) &= \sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot a_{k1} \cdot \det(\mathbf{C}(k;1)) + \\
&\quad (-1)^{i+1} \cdot (a_{i1} + b_{i1}) \cdot \det(\mathbf{C}(i;1)) \\
&= \sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot a_{k1} \cdot (c \cdot \det(\mathbf{A}(k;1)) + \det(\mathbf{B}(k;1))) \\
&\quad + (-1)^{i+1} \cdot c \cdot a_{i1} \cdot \det(\mathbf{A}(i;1)) + (-1)^{i+1} \cdot b_{i1} \cdot \det(\mathbf{B}(i;1)) \\
&= \sum_{k=1;k\neq i}^{n} c \cdot (-1)^{k+1} \cdot a_{k1} \cdot \det(\mathbf{A}(k;1)) + (-1)^{i+1} \cdot c \cdot a_{i1} \cdot \det(\mathbf{A}(i;1)) \\
&\quad + \sum_{k=1;k\neq i}^{n} (-1)^{k+1} \cdot b_{k1} \cdot \det(\mathbf{B}(k;1))) + (-1)^{i+1} \cdot b_{i1} \cdot \det(\mathbf{B}(i;1)) \\
&= c \cdot \det(\mathbf{A}) + \det(\mathbf{B}).
\end{aligned}$$

This concludes the induction step and hence the induction proof. □

|||| Corollary 8.13

Let $\mathbf{A} \in \mathbb{F}^{n\times n}$ be given and suppose that $\mathbf{C}$ is obtained from $\mathbf{A}$ by applying the elementary row operation $R_i \leftarrow c \cdot R_i$ on $\mathbf{A}$, for some i and some $c \in \mathbb{F}$. Then $\det(\mathbf{C}) = c \cdot \det(\mathbf{A})$.

Proof. If we choose $\mathbf{b}_i = \mathbf{0}$ in Theorem 8.12, we find that $\det(\mathbf{C}) = c \cdot \det(\mathbf{A}) + \det(\mathbf{B})$, where $\mathbf{B}$ is a matrix whose i-th row is the zero row. Lemma 8.10, implies that $\det(\mathbf{B}) = 0$. Hence the corollary follows. □

Investigating the effect of the remaining two types of elementary row operation turns out to be more elaborate. What turns out to happen is the following:

$$\text{Applying } R_i \leftrightarrow R_j \text{ on a square matrix, changes the sign of the determinant.} \tag{8-2}$$

$$\text{Applying } R_i \leftarrow R_i + c \cdot R_j \text{ on a square matrix, does not affect the determinant.} \tag{8-3}$$

|||| Example 8.14

As in Example 7.21, let $\mathbb{F} = \mathbb{R}$ and

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}.$$

Compute the determinant of $\mathbf{A}$ using elementary row operations.

Answer: From Example 7.21 we can read off that:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix} \underset{R_2 \leftarrow R_2 - 4 \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 5 & 7 & 9 \end{bmatrix} \\ &\underset{R_3 \leftarrow R_3 - 5 \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -3 & -6 \end{bmatrix} \underset{R_3 \leftarrow R_3 - R_2}{\longrightarrow} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Using equation (8-3) three times, we may conclude that

$$\det(\mathbf{A}) = \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix}\right).$$

Now note that the matrix on the right-hand side is an upper triangular matrix. Hence using Theorem 8.8, we obtain that

$$\det(\mathbf{A}) = \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix}\right) = 1 \cdot (-3) \cdot 0 = 0.$$

In the rest of this section, we will prove the validity of equations (8-2) and (8-3). A reader willing to accept their validity without proof can directly proceed to Section 8.3. A reader who wants to read the proof of of equations (8-2) and (8-3) is invited to do so, but on a first reading it may be best to read Section 8.3 first.

We start with two lemmas.

|||| Lemma 8.15

Assume that $n \geq 2$ and let a square matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ be given. Further, denote by $\mathbf{B} \in \mathbb{F}^{n \times n}$ a matrix obtained from $\mathbf{A}$ by interchanging two consecutive rows of $\mathbf{A}$. Then $\det(\mathbf{B}) = -\det(\mathbf{A})$.

Proof. We prove this using induction on n.

If $n = 2$, we have
$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix},$$
implying $\det(\mathbf{A}) = a_{11} \cdot a_{22} - a_{21} \cdot a_{12}$ and $\det(\mathbf{B}) = a_{21} \cdot a_{12} - a_{11} \cdot a_{22}$. Hence $\det(\mathbf{B}) = -\det(\mathbf{A})$.

Now let $n \geq 3$ and assume that the lemma holds for $n-1$. Let us denote the two rows of $\mathbf{A}$ that are interchanged by R_i and R_{i+1}. Then, we see that $\mathbf{A}(i;1) = \mathbf{B}(i+1;1)$ and $\mathbf{A}(i+1;1) = \mathbf{B}(i;1)$. Further for $k \neq i$ and $k \neq i+1$, we have that $\mathbf{B}(k;1)$ can be obtained from $\mathbf{A}(k;1)$ by interchanging two consecutive rows. Hence for such k, we have $\det(\mathbf{B}(i;1)) = -\det(\mathbf{A}(i;1))$ from the induction hypothesis. Putting all this together, we find:

$$\begin{aligned}
\det(\mathbf{B}) &= \sum_{k=1;k\neq i;k\neq i+1}^{n} (-1)^{k+1} \cdot a_{k\,1} \cdot \det(\mathbf{B}(k;1)) \\
&\quad +(-1)^{i+1} \cdot a_{i+1\,1} \cdot \det(\mathbf{B}(i;1)) + (-1)^{i+2} \cdot a_{i\,1} \cdot \det(\mathbf{B}(i+1;1)) \\
&= -\sum_{k=1;k\neq i;k\neq i+1}^{n} (-1)^{k+1} \cdot a_{k\,1} \cdot \det(\mathbf{A}(k;1)) \\
&\quad +(-1)^{i+1} \cdot a_{i+1\,1} \cdot \det(\mathbf{A}(i+1;1)) + (-1)^{i+2} \cdot a_{i\,1} \cdot \det(\mathbf{A}(i;1)) \\
&= -\sum_{k=1}^{n} (-1)^{k+1} \cdot a_{k\,1} \cdot \det(\mathbf{A}(k;1)) \\
&= -\det(\mathbf{A}).
\end{aligned}$$

This concludes the induction step and hence the proof. □

|||| Lemma 8.16

Assume that $n \geq 2$ and let a square matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ be given. Assume that two consecutive rows of $\mathbf{A}$ are identical. Then $\det(\mathbf{A}) = 0$.

Proof. This can be shown following the same strategy as in the proof of Lemma 8.15 ☐

The above lemma is just a special case of a more general result:

Proposition 8.17

Assume that $n \geq 2$ and let a square matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ be given. Assume that two rows of $\mathbf{A}$ are identical. Then $\det(\mathbf{A}) = 0$.

Proof. If two consecutive rows of $\mathbf{A}$ are identical, Lemma 8.16 implies $\det(\mathbf{A}) = 0$. Therefore we are left with the case that two rows of $\mathbf{A}$ are identical, but that these are not consecutive. Now let us denote the two given identical rows of $\mathbf{A}$ by R_i and R_j, for some $i > j \geq 1$. We interchange rows R_i and R_{i-1}, thus moving the row R_i up in the matrix. The effect on the determinant is a sign change using Lemma 8.15. In the new matrix, the identical rows are now rows R_j and R_{i-1}. If these rows are consecutive, we stop interchanging rows, but otherwise, we move the lowest of the two identical rows up, one row at the time. Therefore, we end up with a matrix $\mathbf{B}$ with two consecutive rows. Moreover, using Lemma 8.15 each time we interchange to consecutive rows, we know that $\det(\mathbf{B}) = \pm\det(\mathbf{A})$. On the other hand, $\det(\mathbf{B}) = 0$ by Lemma 8.16. Hence we can conclude that $\det(\mathbf{A}) = 0$. ☐

We now have all the ingredients needed to show the effect of interchanging two rows on the determinant of a square matrix.

Theorem 8.18

Let a square matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$ be given and denote by $\mathbf{B} \in \mathbb{F}^{n\times n}$ a matrix obtained from $\mathbf{A}$ using an elementary operation of the form $R_i \leftrightarrow R_j$ for some integers $i < j$. Then $\det(\mathbf{B}) = -\det(\mathbf{A})$.

Proof. Let us write

$$\mathbf{A} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_{j-1} & - \\ - & \mathbf{a}_j & - \\ - & \mathbf{a}_{j+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix} \text{ and } \mathbf{C} = \begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_i + \mathbf{a}_j & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_{j-1} & - \\ - & \mathbf{a}_j + \mathbf{a}_i & - \\ - & \mathbf{a}_{j+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix}.$$

Applying Theorem 8.12 on row i of $\mathbf{C}$, we see that

$$\det(\mathbf{C}) = \det\left(\begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_i & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_{j-1} & - \\ - & \mathbf{a}_j + \mathbf{a}_i & - \\ - & \mathbf{a}_{j+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix}\right) + \det\left(\begin{bmatrix} - & \mathbf{a}_1 & - \\ & \vdots & \\ - & \mathbf{a}_{i-1} & - \\ - & \mathbf{a}_j & - \\ - & \mathbf{a}_{i+1} & - \\ & \vdots & \\ - & \mathbf{a}_{j-1} & - \\ - & \mathbf{a}_j + \mathbf{a}_i & - \\ - & \mathbf{a}_{j+1} & - \\ & \vdots & \\ - & \mathbf{a}_n & - \end{bmatrix}\right)$$

Now we apply Theorem 8.12 again, but this time for row j in the two determinants on the right-hand side of this equation and use Proposition 8.17 afterwards. Then we obtain that

$$\det(\mathbf{C}) = \det(\mathbf{A}) + \det(\mathbf{B}).$$

However, Proposition 8.17 implies that $\det(\mathbf{C}) = 0$, since rows i and j of $\mathbf{C}$ are identical. Hence $0 = \det(\mathbf{A}) + \det(\mathbf{B})$, which implies what we wanted to show. □

Now that we know the effect of the elementary row operations $R_i \leftarrow c \cdot R_i$ and $R_i \leftrightarrow R_j$ on the determinant, let us also see what happens with the determinant when using an elementary operations of the form $R_i \leftarrow R_i + c \cdot R_j$.

Theorem 8.19

Let $\mathbf{A} \in \mathbb{F}^{n\times n}$ be given and suppose that the matrix $\mathbf{B}$ is obtained from $\mathbf{A}$ by applying the elementary row operation $R_i \leftarrow R_i + c \cdot R_j$ on $\mathbf{A}$, for some distinct row indices i, j, and $c \in \mathbb{F}$. Then $\det(\mathbf{B}) = \det(\mathbf{A})$.

Proof. This follows from Theorem 8.12 and Proposition 8.17. □

8.3 Alternative descriptions of the determinant

In our description of a determinant of a square matrix $\mathbf{A} \in \mathbb{F}^{n\times n}$, the first column of $\mathbf{A}$ played a special role. After all, in the recursive definition, we multiply entries from the first column of $\mathbf{A}$ with the determinants of smaller matrices. These smaller matrices were obtained from $\mathbf{A}$ by deleting the first column and some row. For this reason, one sometimes says that one in Definition 8.2 computes the determinant by expanding it along the first column. More precisely, one often refers to this as the *expansion* or *Laplace expansion* of the determinant along the first column.

One can now ask if there is any reason why the first column is so special. The answer is: it is not! It is possible to compute determinants by expansion along other columns and in fact also by expansion along rows. More precisely, we have the following theorem:

Theorem 8.20

Let $n \geq 2$ and $\mathbf{A} \in \mathbb{F}^{n\times n}$ be a square matrix. Then for any j between 1 and n:

$$\det(\mathbf{A}) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{i\,j} \cdot \det(\mathbf{A}(i;j)). \tag{8-4}$$

Moreover, for any i between 1 and n:

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{i\,j} \cdot \det(\mathbf{A}(i;j)). \tag{8-5}$$

Proof. We will not prove this theorem, but the interested reader can find some remarks at the end of this section explaining the main ideas behind the proof. □

Note that for $j = 1$, equation (8-4) simply becomes the formula given for the determinant given in Definition 8.2. Equation (8-4) describes the Laplace expansion of the determinant along the j-th column, while equation (8-5) describes the Laplace expansion of the determinant along the i-th row. These equations can also be expressed without using the summation sign in the following way:

$$\det(\mathbf{A}) = (-1)^{1+j} \cdot a_{1j} \cdot \det(\mathbf{A}(1;j)) + (-1)^{2+j} \cdot a_{2j} \cdot \det(\mathbf{A}(2;j)) + \\ \cdots + (-1)^{n+j} \cdot a_{nj} \cdot \det(\mathbf{A}(n;j)).$$

and

$$\det(\mathbf{A}) = (-1)^{i+1} \cdot a_{i1} \cdot \det(\mathbf{A}(i;1)) + (-1)^{i+2} \cdot a_{i2} \cdot \det(\mathbf{A}(i;2)) + \\ \cdots + (-1)^{i+n} \cdot a_{in} \cdot \det(\mathbf{A}(i;n)).$$

|||| Example 8.21

As in Example 7.21, let $\mathbb{F} = \mathbb{R}$ and

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}.$$

Compute the determinant of $\mathbf{A}$ using Laplace expansion along the first row.

Answer: First of all, note that

$$\mathbf{A}(1;1) = \begin{bmatrix} 5 & 6 \\ 7 & 9 \end{bmatrix}, \mathbf{A}(1;2) = \begin{bmatrix} 4 & 6 \\ 5 & 9 \end{bmatrix}, \text{ and } \mathbf{A}(1;3) = \begin{bmatrix} 4 & 5 \\ 5 & 7 \end{bmatrix}.$$

Hence using Laplace expansion along the first row, we obtain that

$$\det(\mathbf{A}) = (-1)^{1+1} \cdot 1 \cdot \det\left(\begin{bmatrix} 5 & 6 \\ 7 & 9 \end{bmatrix}\right) + (-1)^{1+2} \cdot 2 \cdot \det\left(\begin{bmatrix} 4 & 6 \\ 5 & 9 \end{bmatrix}\right) + \\ (-1)^{1+3} \cdot 3 \cdot \det\left(\begin{bmatrix} 4 & 5 \\ 5 & 7 \end{bmatrix}\right).$$

Using equation (8-1), we can quickly compute the determinants of 2×2 matrices. Then we obtain that

$$\det(\mathbf{A}) = 1 \cdot (45 - 42) - 2 \cdot (36 - 30) + 3 \cdot (28 - 25) = 3 - 12 + 9 = 0.$$

Theorem 8.20 has a nice consequence involving transpose matrices.

|||| Corollary 8.22

Let $\mathbf{A} \in \mathbb{F}^{n\times n}$ be given. Then $\det(\mathbf{A}) = \det(\mathbf{A}^T)$.

Proof. We use induction on n. If $n = 1$, $\mathbf{A} = \mathbf{A}^T$, so certainly $\det(\mathbf{A}) = \det(\mathbf{A}^T)$. Now assume $n \geq 2$ and that the corollary holds for $n-1$. Note that $\mathbf{A}(j;1)^T = \mathbf{A}^T(1;j)$, so that using the induction hypothesis, we may use that $\det(\mathbf{A}^T(1;j)) = \det(\mathbf{A}(j;1)^T) = \det(\mathbf{A}(j;1))$. Now using Laplace expansion of the determinant of $\mathbf{A}^T$ along the first row, we see that

$$\begin{aligned}\det(\mathbf{A}^T) &= \sum_{j=1}^{n}(-1)^{1+j}\cdot(\mathbf{A}^T)_{1j}\cdot\det(\mathbf{A}^T(1;j))\\ &= \sum_{j=1}^{n}(-1)^{1+j}\cdot a_{j1}\cdot\det(\mathbf{A}(j;1))\\ &= \det(\mathbf{A}),\end{aligned}$$

where in the last equality, we used Definition 8.2. This concludes the induction step and thereby the proof. □

Finally, one very important property of determinants that we want to mention here, is that determinants behave well with respect to matrix multiplication:

|||| Theorem 8.23

Let $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{n\times n}$ be given. Then $\det(\mathbf{A}\cdot\mathbf{B}) = \det(\mathbf{A})\cdot\det(\mathbf{B})$.

The interested reader can find a sketch of the proof at the end of this section, but this is not required reading. This theorem looks innocent, but has a number of consequences that all are quite important for us later on. We formulate them as a number of corollaries.

|||| Corollary 8.24

Let $\mathbf{A} \in \mathbb{F}^{n\times n}$ be given. Then $\mathbf{A}$ has an inverse if and only if $\det(\mathbf{A}) \neq 0$.

Proof. If $\mathbf{A}$ has an inverse $\mathbf{A}^{-1}$, then $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}_n$. Applying Theorem 8.23, we see that $\det(\mathbf{A}) \cdot \det(\mathbf{A}^{-1}) = \det(\mathbf{I}_n) = 1$. For the last equality we used Proposition 8.6. But then $\det(\mathbf{A}) \neq 0$, since otherwise the product $\det(\mathbf{A}) \cdot \det(\mathbf{A}^{-1})$ would be zero.

Conversely, assume that $\det(\mathbf{A}) \neq 0$. If we transform $\mathbf{A}$ using any sequence of elementary row operations to a matrix $\mathbf{B}$ in reduced row echelon form, then Corollary 8.13 and Theorems 8.18, 8.19 imply that $\det(\mathbf{B}) = d \cdot \det(\mathbf{A})$ for some nonzero constant $d \in \mathbb{F}$. Therefore $\det(\mathbf{B}) \neq 0$. This means in particular that $\mathbf{B}$ does not contain a zero row, since otherwise its determinant would be zero by Lemma 8.10. But then $\mathbf{B} = \mathbf{I}_n$, implying that $\mathbf{A}$ has rank n. As observed in equation (7-10) and Corollary 7.25, this implies that $\mathbf{A}$ has an inverse. □

|||| Corollary 8.25

Let $\mathbf{A} \in \mathbb{F}^{n \times n}$ be given. Then the columns of $\mathbf{A}$ are linearly independent if and only if $\det(\mathbf{A}) \neq 0$.

Proof. This follows by combining Theorem 7.8, the previous corollary, and Corollary 7.25. □

|||| Corollary 8.26

Let $\mathbf{A} \in \mathbb{F}^{n \times n}$ be given. Then $\det(\mathbf{A}) \neq 0$ if and only if the homogeneous system of linear equations with coefficient matrix $\mathbf{A}$ only has the zero vector as solution.

Proof. This follows by combining Corollaries 6.30, 7.25, and 8.24. □

We will not prove Theorem 8.23 in detail, but the reader who would like to know more, can read the remainder of this section and get a good impression on why this theorem as well as Theorem 8.20 is true. The remainder of this section can be skipped on a first reading. If a reader is willing to accept the statements of Theorems 8.20 and 8.23 without proof, feel free to continue to the next chapter.

The key to understanding why Theorem 8.20 is true is the following:

|||| Lemma 8.27

Let $f : \mathbb{F}^{n\times n} \to \mathbb{F}$ be a function that satisfies the following two conditions:

1. $f(\mathbf{A}) = 0$ for all square matrices $\mathbf{A} \in \mathbb{F}^{n\times n}$ that have two identical rows.
2. For all matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ as given in Theorem 8.12, it holds that $f(\mathbf{C}) = c \cdot f(\mathbf{A}) + f(\mathbf{B})$.

Then $f(\mathbf{A}) = \det(\mathbf{A}) \cdot f(\mathbf{I}_n)$ for all $\mathbf{A} \in \mathbb{F}^{n\times n}$.

Proof. We only sketch the proof: the two conditions that f satisfies, are enough to deduce exactly how the value of f changes, when a matrix $\mathbf{A}$ is changed using an elementary row operation. In fact, many of the proofs in Section 8.2 can be reused. The two given conditions are also enough to deduce that $f(\mathbf{A}) = 0$ for all $\mathbf{A}$ that have a zero row. The outcome is then that f behaves exactly the same as the determinant under elementary row operations and that both f and the determinant take the value zero for matrices with a zero row.

Given any square matrix $\mathbf{A}$ and a sequence of elementary row operations that transform $\mathbf{A}$ into its reduced row echelon form, say $\mathbf{B}$, one can then compare the values of f and the determinant under these elementary row operations. The outcome is that $f(\mathbf{A}) = d \cdot f(\mathbf{B})$ for some constant $d \in \mathbb{F}$, but also $\det(\mathbf{A}) = d \cdot \det(\mathbf{B})$ for *the same* constant d. If $\mathbf{A}$ has rank strictly less than n, its reduced row echelon form $\mathbf{B}$ contains a zero row. But then $f(\mathbf{B}) = 0$ and $\det(\mathbf{B}) = 0$. If $\mathbf{A}$ has rank n, then $\mathbf{B} = \mathbf{I}_n$. Hence in this case $f(\mathbf{A}) = d \cdot f(\mathbf{I}_n)$, while $\det(\mathbf{A}) = d \cdot \det(\mathbf{I}_n) = d \cdot 1 = d$. In all cases, we see that $f(\mathbf{A}) = \det(\mathbf{A}) \cdot f(\mathbf{I}_n)$. □

Note that the determinant as we defined it in Definition 8.2 satisfies the two conditions from Lemma 8.27, see Proposition 8.17 and Theorem 8.12. To prove that Theorem 8.20 is valid, what one needs to do is to show that the function f one obtains by expanding a determinant along some row or some column, always has the properties mentioned in Lemma 8.27 and that $f(\mathbf{I}_n) = 1$. To a high extent, this can be done similarly to how we showed these things for the determinant defined in Definition 8.2.

Finally, let us give a sketch of the proof of Theorem 8.23:

Proof. To give a proof sketch of Theorem 8.23, we consider the function $f : \mathbb{F}^{n\times n} \to \mathbb{F}$ defined by $f(\mathbf{A}) = \det(\mathbf{A} \cdot \mathbf{B})$ for some arbitrarily chosen $\mathbf{B} \in \mathbb{F}^{n\times n}$. Using Proposition

8.17 and Theorem 8.12, one first shows that f satisfies the conditions in Lemma 8.27. One can then conclude that $f(\mathbf{A}) = \det(\mathbf{A}) \cdot f(\mathbf{I}_n)$ for all $\mathbf{A} \in \mathbb{F}^{n \times n}$. But then $\det(\mathbf{A} \cdot \mathbf{B}) = f(\mathbf{A}) = \det(\mathbf{A}) \cdot f(\mathbf{I}_n) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$. In the last equality, we used that $\mathbf{I}_n \cdot \mathbf{B} = \mathbf{B}$. $\square$