Linear Algebra 2: Lecture 15

Irena Penev

Summer 2023

Throughout this lecture, $\mathbb{F}$ is a fixed field.

1 A brief review of permutations

In this section, we briefly review permutations, focusing on those properties that we need for our study of determinants. A more detailed exposition can be found in section 3 of Lecture Notes 5.

A *permutation* of a set X is any bijection from X to itself. For a positive integer n, S_n denotes the set of all permutations of the set $\{1, \dots, n\}$. S_n is a group under the composition of functions. Recall that any permutation in S_n can be written as a composition of disjoint cycles, i.e. any permutation in S_n has a disjoint cycle decomposition. For instance, we have the following permutation in S_9:

$$\pi \quad = \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 5 & 4 & 1 & 9 & 8 & 7 & 6 & 3 \end{pmatrix} \quad = \quad (125934)(68)(7).$$

When n is clear from context, cycles of length one may be omitted. So, for the permutation π above, we may also write $\pi = (125934)(68)$. The identity permutation in S_n is denoted by 1, i.e. $1 = (1)(2)\dots(n)$. A *transposition* in S_n is a permutation of the form (ij), for some distinct $i, j \in \{1, \dots, n\}$.

The *sign* of a permutation in S_n is $(-1)^{n-k}$, where k is the number of cycles in the disjoint cycle decomposition of that permutation (with cycles of length one included). For example, the sign of the permutation π above is $\text{sgn}(\pi) = (-1)^{9-3} = 1$. A permutation is *even* if its sign is 1, and it is *odd* if its sign is -1. The set of all even permutations in S_n is denoted by A_n; this is a subgroup of S_n, and it is called the *alternating group* of degree n.

Note that the sign of any transposition is -1, i.e. transpositions are odd.

We will also need the following two propositions from Lecture Notes 5.

Proposition 3.2 from Lecture Notes 5. *Let $n \geq 2$ be an integer, and let π be a permutation in S_n. Then $sgn(\pi^{-1}) = sgn(\pi)$.*

Proposition 3.7 from Lecture Notes 5. *Let $n \geq 2$ be an integer, and let $\sigma, \tau \in S_n$. Then $sgn(\sigma \circ \tau) = sgn(\sigma)sgn(\tau)$.*

Since transpositions have sign -1, Proposition 3.7 from Lecture Notes 5 implies that $\operatorname{sgn}(\sigma \circ \tau) = -\operatorname{sgn}(\sigma)$ for any permutation σ and transposition τ in S_n.

2 Determinants

The *determinant* of a matrix $A = [a_{i,j}]_{n\times n}$ in $\mathbb{F}^{n\times n}$, denoted by $\det(A)$ or $|A|$, is defined by

$$\begin{aligned}\det(A) &:= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \\ &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.\end{aligned}$$

Notation: We typically write

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix}$$

instead of

$$\det\left(\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \right)$$

Remark: Only square matrices have determinants. Moreover, the determinant of a matrix in $\mathbb{F}^{n\times n}$ is always a scalar in $\mathbb{F}$.

Proposition 2.1. *We have the following formulas for the determinants of* 1×1, 2×2, *and* 3×3 *matrices (with entries in the field* $\mathbb{F}$*):*

(a) $\begin{vmatrix} a_{1,1} \end{vmatrix} = a_{1,1}$;[1]

(b) $\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$;

(c) $\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = \left\{ \begin{array}{l} \quad a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\ -a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3}. \end{array} \right.$

[1] Be careful not to confuse this with the absolute value! (The notation is admittedly somewhat unfortunate/ambiguous.) If there is any danger of confusion, it is always possible to write $\det\left(\begin{bmatrix} a_{1,1} \end{bmatrix} \right)$ instead of $\begin{vmatrix} a_{1,1} \end{vmatrix}$.

Proof. (a) S_1 has just one element, namely $\sigma_1 = (1)$, with $\text{sgn}(\sigma_1) = 1$. So, we have that

$$\begin{vmatrix} a_{1,1} \end{vmatrix} = \text{sgn}(\sigma_1) a_{1,\sigma_1(1)} = a_{1,1}.$$

(b) S_2 has two elements, listed below, along with their signs.

- $\sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = (1)(2)$, with $\text{sgn}(\sigma_1) = 1$;
- $\sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = (12)$, with $\text{sgn}(\sigma_2) = -1$.

So, we have that

$$\begin{aligned} \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} &= \text{sgn}(\sigma_1) a_{1,\sigma_1(1)} a_{2,\sigma_1(2)} + \text{sgn}(\sigma_2) a_{1,\sigma_2(1)} a_{2,\sigma_2(2)} \\ &= a_{1,1}a_{2,2} - a_{1,2}a_{2,1}. \end{aligned}$$

(c) S_3 has six elements, listed below, along with their signs.

- $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = (1)(2)(3)$, with $\text{sgn}(\sigma_1) = 1$;
- $\sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (123)$, with $\text{sgn}(\sigma_2) = 1$;
- $\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (132)$, with $\text{sgn}(\sigma_3) = 1$;
- $\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (13)(2)$, with $\text{sgn}(\sigma_4) = -1$;
- $\sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (1)(23)$, with $\text{sgn}(\sigma_5) = -1$;
- $\sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (12)(3)$, with $\text{sgn}(\sigma_6) = -1$.

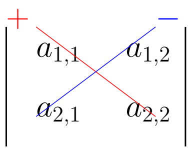
So, we have that

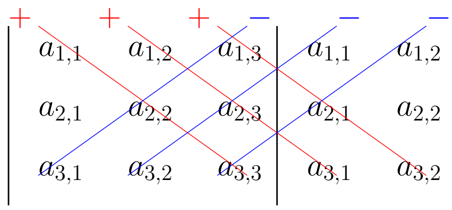
$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = \left\{ \begin{array}{l} \mathrm{sgn}(\sigma_1) a_{1,\sigma_1(1)} a_{2,\sigma_1(2)} a_{3,\sigma_1(3)} \\ +\mathrm{sgn}(\sigma_2) a_{1,\sigma_2(1)} a_{2,\sigma_2(2)} a_{3,\sigma_2(3)} \\ +\mathrm{sgn}(\sigma_3) a_{1,\sigma_3(1)} a_{2,\sigma_3(2)} a_{3,\sigma_3(3)} \\ +\mathrm{sgn}(\sigma_4) a_{1,\sigma_4(1)} a_{2,\sigma_4(2)} a_{3,\sigma_4(3)} \\ +\mathrm{sgn}(\sigma_5) a_{1,\sigma_5(1)} a_{2,\sigma_5(2)} a_{3,\sigma_5(3)} \\ +\mathrm{sgn}(\sigma_6) a_{1,\sigma_6(1)} a_{2,\sigma_6(2)} a_{3,\sigma_6(3)} \end{array} \right.$$

$$= \left\{ \begin{array}{l} a_{1,1} a_{2,2} a_{3,3} \\ +a_{1,2} a_{2,3} a_{3,1} \\ +a_{1,3} a_{2,1} a_{3,2} \\ -a_{1,3} a_{2,2} a_{3,1} \\ -a_{1,1} a_{2,3} a_{3,2} \\ -a_{1,2} a_{2,1} a_{3,3}. \end{array} \right.$$

□

Determinants of 2×2 and 3×3 matrices can be represented schematically, as shown below.

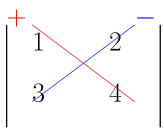




We multiply the entries along each of the red lines and add them up, and then we multiply the entries along each of the blue lines and subtract them. In each case, the result we get is precisely the formula from Proposition 2.1. For example, we can compute the determinant of the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

in $\mathbb{R}^{2\times 2}$ by forming the diagram



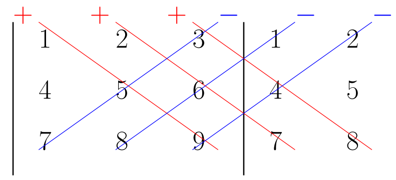
and the computing

$$\det(A) = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = {\color{red}1 \cdot 4} - {\color{blue}2 \cdot 3} = -2.$$

Similarly, we can compute the determinant of the matrix

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

in $\mathbb{R}^{3\times 3}$ by forming the diagram



and then computing

$$\begin{aligned} \det(B) &= \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} \\ &= 1\cdot 5\cdot 9 + 2\cdot 6\cdot 7 + 3\cdot 4\cdot 8 - 3\cdot 5\cdot 7 - 1\cdot 6\cdot 8 - 2\cdot 4\cdot 9 \\ &= 0. \end{aligned}$$

Warning: Do not try this with matrices of larger size!

Theorem 2.2. *For all $A \in \mathbb{F}^{n\times n}$, we have that $det(A) = det(A^T)$.*

Proof. We set $A = [a_{i,j}]_{n\times n}$ and $A^T = [a^T_{i,j}]_{n\times n}$. So, for all $i, j \in \{1,\dots,n\}$, we have $a_{i,j} = a^T_{j,i}$. Now, we compute:

$$\begin{aligned} \det(A^T) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a^T_{i,\sigma(i)} \\ &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{\sigma(i),i} \\ &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^n a_{j,\sigma^{-1}(j)} \\ &\overset{(*)}{=} \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma^{-1}) \prod_{j=1}^n a_{j,\sigma^{-1}(j)} \\ &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \prod_{j=1}^n a_{j,\pi(j)} \\ &= \det(A), \end{aligned}$$

where (*) follows from Proposition 3.2 from Lecture Notes 5. □

2.1 Some matrices whose determinants are zero

Proposition 2.3. *Let $A = [a_{i,j}]_{n\times n}$ be a matrix in $\mathbb{F}^{n\times n}$. If A has a zero row or a zero column,*[2] *then $det(A) = 0$.*

Proof. In view of Theorem 2.2, it suffices to consider the case when A has a zero row.[3] Suppose that that the p-th row of A is a zero row. Then for all $\sigma \in S_n$, we have that $a_{p,\sigma(p)} = 0$. Consequently,

$$\det(A) \quad = \quad \sum_{\sigma\in S_n} \text{sgn}(\sigma)a_{1,\sigma(1)}\dots a_{n,\sigma(n)} \quad = \quad 0,$$

which is what we needed to show. ☐

Proposition 2.4. *Let $A = [a_{i,j}]_{n\times n}$ be a matrix in $\mathbb{F}^{n\times n}$. If A has two identical rows or two identical columns, then $det(A) = 0$.*

Proof. In view of Theorem 2.2, it suffices to consider the case when A has two identical rows.[4] So, suppose that for some distinct $p, q \in \{1, \dots, n\}$, the p-th and q-th row of A are the same. (In particular, $n \geq 2$.) Now, let A_n be the alternating group of degree n, i.e. the group of all even permutations in S_n, and let O_n be the set of all odd permutations in S_n.[5] Obviously,

$$S_n \quad = \quad A_n \cup O_n \qquad \text{and} \qquad A_n \cap O_n \quad = \quad \emptyset.$$

Next, consider the transposition $\tau = (pq)$. By Proposition 3.7 from Lecture Notes 5, for all $\sigma \in S_n$, we have that $\text{sgn}(\sigma\circ\tau) = -\text{sgn}(\sigma)$; it then readily follows that $O_n = \{\sigma\circ\tau \mid \sigma \in A_n\}$,[6] and obviously, for all distinct $\sigma_1, \sigma_2 \in A_n$, we have that $\sigma_1 \circ \tau \neq \sigma_2 \circ \tau$.[7]

---

[2]A *zero row* is a row in which all entries are zero. Similarly, a *zero column* is a column in which all entries are zero.

[3]Let us explain this in more detail. Suppose we have shown that if a matrix in $\mathbb{F}^{n\times n}$ has a zero row, then its determinant is zero. Suppose now that B is a matrix in $\mathbb{F}^{n\times n}$ that has a zero column. Then B^T has a zero row, and we see that $\det(B) \overset{(*)}{=} \det(B^T) \overset{(**)}{=} 0$, where (*) follows from Theorem 2.2, and (**) follows from the fact that B^T has a zero row.

[4]Let us explain this in more detail. Suppose we have shown that if a matrix in $\mathbb{F}^{n\times n}$ has two identical rows, then its determinant is zero. Let us prove this for matrices with two identical columns. Suppose B is a matrix in $\mathbb{F}^{n\times n}$ with two identical columns. Then B^T has two identical rows, and we see that $\det(B) \overset{(*)}{=} \det(B^T) \overset{(**)}{=} 0$, where (*) follows from Theorem 2.2, and (**) follows from the fact that B^T has two identical rows.

[5]Unlike A_n, O_n is not a group.

[6]Let us check this. First, suppose that $\sigma \in A_n$. Then $\text{sgn}(\sigma\circ\tau) \overset{(*)}{=} -\text{sgn}(\sigma) \overset{(**)}{=} -1$, where (*) follows from Proposition 3.7 from Lecture Notes 5, and (**) follows from the fact that σ is even. So, $\sigma\circ\tau \in O_n$.

Conversely, suppose that $\pi \in O_n$. Set $\sigma := \pi\circ\tau$. Then $\text{sgn}(\sigma) = \text{sgn}(\pi\circ\tau) \overset{(*)}{=} -\text{sgn}(\pi) \overset{(**)}{=} 1$, where (*) follows from Proposition 3.7 from Lecture Notes 5, and (**) follows from the fact that π is odd. So, $\sigma \in A_n$. But τ is a transposition, and consequently, $\tau^{-1} = \tau$. So, $\pi = \pi\circ\tau\circ\tau = \sigma\circ\tau$.

[7]This follows from the fact that τ is a bijection. So, if we had that $\sigma_1\circ\tau = \sigma_2\circ\tau$, then we would have $\sigma_1 = \sigma_1\circ\tau\circ\tau^{-1} = \sigma_2\circ\tau\circ\tau^{-1} = \sigma_2$.

Claim. For all $\sigma \in S_n$, we have that $\prod_{i=1}^n a_{i,\sigma(i)} = \prod_{i=1}^n a_{i,\sigma\circ\tau(i)}$.

Proof of the Claim. Fix $\sigma \in S_n$. First, note that

- $a_{p,\sigma(p)} = a_{p,\sigma\circ\tau(q)} \overset{(*)}{=} a_{q,\sigma\circ\tau(q)}$,
- $a_{q,\sigma(q)} = a_{q,\sigma\circ\tau(p)} \overset{(*)}{=} a_{p,\sigma\circ\tau(p)}$,

where in both cases, (*) follows from the fact that the p-th and q-th row of A are the same. So, $a_{p,\sigma(p)}a_{q,\sigma(q)} = a_{p,\sigma\circ\tau(p)}a_{q,\sigma\circ\tau(q)}$. On the other hand, it is clear that for all $i \in \{1,\dots,n\} \setminus \{p,q\}$, we have that $a_{i,\sigma(i)} = a_{i,\sigma\circ\tau(i)}$. It follows that $\prod_{i=1}^n a_{i,\sigma(i)} = \prod_{i=1}^n a_{i,\sigma\circ\tau(i)}$, which is what we needed to show. ♦

We now compute:

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)}\dots a_{n,\sigma(n)} \\
&= \sum_{\sigma\in A_n} \underbrace{\operatorname{sgn}(\sigma)}_{=1} a_{1,\sigma(1)}\dots a_{n,\sigma(n)} + \sum_{\pi\in O_n} \underbrace{\operatorname{sgn}(\pi)}_{=-1} a_{1,\pi(1)}\dots a_{n,\pi(n)} \\
&= \sum_{\sigma\in A_n} a_{1,\sigma(1)}\dots a_{n,\sigma(n)} - \sum_{\pi\in O_n} a_{1,\pi(1)}\dots a_{n,\pi(n)} \\
&= \sum_{\sigma\in A_n} a_{1,\sigma(1)}\dots a_{n,\sigma(n)} - \sum_{\sigma\in A_n} a_{1,\sigma\circ\tau(1)}\dots a_{n,\sigma\circ\tau(n)} \\
&\overset{(*)}{=} \sum_{\sigma\in A_n} a_{1,\sigma(1)}\dots a_{n,\sigma(n)} - \sum_{\sigma\in A_n} a_{1,\sigma(1)}\dots a_{n,\sigma(n)} \\
&= 0,
\end{aligned}$$

where (*) follows from the Claim. □

3 Linearity of determinants in one row or one column

In general, for matrices $A, B \in \mathbb{F}^{n\times n}$ and a scalar $\alpha \in \mathbb{F}$, we have that

$$\det(A+B) \neq \det(A)+\det(B) \qquad \text{and} \qquad \det(\alpha A) \neq \alpha\det(A).$$

We do, however, have the following proposition.

Proposition 3.1. *Let $\mathbf{a}_1,\dots,\mathbf{a}_{p-1},\mathbf{a}_{p+1},\dots,\mathbf{a}_n \in \mathbb{F}^n$. Then both the following hold:*

(a) the function $f_{C_i} : \mathbb{F}^n \to \mathbb{F}$ *given by*

$$f_{C_i}(\mathbf{x}) \quad = \quad \det\left(\begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_{p-1} & \mathbf{x} & \mathbf{a}_{p+1} & \dots & \mathbf{a}_n \end{bmatrix} \right)$$

for all $\mathbf{x} \in \mathbb{F}^n$ *is linear;*

(b) the function $f_{R_i} : \mathbb{F}^n \to \mathbb{F}$ *given by*

$$f_{R_i}(\mathbf{x}) \quad = \quad \det\left(\begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_{p-1}^T \\ \mathbf{x}^T \\ \mathbf{a}_{p+1}^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix} \right)$$

for all $\mathbf{x} \in \mathbb{F}^n$ *is linear.*

Remark: Before reading the proof, the reader might want to take a look at Example 3.2 (below), since it illustrates how Proposition 3.1 can be used in practice.

Proof. Clearly, (b) and Theorem 2.2 imply (a).[8] So, it suffices to prove (b).

We first set up some notation. For each $i \in \{1, \dots, n\} \setminus \{p\}$, set $\mathbf{a}_i = \begin{bmatrix} a_{i,1} & \dots & a_{i,n} \end{bmatrix}^T$, so that $\mathbf{a}_i^T = \begin{bmatrix} a_{i,1} & \dots & a_{i,n} \end{bmatrix}$. Now, let us prove that f_{R_i} is linear.

1. Fix $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, and set $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$.

[8] Details?

We compute:

$$\begin{aligned}
f_{R_i}(\mathbf{x}+\mathbf{y}) &= \det\left(\begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_{p-1}^T \\ (\mathbf{x}+\mathbf{y})^T \\ \mathbf{a}_{p+1}^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}\right) \\
&= \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{p-1,1} & \cdots & a_{p-1,n} \\ x_1+y_1 & \cdots & x_n+y_n \\ a_{p+1,1} & \cdots & a_{p+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{p-1,\sigma(p-1)} \left(x_{\sigma(p)} + y_{\sigma(p)}\right) a_{p+1,\sigma(p+1)} \cdots a_{n,\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{p-1,\sigma(p-1)} x_{\sigma(p)} a_{p+1,\sigma(p+1)} \cdots a_{n,\sigma(n)} \\
&\quad + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{p-1,\sigma(p-1)} y_{\sigma(p)} a_{p+1,\sigma(p+1)} \cdots a_{n,\sigma(n)} \\
&= \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{p-1,1} & \cdots & a_{p-1,n} \\ x_1 & \cdots & x_n \\ a_{p+1,1} & \cdots & a_{p+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix} + \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{p-1,1} & \cdots & a_{p-1,n} \\ y_1 & \cdots & y_n \\ a_{p+1,1} & \cdots & a_{p+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix} \\
&= \det\left(\begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_{p-1}^T \\ \mathbf{x}^T \\ \mathbf{a}_{p+1}^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}\right) + \det\left(\begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_{p-1}^T \\ \mathbf{y}^T \\ \mathbf{a}_{p+1}^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}\right) \\
&= f_{R_i}(\mathbf{x}) + f_{R_i}(\mathbf{y}).
\end{aligned}$$

2. Fix $\mathbf{x} \in \mathbb{F}^n$ and $\alpha \in \mathbb{F}$, and set $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$. We compute:

$$\begin{aligned}
f_{R_i}(\alpha\mathbf{x}) &= \det\left(\begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_{p-1}^T \\ \alpha\mathbf{x}^T \\ \mathbf{a}_{p+1}^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}\right) \\
&= \begin{vmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{p-1,1} & \dots & a_{p-1,n} \\ \alpha x_1 & \dots & \alpha x_n \\ a_{p+1,1} & \dots & a_{p+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{vmatrix} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \dots a_{p-1,\sigma(p-1)} \left(\alpha x_{\sigma(p)}\right) a_{p+1,\sigma(p+1)} \dots a_{n,\sigma(n)} \\
&= \alpha \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \dots a_{p-1,\sigma(p-1)} x_{\sigma(p)} a_{p+1,\sigma(p+1)} \dots a_{n,\sigma(n)} \\
&= \alpha \begin{vmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{p-1,1} & \dots & a_{p-1,n} \\ x_1 & \dots & x_n \\ a_{p+1,1} & \dots & a_{p+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{vmatrix} \\
&= \alpha \det\left(\begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_{p-1}^T \\ \mathbf{x}^T \\ \mathbf{a}_{p+1}^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}\right) \\
&= \alpha f_{R_i}(\mathbf{x}).
\end{aligned}$$

By 1. and 2., f_{R_i} is linear, i.e. (b) holds. $\square$

Example 3.2. *By Proposition 3.1, we have the following (entries are understood to be in $\mathbb{R}$, and the row/column being manipulated is in red to facilitate reading):*

- $\begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 4 \\ 0 & 1 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 4 \\ 0 & -2 & 5 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 0 & 3 & 5 \end{vmatrix}$

- $\begin{vmatrix} 3 & 2 & 4 \\ 6 & -1 & 0 \\ -3 & 0 & 5 \end{vmatrix} = 3 \begin{vmatrix} 1 & 2 & 4 \\ 2 & -1 & 0 \\ -1 & 0 & 5 \end{vmatrix}$

- $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 7 & 3 & -2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 4 & 4 & -2 \end{vmatrix} + \begin{vmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & -1 & 0 \end{vmatrix}$

- $\begin{vmatrix} 2 & -2 & 4 \\ 1 & 0 & -2 \\ 2 & 1 & 4 \end{vmatrix} = 2 \begin{vmatrix} 1 & -1 & 2 \\ 1 & 0 & -2 \\ 2 & 1 & 4 \end{vmatrix}$

4 Computing determinants via elementary row and column operations

In this section, we examine how performing elementary row and column operations affects the value of the determinant,[9] and how we can use these operations to compute the determinant of a square matrix.

4.1 Computing the determinant of upper-triangular and lower-triangular matrices

Given a square matrix $A = [a_{i,j}]_{n\times n}$ in $\mathbb{F}^{n\times n}$, we say that

- A is *upper triangular* if all entries of A below the main diagonal are zero, i.e. if for all $i, j \in \{1, \dots, n\}$ such that $i > j$, we have that $a_{i,j} = 0$;
- A is *lower triangular* if all entries of A above the main diagonal are zero, i.e. if for all $i, j \in \{1, \dots, n\}$ such that $i < j$, we have that $a_{i,j} = 0$;
- A is *triangular* if it is upper triangular or lower triangular.

[9] We studied elementary row operations in Linear Algebra 1. Elementary column operations are analogous, only they are performed on columns, rather than rows.

A schematic representation of an upper triangular and a lower triangular matrix is given below (*'s represent arbitrary elements of the field $\mathbb{F}$, and the main diagonal is in red in both cases).

$$\begin{bmatrix} * & * & * & \dots & * & * \\ 0 & * & * & \dots & * & * \\ 0 & 0 & * & \dots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & * & * \\ 0 & 0 & 0 & \dots & 0 & * \end{bmatrix} \qquad \begin{bmatrix} * & 0 & 0 & \dots & 0 & 0 \\ * & * & 0 & \dots & 0 & 0 \\ * & * & * & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & * & \dots & * & 0 \\ * & * & * & \dots & * & * \end{bmatrix}$$

upper triangular matrix lower triangular matrix

Note that any square matrix in row echelon form is in fact an upper triangular matrix. (However, not all upper triangular matrices are in row echelon form.) So, the row reduction algorithm performed on a square matrix will, in particular, yield an upper triangular matrix.

It turns out that the determinant of any triangular matrix is particularly easy to compute, as we now show.

Proposition 4.1. *Let $A = [a_{i,j}]_{n \times n}$ be a triangular matrix in $\mathbb{F}^{n \times n}$. Then*

$$det(A) \quad = \quad \prod_{i=1}^{n} a_{i,i} \quad = \quad a_{1,1}a_{2,2}\dots a_{n,n},$$

that is, $det(A)$ is equal to the product of entries on the main diagonal of A.

Proof. Note that the transpose of a lower triangular matrix is an upper triangular matrix, and moreover, the main diagonal remains unchanged when we take the transpose of a square matrix. So, in view of Theorem 2.2, it suffices to prove the result for the case when A is upper triangular. Now, note that for all $\sigma \in S_n \setminus \{1\}$,[10] there exists some $i \in \{1, \dots, n\}$ such that $i > \sigma(i)$, and consequently, $a_{i,\sigma(i)} = 0$ (since A is upper triangular). It follows that for all $\sigma \in S_n \setminus \{1\}$, we have that $a_{1,\sigma(1)}a_{2,\sigma(2)} \dots a_{n,\sigma(n)} = 0$, and consequently,

$$\begin{aligned} \det(A) \quad &= \quad \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)} \\ &= \quad \text{sgn}(1) a_{1,1} a_{2,2} \dots a_{n,n} \\ &= \quad a_{1,1} a_{2,2} \dots a_{n,n}. \end{aligned}$$

This completes the argument. $\square$

For example, we can compute the determinants of the following matrices in $\mathbb{R}^{3 \times 3}$:

[10] Recall that 1 is the identity permutation in S_n.

- $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} = 1 \cdot 4 \cdot 6 = 24;$
- $\begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{vmatrix} = 1 \cdot 3 \cdot 6 = 18.$

4.2 Determinants and elementary row and column operations

Theorem 4.2. *Let $A = [a_{i,j}]_{n\times n}$ be a matrix in $\mathbb{F}^{n\times n}$. Then all the following hold:*

(a) if a matrix B is obtained by swapping two rows or swapping two columns of A, then

$$det(B) \quad = \quad -det(A);$$

(b) if a matrix B is obtained by multiplying some row or some column of A by a scalar $\alpha \in \mathbb{F} \setminus \{0\}$, then

$$det(B) \quad = \quad \alpha det(A) \qquad \textit{and} \qquad det(A) \quad = \quad \alpha^{-1} det(B);$$

(c) if a matrix B is obtained from A by adding a scalar multiple of one row (resp. column) of A to another row (resp. column) of A, then

$$det(B) \quad = \quad det(A).$$

Proof. In view of Theorem 2.2, it suffices to prove the result for row operations only.

(a) Fix distinct indices $p, q \in \{1, \dots, n\}$, and suppose that B is obtained by swapping rows p and q of A ("$R_p \leftrightarrow R_q$"). Set $B = [b_{i,j}]_{n\times n}$, so that

- for all $j \in \{1, \dots, n\}$, we have that $b_{p,j} = a_{q,j}$ and $b_{q,j} = a_{p,j}$;
- for all $i \in \{1, \dots, n\} \setminus \{p, q\}$ and $j \in \{1, \dots, n\}$, we have that $b_{i,j} = a_{i,j}$.

Next, consider the transposition $\tau = (pq)$ in S_n.

> **Claim.** For all $\sigma \in S_n$, we have that $\prod\limits_{i=1}^{n} b_{i,\sigma(i)} = \prod\limits_{i=1}^{n} a_{i,\sigma\circ\tau(i)}$.

Proof of the Claim. First, we note that

- $b_{p,\sigma(p)} = a_{q,\sigma(p)} = a_{q,\sigma\circ\tau(q)}$;
- $b_{q,\sigma(q)} = a_{p,\sigma(q)} = a_{p,\sigma\circ\tau(p)}$.

So, $b_{p,\sigma(p)} b_{q,\sigma(q)} = a_{p,\sigma\circ\tau(p)} a_{q,\sigma\circ\tau(q)}$. On the other hand, for all $i \in \{1, \dots, n\} \setminus \{p, q\}$, we have that $b_{i,\sigma(i)} = a_{i,\sigma\circ\tau(i)}$. It follows that $\prod\limits_{i=1}^{n} b_{i,\sigma(i)} = \prod\limits_{i=1}^{n} a_{i,\sigma\circ\tau(i)}$, which is what we needed to show. ♦

We now compute:

$$\begin{aligned}
\det(B) &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \textstyle\prod_{i=1}^n b_{i,\sigma(i)} \\
&\overset{(*)}{=} \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \textstyle\prod_{i=1}^n a_{i,\sigma\circ\tau(i)} \\
&\overset{(**)}{=} \sum_{\sigma \in S_n} \Big(-\mathrm{sgn}(\sigma\circ\tau)\Big) \textstyle\prod_{i=1}^n a_{i,\sigma\circ\tau(i)} \\
&= -\sum_{\sigma \in S_n} \mathrm{sgn}(\sigma\circ\tau) \textstyle\prod_{i=1}^n a_{i,\sigma\circ\tau(i)} \\
&= -\sum_{\pi \in S_n} \mathrm{sgn}(\pi) \textstyle\prod_{i=1}^n a_{i,\pi(i)} \\
&= -\det(A),
\end{aligned}$$

where (*) follows from the Claim, and (**) follows from Proposition 3.7 from Lecture Notes 5. This proves (a).

(b) Fix an index $p \in \{1,\dots,n\}$ and a scalar $\alpha \in \mathbb{F}\setminus\{0\}$, and suppose that B is obtained by multiplying the p-th row of A by α ("$R_p \to \alpha R_p$"). Set $B = [b_{i,j}]_{n\times n}$, so that

- for all $j \in \{1,\dots,n\}$, we have that $b_{p,j} = \alpha a_{p,j}$;
- for all $i \in \{1,\dots,n\}\setminus\{p\}$ and $j \in \{1,\dots,n\}$, we have that $b_{i,j} = a_{i,j}$.

We now compute:

$$\begin{aligned}
\det(B) &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) b_{1,\sigma(1)} \dots b_{n,\sigma(n)} \\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \dots a_{p-1,\sigma(p-1)} \Big(\alpha a_{p,\sigma(p)}\Big) a_{p+1,\sigma(p+1)} \dots a_{n,\sigma(n)} \\
&= \alpha \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \dots a_{n,\sigma(n)} \\
&= \alpha \det(A).
\end{aligned}$$

Since $\alpha \neq 0$, we deduce that $\det(A) = \alpha^{-1}\det(B)$. This proves (b).

(c) Fix distinct indices $p, q \in \{1,\dots,n\}$ and a scalar $\alpha \in \mathbb{F}$, and suppose that B is obtained by adding α times row p to row q ("$R_q \to R_q + \alpha R_p$"). Set $B = [b_{i,j}]_{n\times n}$, so that

- for all $j \in \{1,\dots,n\}$, we have that $b_{q,j} = a_{q,j} + \alpha a_{p,j}$;

- for all $i \in \{1, \dots, n\} \setminus \{q\}$ and $j \in \{1, \dots, n\}$, we have that $b_{i,j} = a_{i,j}$.

We now compute (the q-th row is in red for emphasis):

$$\begin{aligned}
\det(B) \quad &= \quad \begin{vmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{q-1,1} & \dots & a_{q-1,n} \\ a_{q,1} + \alpha a_{p,1} & \dots & a_{q,n} + \alpha a_{p,n} \\ a_{q+1,1} & \dots & a_{q+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{vmatrix} \\
&\overset{(*)}{=} \quad \underbrace{\begin{vmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{q-1,1} & \dots & a_{q-1,n} \\ a_{q,1} & \dots & a_{q,n} \\ a_{q+1,1} & \dots & a_{q+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{vmatrix}}_{=\det(A)} + \alpha \underbrace{\begin{vmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{q-1,1} & \dots & a_{q-1,n} \\ a_{p,1} & \dots & a_{p,n} \\ a_{q+1,1} & \dots & a_{q+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{vmatrix}}_{\overset{(**)}{=} 0} \\
&= \quad \det(A),
\end{aligned}$$

where (*) follows from the fact that the determinant is linear in the q-th row (by Proposition 3.1), and (**) follows from the fact that a matrix with two identical rows (in this case, the p-th and q-th row) has determinant zero (by Proposition 2.4). □

Example 4.3. *Compute the determinant of the matrix below (with entries understood to be in $\mathbb{R}$).*

$$A \quad = \quad \begin{bmatrix} 2 & 4 & 6 \\ 2 & 4 & 4 \\ 3 & 3 & 7 \end{bmatrix}$$

Solution. We perform elementary row operations on A (keeping track of the way that this changes the value of the determinant, as per Theorem 4.2) until we transform A into a matrix in row echelon form. Square matrices in row echelon form are upper triangular, and so by Proposition 4.1, we can obtain their determinant by multiplying the entries on the main diagonal.

We now compute:

$$\begin{aligned}
\det(A) &= \begin{vmatrix} 2 & 4 & 6 \\ 2 & 4 & 4 \\ 3 & 3 & 7 \end{vmatrix} \\
&\overset{R_2 \to R_2 - R_1}{=} \begin{vmatrix} 2 & 4 & 6 \\ 0 & 0 & -2 \\ 3 & 3 & 7 \end{vmatrix} \\
&\overset{R_2 \leftrightarrow R_3}{=} -\begin{vmatrix} 2 & 4 & 6 \\ 3 & 3 & 7 \\ 0 & 0 & -2 \end{vmatrix} \\
&\overset{R_1 \to \frac{1}{2} R_1}{=} -2\begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 7 \\ 0 & 0 & -2 \end{vmatrix} \\
&\overset{R_2 \to R_1 - 3R_1}{=} -2\begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 0 & -2 \end{vmatrix} \\
&\overset{(*)}{=} (-2)1(-3)(-2) \\
&= -12,
\end{aligned}$$

where (*) follows by taking the determinant of an upper triangular matrix. $\square$

Example 4.4. *Compute the determinant of the matrix below (with entries understood to be in $\mathbb{Z}_3$).*

$$A = \begin{bmatrix} 1 & 2 & 1 & 1 & 2 \\ 1 & 1 & 0 & 2 & 1 \\ 2 & 0 & 1 & 1 & 2 \\ 2 & 2 & 0 & 0 & 1 \\ 1 & 0 & 2 & 1 & 2 \end{bmatrix}$$

Solution. Here, we just notice that the second column is the sum of the first and third. This allows us to turn the second column into a zero column via two elementary column operations, which implies that $\det(A) = 0$. The detailed computation is as follows:

$$\begin{aligned}
\det(A) &= \begin{vmatrix} 1 & 2 & 1 & 1 & 2 \\ 1 & 1 & 0 & 2 & 1 \\ 2 & 0 & 1 & 1 & 2 \\ 2 & 2 & 0 & 0 & 1 \\ 1 & 0 & 2 & 1 & 2 \end{vmatrix} \\
&\overset{C_2 \to C_2 - C_1}{=} \begin{vmatrix} 1 & 1 & 1 & 1 & 2 \\ 1 & 0 & 0 & 2 & 1 \\ 2 & 1 & 1 & 1 & 2 \\ 2 & 0 & 0 & 0 & 1 \\ 1 & 2 & 2 & 1 & 2 \end{vmatrix} \\
&\overset{C_2 \to C_2 - C_3}{=} \begin{vmatrix} 1 & 0 & 1 & 1 & 2 \\ 1 & 0 & 0 & 2 & 1 \\ 2 & 0 & 1 & 1 & 2 \\ 2 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & 1 & 2 \end{vmatrix} \\
&\overset{(*)}{=} 0,
\end{aligned}$$

where (*) follows from the fact that a matrix with a zero column has determinant zero (by Proposition 2.3). We could also have noticed that the matrix in the second line of the computation above has two identical columns, and so by Proposition 2.4, its determinant is zero. Finally, we note that our two elementary column operations could also have been written as "$C_2 \to C_2 + 2C_1$" and "$C_2 \to C_2 + 2C_3$," since in $\mathbb{Z}_3$, we have that $-1 = 2$. $\square$

5 Determinants and matrix invertibility

Theorem 5.1. *Let $A \in \mathbb{F}^{n \times n}$. Then A is invertible if and only if $\det(A) \neq 0$.*

Proof. We can transform A into a matrix in reduced row echelon form via a sequence of elementary row operations. Each elementary row operation has the effect of multiplying the value of the determinant by some non-zero scalar. So, there exists some scalar $\alpha \in \mathbb{F} \setminus \{0\}$ such that $\det(A) = \alpha \det(\text{RREF}(A))$. Therefore, $\det(A) = 0$ if and only if $\det(\text{RREF}(A)) = 0$. Moreover, $\text{RREF}(A)$ is an upper triangular matrix, and so its determinant is zero if and only if at least one entry on its main diagonal is zero. We now have the following sequence of equivalences:

$$\begin{aligned}
\det(A) = 0 &\iff \det(\mathrm{RREF}(A)) = 0 \\
&\iff \text{at least one entry on the main diagonal of } \mathrm{RREF}(A) \text{ is } 0 \\
&\overset{(*)}{\iff} \mathrm{RREF}(A) \neq I_n \\
&\overset{(**)}{\iff} A \text{ is not invertible,}
\end{aligned}$$

where (*) follows from the fact that $\mathrm{RREF}(A)$ is a square matrix in reduced row echelon form, and (**) follows from Corollary 5.1 from Lecture Notes 4. It now obviously follows that A is invertible if and only if $\det(A) \neq 0$, which is what we needed to show. $\square$