2.2. The inverse of a matrix

For real numbers $a, b \in \mathbb{R}$ $(a \neq 0)$ the equation $ax = b$ is easy to solve:

$$ax = b \quad \Big| : a$$
$$x = \frac{b}{a}$$

In other words: $ax = b \overset{a^{-1}\cdot}{\Rightarrow} \underbrace{a^{-1} a}_{=1} x = a^{-1} b \Rightarrow x = a^{-1} b$

($a^{-1} = \frac{1}{a}$ is the multiplicative inverse of a in $\mathbb{R}$)

$\leadsto$ i.e. we solved the equation by multiplying with the multiplicative inverse of a, which is $a^{-1} = \frac{1}{a}$.

Q: Can we adapt this to matrix equations?
i.e., can we solve $Ax = b$ by multiplying with some matrix that takes the role of "an inverse" of A with respect to matrix multiplication?

$\hookrightarrow$ Given A, is there a matrix M such that:

$$Ax = b \overset{M\cdot}{\Rightarrow} \underbrace{MA}_{=I_n} x = Mb \Rightarrow x = Mb$$

$\leadsto$ $A^{-1} := M$ "inverse of A"

Spoiler: not all matrices A have an inverse A^{-1}.

Goal of this section: Find out which matrices A have an inverse A^{-1} and learn how to compute it.

Def. $A \in \mathbb{R}^{n \times n}$ is called invertible if $\exists\, C \in \mathbb{R}^{n \times n}$ such that $AC = CA = I_n$.

If A is invertible and hence such a C exists we will call C the inverse of A and denote it by $C = A^{-1}$.

Ex $A = \begin{pmatrix} 1 & 3 \\ 0 & 1/2 \end{pmatrix}$. Is A invertible?

$$\underbrace{\begin{pmatrix} 1 & 3 \\ 0 & 1/2 \end{pmatrix}}_{=A} \cdot \underbrace{\begin{pmatrix} 1 & -6 \\ 0 & 2 \end{pmatrix}}_{\exists C?} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{I_2}$$

Check: $C \cdot A = I_2$ also?

Ex. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ has no inverse

Q: Can there be several inverses of a given matrix?

Let's check... Assume B and C are both inverses of A.

$\Rightarrow BA = AB = I_n$ and $CA = AC = I_n$

So: $B = B \cdot I_n = B(AC) = \underbrace{(BA)}_{= I_n}C = I_n C = C$

$\Rightarrow$ So the inverse of a matrix (if it exists) is unique.

Ex (Invertibility of 2×2 matrices)

- $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}$ is invertible if and only if $ad - bc \neq 0$
- In case $ad - bc \neq 0$ the inverse of A is:

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$ (Proof: exercise)

Ex Is $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ invertible?

$ad - bc = 1\cdot 4 - 2\cdot 3 = 4 - 6 = -2 \neq 0$

$\Rightarrow A$ invertible ✓.

Thm. 2.5 (Properties of inverse matrices)

(a) If $A \in \mathbb{R}^{n\times n}$ is invertible then also A^{-1} is invertible with inverse $(A^{-1})^{-1} = A$.

(b) If both A and $B \in \mathbb{R}^{n\times n}$ are invertible then also $AB \in \mathbb{R}^{n\times n}$ is invertible with inverse $(AB)^{-1} = B^{-1}A^{-1}$

(c) If $A \in \mathbb{R}^{n\times n}$ is invertible then also A^T is invertible with inverse $(A^T)^{-1} = (A^{-1})^T$.

Proof (a) To check whether A^{-1} is invertible, we have to find out whether there exists a matrix C such that $A^{-1}C = CA^{-1} = I_n$. However, since A^{-1} is the inverse of A, $C = A$ fullfills this requirement. So A is the inverse of A^{-1} (i.e. $(A^{-1})^{-1} = A$) and A^{-1} is invertible.

b.) Have to check that $C = B^{-1}A^{-1}$ satisfies $(AB)C = C(AB) = I_n$:

$$(AB)C = (AB)B^{-1}A^{-1} = A\underbrace{(BB^{-1})}_{=I_n}A^{-1} = A\underbrace{I_n A^{-1}}_{=A^{-1}} = AA^{-1} = I_n$$

$$C(AB) = \ldots \text{ analogously } \ldots = I_n$$

c.) We have to check that $C = (A^{-1})^T$ satisfies $A^TC = CA^T = I_n$ similar as above.

$\square$

Two consequence from Thm 2.5:

For a product of k square matrices A_i:

$$(A_1 \cdot \ldots \cdot A_k)^{-1} = A_k^{-1} \cdot \ldots \cdot A_1^{-1}$$

(This follows from applying property (b) k-times)

Q: How can we compute A^{-1} in practice? (in particular when $n \geq 2$)

↳ Finding a solution C for the matrix-matrix equation $AX = I_n$ will provide a good candidate: $AC = I_n$.

But we will have to make sure that this candidate also satisfies $CA = I_n$. (See Thm. 2.7. (j) and (k).)

For this, we need more tools...

Elementary (row) operations in terms of matrix multiplication :

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = (A_1, \dots, A_n)$$

- Elementary operation 1: Multiply a row by a scalar

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \rightsquigarrow \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{k1} & \dots & \lambda a_{kn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \qquad (\lambda \neq 0)$$

Q: Is there an $M \in \mathbb{R}^{m \times m}$ so that $M \cdot A =$ ↰ ?

→ Yes!

$$M = \begin{array}{cc} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ \\ \leftarrow k \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} \uparrow \\ k \end{matrix} & \end{array}$$

(Verify by computing $M \cdot A$)

Q: Is M invertible ?

→ Yes

$$M^{-1} = \begin{array}{cc} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & \frac{1}{\lambda} & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ \\ \leftarrow k \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} \uparrow \\ k \end{matrix} & \end{array}$$

(Verify that $M \cdot M^{-1} = I_n$)

- **Elementary operation 2:** Exchanging two rows

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{mn} & \cdots & a_{mn} \end{pmatrix} \sim \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{mn} & \cdots & a_{mn} \end{pmatrix} \qquad (j>i)$$

Q: Is there an $M \in \mathbb{R}^{m \times m}$ so that $M \cdot A = \curvearrowleft$?

→ Yes!

$$M = \begin{array}{cl} \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & & & 1 & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & 1 & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow i \\ \\ \\ \leftarrow j \\ \\ \\ \end{array} \\ \begin{array}{ccccccccc} & & & \uparrow & & & \uparrow & & \\ & & & i & & & j & & \end{array} & \end{array} = \begin{pmatrix} e_1^T \\ \vdots \\ e_j^T \\ \vdots \\ e_i^T \\ \vdots \\ e_n^T \end{pmatrix}$$

Q: Is M invertible?

→ Yes: $M^{-1} = M$ (Verify that $M \cdot M = I_n$)

- **Elementary operation 3:** Adding a multiple of a row to another row

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{mn} & \cdots & a_{mn} \end{pmatrix} \underset{\sim}{\xrightarrow{+\lambda\cdot\ldots}} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{j1}+\lambda a_{i1} & \cdots & a_{jn}+\lambda a_{in} \\ \vdots & & \vdots \\ a_{mn} & \cdots & a_{mn} \end{pmatrix} \qquad (j>i)$$

Q: Is there an M ... ? → Yes! $M = \begin{array}{cc} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & \lambda & \ddots & \\ & & & 1 \end{pmatrix} & \leftarrow j \\ \quad\ \uparrow i & \end{array}$

Q: Is M invertible ? → Yes: $M^{-1} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & -\lambda & \ddots & \\ & & & 1 \end{pmatrix}$

Def We call matrices M that represent an elementary operation for row reduction an *elementary matrix*.

Rmk.

- We have seen that all elementary matrices are invertible and that their inverses are elementary as well.
- Notice that invertibility of elementary matrices corresponds to the fact that every elementary operation is reversible by another elementary operation

How to recognize elementary matrices :

- Type 1 : the identity matrix except for one entry on the diagonal replaced by a non-zero scalar.
- Type 2 : the identity matrix except for two rows interchanged
- Type 3 : the identity matrix except for one non-zero entry outside of the diagonal.

Ex.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Type 1 — Type 2 — Not elementary

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{10} & 0 & 1 \end{pmatrix}$$

Type 3 — Not elementary — Type 3

Thm 2.6 $A \in \mathbb{R}^{n \times n}$ is invertible if and only if A can be transformed into I_n by elementary operations.

Proof $\boxed{\Rightarrow}$ Assume that A is invertible $\Rightarrow A^{-1}$ exists.

Consider an equation $Ax = b$ for some $b \in \mathbb{R}^n$

$\Rightarrow A^{-1}Ax = A^{-1}b \Rightarrow x = A^{-1}b$ solution.

Hence, for every $b \in \mathbb{R}^n$, there exists a (unique) solution for $Ax = b$.

So, by Thm 1.4 the reduced echelon form $\tilde{A}$ of A has n pivot elements, that is,

$$\tilde{A} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & 1 \end{pmatrix} = I_n \qquad \text{(Because } \tilde{A} \text{ is square)}$$

$\Rightarrow$ A can be transformed into I_n by elementary operations.

$\boxed{\Leftarrow}$ Now assume that A can be transformed into I_n by elementary operations.

$\Rightarrow$ There exist elementary matrices $E_1, \ldots, E_k$ so that

$$E_1 \cdot \ldots \cdot E_k A = I_n \qquad \left(\begin{array}{l} \Rightarrow CA = I_n \text{ for } C = E_1 \cdot \ldots \cdot E_k \\ \text{but what about } AC = I_n \text{ ?} \end{array} \right)$$

the E_i are invertible

$\Rightarrow$ $A = E_k^{-1} \cdot \ldots \cdot E_1^{-1}$ and A is invertible (by Thm. 2.5 (b) + (c))

and $A^{-1} = E_1 \cdot \ldots \cdot E_k$. $\square$

This proof also tells us how to compute A^{-1} in concrete examples!

$$A^{-1} = E_1 \cdot \ldots \cdot E_k = \underbrace{E_1 \cdot \ldots \cdot E_k}_{\text{row operations that transform } A \text{ into } \tilde{A}} \cdot I_n$$

$\rightsquigarrow$ Find row operations that transform A into $\tilde{A}$ and apply them to I_n (in same order!)

($\rightsquigarrow$ which is the same as solving the equation $AX = I_n$)

Ex. $A = \begin{pmatrix} 0 & 0 & 1 \\ 2 & 5 & 1 \\ 1 & 3 & 1 \end{pmatrix}$ **Q:** Find A^{-1}.

$$(A \mid I_n) = {\scriptstyle -2}\left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 0 & 0 \\ 2 & 5 & 1 & 0 & 1 & 0 \\ 1 & 3 & 1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & -2 \\ 1 & 3 & 1 & 0 & 0 & 1 \end{array}\right) {\scriptstyle +3}$$

$$\sim \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & -2 \\ 1 & 0 & -2 & 0 & 3 & -5 \end{array}\right) {\scriptstyle +1} \; {\scriptstyle +2} \sim \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & -2 \\ 1 & 0 & 0 & 2 & 3 & -5 \end{array}\right) \cdot(-1)$$

$$\sim \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 2 \\ 1 & 0 & 0 & 2 & 3 & -5 \end{array}\right) \sim \left(\underbrace{\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}}_{= \tilde{A} = I_n} \middle| \underbrace{\begin{array}{ccc} 2 & 3 & -5 \\ -1 & -1 & 2 \\ 1 & 0 & 0 \end{array}}_{= A^{-1}} \right)$$

$$\Rightarrow \quad A^{-1} = \begin{pmatrix} 2 & 3 & -5 \\ -1 & -1 & 2 \\ 1 & 0 & 0 \end{pmatrix}$$

Rmk: Compare this to the section about solving $AX = B$. You will notice, it is the same procedure (for $B = I_n$).

In particular: solving the equation $Ax = e_j$ gives the j-th column of A^{-1}.

2.3 Characterization of invertible matrices

Thm 2.7 Let $A \in \mathbb{R}^{n\times n}$. Then the following statements are equivalent:

(a) A is invertible

(b) A can be transformed into I_n by elementary operations.

(c) The reduced echelon form of A has n pivot elements

(d) The equation $Ax = 0$ has $x = 0$ as its only solution

(e) The columns of A are linearly independent

(f) The linear fct. $f(x) = Ax$ is injective

(g) For every $b \in \mathbb{R}^n$, the equation $Ax = b$ has a solution.

(h) The columns of A span $\mathbb{R}^n$

(i) The linear fct. $f(x) = Ax$ is surjective

(j) $\exists\, C \in \mathbb{R}^{n\times n}$ s.t. $CA = I_n$

(k) $\exists\, C \in \mathbb{R}^{n\times n}$ s.t. $AC = I_n$

in both cases: $A^{-1} = C$

(l) A^T is invertible

(m) The rows of A are linearly independent

(n) The rows of A span $\mathbb{R}^n$

Rem: By the equivalence of (j) and (k), it suffices to prove only one of them in order to have both.

Proof.

(a) A is invertible

Thm. 2.6 $\Leftrightarrow$

(b) A can be transformed into I_n by elementary operations.

I_n has n pivot elements $\Leftrightarrow$

(c) The reduces echelon form of A has n pivot elements

Thms 1.15 + 1.13 $\Leftrightarrow$

(d) The equation $Ax = 0$ has $x = 0$ as its only solution

Thm 1.13 $\Leftrightarrow$

(e) The columns of A are linearly independent

Thm 1.13 $\Leftrightarrow$

(f) The linear fct. $f(x) = Ax$ is injective

Thm 1.15 with $n = m$ $\Leftrightarrow$

(g) For every $b \in \mathbb{R}^n$, the equation $Ax = b$ has a solution.

Thm. 1.4 $\Leftrightarrow$

(h) The columns of A span $\mathbb{R}^n$

Thm 1.13 $\Leftrightarrow$

(i) The linear fct. $f(x) = Ax$ is surjective

Thm 1.15 $\Leftrightarrow$

$\tilde{A}$ has n pivot elements $\Rightarrow$ $\tilde{A} = I_n$

Thm 2.6 $\Leftrightarrow$

(a) A is invertible

$\Leftrightarrow$

There exists $C \in \mathbb{R}^{n \times n}$ s.t. $AC = CA = I_n$

$\Leftrightarrow$

$(AC)^T = (CA)^T = I_n^T$

$\Leftrightarrow$

$C^T A^T = A^T C^T = I_n$

$\Leftrightarrow$

(k) A^T is invertible

Since [(a) A is invertible $\iff$ (b) A^T is invertible],

and since the columns of A are the rows of A^T:

(e) The columns of A are linearly independent

(m) The rows of A are linearly independent,

(h) The columns of A span $\mathbb{R}^n$

(n) The rows of A span $\mathbb{R}^n$

Lastly, check that each of (j) and (k) follows from one of the other statements, but also, induces one of the other statements:

(a) A is invertible

(j) $\exists\, C \in \mathbb{R}^{n \times n}$ s.t. $CA = I_n$

(k) $\exists\, C \in \mathbb{R}^{n \times n}$ s.t. $AC = I_n$

But also:

(j) $\exists\, C \in \mathbb{R}^{n \times n}$ s.t. $CA = I_n$

(g) For every $b \in \mathbb{R}^n$, the equation $Ax = b$ has a solution.

Because $Ax = b \overset{j}{\Rightarrow} \underbrace{CA}_{=I_n} x = Cb \Rightarrow x = Cb$ solution

(k) $\exists\, C \in \mathbb{R}^{n \times n}$ s.t. $AC = I_n$

$(AC)^T = I_n^T \Rightarrow C^T A^T = I_n \iff$ (j) for $A^T \Rightarrow$ (g) for A^T

$=$ (a) for A^T

(b) A^T is invertible

$\square$

Invertible linear transformations

(vocabulary explanation :
linear transformations = linear map = linear mapping = linear fct.)

Let $A \in \mathbb{R}^{n\times n}$, $f(x) := Ax$ $\quad (f: \mathbb{R}^n \to \mathbb{R}^n)$

From the Thm. 2.7 we can deduce

$$A \text{ invertible} \iff f \text{ bijective}$$

Formally : $\boxed{\Rightarrow}$ A invertible $\overset{\text{Thm 2.7 (E)+(I)}}{\Longrightarrow} \begin{cases} f(x) = Ax \text{ injective} \\ f(x) = Ax \text{ surjective} \end{cases}$

$\Rightarrow f$ bijective

$\boxed{\Leftarrow}$ f bijective $\Rightarrow f$ injective $\overset{\text{Thm 2.7 (f)}}{\Longrightarrow} A$ invertible

So assume f bijective and since A^{-1} exist.

let us view A^{-1} as a linear map as well :

$h(x) := A^{-1}x$

Q : What is the relationship between the linear maps h and f ?

Claim : $h = f^{-1}$

Proof : To show: $f(x) = y \iff h(y) = x$

$\boxed{\Rightarrow}$ $f(x) = y \Rightarrow Ax = y \Rightarrow \underbrace{A^{-1}A}_{=I_n} x = A^{-1}y \Rightarrow x = A^{-1}y = h(y)$

$\boxed{\Leftarrow}$ Symmetrically : $h(y) = x \Rightarrow f(x) = y$

$\square$

except for uniqueness

We have just proven the following Thm. :

Thm. 2.8 Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map and denote its matrix by A. Then, f is bijective if and only if A is invertible. In that case the linear map $f^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ given by the inverse matrix A^{-1} (i.e. $f^{-1}(y) := A^{-1}y$) satisfies :

$$f \circ f^{-1} = f^{-1} \circ f = id_{\mathbb{R}^n}$$

and this map f^{-1} is the unique map with this property.

We call f^{-1} the inverse map of f.

Proof of uniqueness :

By def. : $f \circ f^{-1} = f^{-1} \circ f = id_{\mathbb{R}^n}$

Assume : $\exists\ h : \mathbb{R}^n \to \mathbb{R}^n$ for which $f \circ h = h \circ f = id_{\mathbb{R}^n}$

Want to show: $f^{-1} = h$, i.e. , $f^{-1}(x) = h(x)\ \ \forall x \in \mathbb{R}^n$

Let $x \in \mathbb{R}^n$: $h(x) = h(\underbrace{f \circ f^{-1}}_{= id_{\mathbb{R}^n}}(x)) = \underbrace{(h \circ f)}_{= id_{\mathbb{R}^n}} \circ f^{-1}(x) = f^{-1}(x)$. $\square$

Q: What about matrices that are not square?

→ Let us try an analog of the definition of invertability:

Let $A \in \mathbb{R}^{m\times n}$. Assume that there exist:

(i) $C_1 \in \mathbb{R}^{n\times m}$ so that $A C_1 = I_m$

(ii) $C_2 \in \mathbb{R}^{n\times m}$ so that $C_2 A = I_n$

Then:

- Assume $Ax = 0 \overset{(ii)}{\Rightarrow} x = I_n x = C_2 A x = C_2 \cdot 0 = 0$

  $\Rightarrow$ $x = 0$ only solution for $Ax = 0 \Rightarrow n \leq m$

- Let $b \in \mathbb{R}^m \overset{(i)}{\Rightarrow} b = I_n b = A C_1 b = A \underbrace{(C_1 b)}_{=x}$

  $\Rightarrow$ For every $b \in \mathbb{R}^n$ there is a solution for $Ax = b \Rightarrow m \leq n$

$\Rightarrow$ (i) and (ii) can only both be satisfied if $m = n$, that is, if A is square.

$\Rightarrow$ So the concept of invertibility (defined by (i) and (ii)) only makes sense for square matrices.

(In some algebra books, you will see the notions of left-invertible matrices (matrices in $\mathbb{R}^{m\times n}$ satisfying (ii)) and right-invertible matrices (matrices in $\mathbb{R}^{m\times n}$ satisfying (i)). In this context we could say:
A matrix that is both left-invertible and right-invertible must be an (invertible) square matrix.)

3. Determinants

We already know:

For a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$: A invertible $\iff ad - bc \neq 0$

We call $\det(A) := ad - bc$ the determinant of A.

Goal: Define a notation of determinant for $A \in \mathbb{R}^{n\times n}$, $n \geq 2$.

3.1 Definition of the determinant

Def. Let $A \in \mathbb{R}^{n\times n}$ for each pair (i,j), define A_{ij} to be the $(n-1)\times(n-1)$ matrix obtained from A by deleting the i-th row and the j-th column.

The determinant of A is defined recursively as follows:

(i) if $n=1$: $A = (a_{11})$ and $\det A := a_{11}$

(ii) if $n>1$:

$$\det A = a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + \dots + (-1)^{n+1} a_{1n} \det(A_{1n})$$

Ex. • $n=2$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ Q: $\det A = ?$

$$A_{11} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (d) \in \mathbb{R}^{1\times 1}, \quad a_{11} = a$$

$$A_{12} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (c) \in \mathbb{R}^{1\times 1}, \quad a_{12} = b$$

$$\det A = a_{11} \cdot \underbrace{\det(d)}_{\substack{= d \\ (i)}} - a_{12} \underbrace{\det(c)}_{\substack{= c \\ (i)}} = a\cdot d - b\cdot c$$

- $n=3$, $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ $\qquad$ Q: $\det A = ?$

$$\begin{pmatrix} a_{11} & \cancel{a_{12}} & \cancel{a_{13}} \\ \cancel{a_{21}} & a_{22} & a_{23} \\ \cancel{a_{31}} & a_{32} & a_{33} \end{pmatrix} = A_{11} \qquad \begin{pmatrix} \cancel{a_{11}} & a_{12} & \cancel{a_{13}} \\ a_{21} & \cancel{a_{22}} & a_{23} \\ a_{31} & \cancel{a_{32}} & a_{33} \end{pmatrix} = A_{12} \qquad \begin{pmatrix} \cancel{a_{11}} & \cancel{a_{12}} & a_{13} \\ a_{21} & a_{22} & \cancel{a_{23}} \\ a_{31} & a_{32} & \cancel{a_{33}} \end{pmatrix} = A_{13}$$

$$\det A = a_{11} \underbrace{\det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}}_{= a_{22}a_{33} - a_{32}a_{23}} - a_{12} \underbrace{\det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}}_{= a_{21}a_{33} - a_{23}a_{31}} + a_{13} \underbrace{\det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}}_{= a_{21}a_{32} - a_{22}a_{31}}$$

determinant formula for $n=2$

We notice: computing the determinant of an $n\times n$-matrix by definition is not very efficient

$\rightsquigarrow$ in general requires $n!$ multiplication (e.g. $25! \approx 1.5 \cdot 10^{25}$ ☹)

$\rightsquigarrow$ We want a more efficient method! Moreover, we want to understand the connection between the determinant and the invertability of $n\times n$ matrices.

3.2. Properties of the determinant

Thm. 3.1 Let $A \in \mathbb{R}^{n\times n}$, $A = (A_1, \dots, A_n)$. Then:

(i) $\det(I_n) = 1$

(ii) For $k \in \{1, \dots, n\}$, define the function $f: \mathbb{R}^n \to \mathbb{R}$ by
$$f(x) = \det(A_1, \dots, A_{(k-1)}, x, A_{(k+1)}, \dots, A_n)$$
This function is linear.

Replace k-th column of A by the vector x, then take its determinant.

(iii) Let $\tilde{A}$ be the matrix obtained from A by interchanging two columns. Then $\det A = -\det \tilde{A}$.

Proof

(i) $I_n = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}$

i.e. $I_n = A$ where $a_{ii} = 1$ and $a_{ij} = 0$ for $i \neq j$

$$\det I_n = \underbrace{a_{11}}_{=1} \det(A_{11}) - \underbrace{a_{12}}_{=0} \det(A_{12}) + \ldots + (-1)^{n+1} \underbrace{a_{1n}}_{=0} \det(A_{1n}) = \det(A_{11})$$

$$\Rightarrow \det I_n = \det(A_{11})$$

$$A_{11} = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \\ a_{n2} & \cdots & a_{nn} \end{pmatrix} \Rightarrow A_{11} = I_{n-1} \text{ the } (n-1)\times(n-1) \text{ identity matrix}$$

$$\Rightarrow \det I_n = \det I_{n-1} \quad \forall n \in \mathbb{N} \ (n \geq 2)$$

$$\Rightarrow \det I_n = \det I_{n-1} = \ldots = \det I_2 = \det I_1 = 1$$

$$I_1 = [1] \Rightarrow \det I_1 = 1$$

(This was an example of an inductive proof! Meaning:

- We related the desired statement for n to the analogous statement for $n-1$.
- Then we treated case $n=1$.
- In consequence, we obtained the desired statement for all n.)

(ii)

$$x \in \mathbb{R}^n, \quad f(x) = \det \underbrace{\begin{array}{c} \\ \end{array}\!\!\begin{array}{c} \begin{array}{ccccc} & & & k\text{-th column} & \\ & & & \downarrow & \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & x_1 & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & x_n & \cdots & a_{nn} \end{pmatrix} \end{array}}_{=: A(x)}$$

We have to verify the <u>two conditions</u> from the definition of linearity.

or equivalently we can verify the combined condition:

$$f(x + \lambda y) = f(x) + \lambda f(y).$$

We again do this by an inductive proof.

$$f(x) = \det A(x) = a_{11} \det (A(x))_{11} - a_{12} \det (A(x))_{12} \ldots (-1)^{k+1} x_1 \det (A(x))_{1k}$$

$$\ldots (-1)^{n+1} a_{1n} \det (A(x))_{1n}$$

What do the matrices $(A(x))_{1j}$ look like?

- they are all $(n-1) \times (n-1)$ matrices

- $$(A(x))_{1k} = \begin{pmatrix} a_{11} & a_{12} & \cdots & x_1 & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & x_n & \cdots & a_{nn} \end{pmatrix} = A_{1k}$$

- $$\underset{\substack{\uparrow \\ j \neq k}}{(A(x))_{1j}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & x_1 & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & x_n & \cdots & a_{nn} \end{pmatrix} = A_{1j}(\tilde{x})$$

  (e.g. $j = 2$)

  with $\tilde{x}$ in the $(k-1)$-st (or k-th) column where

  $$\tilde{x} = \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n-1}$$

$$\Rightarrow \det A(x) = a_{11} \det A_{11}(\tilde{x}) - a_{12} \det A_{12}(\tilde{x}) \qquad (*)$$

$$\ldots (-1)^{k+1} x_1 \det A_{1k} \ldots (-1)^{n+1} a_{1n} \det A_{1n}(\tilde{x})$$

We want: $\overbrace{\det A(x+\lambda y)}^{f(x+\lambda y)} = \overbrace{\det A(x)}^{=f(x)} + \overbrace{\lambda \det A(y)}^{\lambda f(y)}$ $\left[\Leftrightarrow f(x+\lambda y) = f(x) + \lambda f(y)\right]$

Observe that $\overbrace{f(x) \text{ consists}}^{\text{from } (*)}$ of $n-1$ additive terms of the form

$\mu \cdot B(\tilde{x})$ (where λ is a scalar and $B \in \mathbb{R}^{(n-1)\times(n-1)}$)

and one term of the form $x_1 \cdot \tilde{\mu}$ (where $\tilde{\mu}$ is a scalar).

- $h(x_1) := x_1 \cdot \tilde{\mu}$ is linear (easy to check)
- $g(\tilde{x}) = B(\tilde{x})$ is the same object as $f(x) = A(x)$ but for size $(n-1)$ instead instead of n.

  So by inductive assumption, we may assume g is linear.

If h and g are linear, then f is linear. This concludes the inductive step $((n-1) \rightsquigarrow n)$.

So to finish this induction proof, we only need to check case $n=1$:

For $n=1$: $A(x) = (x)$, $x \in \mathbb{R}$

$$\det A(x+\lambda y) = \det \underbrace{(x+\lambda y)}_{1\times 1-\text{matrix}} = x + \lambda y = \det(x) + \lambda \det(y)$$

(iii) Similar to (ii) but easier. (Omitted)

$\square$

Lemma 3.2 Let $A = (A_1, \dots, A_n) \in \mathbb{R}^{n \times n}$

(a) If $A_i = A_j$ for some $i \neq j$ then $\det A = 0$

(b) If A_j is a linear combination of other columns then $\det A = 0$

(c) If A is not invertible $\Rightarrow \det A = 0$

Proof (a) If $A_i = A_j$ then swapping the i-th and j-th column of A does not change A.

$$\underset{\text{Thm 3.2.(iii)}}{\Rightarrow} \quad \det A = -\det A \Rightarrow \det A = 0.$$

(b) $A = (A_1 \dots \boxed{A_k} \dots, A_n)$, $A_k = \underbrace{\lambda_1 A_1 + \dots + \lambda_n A_n}_{= \text{ linear combo of all other } A_i}$

k-th column

$\Rightarrow A = A(x)$

where $x = \lambda_1 A_1 + \dots + \lambda_n A_n$ in k-th column.

Notation from Thm. 3.1

Notation from Thm 3.1.

$$\Rightarrow \det A = \det A(x) = f(x) = f(\lambda_1 A_1 + \dots + \lambda_n A_n)$$

$$= \lambda_1 f(A_1) + \dots + \lambda_n f(A_n) \quad \text{!}$$

Thm 3.1: f is linear

$$= \lambda_1 \det(\underbrace{A(A_1)}) + \dots + \lambda_n \det(\underbrace{A(A_n)})$$

What are these matrices?

$A(A_i) = A(x)$, $x = A_i$ meaning : A with k-th column replaced by A_i

e.g. $A(A_1) = (A_1, A_2 \dots A_1 \dots A_n)$

$A(A_2) = (A_1, A_2 \dots A_2 \dots A_n)$

$\vdots$

$A(A_n) = (A_1, A_2 \dots A_n \dots A_n)$

All of these matrices have a double column (since $A(A_k)$ is not in the list)

$\Rightarrow \det A(A_i) = 0$ by lemma 3.2.(a)

$\Rightarrow \det A = \lambda_1 \cdot 0 + \dots + \lambda_n \cdot 0 = 0$

(c) If A is not invertible $\Rightarrow$ columns are linearly dependent

$\Rightarrow \exists$ column that can be expressed as a linear combination of the other columns

$\underset{(b)}{\Rightarrow} \det A = 0$ $\square$