Mathematics 1 - Linear Algebra

Lecture 07 – §3.6 Computing with matrices

Sabine Le Borne

Preview

Matrices - what we already know:

- operation $\mathbf{A} \cdot \mathbf{v}$
- LES $\mathbf{A}\mathbf{x} = \mathbf{b}$, existence and uniqueness of solutions, solution by Gauss elimination
- $\mathrm{Ran}(\mathbf{A})$, $\mathrm{Ker}(\mathbf{A})$, $\mathrm{rank}(\mathbf{A})$

Matrices - what we learn today:

- dimension formula for $\mathrm{Ran}(\mathbf{A})$, $\mathrm{Ker}(\mathbf{A})$, $\mathrm{rank}(\mathbf{A})$
- operations $\alpha \cdot \mathbf{A}$, $\mathbf{A} + \mathbf{B}$, $\mathbf{A} \cdot \mathbf{B}$, $\mathbf{A}^k$ (for matrices of compatible sizes)
- inverse mapping $\mathbf{A}^{-1}$ for bijective $\mathbf{A}$

Definition & Theorem 3.53 (row and column space have the same dimension)

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. We define

row space of $\mathbf{A}$, $\mathrm{RS}(\mathbf{A}) \subset \mathbb{R}^n$	:	the subspace spanned by the rows of $\mathbf{A}$
column space of $\mathbf{A}$, $\mathrm{CS}(\mathbf{A}) \subset \mathbb{R}^m$	:	the subspace spanned by the columns of $\mathbf{A}$ $(= \mathrm{Ran}(\mathbf{A}))$

One can show (however, this is difficult and hence omitted here) that

$$\dim(\mathrm{RS}(\mathbf{A})) \;=\; \dim(\mathrm{CS}(\mathbf{A})),$$

i.e., $\mathbf{A}$ has the same number of linearly independent rows as columns.

Reminder. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and let $\mathbf{A} \rightsquigarrow \mathbf{A}'$ be a transition of $\mathbf{A}$ to row echelon form $\mathbf{A}'$.

$$\begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \left(\begin{array}{ccccc|c} \boxed{1} & 2 & 0 & 1 & 0 & 3 \\ 0 & 0 & \boxed{2} & -1 & 4 & 2 \\ 0 & 0 & 0 & \boxed{4} & 8 & 8 \end{array}\right) \end{array} \quad \rightsquigarrow \quad \mathcal{L} = \left\{ \begin{pmatrix} 1\\0\\2\\2\\0 \end{pmatrix} + x_2 \begin{pmatrix} -2\\1\\0\\0\\0 \end{pmatrix} + x_5 \begin{pmatrix} 2\\0\\-3\\-2\\1 \end{pmatrix} : x_2, x_5 \in \mathbb{R} \right\}$$

The number of pivot elements (dependent variables) in $\mathbf{A}'$ is $\mathrm{rank}(\mathbf{A})$.
There are $k = n - \mathrm{rank}(\mathbf{A})$ free variables, the kernel is spanned by k vectors (Thm. 3.26).

Dimension formula

Theorem 3.54 (dimension of range and kernel, dimension formula)

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and let $k \in \mathbb{N}_0$ such that $\operatorname{rank}(\mathbf{A}) + k = n$. Then there holds

a) $\operatorname{rank}(\mathbf{A}) \overset{a_1)}{=} \dim(\mathrm{RS}(\mathbf{A})) \overset{a_2)}{=} \dim(\mathrm{CS}(\mathbf{A})) \overset{a_3)}{=} \dim(\mathrm{Ran}(\mathbf{A}))$,

b) $k = \dim(\mathrm{Ker}(\mathbf{A}))$.

We obtain the so-called *dimension formula*:

$$\dim(\mathrm{Ran}(\mathbf{A})) + \dim(\mathrm{Ker}(\mathbf{A})) \;=\; \dim(\mathbb{R}^n).$$

Proof. a_1) It is easily shown that admissible elimination steps do not change the row space. For an elimination $\mathbf{A} \rightsquigarrow \mathbf{A}'$ to row echelon form, there holds $\mathrm{RS}(\mathbf{A}) = \mathrm{RS}(\mathbf{A}')$ and hence also

$$\begin{aligned} \dim(\mathrm{RS}(\mathbf{A})) = \dim(\mathrm{RS}(\mathbf{A}')) &= \text{number of non-zero rows in } \mathbf{A}' \\ &= \text{number of pivot elements of } \mathbf{A}' = \operatorname{rank}(\mathbf{A}). \end{aligned}$$

a_2) see Theorem 3.53.

a_3) $\mathrm{CS}(\mathbf{A}) = \mathrm{Ran}(\mathbf{A})$ by definition.

Dimension formula

Theorem 3.54 (dimension of range and kernel, dimension formula)

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and let $k \in \mathbb{N}_0$ such that $\mathrm{rank}(\mathbf{A}) + k = n$. Then there holds

a) $\mathrm{rank}(\mathbf{A}) \overset{a_1)}{=} \dim(\mathrm{RS}(\mathbf{A})) \overset{a_2)}{=} \dim(\mathrm{CS}(\mathbf{A})) \overset{a_3)}{=} \dim(\mathrm{Ran}(\mathbf{A}))$,

b) $k = \dim(\mathrm{Ker}(\mathbf{A}))$.

We obtain the so-called *dimension formula*:

$$\dim(\mathrm{Ran}(\mathbf{A})) + \dim(\mathrm{Ker}(\mathbf{A})) \;=\; \dim(\mathbb{R}^n).$$

Proof (continued). b) Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be the vectors which result from the Gauß algorithm and span the kernel of $\mathbf{A}$ (see Theorem 3.26c)).

A closer look at the Gauß algorithm reveals that the vector in $\mathcal{K} := (\mathbf{v}_1, \ldots, \mathbf{v}_k)$ that is multiplied with the free variable x_j has a 1 in its j-th position whereas all other vectors in $\mathcal{K}$ have a 0 in this position.

This structure implies that the family $\mathcal{K}$ is linearly independent. Hence $\mathcal{K}$ is a basis of $\mathrm{Ker}(\mathbf{A})$ and there holds $k = |\mathcal{K}| = \dim(\mathrm{Ker}(\mathbf{A}))$.

Using a) and b) in $\mathrm{rank}(\mathbf{A}) + k = n$ leads to the dimension formula. $\square$

Example for the application of the dimension formula

Example

Determine the dimension of

$$V := \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : x_1 + 3x_2 - 5x_3 = 0 \right\}.$$

We rewrite V as follows (which also shows that V is indeed a subspace):

$$V = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : \underbrace{\begin{pmatrix} 1 & 3 & -5 \end{pmatrix}}_{=:\mathbf{A}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \right\} = \left\{ \mathbf{x} \in \mathbb{R}^3 : \mathbf{A}\mathbf{x} = 0 \right\} = \mathrm{Ker}(\mathbf{A}).$$

$\mathbf{A}$ is already in row echelon form and we read off that $\mathrm{rank}(\mathbf{A}) = 1$. It follows that

$$\dim(V) = \dim(\mathrm{Ker}(\mathbf{A})) = 3 - \mathrm{rank}(\mathbf{A}) = 3 - 1 = 2.$$

Geometric view: V describes a plane in $\mathbb{R}^3$, hence $\dim(V) = 2$.

Computing with matrices

Matrix: introduced for the compact notation of linear systems of equations
Matrix-vector multiplication: Mapping $f_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m,\ \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$

Questions: (How) can matrices be added or multiplied, and
what does this imply for the associated mappings?

Definition 3.57 (Matrix + Matrix = Matrix)

Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$. Then we define the matrix $\mathbf{A}+\mathbf{B} \in \mathbb{R}^{m \times n}$ via

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{\mathbf{A}} + \underbrace{\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}}_{\mathbf{B}} := \underbrace{\begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}}_{\mathbf{A}+\mathbf{B}}.$$

There holds for all $\mathbf{x} \in \mathbb{R}^n$

$$f_{\mathbf{A}+\mathbf{B}}(\mathbf{x}) = f_{\mathbf{A}}(\mathbf{x}) + f_{\mathbf{B}}(\mathbf{x}), \qquad \text{hence} \qquad (\mathbf{A}+\mathbf{B})\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}.$$

Computing with matrices

Definition 3.58 (number · matrix = matrix)

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\alpha \in \mathbb{R}$. Then we define the matrix $\alpha\cdot\mathbf{A} \in \mathbb{R}^{m\times n}$ via

$$\alpha\cdot \underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}}_{\mathbf{A}} := \underbrace{\begin{pmatrix} \alpha a_{11} & \cdots & \alpha a_{1n} \\ \vdots & & \vdots \\ \alpha a_{m1} & \cdots & \alpha a_{mn} \end{pmatrix}}_{\alpha\cdot\mathbf{A}}.$$

There holds for all $\mathbf{x} \in \mathbb{R}^n$

$$f_{\alpha\cdot\mathbf{A}}(\mathbf{x}) = \alpha f_{\mathbf{A}}(\mathbf{x}), \qquad \text{hence} \qquad (\alpha\cdot\mathbf{A})\mathbf{x} = \alpha\mathbf{A}\mathbf{x}.$$

Examples

$$2\cdot\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 2+1 & 4+0 \\ 6+2 & 8-1 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 8 & 7 \end{pmatrix},$$

$$2\cdot\left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) + \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 2\cdot\begin{pmatrix} 5 \\ 11 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 11 \\ 22 \end{pmatrix}, \quad \begin{pmatrix} 3 & 4 \\ 8 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 11 \\ 22 \end{pmatrix}.$$

Computing with matrices

Theorem (Rules for computing with $+$ und $\alpha\cdot$ for matrices)

For $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{m\times n}$ and $\alpha, \beta \in \mathbb{R}$ there hold:

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$, — commutative law
2. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$, — associative law
3. $\mathbf{A} + \mathbf{0} = \mathbf{A}$, — neutral element $\mathbf{0}$
4. $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$, — inverse element $-\mathbf{A}$
5. $(\alpha \cdot \beta) \cdot \mathbf{A} = \alpha \cdot (\beta \cdot \mathbf{A})$,
6. $1 \cdot \mathbf{A} = \mathbf{A}$,
7. $(\alpha + \beta) \cdot \mathbf{A} = \alpha \cdot \mathbf{A} + \beta \cdot \mathbf{A}$,
8. $\alpha \cdot (\mathbf{A} + \mathbf{B}) = \alpha \cdot \mathbf{A} + \alpha \cdot \mathbf{B}$.

Here, $\mathbf{0}$ denotes the zero matrix in $\mathbb{R}^{m\times n}$ and $-\mathbf{A}$ denotes the matrix $(-1) \cdot \mathbf{A}$.

Proof. All propositions above follow from elementwise considerations and applying respective properties in $\mathbb{R}$. $\square$

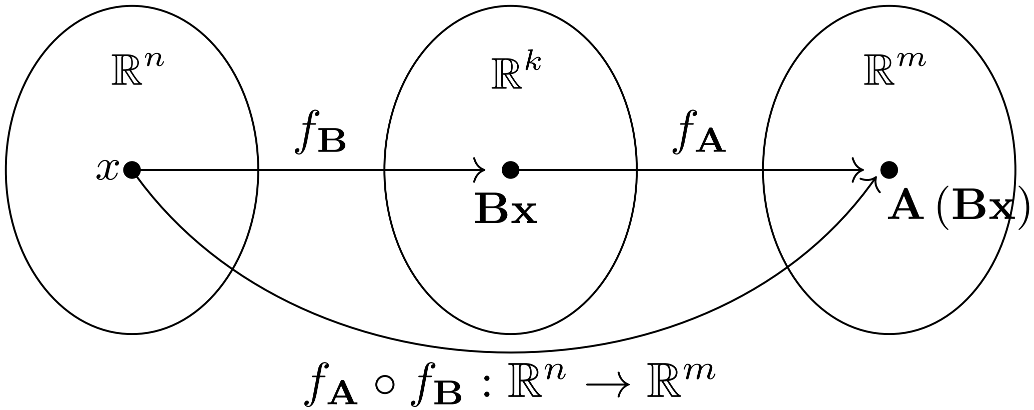
Computing with matrices

Matrix product: Let $\mathbf{A} \in \mathbb{R}^{m\times k}$ and $\mathbf{B} \in \mathbb{R}^{k\times n}$ with corresponding mappings

$$f_{\mathbf{A}} : \mathbb{R}^k \to \mathbb{R}^m,\ \mathbf{x} \mapsto \mathbf{A}\mathbf{x}, \qquad f_{\mathbf{B}} : \mathbb{R}^n \to \mathbb{R}^k,\ \mathbf{x} \mapsto \mathbf{B}\mathbf{x}.$$

Then we can concatenate $f_{\mathbf{B}}$ and $f_{\mathbf{A}}$:

$$f_{\mathbf{A}} \circ f_{\mathbf{B}} : \mathbb{R}^n \to \mathbb{R}^m,\ \mathbf{x} \mapsto f_{\mathbf{A}}\left(f_{\mathbf{B}}(\mathbf{x})\right) = f_{\mathbf{A}}\left(\mathbf{B}\mathbf{x}\right) = \mathbf{A}\left(\mathbf{B}\mathbf{x}\right).$$



Question: Does there exist a matrix $\mathbf{M} \in \mathbb{R}^{m\times n}$ such that

$$f_{\mathbf{A}} \circ f_{\mathbf{B}} = f_{\mathbf{M}},\ \text{i.e., } \mathbf{A}\left(\mathbf{B}\mathbf{x}\right) = \mathbf{M}\mathbf{x} \text{ for all } \mathbf{x} \in \mathbb{R}^n?$$

Computing with matrices

Let $\mathbf{A} \in \mathbb{R}^{m\times k}$, $\mathbf{B} \in \mathbb{R}^{k\times n}$ with corresponding mappings

$$f_{\mathbf{A}} : \mathbb{R}^k \to \mathbb{R}^m,\ \mathbf{x} \mapsto \mathbf{A}\mathbf{x},$$
$$f_{\mathbf{B}} : \mathbb{R}^n \to \mathbb{R}^k,\ \mathbf{x} \mapsto \mathbf{B}\mathbf{x}.$$

Question: Does there exist a matrix $\mathbf{M} \in \mathbb{R}^{m\times n}$ such that

$$f_{\mathbf{A}} \circ f_{\mathbf{B}} = f_{\mathbf{M}},\ \text{i.e., } \mathbf{A}(\mathbf{B}\mathbf{x}) = \mathbf{M}\mathbf{x} \text{ for all } \mathbf{x} \in \mathbb{R}^n?$$

Answer: Yes!

$$\begin{aligned}
\mathbf{A}(\mathbf{B}\mathbf{x}) &= \mathbf{A}\left(\begin{pmatrix} | & & | \\ \mathbf{b}_1 & \cdots & \mathbf{b}_n \\ | & & | \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \mathbf{A}\left(x_1 \begin{pmatrix} | \\ \mathbf{b}_1 \\ | \end{pmatrix} + \cdots + x_n \begin{pmatrix} | \\ \mathbf{b}_n \\ | \end{pmatrix}\right) \\
&= x_1\mathbf{A}\begin{pmatrix} | \\ \mathbf{b}_1 \\ | \end{pmatrix} + \cdots + x_n\mathbf{A}\begin{pmatrix} | \\ \mathbf{b}_n \\ | \end{pmatrix} = x_1\begin{pmatrix} | \\ \mathbf{A}\mathbf{b}_1 \\ | \end{pmatrix} + \cdots + x_n\begin{pmatrix} | \\ \mathbf{A}\mathbf{b}_n \\ | \end{pmatrix} \\
&= \begin{pmatrix} | & & | \\ \mathbf{A}\mathbf{b}_1 & \cdots & \mathbf{A}\mathbf{b}_n \\ | & & | \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} =: \mathbf{M}\mathbf{x}.
\end{aligned}$$

Computing with matrices

Matrix product: three methods/viewpoints

Let $\mathbf{A} \in \mathbb{R}^{m\times k}$ and $\mathbf{B} \in \mathbb{R}^{k\times n}$. Then the product $\mathbf{AB} \in \mathbb{R}^{m\times n}$ is given by

$$\begin{aligned}
\mathbf{AB} = \mathbf{A}\begin{pmatrix} | & & | \\ \mathbf{b}_1 & \cdots & \mathbf{b}_n \\ | & & | \end{pmatrix} &:= \begin{pmatrix} | & & | \\ \mathbf{A}\mathbf{b}_1 & \cdots & \mathbf{A}\mathbf{b}_n \\ | & & | \end{pmatrix} \\
&= \begin{pmatrix} \text{—}\mathbf{a}_1\text{—} \\ \vdots \\ \text{—}\mathbf{a}_m\text{—} \end{pmatrix} \begin{pmatrix} | & & | \\ \mathbf{b}_1 & \cdots & \mathbf{b}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1\mathbf{b}_1 & \cdots & \mathbf{a}_1\mathbf{b}_n \\ \vdots & & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \cdots & \mathbf{a}_m\mathbf{b}_n \end{pmatrix} \\
&= \begin{pmatrix} \text{—}\mathbf{a}_1\text{—} \\ \vdots \\ \text{—}\mathbf{a}_m\text{—} \end{pmatrix} \mathbf{B} = \begin{pmatrix} \text{—}\mathbf{a}_1\mathbf{B}\text{—} \\ \vdots \\ \text{—}\mathbf{a}_m\mathbf{B}\text{—} \end{pmatrix}.
\end{aligned}$$

- Each column of $\mathbf{AB}$ is a linear combination of the columns of $\mathbf{A}$.
- Each entry $(\mathbf{AB})_{ij}$ is given by $\mathbf{a}_i\mathbf{b}_j = a_{i1}b_{1j} + \ldots + a_{ik}b_{kj} = \sum_{\ell=1}^{k} a_{i\ell}b_{\ell j}$.
- Each row of $\mathbf{AB}$ is a linear combination of the rows of $\mathbf{B}$.

Computing with matrices

Let $\mathbf{A} \in \mathbb{R}^{m\times k}$ and $\mathbf{B} \in \mathbb{R}^{k\times n}$. Then the product $\mathbf{AB} \in \mathbb{R}^{m\times n}$ is given by

$$\mathbf{AB} = \begin{pmatrix} | & & | \\ \mathbf{A}\mathbf{b}_1 & \cdots & \mathbf{A}\mathbf{b}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1\mathbf{b}_1 & \cdots & \mathbf{a}_1\mathbf{b}_n \\ \vdots & & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \cdots & \mathbf{a}_m\mathbf{b}_n \end{pmatrix} = \begin{pmatrix} —\mathbf{a}_1\mathbf{B}— \\ \vdots \\ —\mathbf{a}_m\mathbf{B}— \end{pmatrix}.$$

Example for a matrix product

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{\begin{pmatrix} —\mathbf{a}_1— \\ —\mathbf{a}_2— \end{pmatrix}} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}}_{\begin{pmatrix} | & | & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \\ | & | & | \end{pmatrix}} = \left(\underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} 1 \\ 4 \end{pmatrix}}_{\mathbf{b}_1} \quad \underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} 2 \\ 5 \end{pmatrix}}_{\mathbf{b}_2} \quad \underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} 3 \\ 6 \end{pmatrix}}_{\mathbf{b}_3} \right)$$

$$= \underbrace{\begin{pmatrix} \underbrace{\begin{matrix} 1\cdot 1 + 2\cdot 4 \\ 3\cdot 1 + 4\cdot 4 \end{matrix}}_{\mathbf{A}\mathbf{b}_1} & \underbrace{\begin{matrix} 1\cdot 2 + 2\cdot 5 \\ 3\cdot 2 + 4\cdot 5 \end{matrix}}_{\mathbf{A}\mathbf{b}_2} & \underbrace{\begin{matrix} 1\cdot 3 + 2\cdot 6 \\ 3\cdot 3 + 4\cdot 6 \end{matrix}}_{\mathbf{A}\mathbf{b}_3} \end{pmatrix}}_{\begin{pmatrix} \mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \mathbf{a}_1\mathbf{b}_3 \\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \mathbf{a}_2\mathbf{b}_3 \end{pmatrix}} = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{pmatrix}$$

Computing with matrices

Examples for matrix products

$$\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}, \qquad \begin{pmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{pmatrix},$$

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix} = \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}, \qquad \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{pmatrix} = \text{undefined,}$$

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{pmatrix} = \begin{pmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{pmatrix}, \qquad \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix} = \begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}.$$

For all $\mathbf{A} \in \mathbb{R}^{m\times n}$ and the identity matrix $\mathbf{I}_k = \begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & 1\end{pmatrix} \in \mathbb{R}^{k\times k}$ there holds:

$$\mathbf{A}\cdot\mathbf{I}_n = \mathbf{A} \quad \text{and} \quad \mathbf{I}_m\cdot\mathbf{A} = \mathbf{A}.$$

Computing with matrices

$$\mathbf{A}\cdot\mathbf{B} := \begin{pmatrix}1 & 2\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & 2\\ 0 & 5\end{pmatrix} = \begin{pmatrix}0 & 12\\ 0 & 0\end{pmatrix}, \qquad \mathbf{B}\cdot\mathbf{A} = \begin{pmatrix}0 & 2\\ 0 & 5\end{pmatrix}\begin{pmatrix}1 & 2\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

In general, matrix products do **not** satisfy the commutative law:

$$\mathbf{AB} \neq \mathbf{BA}.$$

Theorem (Rules for computing $+$, $\cdot$ and $\alpha\cdot$ with matrices)

a) For all $\mathbf{A},\mathbf{B} \in \mathbb{R}^{m\times n}$ and $\mathbf{C} \in \mathbb{R}^{n\times r}$ and $\mathbf{D} \in \mathbb{R}^{\ell\times m}$ there holds:

$$(\mathbf{A}+\mathbf{B})\cdot\mathbf{C} = \mathbf{A}\cdot\mathbf{C} + \mathbf{B}\cdot\mathbf{C} \quad \text{and} \quad \mathbf{D}\cdot(\mathbf{A}+\mathbf{B}) = \mathbf{D}\cdot\mathbf{A} + \mathbf{D}\cdot\mathbf{B}.$$

b) For all $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times r}$ and $\alpha \in \mathbb{R}$ there holds:

$$\alpha\cdot(\mathbf{A}\cdot\mathbf{B}) = (\alpha\cdot\mathbf{A})\cdot\mathbf{B} = \mathbf{A}\cdot(\alpha\cdot\mathbf{B}).$$

c) For all $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times r}$ and $\mathbf{C} \in \mathbb{R}^{r\times\ell}$ there holds the associative law:

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

Definition 3.62 (powers of square(!) matrices)

For a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ we set $\mathbf{A}^0 := \mathbf{I}_n$ and define for $k \in \mathbb{N}$

$$\mathbf{A}^k := \underbrace{\mathbf{A} \cdot \mathbf{A} \cdot \ldots \cdot \mathbf{A}}_{k} \in \mathbb{R}^{n\times n}.$$

True or false?

Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, $\mathbf{C} \in \mathbb{R}^{n\times r}, \mathbf{D} \in \mathbb{R}^{n\times n}$ and $\alpha, \beta \in \mathbb{R}$. Let $\mathbf{I}_n \in \mathbb{R}^{n\times n}$ be the identity matrix. Then there holds:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 4 \end{pmatrix}$$

$$\alpha(\mathbf{C} + \mathbf{A}) = \alpha\mathbf{C} + \alpha\mathbf{A}$$

$$(\alpha + \beta)(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \beta\mathbf{B}$$

$$(\alpha\mathbf{A})\mathbf{C} = \mathbf{A}(\alpha\mathbf{C})$$

$$(\alpha + \mathbf{A})\mathbf{C} = \alpha\mathbf{C} + \mathbf{A}\mathbf{C}$$

$$\mathbf{B} = \mathbf{B}I_n$$

$$\mathbf{D}(\mathbf{D} + \mathbf{I}_n) = (\mathbf{D} + \mathbf{I}_n)\mathbf{D}$$

The inverse matrix

Recall (inverse function):
Let $f : X \to Y$ be a bijective function. Then we define its inverse function

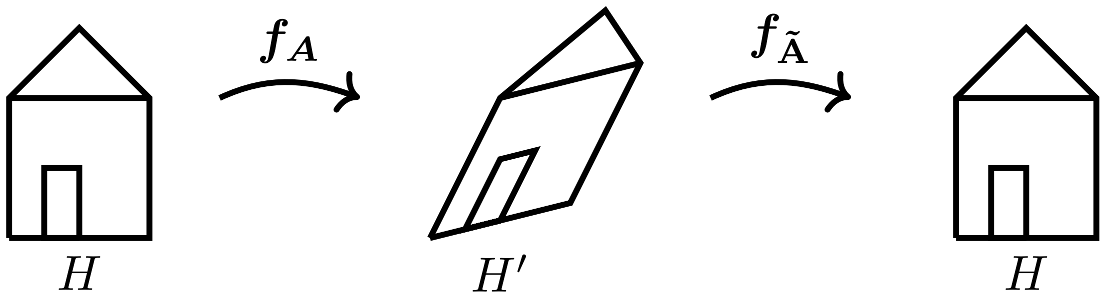
$$f^{-1} : Y \to X, \quad y \mapsto \text{the unique } x \text{ with } f(x) = y.$$

If $f : X \to Y$ is bijective and $f^{-1} : Y \to X$ is its inverse function, then we have

$$f \circ f^{-1} = \mathrm{id}_Y \quad \text{and} \quad f^{-1} \circ f = \mathrm{id}_X.$$

Question: Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ and the associated mapping $f_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m,\ \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$, is there a matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{n\times m}$ such that $f_{\tilde{\mathbf{A}}}$ is the inverse function of $f_{\mathbf{A}}$?

In terms of our house H, there should hold $f_{\tilde{\mathbf{A}}}(f_{\mathbf{A}}(H)) = H$:



The matrix $\tilde{\mathbf{A}}$ would have to satisfy $\tilde{\mathbf{A}}\mathbf{A} = \mathbf{I}_n$ and $\mathbf{A}\tilde{\mathbf{A}} = \mathbf{I}_m$.

The inverse matrix

Definition 3.65 (invertible matrix, $\mathbf{A}^{-1}$)

We call a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ *invertible* or *regular*, if its mapping $f_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^n$ is invertible. Otherwise we call $\mathbf{A}$ *singular*.
A matrix $\tilde{\mathbf{A}}$ with $f_{\tilde{\mathbf{A}}} = (f_{\mathbf{A}})^{-1}$ is called the *inverse matrix* of $\mathbf{A}$ and is denoted by $\mathbf{A}^{-1}$.

Hence there holds $f_{\mathbf{A}^{-1}} \circ f_{\mathbf{A}} = \mathrm{id}$ and $f_{\mathbf{A}} \circ f_{\mathbf{A}^{-1}} = \mathrm{id}$, i.e., $f_{\mathbf{A}^{-1}} = (f_{\mathbf{A}})^{-1}$.

Expressed with matrices and vectors:

$$\mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) = \mathbf{x} \qquad \text{and} \qquad \mathbf{A}(\mathbf{A}^{-1}\mathbf{x}) = \mathbf{x} \qquad \text{for all} \qquad \mathbf{x} \in \mathbb{R}^n.$$

In short: $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ and $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

$$\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -1 \end{pmatrix} \stackrel{?}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 2 & 4 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 2 \\ 0 & -1 \end{pmatrix} \stackrel{?}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \text{does not} \\ \text{exist} \end{pmatrix} \stackrel{?}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 2 & 4 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 0.1 & 0.4 \\ 0.2 & -0.2 \end{pmatrix} \stackrel{?}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Now check that also $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

The inverse matrix

The inverse matrix $\mathbf{A}^{-1}$ only exists if $\mathbf{A}$ is square and has full rank.

Uniqueness of the inverse matrix

The inverse matrix of a regular matrix is uniquely determined.

Proof. Let $\mathbf{B}$ and $\mathbf{C}$ both be inverse matrices of the regular matrix $\mathbf{A}$. Then there holds

$$\mathbf{BA} = \mathbf{I} = \mathbf{AB} \qquad \text{und} \qquad \mathbf{CA} = \mathbf{I} = \mathbf{AC}.$$

It follows that

$$\mathbf{B} = \mathbf{B}\underbrace{(\mathbf{AC})}_{\mathbf{I}} \overset{\text{Thm. 3.61c)}}{=} \underbrace{(\mathbf{BA})}_{\mathbf{I}}\mathbf{C} = \mathbf{C},$$

hence $\mathbf{B} = \mathbf{C}$. $\square$

The inverse matrix

Theorem 3.66 (Some rules for computing with $\mathbf{A}^{-1}$)

Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ be invertible and let $\alpha \in \mathbb{R}$ be nonzero. Then there holds

a) $\mathbf{A}^{-1}$ is invertible and $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

b) $\mathbf{AB}$ is invertible and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

c) $\alpha\mathbf{A}$ is invertible and $(\alpha\mathbf{A})^{-1} = \frac{1}{\alpha}\mathbf{A}^{-1}$.

Proof. We check that the product of the matrix with its suggested inverse yields $\mathbf{I}$ (in both orders):

a) $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$, hence $\mathbf{A}$ is the inverse of $\mathbf{A}^{-1}$.

b) $(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{B}\mathbf{B}^{-1})\mathbf{A}^{-1} = \mathbf{A}\mathbf{I}\mathbf{A}^{-1} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.
Analogously for $(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB})$.

c) $(\alpha\mathbf{A})(\frac{1}{\alpha}\mathbf{A}^{-1}) = \alpha\frac{1}{\alpha}\mathbf{A}\mathbf{A}^{-1} = 1\mathbf{I} = \mathbf{I}$. Analogously for $(\frac{1}{\alpha}\mathbf{A}^{-1})(\alpha\mathbf{A}) = \mathbf{I}$. $\square$

The inverse matrix

Theorem 3.67 (testing for the inverse)

For two square matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ there holds:

$$\mathbf{AB} = \mathbf{I} \qquad \Longleftrightarrow \qquad \mathbf{BA} = \mathbf{I}.$$

Hence one has to check only one of the two equations in order to show $\mathbf{B} = \mathbf{A}^{-1}$.

Proof. We begin with $\boxed{\Rightarrow}$. Assume that $\mathbf{AB} = \mathbf{I}$. Then there holds for all $\mathbf{x} \in \mathrm{Ker}(\mathbf{B})$

$$\mathbf{x} = \mathbf{Ix} = \underbrace{(\mathbf{AB})}_{\mathbf{I}}\mathbf{x} = \mathbf{A}\underbrace{(\mathbf{Bx})}_{\mathbf{o}} = \mathbf{Ao} = \mathbf{o},$$

i.e., $\mathrm{Ker}(\mathbf{B}) = \{\mathbf{o}\}$. Hence $f_{\mathbf{B}}$ and also $\mathbf{B}$ are invertible, i.e., there exists $\mathbf{B}^{-1}$. Multiplying both sides of the equation $\mathbf{AB} = \mathbf{I}$ from the left with $\mathbf{B}$ and from the right with $\mathbf{B}^{-1}$ leads to

$$\mathbf{B}(\mathbf{AB})\mathbf{B}^{-1} = \mathbf{BIB}^{-1}, \quad \text{i.e.,} \quad (\mathbf{BA})\underbrace{(\mathbf{BB}^{-1})}_{\mathbf{I}} = \underbrace{(\mathbf{BB}^{-1})}_{\mathbf{I}}$$

and hence the second equation $\mathbf{BA} = \mathbf{I}$. The opposite direction $\boxed{\Leftarrow}$ is shown analogously. $\square$

The inverse matrix

Question 1: Given two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$, how can we test whether $\mathbf{B}$ is the inverse of $\mathbf{A}$?

Check whether $\mathbf{AB} = \mathbf{I}$ is true.

Question 2: Given a regular matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, how can we find its inverse $\mathbf{A}^{-1} \in \mathbb{R}^{n\times n}$?

We denote the columns of $\mathbf{A}^{-1}$ by $\mathbf{s}_1, \dots, \mathbf{s}_n$. Then there must hold

$$\begin{pmatrix} | & & | \\ \mathbf{e}_1 & \cdots & \mathbf{e}_n \\ | & & | \end{pmatrix} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{A}\begin{pmatrix} | & & | \\ \mathbf{s}_1 & \cdots & \mathbf{s}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ \mathbf{A}\mathbf{s}_1 & \cdots & \mathbf{A}\mathbf{s}_n \\ | & & | \end{pmatrix},$$

hence $\mathbf{A}\mathbf{s}_1 = \mathbf{e}_1, \quad \mathbf{A}\mathbf{s}_2 = \mathbf{e}_2, \quad \dots, \quad \mathbf{A}\mathbf{s}_n = \mathbf{e}_n.$

The inverse $\mathbf{A}^{-1}$ can be computed by the (simultaneous) solution of n systems of equations, e.g., using Gauss elimination:

$$(\mathbf{A} \mid \mathbf{e}_1 \cdots \mathbf{e}_n) \rightsquigarrow (\mathbf{A}' \mid \mathbf{f}_1 \cdots \mathbf{f}_n) \rightsquigarrow (\mathbf{I} \mid \mathbf{s}_1 \cdots \mathbf{s}_n) = (\mathbf{I} \mid \mathbf{A}^{-1})$$

Addition to Theorem 3.18: The following admissible operation does not change the solution of an LES: (iii) Multiplication of a row with an $\alpha \in \mathbb{R} \setminus \{0\}$.

The inverse matrix

Example for computing the inverse $\mathbf{A}^{-1}$

$$\underbrace{\left(\begin{array}{rrr|rrr} 2 & -1 & 0 & 1 & 0 & 0 \\ \boxed{1} & 2 & -2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right)}_{(\mathbf{A}|\ \mathbf{I})} \rightsquigarrow \left(\begin{array}{rrr|rrr} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \boxed{2.5} & -2 & -0.5 & 1 & 0 \\ 0 & \boxed{-1} & 1 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{rrr|rrr} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \\ 0 & \boxed{2.5} & -2 & -0.5 & 1 & 0 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{rrr|rrr} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \\ 0 & 0 & \boxed{0.5} & -0.5 & 1 & 2.5 \end{array}\right) \rightsquigarrow \left(\begin{array}{rrr|rrr} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & \boxed{1} & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 2 & 5 \end{array}\right) \rightsquigarrow \left(\begin{array}{rrr|rrr} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \boxed{-1} & 0 & 1 & -2 & -4 \\ 0 & 0 & 1 & -1 & 2 & 5 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{rrr|rrr} 2 & \boxed{-1} & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 2 & 4 \\ 0 & 0 & 1 & -1 & 2 & 5 \end{array}\right) \rightsquigarrow \left(\begin{array}{rrr|rrr} \boxed{2} & 0 & 0 & 0 & 2 & 4 \\ 0 & 1 & 0 & -1 & 2 & 4 \\ 0 & 0 & 1 & -1 & 2 & 5 \end{array}\right) \rightsquigarrow \underbrace{\left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & -1 & 2 & 4 \\ 0 & 0 & 1 & -1 & 2 & 5 \end{array}\right)}_{(\mathbf{I}\ |\mathbf{A}^{-1})}$$

Finally, check that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

The inverse matrix

Task: Solve the system of equations $\mathbf{Ax} = \mathbf{b}$.

Solution 1: Gauss elimination: $(\mathbf{A}|\mathbf{b}) \rightsquigarrow (\mathbf{A}'|\mathbf{b}')$ and backward elimination.

Solution 2: Compute the inverse $\mathbf{A}^{-1}$ which leads to

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

Solution 2 is considerably more work intensive than **Solution 1** but may be preferable for solving several linear systems with the same matrix $\mathbf{A}$ but different right hand sides $\mathbf{b}_i$. However, in this case one rather uses

Solution 3: Compute a so-called LU factorization of $\mathbf{A}$ (comes later!) and then compute $\mathbf{x}$ by forward and backward substitution.