

Operations with matrices

Definitions: For any $m, n \in \mathbb{N}$ we define the *zero matrix* $\mathbf{0} \in \mathbb{R}^{m \times n}$ satisfying $\forall i, j : (\mathbf{0})_{i,j} = 0$. For $n \in \mathbb{N}$ the *identity matrix* \mathbf{I}_n is defined as $(\mathbf{I}_n)_{i,j} = 1$ when $i = j$ and $(\mathbf{I}_n)_{i,j} = 0$ otherwise.

The *main diagonal* of a square matrix \mathbf{A} means the elements $a_{i,i}$.

The *transpose* of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is $\mathbf{A}^T \in \mathbb{R}^{n \times m}$ defined as $(\mathbf{A}^T)_{i,j} = a_{j,i}$. A square matrix \mathbf{A} is *symmetric* if $\mathbf{A}^T = \mathbf{A}$.

The *sum* of matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ is $(\mathbf{A} + \mathbf{B}) \in \mathbb{R}^{m \times n}$ defined as

$$(\mathbf{A} + \mathbf{B})_{i,j} = a_{i,j} + b_{i,j}.$$

The *α -multiple* of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\alpha \in \mathbb{R}$ is $(\alpha\mathbf{A}) \in \mathbb{R}^{m \times n}$ s.t.

$$(\alpha\mathbf{A})_{i,j} = \alpha a_{i,j}.$$

For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times p}$ the *product* $(\mathbf{AB}) \in \mathbb{R}^{m \times p}$ is defined as

$$(\mathbf{AB})_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}.$$

Example of matrix product

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 4 & 0 \\ 0 & 0 & 1 & 3 \\ 3 & 1 & 2 & 0 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \\ 2 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{AB} = \begin{pmatrix} 9 & 8 \\ 2 & 3 \\ 7 & 9 \end{pmatrix}$$

Mnemotechnically:

					1	2				1	.
					0	3				0	.
					2	0				2	.
					0	1				0	.
					<hr/>					<hr/>	
				1	2	4	0	9	8	.	.
				0	0	1	3	2	3	.	.
				3	1	2	0	7	9	.	.
										.	.
										.	.
										7	.

$$3 \cdot 1 + 1 \cdot 0 + 2 \cdot 2 + 0 \cdot 0 = 7$$

Observation: The matrix product \mathbf{AB} for $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times p}$ requires mnp multiplications and $m(n-1)p$ additions $\approx mnp$ arithmetic operations.

Prove or solve on your own

Proposition: If the results of operations are defined then:

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

$$\mathbf{A} + \mathbf{0} = \mathbf{A}$$

$$\exists! \mathbf{B} : \mathbf{A} + \mathbf{B} = \mathbf{A}$$

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

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

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$$

$$(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T$$

$$(\mathbf{A}^T)^T = \mathbf{A}$$

$$\text{Pf: } ((\mathbf{A}^T)^T)_{i,j} = (\mathbf{A}^T)_{j,i} = a_{i,j}$$

Find square matrices \mathbf{A}, \mathbf{B} such that $\mathbf{AB} \neq \mathbf{BA}$.

Show that the matrix \mathbf{AA}^T is symmetric for any \mathbf{A} .

Show that for any $\mathbf{A} \in \mathbb{R}^{m \times n}$ it holds that $\mathbf{I}_m \mathbf{A} = \mathbf{A} \mathbf{I}_n = \mathbf{A}$.

Derive the following rules for the products of block matrices.

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{pmatrix} \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix} = \mathbf{A}_1 \mathbf{B}_1 + \mathbf{A}_2 \mathbf{B}_2$$

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{pmatrix} \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_3 & \mathbf{B}_4 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 \mathbf{B}_1 + \mathbf{A}_2 \mathbf{B}_3 & \mathbf{A}_1 \mathbf{B}_2 + \mathbf{A}_2 \mathbf{B}_4 \\ \mathbf{A}_3 \mathbf{B}_1 + \mathbf{A}_4 \mathbf{B}_3 & \mathbf{A}_3 \mathbf{B}_2 + \mathbf{A}_4 \mathbf{B}_4 \end{pmatrix}$$

Proposition: If the results of the operations are defined then:

1. $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$
2. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
3. $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
4. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$

Proof: 1. \mathbf{A} is of order $m \times n \implies \mathbf{B}$ is of order $n \times p$

$$\begin{aligned} ((\mathbf{AB})^T)_{i,j} &= (\mathbf{AB})_{j,i} = \sum_{k=1}^n a_{j,k} b_{k,i} = \\ &= \sum_{k=1}^n b_{k,i} a_{j,k} = \sum_{k=1}^n (\mathbf{B}^T)_{i,k} (\mathbf{A}^T)_{k,j} = (\mathbf{B}^T \mathbf{A}^T)_{i,j} \end{aligned}$$

2. \mathbf{A} of order $m \times n \implies \mathbf{B}$ of order $n \times p \implies \mathbf{C}$ of order $p \times q$

$$\begin{aligned} ((\mathbf{AB})\mathbf{C})_{i,j} &= \sum_{k=1}^p (\mathbf{AB})_{i,k} c_{k,j} = \sum_{k=1}^p \left(\sum_{l=1}^n a_{i,l} b_{l,k} \right) c_{k,j} = \\ &= \sum_{k=1}^p \sum_{l=1}^n a_{i,l} b_{l,k} c_{k,j} = \sum_{l=1}^n \sum_{k=1}^p a_{i,l} b_{l,k} c_{k,j} = \\ &= \sum_{l=1}^n a_{i,l} \left(\sum_{k=1}^p b_{l,k} c_{k,j} \right) = \sum_{l=1}^n a_{i,l} (\mathbf{BC})_{l,j} = (\mathbf{A}(\mathbf{BC}))_{i,j} \end{aligned}$$

Proposition: If the results of the operations are defined then:

1. $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$
2. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
3. $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
4. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$

Proof: 3. $\mathbf{A} \in \mathbb{R}^{m \times n} \implies \mathbf{B} \in \mathbb{R}^{m \times n}, \mathbf{C} \in \mathbb{R}^{n \times p}$

$$\begin{aligned} ((\mathbf{A} + \mathbf{B})\mathbf{C})_{i,j} &= \sum_{k=1}^n (\mathbf{A} + \mathbf{B})_{i,k} c_{k,j} = \sum_{k=1}^n (a_{i,k} + b_{i,k}) c_{k,j} = \\ &= \sum_{k=1}^n (a_{i,k} c_{k,j} + b_{i,k} c_{k,j}) = \sum_{k=1}^n a_{i,k} c_{k,j} + \sum_{k=1}^n b_{i,k} c_{k,j} = \\ &= (\mathbf{AC})_{i,j} + (\mathbf{BC})_{i,j} = (\mathbf{AC} + \mathbf{BC})_{i,j} \end{aligned}$$

4. $\mathbf{A} \in \mathbb{R}^{m \times n} \implies \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times p}$

$$\begin{aligned} (\mathbf{A}(\mathbf{B} + \mathbf{C}))_{i,j} &= \sum_{k=1}^n a_{i,k} (\mathbf{B} + \mathbf{C})_{k,j} = \sum_{k=1}^n a_{i,k} (b_{k,j} + c_{k,j}) = \\ &= \sum_{k=1}^n (a_{i,k} b_{k,j} + a_{i,k} c_{k,j}) = \sum_{k=1}^n a_{i,k} b_{k,j} + \sum_{k=1}^n a_{i,k} c_{k,j} = \\ &= (\mathbf{AB})_{i,j} + (\mathbf{AC})_{i,j} = (\mathbf{AB} + \mathbf{AC})_{i,j} \end{aligned}$$

Example for the proof of the product associativity

$$\begin{aligned}
 ((\mathbf{AB})\mathbf{C})_{i,j} &= \sum_{k=1}^p (\mathbf{AB})_{i,k} c_{k,j} = \sum_{k=1}^p \left(\sum_{l=1}^n a_{i,l} b_{l,k} \right) c_{k,j} = \\
 &= \sum_{k=1}^p \sum_{l=1}^n a_{i,l} b_{l,k} c_{k,j} = \sum_{l=1}^n \sum_{k=1}^p a_{i,l} b_{l,k} c_{k,j} = \\
 &= \sum_{l=1}^n a_{i,l} \left(\sum_{k=1}^p b_{l,k} c_{k,j} \right) = \sum_{l=1}^n a_{i,l} (\mathbf{BC})_{l,j} = (\mathbf{A}(\mathbf{BC}))_{i,j}
 \end{aligned}$$

$m = q = i = j = 1, n = p = 2$:

				6
				7
		0	1	6
		2	3	7
4	5	10	19	193

				6
				7
		0	1	7
		2	3	33
4	5			193

$$\begin{aligned}
 193 &= 10 \cdot 6 + 19 \cdot 7 = (4 \cdot 0 + 5 \cdot 2) \cdot 6 + (4 \cdot 1 + 5 \cdot 3) \cdot 7 = \\
 &= (4 \cdot 0 \cdot 6 + 5 \cdot 2 \cdot 6) + (4 \cdot 1 \cdot 7 + 5 \cdot 3 \cdot 7) = (4 \cdot 0 \cdot 6 + 4 \cdot 1 \cdot 7) + (5 \cdot 2 \cdot 6 + 5 \cdot 3 \cdot 7) = \\
 &= 4 \cdot (0 \cdot 6 + 1 \cdot 7) + 5 \cdot (2 \cdot 6 + 3 \cdot 7) = 4 \cdot 7 + 5 \cdot 33 = 193
 \end{aligned}$$

Efficiency of the product evaluation

$$\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{n \times p}, \mathbf{C} \in \mathbb{R}^{p \times q} : (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$

	B	C
A	AB	(AB)C

$\mathbf{AB} \approx mnp$ arithmetic operations,
 $(\mathbf{AB})\mathbf{C} \approx mpq$ operations,
 in total $\approx mp(n + q)$ operations.

	C
B	BC
A	A(BC)

$\mathbf{BC} \approx npq$ operations,
 $\mathbf{A}(\mathbf{BC}) \approx mnq$ operations,
 in total $\approx nq(m + p)$ operations.

When $m \ll n, p, q$, we get $mp(n + q) \ll nq(m + p)$.

					4
		2	-3	3	1
		-1	2	-1	-1
1	2	0	1	1	0
-2	-4	0	-2	-2	0

				4
				1
				-1
	2	-3	3	2
	-1	2	-1	-1
		1	2	0
		-2	-4	0

The first way has *more* intermediate values than the second.

Though the final result is the same in both ways, a suitable order of evaluation may affect the overall computational complexity.

Elementary matrices

Observation: Let B be the matrix obtained from A by:

- ▶ the multiplication of the i -th row by $t \neq 0$. Then $B = EA$ where $e_{i,i} = t$, $e_{k,k} = 1$ for $k \neq i$ and $e_{k,l} = 0$ for $k \neq l$.
- ▶ the adding the j -th row to the i -th. Then $B = EA$ where $e_{i,j} = 1$, $e_{k,k} = 1$, for all k , and $e_{k,l} = 0$ for $i \neq k \neq l \neq j$.

Such matrices E are called *elementary* matrices.

Example:

				3	6	...	$a_{1,k}$	
				7	1	...	$a_{2,k}$	$i = 2$
				2	4	...	$a_{3,k}$	
				5	3	...	$a_{4,k}$	$j = 4$
	1	0	0	0	3	6	...	$a_{1,k}$
	0	4	0	0	28	4	...	$4 \cdot a_{2,k}$
	0	0	1	0	2	4	...	$a_{3,k}$
	0	0	0	1	5	3	...	$a_{4,k}$
	1	0	0	0	3	6	...	$a_{1,k}$
	0	1	0	1	12	4	...	$a_{2,k} + a_{4,k}$
	0	0	1	0	2	4	...	$a_{3,k}$
	0	0	0	1	5	3	...	$a_{4,k}$

$t = 4$

Inverse matrix

Definition: If for $\mathbf{A} \in \mathbb{R}^{n \times n}$ there exists $\mathbf{B} \in \mathbb{R}^{n \times n}$ such that $\mathbf{AB} = \mathbf{I}_n$, then \mathbf{B} is called the *inverse matrix* and denoted by \mathbf{A}^{-1} . If \mathbf{A} has an inverse, then it is called *regular*, otherwise it is *singular*.

Theorem: For $\mathbf{A} \in \mathbb{R}^{n \times n}$ the following are equivalent:

1. \mathbf{A} is regular, i.e. $\exists \mathbf{B} \in \mathbb{R}^{n \times n} : \mathbf{AB} = \mathbf{I}_n$.
2. $\text{rank}(\mathbf{A}) = n$.
3. $\mathbf{A} \sim \mathbf{I}_n$.
4. The system $\mathbf{Ax} = \mathbf{0}$ has only trivial solution $\mathbf{x} = \mathbf{0}$.

Proof: 2. \Leftrightarrow 4. was already proved.

2. \Rightarrow 3. by Gauss-Jordan elimination, 2. \Leftarrow 3. trivially.

2. \Rightarrow 1. Denote $\mathbf{I}_n = (\mathbf{e}^1 | \dots | \mathbf{e}^n)$. For $i = 1, \dots, n$ consider systems $\mathbf{Ax}^i = \mathbf{e}^i$. Since $\text{rank}(\mathbf{A}) = n$ we get $\mathbf{B} = (\mathbf{x}^1 | \dots | \mathbf{x}^n)$.

1. \Rightarrow 2. If $\text{rank}(\mathbf{A}) < n$ then for some i , the i -th row of \mathbf{A} can be eliminated by the other rows, thus $\mathbf{Ax}^i = \mathbf{e}^i$ has no solution as the only 1 in \mathbf{e}^i cannot be eliminated by zeroes.

Inverse matrix

Corollary: If the inverse matrix \mathbf{A}^{-1} exists, it is unique.

Theorem: The inverse matrix satisfies: $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$.

Proof: We first show that \mathbf{A}^{-1} is regular:

If $\mathbf{A}^{-1}\mathbf{x} = \mathbf{0}$ has a solution then $\mathbf{x} = \mathbf{I}_n\mathbf{x} = \mathbf{A}\mathbf{A}^{-1}\mathbf{x} = \mathbf{A}\mathbf{0} = \mathbf{0}$.

Hence there exists $(\mathbf{A}^{-1})^{-1}$ and we get:

$$\begin{aligned}\mathbf{A}^{-1}\mathbf{A} &= \mathbf{A}^{-1}\mathbf{A}\mathbf{I}_n = (\mathbf{A}^{-1}\mathbf{A})(\mathbf{A}^{-1}(\mathbf{A}^{-1})^{-1}) = \\ \mathbf{A}^{-1}(\mathbf{A}\mathbf{A}^{-1})(\mathbf{A}^{-1})^{-1} &= \mathbf{A}^{-1}\mathbf{I}_n(\mathbf{A}^{-1})^{-1} = \mathbf{A}^{-1}(\mathbf{A}^{-1})^{-1} = \mathbf{I}_n\end{aligned}$$

Corollary: If $\mathbf{BA} = \mathbf{I}_n$ then $\mathbf{A}^{-1} = \mathbf{B}$.

Proof: $\mathbf{BA} = \mathbf{I}_n \implies \mathbf{AB} = \mathbf{I}_n \implies \mathbf{A}^{-1} = \mathbf{B}$

Inverse matrix calculation

- ▶ Assemble $(\mathbf{A}|\mathbf{I}_n)$ and by Gauss-Jordan elimination get $(\mathbf{I}_n|\mathbf{B})$.
- ▶ If this process fails, then \mathbf{A} is singular.
- ▶ Denote $\mathbf{E}_1, \dots, \mathbf{E}_k$ the elementary matrices of the applied elementary transforms. Then the left side of $(\mathbf{A}|\mathbf{I}_n) \sim \sim (\mathbf{I}_n|\mathbf{B})$ yields $\mathbf{E}_k \cdots \mathbf{E}_1 \mathbf{A} = \mathbf{I}_n$, the right side yields $\mathbf{E}_k \cdots \mathbf{E}_1 \mathbf{I}_n = \mathbf{B}$ thus $\mathbf{B}\mathbf{A} = \mathbf{I}_n$ and therefore $\mathbf{A}^{-1} = \mathbf{B}$.
(Indeed the columns of \mathbf{B} are solutions of $\mathbf{A}\mathbf{x}^i = \mathbf{e}^i$)

Example:

$$(\mathbf{A}|\mathbf{I}_n) = \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 3 & 4 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 5 & 6 & 3 & -1 & 0 \end{array} \right) \sim$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & -3 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 3 & -1 & -5 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -1 & -8 \\ 0 & 1 & 0 & -3 & 1 & 6 \\ 0 & 0 & 1 & 3 & -1 & -5 \end{array} \right) = (\mathbf{I}_n|\mathbf{A}^{-1})$$

Correctness test: $\begin{pmatrix} 1 & 3 & 2 \\ 3 & 4 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 4 & -1 & -8 \\ -3 & 1 & 6 \\ 3 & -1 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Matrix equations

Observation: If R is regular, then: $A = B \iff AR = BR$.

Proof: \Rightarrow trivially, \Leftarrow : $A = AI_n = ARR^{-1} = BRR^{-1} = B$.

Exercises: Prove that regular matrices A, B of the same order satisfy:

▶ $(A^{-1})^{-1} = A$

▶ AB is regular

▶ $(AB)^{-1} = B^{-1}A^{-1}$

▶ $(A^T)^{-1} = (A^{-1})^T$

Matrix equations

$$\mathbf{A} + \mathbf{X} = \mathbf{B} \quad \implies \quad \mathbf{X} = \mathbf{B} - \mathbf{A} = \mathbf{B} + (-1)\mathbf{A}$$

$$a\mathbf{X} = \mathbf{B} \quad \implies \quad \mathbf{X} = \frac{1}{a}\mathbf{B}$$

$$\mathbf{A}\mathbf{X} = \mathbf{B} \quad \implies \quad \mathbf{X} = \mathbf{A}^{-1}\mathbf{B}, \text{ for regular } \mathbf{A}$$

$$\mathbf{X}\mathbf{A} = \mathbf{B} \quad \implies \quad \mathbf{X} = \mathbf{B}\mathbf{A}^{-1}, \text{ for regular } \mathbf{A}$$

Beware, products $\mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{B}\mathbf{A}^{-1}$ could be distinct.

$$\text{Equations } \begin{pmatrix} 9 & 2 \\ 4 & 1 \end{pmatrix} \mathbf{X} = \begin{pmatrix} 3 & -1 \\ 6 & -2 \end{pmatrix} \text{ and } \mathbf{X}' \begin{pmatrix} 9 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ 6 & -2 \end{pmatrix}$$

$$\text{have different solutions } \mathbf{X} = \begin{pmatrix} -9 & 3 \\ 42 & -14 \end{pmatrix} \text{ and } \mathbf{X}' = \begin{pmatrix} 7 & -15 \\ 14 & -30 \end{pmatrix}$$

Test:

		-9	3		9	2
		42	-14		4	1
	9	2	3	-1	7	-15
	4	1	6	-2	14	-30
					3	-1
					6	-2