Operations with matrices

Definition: For any $m, n \in \mathbb{N}$ we define the *zero matrix* $\mathbf{0}_{m,n} \in \mathbb{R}^{m \times n}$ satisfying $\forall i, j : (\mathbf{0}_{m,n})_{ij} = 0$. Denoted also as $\mathbf{0}$.

For $n \in \mathbb{N}$ the *identity matrix* $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ is defined as $(\mathbf{I}_n)_{ij} = 1$ when $i = j$ and $(\mathbf{I}_n)_{ij} = 0$ otherwise. Denoted also as $\mathbf{I}$.

The *main diagonal* of a square matrix $\boldsymbol{A}$ means the elements a_{ii}.

Example:

$$\mathbf{0}_{2,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Definition: The *transpose* of a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is the matrix $\boldsymbol{A}^\mathsf{T} \in \mathbb{R}^{n \times m}$ defined as $(\boldsymbol{A}^\mathsf{T})_{ij} = a_{ji}$.

A square matrix $\boldsymbol{A}$ is *symmetric* if $\boldsymbol{A}^\mathsf{T} = \boldsymbol{A}$, i.e. $(\boldsymbol{A}^\mathsf{T})_{ij} = a_{ji} = a_{ij}$.

Example:

The transpose matrix to $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ is $\boldsymbol{A}^\mathsf{T} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

The matrix $\boldsymbol{A} = \begin{pmatrix} -3 & 7 \\ 7 & 0 \end{pmatrix}$ is symmetric since it satisfies $\boldsymbol{A} = \boldsymbol{A}^\mathsf{T}$.

Operations with matrices

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

$$(\boldsymbol{A}+\boldsymbol{B})_{ij} = a_{ij} + b_{ij}.$$

Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 7 & 0 & -2 \\ 1 & -5 & 1 \end{pmatrix} = \begin{pmatrix} 1+7 & 2+0 & 3-2 \\ 4+1 & 5-5 & 6+1 \end{pmatrix} = \begin{pmatrix} 8 & 2 & 1 \\ 5 & 0 & 7 \end{pmatrix}$$

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

$$(t\boldsymbol{A})_{ij} = t a_{ij}.$$

Example:

$$3\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 & 3 \cdot 2 & 3 \cdot 3 \\ 3 \cdot 4 & 3 \cdot 5 & 3 \cdot 6 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{pmatrix}$$

Matrix product

Definition:

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

$$(\boldsymbol{AB})_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Example:

For $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 4 & 0 \\ 0 & 0 & 1 & 3 \\ 3 & 1 & 2 & 0 \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \\ 2 & 0 \\ 0 & 1 \end{pmatrix}$ we get $\boldsymbol{AB} = \begin{pmatrix} 9 & 8 \\ 2 & 3 \\ 7 & 9 \end{pmatrix}$.

Mnemotechnically:

	$\boldsymbol{B}$
$\boldsymbol{A}$	$\boldsymbol{AB}$

				1	2
				0	3
				2	0
				0	1
1	2	4	0	9	8
0	0	1	3	2	3
3	1	2	0	7	9

				1	.
				0	.
				2	.
				0	.
.	.	.	.	.	.
.	.	.	.	.	.
3	1	2	0	7	.

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

Matrix product

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

$$(\boldsymbol{AB})_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Observation: The matrix product $\boldsymbol{AB}$ for $\boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{B} \in \mathbb{R}^{n \times p}$ requires mnp multiplications and $m(n-1)p$ additions, in total $mnp + m(n-1)p = 2mnp - mp \approx mnp$ arithmetic operations.

We read $\approx$ as "asymptotically" or "approximately" and it means only the largest term(s) without any multiplicative constant(s).

Warning: For matrices, *multiple* and *product* mean different things!

The matrix product is used in the notation $\boldsymbol{Ax} = \boldsymbol{b}$, where the vector $\boldsymbol{x}$ is considered as a matrix with a single column.

Prove or solve on your own

Proposition: If the results of operations are defined then:

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

$\boldsymbol{A}+\mathbf{0}=\boldsymbol{A}$ $\quad$ $\exists!\boldsymbol{B}:\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{A}$

$s(t\boldsymbol{A})=(st)\boldsymbol{A}$ $\quad$ $(s+t)\boldsymbol{A}=s\boldsymbol{A}+t\boldsymbol{A}$

$(\boldsymbol{A}+\boldsymbol{B})^{\mathsf{T}}=\boldsymbol{A}^{\mathsf{T}}+\boldsymbol{B}^{\mathsf{T}}$ $\quad$ $(t\boldsymbol{A})^{\mathsf{T}}=t\boldsymbol{A}^{\mathsf{T}}$

$(\boldsymbol{A}^{\mathsf{T}})^{\mathsf{T}}=\boldsymbol{A}$ $\quad$ Pf: $((A^{\mathsf{T}})^{\mathsf{T}})_{ij}=(A^{\mathsf{T}})_{ji}=a_{ij}$

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

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

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

Derive the following rules for the products of block matrices.

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

Hint: split the sum.

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

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

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

Proof:

1. $\boldsymbol{A}$ is of order $m \times n \Longrightarrow \boldsymbol{B}$ is of order $n \times p$

$$((\boldsymbol{AB})^{\mathsf{T}})_{ij} = (\boldsymbol{AB})_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki} =$$

$$= \sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} (\boldsymbol{B}^{\mathsf{T}})_{ik} (\boldsymbol{A}^{\mathsf{T}})_{kj} = (\boldsymbol{B}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}})_{ij}$$

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

$$((\boldsymbol{AB})\boldsymbol{C})_{ij} = \sum_{k=1}^{p} (\boldsymbol{AB})_{ik} c_{kj} = \sum_{k=1}^{p} \left(\sum_{l=1}^{n} a_{il} b_{lk} \right) c_{kj} =$$

$$= \sum_{k=1}^{p} \sum_{l=1}^{n} a_{il} b_{lk} c_{kj} = \sum_{l=1}^{n} \sum_{k=1}^{p} a_{il} b_{lk} c_{kj} =$$

$$= \sum_{l=1}^{n} a_{il} \left(\sum_{k=1}^{p} b_{lk} c_{kj} \right) = \sum_{l=1}^{n} a_{il} (\boldsymbol{BC})_{lj} = (\boldsymbol{A}(\boldsymbol{BC}))_{ij}$$

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

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

3. $\boldsymbol{A} \in \mathbb{R}^{m\times n} \Longrightarrow \boldsymbol{B} \in \mathbb{R}^{m\times n}, \boldsymbol{C} \in \mathbb{R}^{n\times p}$

$$((\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C})_{ij} = \sum_{k=1}^{n} (\boldsymbol{A}+\boldsymbol{B})_{ik} c_{kj} = \sum_{k=1}^{n} (a_{ik}+b_{ik}) c_{kj} =$$
$$= \sum_{k=1}^{n} (a_{ik}c_{kj} + b_{ik}c_{kj}) = \sum_{k=1}^{n} a_{ik}c_{kj} + \sum_{k=1}^{n} b_{ik}c_{kj} =$$
$$= (\boldsymbol{AC})_{ij} + (\boldsymbol{BC})_{ij} = (\boldsymbol{AC}+\boldsymbol{BC})_{ij}$$

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

$$(\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}))_{ij} = \sum_{k=1}^{n} a_{ik} (\boldsymbol{B}+\boldsymbol{C})_{kj} = \sum_{k=1}^{n} a_{ik}(b_{kj}+c_{kj}) =$$
$$= \sum_{k=1}^{n} (a_{ik}b_{kj} + a_{ik}c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj} =$$
$$= (\boldsymbol{AB})_{ij} + (\boldsymbol{AC})_{ij} = (\boldsymbol{AB}+\boldsymbol{AC})_{ij}$$

Example for the proof of the product associativity

$$((\boldsymbol{AB})\boldsymbol{C})_{ij} = \sum_{k=1}^{p} (\boldsymbol{AB})_{ik} c_{kj} = \sum_{k=1}^{p} \left(\sum_{l=1}^{n} a_{il} b_{lk} \right) c_{kj} =$$

$$= \sum_{k=1}^{p} \sum_{l=1}^{n} a_{il} b_{lk} c_{kj} = \sum_{l=1}^{n} \sum_{k=1}^{p} a_{il} b_{lk} c_{kj} =$$

$$= \sum_{l=1}^{n} a_{il} \left(\sum_{k=1}^{p} b_{lk} c_{kj} \right) = \sum_{l=1}^{n} a_{il} (\boldsymbol{BC})_{lj} = (\boldsymbol{A}(\boldsymbol{BC}))_{ij}$$

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

		0	1	6
		2	3	7
4	5	10	19	193

		6
		7
0	1	7
2	3	33
4	5	193

$$193 = 10 \cdot 6 + 19 \cdot 7 = (4 \cdot 0 + 5 \cdot 2) \cdot 6 + (4 \cdot 1 + 5 \cdot 3) \cdot 7 =$$

$$(4 \cdot 0 \cdot 6 + 5 \cdot 2 \cdot 6) + (4 \cdot 1 \cdot 7 + 5 \cdot 3 \cdot 7) = (4 \cdot 0 \cdot 6 + 4 \cdot 1 \cdot 7) + (5 \cdot 2 \cdot 6 + 5 \cdot 3 \cdot 7) =$$

$$4 \cdot (0 \cdot 6 + 1 \cdot 7) + 5 \cdot (2 \cdot 6 + 3 \cdot 7) = 4 \cdot 7 + 5 \cdot 33 = 193$$

Efficiency of the product evaluation

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

	B	**C**
A	**AB**	**(AB)C**

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

	C
B	**BC**
A	**A(BC)**

$\boldsymbol{BC} \approx npq$ operations,
$\boldsymbol{A}(\boldsymbol{BC}) \approx mnq$ 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 $\boldsymbol{B}$ be the matix obtained from $\boldsymbol{A}$ by:

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

Such matrices $\boldsymbol{E}$ are called *elementary* matrices.

Example:

$$A = \begin{pmatrix} 3 & 6 & \ldots & a_{1k} \\ 7 & 1 & \ldots & a_{2k} \\ 2 & 4 & \ldots & a_{3k} \\ 5 & 3 & \ldots & a_{4k} \end{pmatrix} \qquad \begin{matrix} \\ i = 2 \\ \\ j = 4 \end{matrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 3 & 6 & \ldots & a_{1k} \\ 28 & 4 & \ldots & 4 \cdot a_{2k} \\ 2 & 4 & \ldots & a_{3k} \\ 5 & 3 & \ldots & a_{4k} \end{pmatrix} \qquad t = 4$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 3 & 6 & \ldots & a_{1k} \\ 12 & 4 & \ldots & a_{2k} + a_{4k} \\ 2 & 4 & \ldots & a_{3k} \\ 5 & 3 & \ldots & a_{4k} \end{pmatrix}$$

Questions to understand the lecture topic

- Which of the matrix identities would become invalid if *product* of individual numbers was not commutative?
- Which would become invalid if *sum* was not commutative?
- What are the assumptions for block sizes in the rules for block matrix multiplication?
- What do the elementary matrices for the remaining elementary operations look like: adding t times the ith row to the jth row and swapping two rows?
- What does the product with the elementary matrix *from the right*, ie. $\boldsymbol{AE}$, yield?