Algoma UNIVERSITY

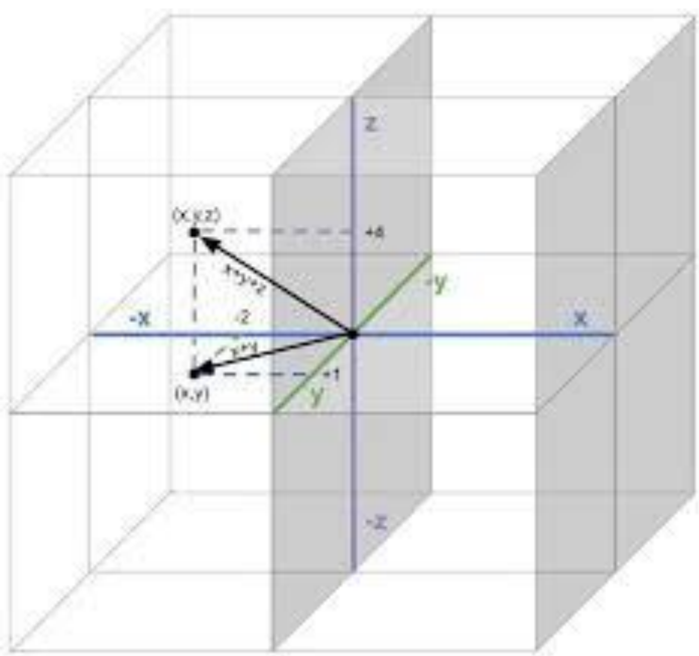
School of Computer Science & Technology - Algoma

Linear Algebra

Math 1057

Instructor

Dr. Mortaza Baky Haskuee



9 Matrix Algebra 75

Lecture 9

Matrix Algebra

9.1 Sums of Matrices

We begin with the definition of matrix addition.

Definition 9.1: Given matrices

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix},$$

both of the same dimension $m \times n$, the sum $\mathbf{A} + \mathbf{B}$ is defined as

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}.$$

Next is the definition of scalar-matrix multiplication.

Definition 9.2: For a scalar α we define $\alpha\mathbf{A}$ by

$$\alpha\mathbf{A} = \alpha \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix}.$$

Example 9.3. Given $\mathbf{A}$ and $\mathbf{B}$ below, find $3\mathbf{A} - 2\mathbf{B}$.

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 5 \\ 0 & -3 & 9 \\ 4 & -6 & 7 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 5 & 0 & -11 \\ 3 & -5 & 1 \\ -1 & -9 & 0 \end{bmatrix}$$

Solution. We compute:

$$3\mathbf{A} - 2\mathbf{B} = \begin{bmatrix} 3 & -6 & 15 \\ 0 & -9 & 27 \\ 12 & -18 & 21 \end{bmatrix} - \begin{bmatrix} 10 & 0 & -22 \\ 6 & -10 & 2 \\ -2 & -18 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -7 & -6 & 37 \\ -6 & 1 & 25 \\ 14 & 0 & 21 \end{bmatrix}$$

□

Below are some basic algebraic properties of matrix addition/scalar multiplication.

Theorem 9.4: Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$ be matrices of the same size and let α, β be scalars. Then

(a) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$

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

(c) $\mathbf{A} + \mathbf{0} = \mathbf{A}$

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

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

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

9.2 Matrix Multiplication

Let $\mathsf{T}_\mathbf{B} : \mathbb{R}^p \to \mathbb{R}^n$ and let $\mathsf{T}_\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^m$ be linear mappings. If $\mathbf{x} \in \mathbb{R}^p$ then $\mathsf{T}_\mathbf{B}(\mathbf{x}) \in \mathbb{R}^n$ and thus we can apply $\mathsf{T}_\mathbf{A}$ to $\mathsf{T}_\mathbf{B}(\mathbf{x})$. The resulting vector $\mathsf{T}_\mathbf{A}(\mathsf{T}_\mathbf{B}(\mathbf{x}))$ is in $\mathbb{R}^m$. Hence, each $\mathbf{x} \in \mathbb{R}^p$ can be mapped to a point in $\mathbb{R}^m$, and because $\mathsf{T}_\mathbf{B}$ and $\mathsf{T}_\mathbf{A}$ are linear mappings the resulting mapping is also linear. This resulting mapping is called the **composition** of $\mathsf{T}_\mathbf{A}$ and $\mathsf{T}_\mathbf{B}$, and is usually denoted by $\mathsf{T}_\mathbf{A} \circ \mathsf{T}_\mathbf{B} : \mathbb{R}^p \to \mathbb{R}^m$ (see Figure 9.1). Hence,

$$(\mathsf{T}_\mathbf{A} \circ \mathsf{T}_\mathbf{B})(\mathbf{x}) = \mathsf{T}_\mathbf{A}(\mathsf{T}_\mathbf{B}(\mathbf{x})).$$

Because $(\mathsf{T}_\mathbf{A} \circ \mathsf{T}_\mathbf{B}) : \mathbb{R}^p \to \mathbb{R}^m$ is a linear mapping it has an associated standard matrix, which we denote for now by $\mathbf{C}$. From Lecture 8, to compute the standard matrix of any linear mapping, we must compute the images of the standard unit vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p$ under the linear mapping. Now, for any $\mathbf{x} \in \mathbb{R}^p$,

$$\mathsf{T}_\mathbf{A}(\mathsf{T}_\mathbf{B}(\mathbf{x})) = \mathsf{T}_\mathbf{A}(\mathbf{B}\mathbf{x}) = \mathbf{A}(\mathbf{B}\mathbf{x}).$$

Applying this to $\mathbf{x} = \mathbf{e}_i$ for all $i = 1, 2, \ldots, p$, we obtain the standard matrix of $\mathsf{T}_\mathbf{A} \circ \mathsf{T}_\mathbf{B}$:

$$\mathbf{C} = \begin{bmatrix} \mathbf{A}(\mathbf{B}\mathbf{e}_1) & \mathbf{A}(\mathbf{B}\mathbf{e}_2) & \cdots & \mathbf{A}(\mathbf{B}\mathbf{e}_p) \end{bmatrix}.$$

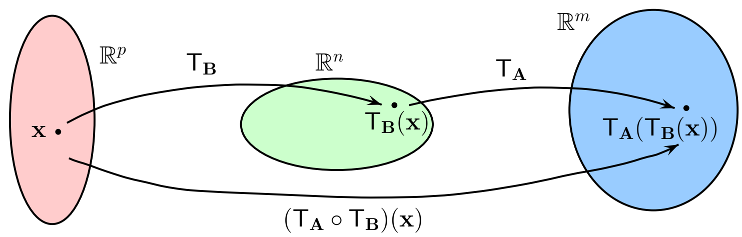


Figure 9.1: Illustration of the composition of two mappings.

Now $\mathbf{B}\mathbf{e}_1$ is

$$\mathbf{B}\mathbf{e}_1 = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{bmatrix} \mathbf{e}_1 = \mathbf{b}_1.$$

And similarly $\mathbf{B}\mathbf{e}_i = \mathbf{b}_i$ for all $i = 1, 2, \ldots, p$. Therefore,

$$\mathbf{C} = \begin{bmatrix} \mathbf{A}\mathbf{b}_1 & \mathbf{A}\mathbf{b}_2 & \cdots & \mathbf{A}\mathbf{b}_p \end{bmatrix}$$

is the standard matrix of $\mathsf{T}_\mathbf{A} \circ \mathsf{T}_\mathbf{B}$. This computation motivates the following definition.

Definition 9.5: For $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$, with $\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \cdots & \mathbf{b}_p \end{bmatrix}$, we define the product $\mathbf{AB}$ by the formula

$$\mathbf{AB} = \begin{bmatrix} \mathbf{A}\mathbf{b}_1 & \mathbf{A}\mathbf{b}_2 & \cdots & \mathbf{A}\mathbf{b}_p \end{bmatrix}.$$

The product $\mathbf{AB}$ is defined only when the number of columns of $\mathbf{A}$ equals the number of rows of $\mathbf{B}$. The following diagram is useful for remembering this:

$$(m \times n) \cdot (n \times p) \to m \times p$$

From our definition of $\mathbf{AB}$, the standard matrix of the composite mapping $\mathsf{T}_\mathbf{A} \circ \mathsf{T}_\mathbf{B}$ is

$$\mathbf{C} = \mathbf{AB}.$$

In other words, composition of linear mappings corresponds to matrix multiplication.

Example 9.6. For $\mathbf{A}$ and $\mathbf{B}$ below compute $\mathbf{AB}$ and $\mathbf{BA}$.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & -3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -4 & 2 & 4 & -4 \\ -1 & -5 & -3 & 3 \\ -4 & -4 & -3 & -1 \end{bmatrix}$$

Solution. First $\mathbf{AB} = [\mathbf{Ab}_1\ \ \mathbf{Ab}_2\ \ \mathbf{Ab}_3\ \ \mathbf{Ab}_4]$:

$$\begin{aligned}
\mathbf{AB} &= \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} -4 & 2 & 4 & -4 \\ -1 & -5 & -3 & 3 \\ -4 & -4 & -3 & -1 \end{bmatrix} \\
&= \left[\begin{matrix} 2 \\ 7 \end{matrix} \right. \\
&= \left[\begin{matrix} 2 & 0 \\ 7 & 9 \end{matrix} \right. \\
&= \left[\begin{matrix} 2 & 0 & 4 \\ 7 & 9 & 10 \end{matrix} \right. \\
&= \begin{bmatrix} 2 & 0 & 4 & 4 \\ 7 & 9 & 10 & 2 \end{bmatrix}
\end{aligned}$$

On the other hand, $\mathbf{BA}$ is not defined! $\mathbf{B}$ has 4 columns and $\mathbf{A}$ has 2 rows. □

Example 9.7. For $\mathbf{A}$ and $\mathbf{B}$ below compute $\mathbf{AB}$ and $\mathbf{BA}$.

$$\mathbf{A} = \begin{bmatrix} -4 & 4 & 3 \\ 3 & -3 & -1 \\ -2 & -1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -1 & -1 & 0 \\ -3 & 0 & -2 \\ -2 & 1 & -2 \end{bmatrix}$$

Solution. First $\mathbf{AB} = [\mathbf{Ab}_1\ \ \mathbf{Ab}_2\ \ \mathbf{Ab}_3]$:

$$\begin{aligned}
\mathbf{AB} &= \begin{bmatrix} -4 & 4 & 3 \\ 3 & -3 & -1 \\ -2 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & 0 \\ -3 & 0 & -2 \\ -2 & 1 & -2 \end{bmatrix} \\
&= \left[\begin{matrix} -14 \\ 8 \\ 3 \end{matrix} \right. \\
&= \left[\begin{matrix} -14 & 7 \\ 8 & -4 \\ 3 & 3 \end{matrix} \right. \\
&= \begin{bmatrix} -14 & 7 & -14 \\ 8 & -4 & 8 \\ 3 & 3 & 0 \end{bmatrix}
\end{aligned}$$

Next $\mathbf{BA} = [\mathbf{Ba}_1 \;\; \mathbf{Ba}_2 \;\; \mathbf{Ba}_3]$:

$$\begin{aligned}
\mathbf{BA} &= \begin{bmatrix} -1 & -1 & 0 \\ -3 & 0 & -2 \\ -2 & 1 & -2 \end{bmatrix} \begin{bmatrix} -4 & 4 & 3 \\ 3 & -3 & -1 \\ -2 & -1 & 1 \end{bmatrix} \\
&= \left[\begin{matrix} 1 \\ 16 \\ 15 \end{matrix}\right. \\
&= \left[\begin{matrix} 1 & -1 \\ 16 & -10 \\ 15 & -9 \end{matrix}\right. \\
&= \begin{bmatrix} 1 & -1 & -2 \\ 16 & -10 & -11 \\ 15 & -9 & -9 \end{bmatrix}
\end{aligned}$$

On the other hand:

$$\mathbf{AB} = \begin{bmatrix} -14 & 7 & -14 \\ 8 & -4 & 8 \\ 3 & 3 & 0 \end{bmatrix}$$

Therefore, in general $\mathbf{AB} \neq \mathbf{BA}$, i.e., matrix multiplication is not commutative. $\square$

An important matrix that arises frequently is the **identity matrix $\mathbf{I}_n \in \mathbb{R}^{n\times n}$ of size** n:

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

You should verify that for any $\mathbf{A} \in \mathbb{R}^{n\times n}$ it holds that $\mathbf{AI}_n = \mathbf{I}_n\mathbf{A} = \mathbf{A}$. Below are some basic algebraic properties of matrix multiplication.

Theorem 9.8: Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$ be matrices, of appropriate dimensions, and let α be a scalar. Then

(1) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$

(2) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC}$

(3) $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$

(4) $\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$

(5) $\mathbf{I}_n\mathbf{A} = \mathbf{AI}_n = \mathbf{A}$

If $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a square matrix, the kth power of $\mathbf{A}$ is

$$\mathbf{A}^k = \underbrace{\mathbf{AAA}\cdots\mathbf{A}}_{k \text{ times}}$$

Example 9.9. Compute $\mathbf{A}^3$ if

$$\mathbf{A} = \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix}.$$

Solution. Compute $\mathbf{A}^2$:

$$\mathbf{A}^2 = \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 7 & -6 \\ -2 & 3 \end{bmatrix}$$

And then $\mathbf{A}^3$:

$$\mathbf{A}^3 = \mathbf{A}^2\mathbf{A} = \begin{bmatrix} 7 & -6 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -20 & 21 \\ 7 & -6 \end{bmatrix}$$

We could also do:

$$\mathbf{A}^3 = \mathbf{A}\mathbf{A}^2 = \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 7 & -6 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} -20 & 21 \\ 7 & -6 \end{bmatrix}.$$

□

9.3 Matrix Transpose

We begin with the definition of the transpose of a matrix.

Definition 9.10: Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, the **transpose** of $\mathbf{A}$ is the matrix $\mathbf{A}^T$ whose ith column is the ith row of $\mathbf{A}$.

If $\mathbf{A}$ is $m \times n$ then $\mathbf{A}^T$ is $n \times m$. For example, if

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 8 & -7 & -4 \\ -4 & 6 & -10 & -9 & 6 \\ 9 & 5 & -2 & -3 & 5 \\ -8 & 8 & 4 & 7 & 7 \end{bmatrix}$$

then

$$\mathbf{A}^T = \begin{bmatrix} 0 & -4 & 9 & -8 \\ -1 & 6 & 5 & 8 \\ 8 & -10 & -2 & 4 \\ -7 & -9 & -3 & 7 \\ -4 & 6 & 5 & 7 \end{bmatrix}.$$

Example 9.11. Compute $(\mathbf{AB})^T$ and $\mathbf{B}^T\mathbf{A}^T$ if

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & 0 \\ 3 & -1 & -3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -2 & 1 & 2 \\ -1 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Solution. Compute $\mathbf{AB}$:

$$\begin{aligned}\mathbf{AB} &= \begin{bmatrix} -2 & 1 & 0 \\ 3 & -1 & -3 \end{bmatrix} \begin{bmatrix} -2 & 1 & 2 \\ -1 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 3 & -4 & -4 \\ -5 & 5 & 9 \end{bmatrix}\end{aligned}$$

Next compute $\mathbf{B}^T\mathbf{A}^T$:

$$\begin{aligned}\mathbf{B}^T\mathbf{A}^T &= \begin{bmatrix} -2 & -1 & 0 \\ 1 & -2 & 0 \\ 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 1 & -1 \\ 0 & -3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & -5 \\ -4 & 5 \\ -4 & 9 \end{bmatrix} = (\mathbf{AB})^T\end{aligned}$$

□

The following theorem summarizes properties of the transpose.

Theorem 9.12: Let $\mathbf{A}$ and $\mathbf{B}$ be matrices of appropriate sizes. The following hold:
(1) $(\mathbf{A}^T)^T = \mathbf{A}$
(2) $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
(3) $(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T$
(4) $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

A consequence of property (4) is that

$$(\mathbf{A}_1\mathbf{A}_2 \ldots \mathbf{A}_k)^T = \mathbf{A}_k^T\mathbf{A}_{k-1}^T \cdots \mathbf{A}_2^T\mathbf{A}_1^T$$

and as a special case

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

Example 9.13. Let $\mathsf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear mapping that first contracts vectors by a factor of $k = 3$ and then rotates by an angle θ. What is the standard matrix $\mathbf{A}$ of T?

Solution. Let $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ denote the standard unit vectors in $\mathbb{R}^2$. From Lecture 8, the standard matrix of T is $\mathbf{A} = \begin{bmatrix}\mathsf{T}(\mathbf{e}_1) & \mathsf{T}(\mathbf{e}_2)\end{bmatrix}$. Recall that the standard matrix of a rotation by θ is

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Contracting $\mathbf{e}_1$ by a factor of $k = 3$ results in $(\frac{1}{3}, 0)$ and then rotation by θ results in

$$\begin{bmatrix} \frac{1}{3}\cos(\theta) \\ \frac{1}{3}\sin(\theta) \end{bmatrix} = \mathsf{T}(\mathbf{e}_1).$$

Contracting $\mathbf{e}_2$ by a factor of $k = 3$ results in $(0, \frac{1}{3})$ and then rotation by θ results in

$$\begin{bmatrix} -\frac{1}{3}\sin(\theta) \\ \frac{1}{3}\cos(\theta) \end{bmatrix} = \mathsf{T}(\mathbf{e}_2).$$

Therefore,

$$\mathbf{A} = \begin{bmatrix} \mathsf{T}(\mathbf{e}_1) & \mathsf{T}(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} \frac{1}{3}\cos(\theta) & -\frac{1}{3}\sin(\theta) \\ \frac{1}{3}\sin(\theta) & \frac{1}{3}\cos(\theta) \end{bmatrix}$$

On the other hand, the standard matrix corresponding to a contraction by a factor $k = \frac{1}{3}$ is

$$\begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

Therefore,

$$\underbrace{\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}}_{\text{rotation}} \underbrace{\begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}}_{\text{contraction}} = \begin{bmatrix} \frac{1}{3}\cos(\theta) & -\frac{1}{3}\sin(\theta) \\ \frac{1}{3}\sin(\theta) & \frac{1}{3}\cos(\theta) \end{bmatrix} = \mathbf{A}$$

□

After this lecture you should know the following:

- know how to add and multiply matrices
- that matrix multiplication corresponds to composition of linear mappings
- the algebraic properties of matrix multiplication (Theorem 9.8)
- how to compute the transpose of a matrix
- the properties of matrix transposition (Theorem 9.12)