23. Session: Linear Transformations and Matrices

Recall: $\mathbb{R}^n$ is the set of vectors with n coordinates.

Def. A function or transformation f from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule, that assigns to each vector $\vec{x}$ in $\mathbb{R}^n$ a unique vector $f(\vec{x})$ in $\mathbb{R}^m$.

Ex. $f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 3x_2 + 7x_3 \\ x_1^2 + 2x_3 \end{bmatrix}$

transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ $\quad$ ($\mathbb{R}^3 \to \mathbb{R}^2$).

Ex. $f(x) = 2x^2 + 5x - 3$, $\quad \mathbb{R} \to \mathbb{R}$

Def. Let A be an $m \times n$-matrix. Then the transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ given by

$$T_A(\vec{x}) = A \cdot \vec{x}$$

is called the matrix transformation induced by A.

Ex.

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & -1 \end{bmatrix} \qquad 3 \times 2\text{-matrix}$$

The matrix transformation induced by A:

$$T_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 + x_2 \\ x_1 - x_2 \end{bmatrix}, \quad \mathbb{R}^2 \to \mathbb{R}^3$$

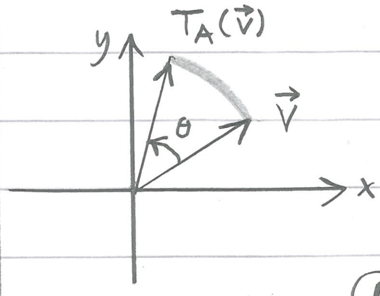
An example:

$$T_A\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}.$$

Ex. $A = R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

2×2-rotation matrix

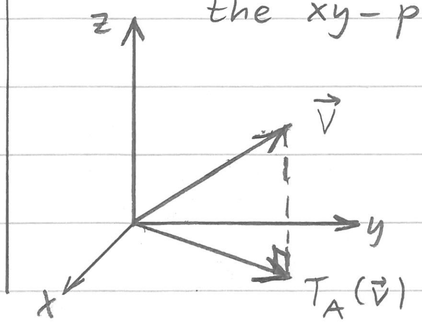
$$T_A\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbb{R}^2 \to \mathbb{R}^2$$



Ex. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$$T_A\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

Ortogonal projection on the xy-plane



! Def. A transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is said to be linear if

(1) $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$,

(2) $T(c\vec{u}) = c\,T(\vec{u})$

for all vectors $\vec{u}$, $\vec{v}$ in $\mathbb{R}^n$ and all scalars c.

Theorem: Matrix transformations are linear.

Proof:

(1) $T_A(\vec{u}+\vec{v}) = A\cdot(\vec{u}+\vec{v}) = A\cdot\vec{u} + A\cdot\vec{v} = T_A(\vec{u}) + T_A(\vec{v})$ OK

(2) $T_A(c\vec{u}) = A\cdot(c\vec{u}) = c(A\cdot\vec{u}) = c\,T_A(\vec{u})$ OK q.e.d.

Theorem: If T is a linear transformation, then

(1) $T(\vec{0}) = \vec{0}$

(2) $T(a\vec{u} + b\vec{v}) = a\,T(\vec{u}) + b\,T(\vec{v})$.

Proof:

(1) $T(\vec{0}) = T(0\cdot\vec{u}) = 0\cdot T(\vec{u}) = \vec{0}$ OK

(2) $T(a\vec{u} + b\vec{v}) = T(a\vec{u}) + T(b\vec{v}) = a\,T(\vec{u}) + b\,T(\vec{v})$ OK

q.e.d.

Ex. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+2 \\ y+1 \end{bmatrix}$ is not linear since $T\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Question: How can one see from the matrix what a matrix transformation will do?

Remark: A linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is determined by the images of the standard basis vectors

$$T(\vec{e}_1),\ T(\vec{e}_2),\ \dots,\ T(\vec{e}_n).$$

Ex. T linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^m$.

$\begin{bmatrix} x \\ y \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \end{bmatrix} = x\vec{e}_1 + y\vec{e}_2$,

so we have

$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = T(x\vec{e}_1 + y\vec{e}_2) = x\,T(\vec{e}_1) + y\,T(\vec{e}_2)$.

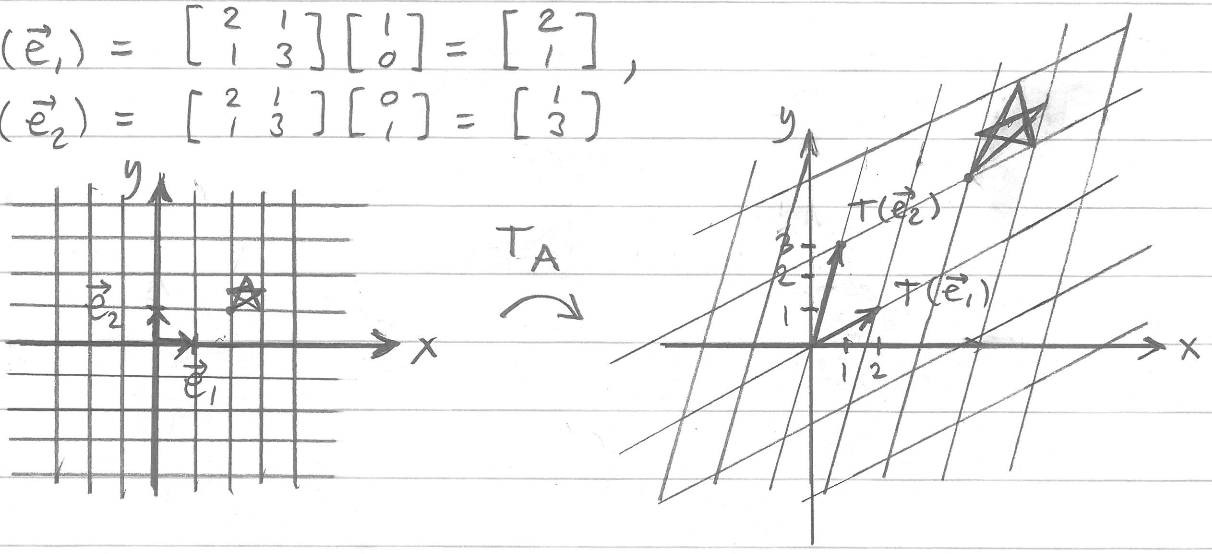
Thus, the two vectors $T(\vec{e}_1)$, $T(\vec{e}_2)$ determine the transformation.

Answer to question: "Plot the column vectors of the matrix, and form a grid"

Ex. $A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$

$T_A(\vec{e}_1) = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$,

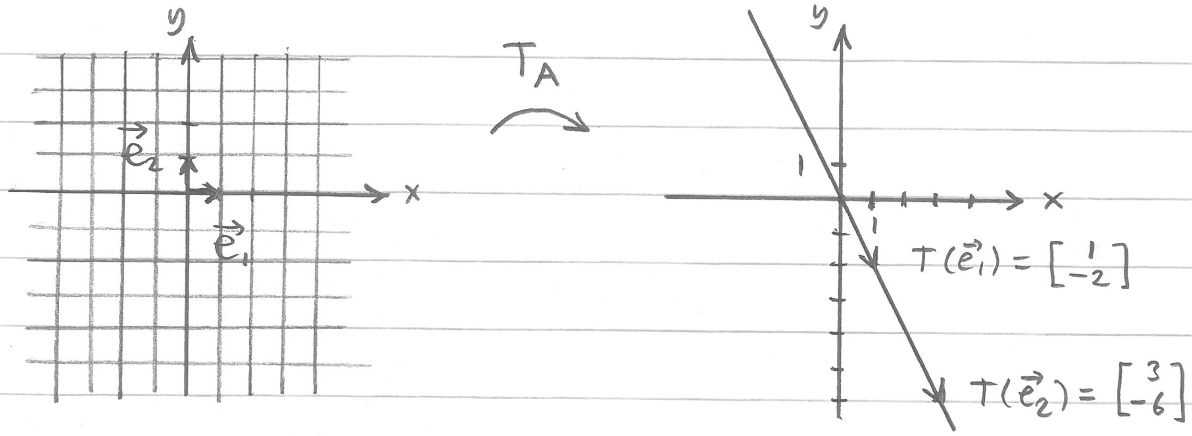
$T_A(\vec{e}_2) = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$



Ex. The non-invertible 2×2-matrix from last Lecture:

$$A = \begin{bmatrix} 1 & 3 \\ -2 & -6 \end{bmatrix}$$

$T_A(\vec{e}_1) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $T_A(\vec{e}_2) = \begin{bmatrix} 3 \\ -6 \end{bmatrix}$. Parallel vectors.



<u>Theorem</u>: Let T be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. Then T is the matrix transformation induced by the $m \times n$-matrix

$$A = [\,T(\vec{e}_1) \;\; T(\vec{e}_2) \;\; \dots \;\; T(\vec{e}_n)\,].$$

Thus, $T(\vec{x}) = A \cdot \vec{x}$. The matrix A is called the <u>standard matrix</u> of T.

<u>Proof</u>.

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n .$$

$$\begin{aligned}
T(\vec{x}) &= T(x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n) \\
&= T(x_1\vec{e}_1) + T(x_2\vec{e}_2) + \dots + T(x_n\vec{e}_n) \\
&= x_1 T(\vec{e}_1) + x_2 T(\vec{e}_2) + \dots + x_n T(\vec{e}_n) \\
&= [\,T(\vec{e}_1) \;\; T(\vec{e}_2) \;\; \dots \;\; T(\vec{e}_n)\,]\,\vec{x} = A\vec{x}
\end{aligned}$$

q.e.d.

<u>Theorem</u>: If A is an $m \times n$-matrix and B an $n \times p$-matrix, then

$$T_A \circ T_B = T_{A \cdot B}$$

If C is an invertible $n \times n$-matrix, then

$$T_C^{-1} = T_{C^{-1}} .$$