Lecture 4 : Linear transformations, Matrix algebra
(book: 1.8, 1.9, 2.1).

Previous lecture: homogeneous / nonhomogeneous SLE
+ linear independence.

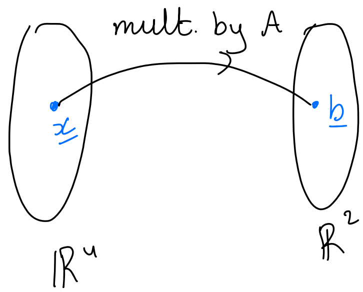
Recall the matrix-vector product:

$$\underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}}_{\underline{x}} = \underbrace{\begin{bmatrix} 13 \\ 21 \end{bmatrix}}_{\underline{b}}$$

And recall the properties:
* $A(\underline{u}+\underline{v}) = A\underline{u} + A\underline{v}$
* $A(c \cdot \underline{u}) = c \cdot (A\underline{u})$.

Multiplication by A transforms $\underline{x}$ into $\underline{b}$.

Schematic



Transformation / function / mapping

$T : \underbrace{\mathbb{R}^m}_{\text{domain}} \to \underbrace{\mathbb{R}^n}_{\text{codomain}}$.

$T(\underline{x}) = \underline{y}$ ⇐ output

the transformation operator. (T)

input ($\underline{x}$)

image : $T(\underline{x})$

range: set of all images.

A transformation is linear if:
* $T(\underline{u}+\underline{v}) = T(\underline{u}) + T(\underline{v})$
* $T(c \cdot \underline{u}) = c \cdot T(\underline{u})$.

} $T(c \cdot \underline{u} + d \cdot \underline{v}) = c \cdot T(\underline{u}) + d \cdot T(\underline{v})$.

If T is linear, then $T(\underline{0}) = \underline{0}$.

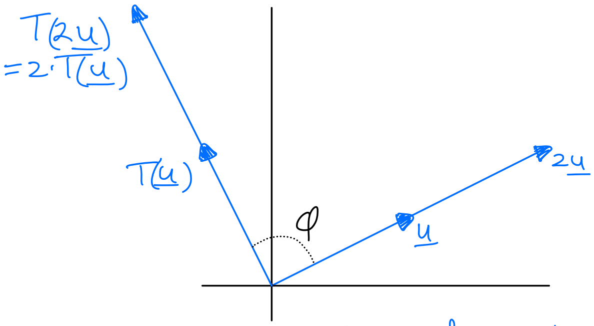
$T : \mathbb{R}^2 \to \mathbb{R}^2$ $\qquad T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x^2 \\ y^2 \end{bmatrix}$.

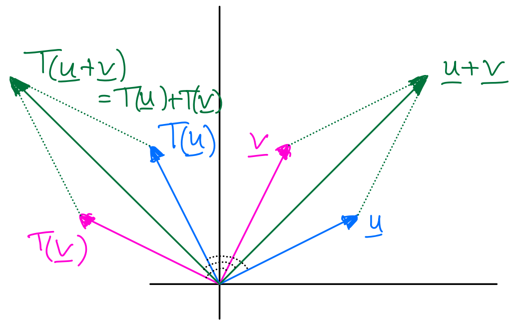
$\underline{u} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ $\quad \underline{v} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$ $\quad T(\underline{u}) = \begin{bmatrix} 9 \\ 9 \end{bmatrix}$ $\quad T(\underline{v}) = \begin{bmatrix} 16 \\ 16 \end{bmatrix}$ $\quad T(\underline{u}) + T(\underline{v}) = \begin{bmatrix} 25 \\ 25 \end{bmatrix}$

$T(\underline{u}+\underline{v}) = T\left(\begin{bmatrix} 7 \\ 7 \end{bmatrix}\right) = \begin{bmatrix} 49 \\ 49 \end{bmatrix}$ $\quad \neq$

So, T is not linear.

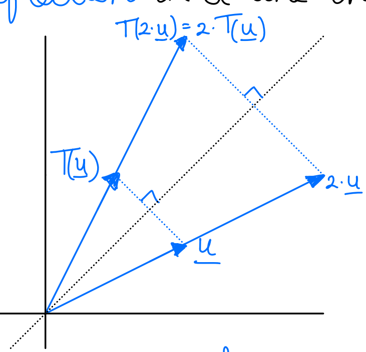
Example: rotation about the origin through an angle φ

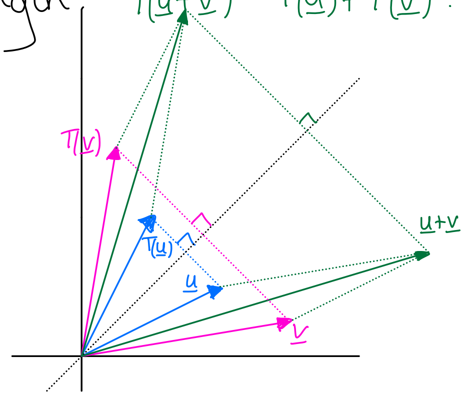




So, it is a linear transformation.

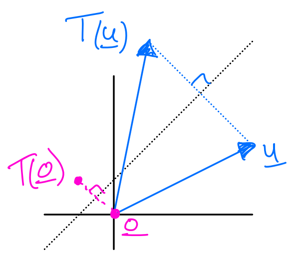
Example: reflection in a line through the origin. $T(\underline{u}+\underline{v}) = T(\underline{u})+T(\underline{v})$.





So, it is a linear transformation.

Example: reflection in a line not through the origin



not a linear transformation
because $T(\underline{o}) \neq \underline{o}$.

Let's go back to the transformation of a matrix-vector product.

$$\underset{m\times n}{A}\,\underset{n\times 1}{\underline{u}} = \underset{m\times 1}{\underline{b}} \qquad \mathbb{R}^n \to \mathbb{R}^m \qquad T(\underline{u}) = A\underline{u}.$$

Is a matrix transformation linear?

$* \; T(\underline{u}+\underline{v}) = A(\underline{u}+\underline{v}) = A\underline{u}+A\underline{v} = T(\underline{u})+T(\underline{v})$

$* \; T(c\cdot\underline{u}) = A(c\cdot\underline{u}) = c\cdot(A\underline{u}) = c\cdot T(\underline{u})$.

follow from the properties of a matrix-vector product.

$\Rightarrow$ Every matrix transformation is a linear transformation.
The opposite is also true (at least, in the context: $\mathbb{R}^n \to \mathbb{R}^m$)

Theorem: Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. There is a unique matrix A such that for $\underline{x} \in \mathbb{R}^n$ $T(\underline{x}) = A\underline{x}$.

Proof: Let $\underline{x} \in \mathbb{R}^n$.

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = x_1 \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\underline{e}_1} + x_2 \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\underline{e}_2} + x_3 \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}}_{\underline{e}_3} + \dots + x_n \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}}_{\underline{e}_n}.$$

Then, $T(\underline{x}) = T(x_1 \cdot \underline{e}_1 + x_2 \cdot \underline{e}_2 + x_3 \cdot \underline{e}_3 + \dots + x_n \cdot \underline{e}_n)$

$$\overset{\text{linearity}}{=} T(x_1 \underline{e}_1) + T(x_2 \underline{e}_2) + T(x_3 \underline{e}_3) + \dots T(x_n \cdot \underline{e}_n)$$

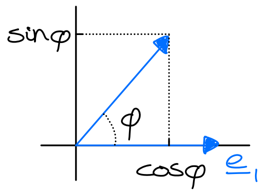
$$\overset{\text{linearity}}{=} x_1 T(\underline{e}_1) + x_2 T(\underline{e}_2) + x_3 T(\underline{e}_3) + \dots + x_n T(\underline{e}_n).$$

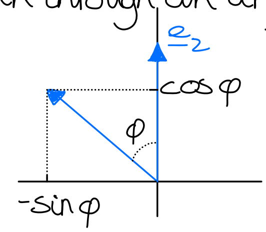
$$= \underbrace{\begin{bmatrix} | & | & | & & | \\ T(\underline{e}_1) & T(\underline{e}_2) & T(\underline{e}_3) & \dots & T(\underline{e}_n) \\ | & | & | & & | \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_{\underline{x}} = A\underline{x} \qquad \square$$

Uniqueness of A? DIY (exc. 41 Ch 1.9).

Standard matrix for the linear transformation T: $[T(\underline{e}_1) \dots T(\underline{e}_n)]$

Example: rotation about the origin through an angle φ $\quad T: \mathbb{R}^2 \to \mathbb{R}^2$





So, $T(\underline{e}_1) = \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}$ $\qquad$ So, $T(\underline{e}_2) = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix}$.

So, $A = [T(\underline{e}_1) \;\; T(\underline{e}_2)] = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$

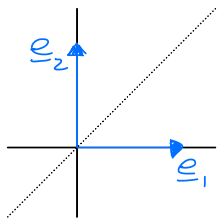
Now it's easy to get the image of $\underline{x} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, namely $T\left(\begin{bmatrix} 2 \\ -3 \end{bmatrix}\right) = \begin{bmatrix} 2\cos\varphi + 3\sin\varphi \\ 2\sin\varphi - 3\cos\varphi \end{bmatrix}$

Example: Suppose the standard matrix is $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

What is the geometric interpretation?

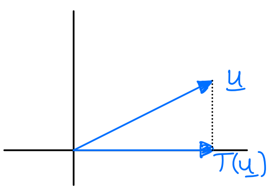
$T(\underline{e}_1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \underline{e}_2$

$T(\underline{e}_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \underline{e}_1$



Hence, the transformation is a reflection in the line $y = x$.

Example: Projection onto the x_1-axis



This is a linear transformation (DIY!) with standard matrix

$$A = \left[T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) \;\; T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)\right] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

So, indeed $T\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

Surjectivity: A transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is surjective / onto if each $\underline{b} \in \mathbb{R}^m$ is the image of at least one $\underline{x} \in \mathbb{R}^n$ (range = codomain).

Injectivity: A transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is injective / one-to-one if each $\underline{b} \in \mathbb{R}^m$ is the image of at most one $\underline{x} \in \mathbb{R}^n$.

Theorem: Let $\mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.
T is injective $\iff$ $T(\underline{x}) = \underline{0}$ has only the trivial solution.

Proof:
$(\Longrightarrow)$ Assume T is injective.
Since T is linear, we have $T(\underline{0}) = \underline{0}$.
Since T is injective, $T(\underline{x}) = \underline{0}$ has at most one solution.
So, $T(\underline{x}) = \underline{0}$ has only the trivial solution.
$(\Longleftarrow)$ By contrapositive.
Assume T is not injective.
So, there is a $\underline{b} \in \mathbb{R}^m$ that is the image of at least two vectors in $\mathbb{R}^n$.
So, $T(\underline{u}) = \underline{b}$ and $T(\underline{v}) = \underline{b}$ with $\underline{u} \neq \underline{v}$.
So, $T(\underline{u} - \underline{v}) \underset{\substack{\text{because } T \\ \text{is linear}}}{=} T(\underline{u}) - T(\underline{v}) = \underline{b} - \underline{b} = \underline{0}$

Note $\underline{u} - \underline{v} \neq \underline{0}$ because $\underline{u} \neq \underline{v}$.
Hence, $T(\underline{x}) = \underline{0}$ has also a nontrivial solution. $\square$

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with standard matrix A.

T is injective
$\iff$ $T(\underline{x}) = \underline{0}$ has only the trivial solution.
$\iff$ $A\underline{x} = \underline{0}$ has only the trivial solution.
$\iff$ A has a pivot in every column.

T is surjective
$\iff$ for each $\underline{b} \in \mathbb{R}^m$, $T(\underline{x}) = \underline{b}$ has a solution.
$\iff$ for each $\underline{b} \in \mathbb{R}^m$ $A\underline{x} = \underline{b}$ has a solution.
$\iff$ A has a pivot in every row.

Matrix Algebra

* equality: same size and corresponding entries are equal.
* addition: $\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}+\begin{bmatrix}0&1\\1&0\\2&-1\end{bmatrix}=\begin{bmatrix}1&3\\4&4\\7&5\end{bmatrix}$ Note: for $A+B$ to be defined, we need size A = size B.
* scaling: $2\cdot\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}=\begin{bmatrix}2&4\\6&8\\10&12\end{bmatrix}$
* matrix-matrix product: it works as a sequence of matrix-vector products.

$$\underbrace{\begin{bmatrix}2&3\\1&-5\end{bmatrix}}_{A}\begin{bmatrix}\underset{\underline{b}_1}{\begin{matrix}4\\1\end{matrix}}&\underset{\underline{b}_2}{\begin{matrix}3\\-2\end{matrix}}&\underset{\underline{b}_3}{\begin{matrix}6\\3\end{matrix}}\end{bmatrix}=\begin{bmatrix}A\underline{b}_1 & A\underline{b}_2 & A\underline{b}_3\end{bmatrix}=\begin{bmatrix}11&0&21\\-1&13&-9\end{bmatrix}$$

Each column of AB is a linear combination of the columns of A with the entries of the corresponding column of B being the weights.

$$\begin{bmatrix}11\\-1\end{bmatrix}=4\cdot\begin{bmatrix}2\\1\end{bmatrix}+1\cdot\begin{bmatrix}3\\-5\end{bmatrix}$$

Note: we need # columns of A = # rows of B.

$$\left.\begin{matrix}A: & m\times n\\ B: & n\times p\end{matrix}\right\}\quad C=AB \quad m\times p$$

In general $AB\neq BA$.

Transpose: $\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}^T=\begin{bmatrix}1&4\\2&5\\3&6\end{bmatrix}$

A matrix is symmetric if $A^T=A$ $\quad\begin{bmatrix}1&1\\1&1\end{bmatrix}\quad\begin{bmatrix}1&4&5\\4&2&6\\5&6&3\end{bmatrix}$

identity matrix $\quad I_2=\begin{bmatrix}1&0\\0&1\end{bmatrix}\qquad I_4=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$

$AI_4=A \qquad I_4A=A$

Power of a square $(n\times n)$ matrix $\quad A^k=\underbrace{A\cdot A\cdot A\cdot\ldots\cdot A}_{k\text{ times}}.$

$A^0=I_n.$

Composition of linear transformation.

$T_1 : \mathbb{R}^n \to \mathbb{R}^m$ with standard matrix A $m \times n$
$T_2 \quad \mathbb{R}^m \to \mathbb{R}^q$ $\qquad B$ $q \times m$

$$T = T_2 \circ T_1 = T_2(T_1)$$

Then, $T : \mathbb{R}^n \to \mathbb{R}^q$ with standard matrix C $q \times n$
where $C = BA$

Let A, B, and C be matrices of the same size, and let r and s be scalars.

a. $A + B = B + A$
b. $(A + B) + C = A + (B + C)$
c. $A + 0 = A$
d. $r(A + B) = rA + rB$
e. $(r + s)A = rA + sA$
f. $r(sA) = (rs)A$

Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

a. $A(BC) = (AB)C$ (associative law of multiplication)
b. $A(B + C) = AB + AC$ (left distributive law)
c. $(B + C)A = BA + CA$ (right distributive law)
d. $r(AB) = (rA)B = A(rB)$ for any scalar r
e. $I_m A = A = A I_n$ (identity for matrix multiplication)

WARNINGS:

1. In general, $AB \neq BA$.

2. The cancellation laws do *not* hold for matrix multiplication. That is, if $AB = AC$, then it is *not* true in general that $B = C$. (See Exercise 10.)

3. If a product AB is the zero matrix, you *cannot* conclude in general that either $A = 0$ or $B = 0$. (See Exercise 12.)

Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^T)^T = A$
b. $(A + B)^T = A^T + B^T$
c. For any scalar r, $(rA)^T = rA^T$
d. $(AB)^T = B^T A^T$

The transpose of a product of matrices equals the product of their transposes in the *reverse* order.