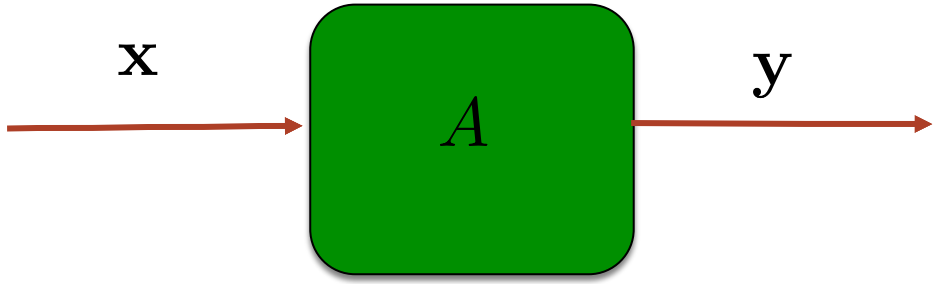
$$\mathbf{y} = A\mathbf{x}$$



$\mathbf{x} \in \mathcal{R}^n$ $\qquad$ $\mathbf{y} \in \mathcal{R}^m$

$$A \in \mathcal{M}_{m \times n}$$

An *m* x *n* **matrix** can be viewed as a **function** that maps an *n*-**component vector** **x** to an *m*-**component vector** **y**.

This type of **functions** is associated with the concept of **linear transformations**, one of the two major concepts in the course of linear algebra.

Set

Set of Vectors

Vector Space

Function

A mapping from one set to another set

Matrix

A mapping of one vector to another vector (with a matrix-vector product).

Linear Transformation

A mapping from one set of vectors to another set of vectors

2.7 Linear Transformations and Matrices

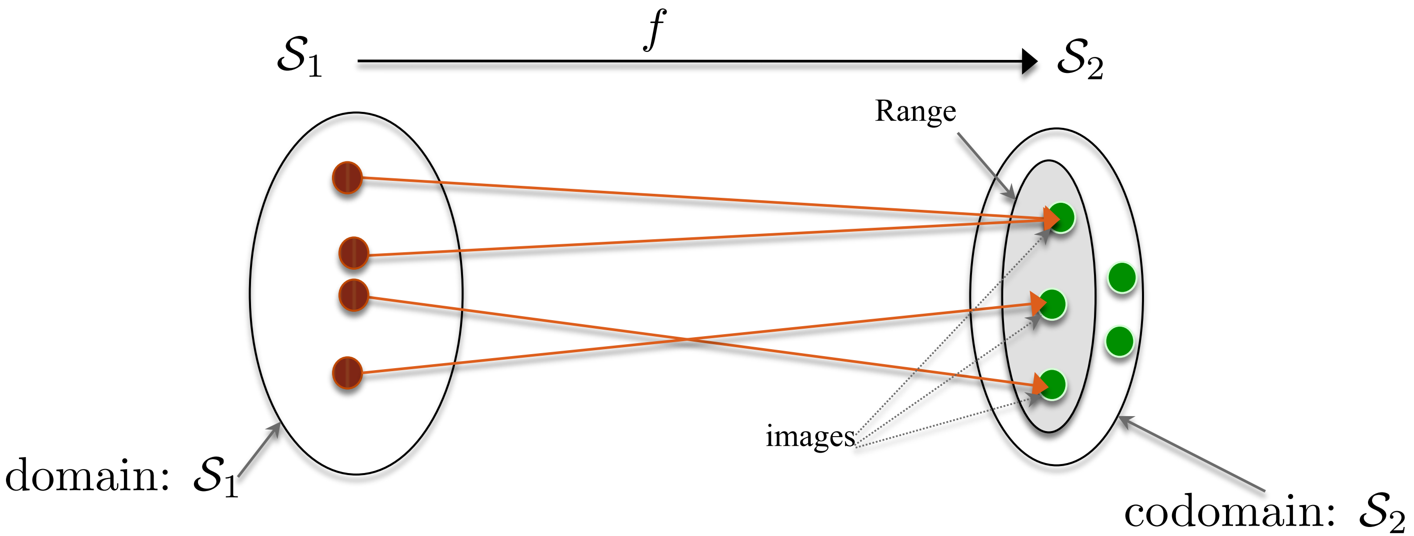
Definitions (Function)

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two sets. A **function** f from $\mathcal{S}_1$ to $\mathcal{S}_2$, written

$$f : \mathcal{S}_1 \to \mathcal{S}_2,$$

is a mapping that assigns an element in $\mathcal{S}_2$ to **each** element in $\mathcal{S}_1$.
For any element $a \in \mathcal{S}_1$, there exists a unique element $b \in \mathcal{S}_2$ such that $f(a) = b$.
This is called the **image** of a.

The set $\mathcal{S}_1$ is called the **domain** of the function f, and the set $\mathcal{S}_2$ is called the **codomain** of f.
The vector $f(\mathbf{v})$ is called the **image** of $\mathbf{v}$ (under f).
The **range** of f is the subset of $\mathcal{S}_2$ that contains **all** images: $\{f(a) : a \in \mathcal{S}_1\}$



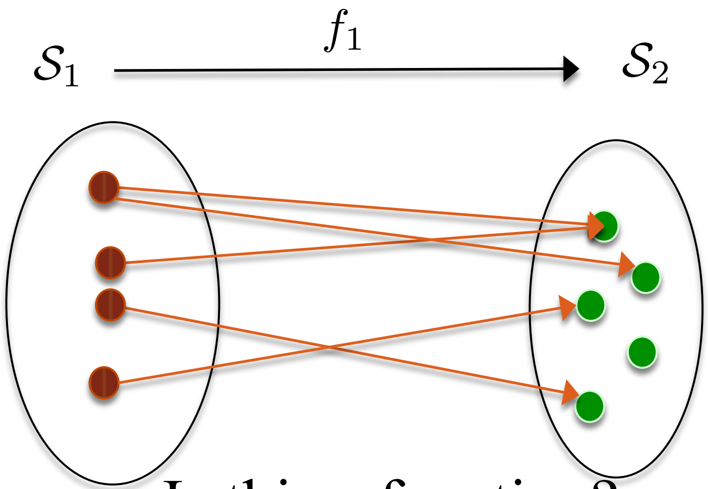
Definitions (Function)

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two sets. A **function** f from $\mathcal{S}_1$ to $\mathcal{S}_2$, written

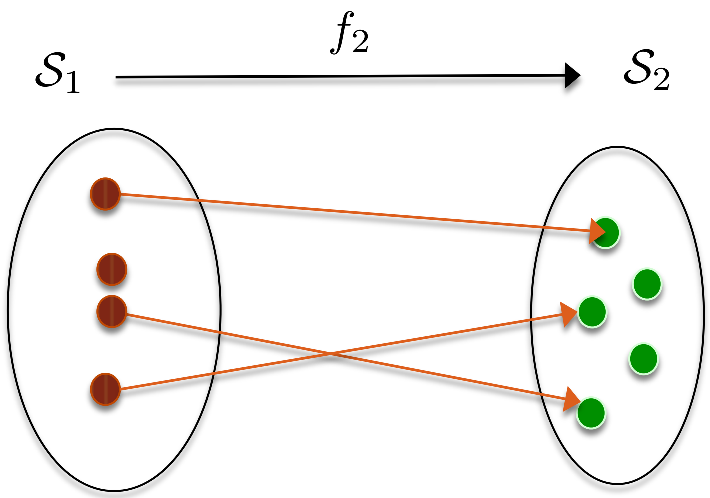
$$f : \mathcal{S}_1 \to \mathcal{S}_2,$$

is a mapping that assigns an element in $\mathcal{S}_2$ to **each** element in $\mathcal{S}_1$.
For any element $a \in \mathcal{S}_1$, there exists a unique element $b \in \mathcal{S}_2$ such that $f(a) = b$.
This is called the **image** of a.

The set $\mathcal{S}_1$ is called the **domain** of the function f, and the set $\mathcal{S}_2$ is called the **codomain** of f.
The vector $f(\mathbf{v})$ is called the **image** of $\mathbf{v}$ (under f).
The **range** of f is the subset of $\mathcal{S}_2$ that contains **all** images: $\{f(a) : a \in \mathcal{S}_1\}$



Is this a function?



Definitions (Functions on subsets of $\mathcal{R}^n$)

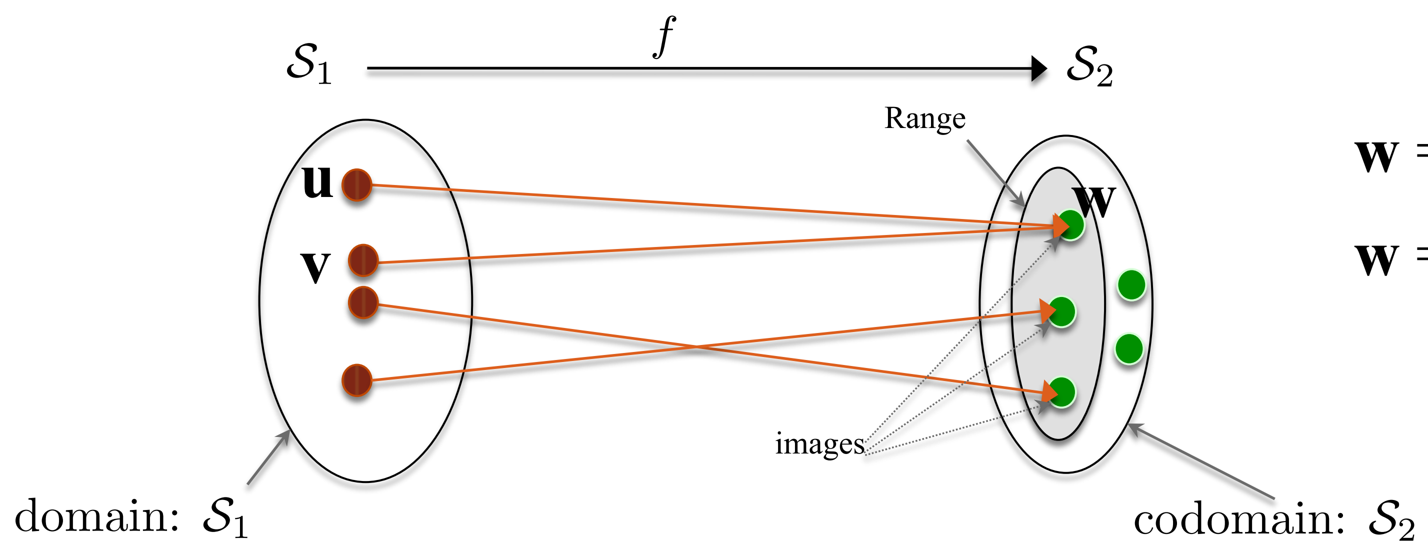
Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be subsets of $\mathcal{R}^n$ and $\mathcal{R}^m$, respectively.

A **function** f from $\mathcal{S}_1$ to $\mathcal{S}_2$, written $f : \mathcal{S}_1 \to \mathcal{S}_2$, is a rule that assigns to each vector $\mathbf{v}$ in $\mathcal{S}_1$ a unique vector $f(\mathbf{v})$ in $\mathcal{S}_2$.

The vector $f(\mathbf{v})$ is called the **image** of $\mathbf{v}$ (under f).

The set $\mathcal{S}_1$ is called the **domain** of a function f, and the set $\mathcal{S}_2$ is called the **codomain** of f.

The **range** of f is defined to be the set of images $f(\mathbf{v})$ for all $\mathbf{v}$ in $\mathcal{S}_1$.



$$\mathbf{w} = f(\mathbf{u})$$

$$\mathbf{w} = f(\mathbf{v})$$

Example:

$$f : \mathcal{R}^3 \to \mathcal{R}^2 \text{ with } f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1^2 \end{bmatrix}$$

$$f\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right) =$$

$$f\left(\begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}\right) =$$

$$f\left(\begin{bmatrix} \\ \\ \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Example:

$$T_A : \mathcal{R}^2 \to \mathcal{R}^3 \text{ with } T_A(\mathbf{x}) = A\mathbf{x} \text{ where } A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & -1 \end{bmatrix}$$

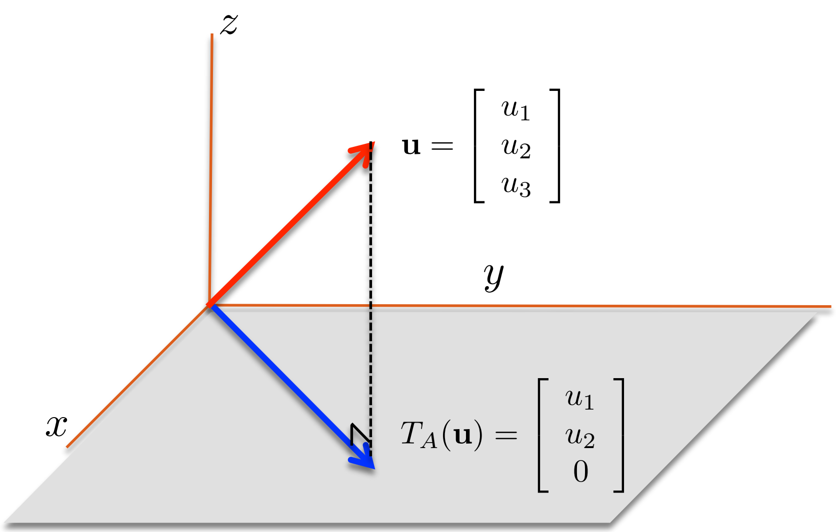
$$\Rightarrow T_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 + x_2 \\ x_1 - x_2 \end{bmatrix}$$

Definition

Let A be an $m \times n$ matrix. The function $T_A : \mathcal{R}^n \to \mathcal{R}^m$ defined by $T_A(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$ in $\mathcal{R}^n$ is called the **matrix transformation induced by A**.

Example:

$$T_A : \mathcal{R}^3 \to \mathcal{R}^3 \text{ with } T_A\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

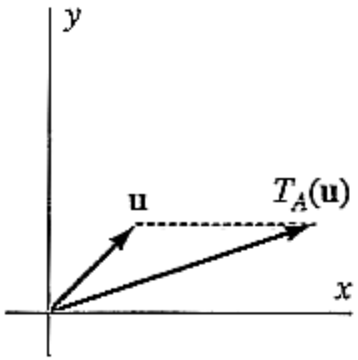


$T_A(\mathbf{u})$ is the projection of $\mathbf{u} \in \mathcal{R}^3$ on to the *xy*-plane, which is the range of T_A.

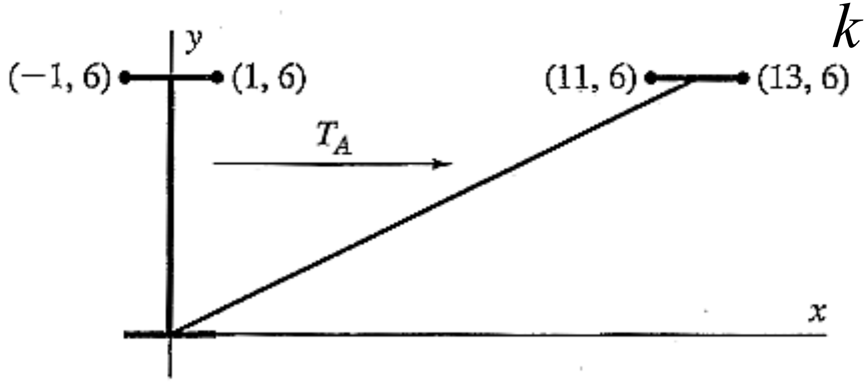
Example:

$$T_A : \mathcal{R}^2 \rightarrow \mathcal{R}^2 \text{ with } T_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + kx_2 \\ x_2 \end{bmatrix}$$ shear transformation

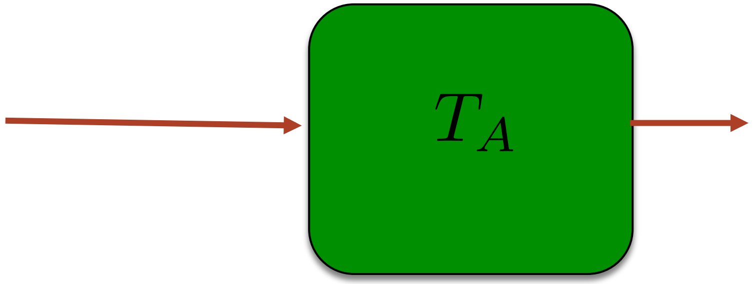
$$= \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$



$k = 2$



$\mathbf{x} \in \mathcal{R}_n$



$$\begin{aligned} \mathbf{y} &= T_A(\mathbf{x}) \\ &= A\mathbf{x} \in \mathcal{R}^m \end{aligned}$$

Question: Is any **matrix transformation** (induced by an $m \times n$ matrix) a **function** (from $\mathcal{R}^n$ to $\mathcal{R}^m$)?

Question: Can any function from $\mathcal{R}^n$ to $\mathcal{R}^m$ be written as a matrix transformation induced by an $m \times n$ matrix?

Think: What kind of properties does a matrix transformation, a special case of functions from $\mathcal{R}^n$ to $\mathcal{R}^m$, possess?

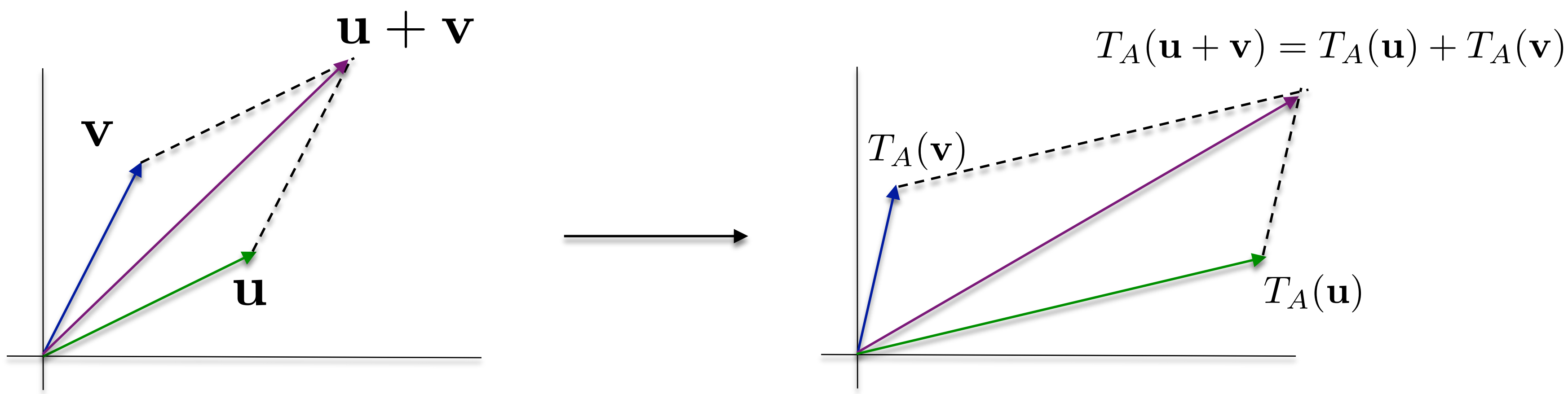
For example, can we express $T_A(\mathbf{u} + \mathbf{v})$ in terms of $T_A(\mathbf{u})$ and $T_A(\mathbf{v})$, for any $\mathbf{u}, \mathbf{v} \in \mathcal{R}^n$?

Theorem 2.7

For any $m \times n$ matrix A and any vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{R}^n$, the following statements are true:

(a) $T_A(\mathbf{u} + \mathbf{v}) = T_A(\mathbf{u}) + T_A(\mathbf{v})$

(b) $T_A(c\mathbf{u}) = cT_A(\mathbf{u})$ for every scalar c.



Theorem 2.7

For any $m \times n$ matrix A and any vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{R}^n$, the following statements are true:

(a) $T_A(\mathbf{u} + \mathbf{v}) = T_A(\mathbf{u}) + T_A(\mathbf{v})$

(b) $T_A(c\mathbf{u}) = cT_A(\mathbf{u})$ for every scalar c.

Proof

Definition

A function T from $\mathcal{R}^n$ to $\mathcal{R}^m$, written $T : \mathcal{R}^n \to \mathcal{R}^m$, is called a **linear transformation** (or simply **linear**) if, for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{R}^n$ and all scalars c, both of the following conditions hold:

(i) $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$.
(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$.

In the case of (i), we say that T **preserves vector addition**.

In the case of (ii), we say that T **preserves scalar multiplication**.

From Theorem 2.7, any **matrix transformation** (induced by any matrix A) must be a **linear transformation**!

Can you find a function from $\mathcal{R}^n$ to $\mathcal{R}^m$ that satisfies condition (ii) but not condition (i)?

Can you find a function from $\mathcal{R}^n$ to $\mathcal{R}^m$ that satisfies condition (i) but not condition (ii)?

There are functions from $\mathcal{R}^n$ to $\mathcal{R}^m$ satisfying condition (ii) of the definition of linear transformation but not (i). For example,

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{cases} \dfrac{xy^2}{x^2 + y^2} & \text{for } \begin{bmatrix} x \\ y \end{bmatrix} \neq \mathbf{0} \\ 0 & \text{otherwise} \end{cases}$$

If a function $T: \mathcal{R}^n \to \mathcal{R}^m$ satisfies condition (i), then for all integers $k, l > 0$ and $\mathbf{x} \in \mathcal{R}^n$, $T(k\mathbf{x}) = T(\mathbf{x} + \cdots + \mathbf{x}) = T(\mathbf{x}) + \cdots + T(\mathbf{x}) = kT(\mathbf{x})$ and $T(\mathbf{x}) = T((1/l)\mathbf{x} + \cdots + (1/l)\mathbf{x}) = T((1/l)\mathbf{x}) + \cdots + T((1/l)\mathbf{x}) = lT((1/l)\mathbf{x})$, or $(1/l)T(\mathbf{x}) = T((1/l)\mathbf{x})$. Also, $T(\mathbf{0}) = T(\mathbf{0} + \mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$, so $T(\mathbf{0}) = \mathbf{0}$. And $T(\mathbf{0}) = T(\mathbf{x} + (-1)\mathbf{x}) = T(\mathbf{x}) + T(-\mathbf{x})$, so $T(-\mathbf{x}) = -T(\mathbf{x})$. Therefore $T(r\mathbf{x}) = rT(\mathbf{x})$ for all rational number r. To further imply $T(c\mathbf{x}) = cT(\mathbf{x})$ for all real number c, "continuity" of the function T must be assumed.

Consider the following functions:
(1) the identity transformation $I : \mathcal{R}^n \rightarrow \mathcal{R}^n$ with $I(\mathbf{x}) = \mathbf{x}$**,** $\forall \mathbf{x} \in \mathcal{R}^n$.
(2) the zero transformation $T_0 : \mathcal{R}^n \rightarrow \mathcal{R}^m$ with $T_0(\mathbf{x}) = \mathbf{0}$**,** $\forall \mathbf{x} \in \mathcal{R}^n$.

Question: Is each of them a linear transformation?

Question: Is each of them a matrix transformation?

Think: Can any linear transformation be written as a matrix transformation induced by an m x n matrix A?

Theorem 2.8

For any linear transformation $T : \mathcal{R}^n \to \mathcal{R}^m$, the following statements are true:
(a) $T(\mathbf{0}) = \mathbf{0}$.
(b) $T(-\mathbf{u}) = -T(\mathbf{u})$ for all vectors $\mathbf{u}$ in $\mathcal{R}^n$.
(c) $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{R}^n$.
(d) $T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$ for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{R}^n$ and all scalars a and b.

Proof

Theorem 2.8

For any linear transformation $T : \mathcal{R}^n \to \mathcal{R}^m$, the following statements are true:
(a) $T(\mathbf{0}) = \mathbf{0}$.
(b) $T(-\mathbf{u}) = -T(\mathbf{u})$ for all vectors $\mathbf{u}$ in $\mathcal{R}^n$.
(c) $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{R}^n$.
(d) $T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$ for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{R}^n$ and all scalars a and b.

Corollary:

Let $T : \mathcal{R}^n \to \mathcal{R}^m$ be a linear transformation. If $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ are vectors in $\mathcal{R}^n$ and $a_1, a_2, \cdots, a_k$ are scalars, then

$$T(a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k) = a_1T(\mathbf{u}_1) + a_2T(\mathbf{u}_2) + \cdots + a_kT(\mathbf{u}_k).$$

Note: $T : \mathcal{R} \to \mathcal{R}$ with $T(x) = 2x + 3$ is not a linear transformation, but is an affine transformation.

Example:

Suppose $T : \mathcal{R}^2 \rightarrow \mathcal{R}^2$ is a linear transformation such that

$$T\left(\begin{bmatrix}1\\1\end{bmatrix}\right)=\begin{bmatrix}2\\3\end{bmatrix} \text{ and } T\left(\begin{bmatrix}1\\-1\end{bmatrix}\right)=\begin{bmatrix}4\\-1\end{bmatrix}$$

(a) $T\left(\begin{bmatrix}3\\3\end{bmatrix}\right) =$

(b) Find $T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)$ and $T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)$, and then determine $T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)$.

Example:

$$T : \mathcal{R}^2 \to \mathcal{R}^2 \quad \text{with } T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 - x_2 \\ x_1 \end{bmatrix} \text{ is linear. Why?}$$

(a) $T(\mathbf{u} + \mathbf{v}) =$

$$= T(\mathbf{u}) + T(\mathbf{v})$$

(b) $T(c\mathbf{u}) =$

$$= cT(\mathbf{u})$$

another way of verifying the linearity:

show ∃ a matrix A such that $T = T_A$, and in this case

$$T_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \underbrace{\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_2 \\ x_2 \end{bmatrix} = T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)$$

Observations:

1. $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 - x_2 \\ x_1 \end{bmatrix}$ is a linear transformation.

2. It can be written as a matrix transformation induced by an *2* x *2* matrix *A*. $\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$

Think again:

Can any linear transformation $T : \mathcal{R}^n \to \mathcal{R}^m$ be written as a matrix transformation induced by an *m* x *n* matrix *A*?

If the answer is yes, then:

How can we determine the *m* x *n* matrix *A*, given the linear transformation $T : \mathcal{R}^n \to \mathcal{R}^m$?

Theorem 2.9

Let $T : \mathcal{R}^n \to \mathcal{R}^m$ be linear. Then there is a unique $m \times n$ matrix

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix},$$

whose columns are the images under T of the standard vectors for $\mathcal{R}^n$, such that $T(\mathbf{v}) = A\mathbf{v}$ for all $\mathbf{v}$ in $\mathcal{R}^n$.

Proof

Theorem 2.9

Let $T : \mathcal{R}^n \to \mathcal{R}^m$ be linear. Then there is a unique $m \times n$ matrix

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix},$$

whose columns are the images under R of the standard vectors for $\mathcal{R}^n$, such that $T(\mathbf{v}) = A\mathbf{v}$ for all $\mathbf{v}$ in $\mathcal{R}^n$.

Definition

Let $T : \mathcal{R}^n \to \mathcal{R}^m$ be a linear transformation. We call the $m \times n$ matrix

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$$

the **standard matrix** of T.

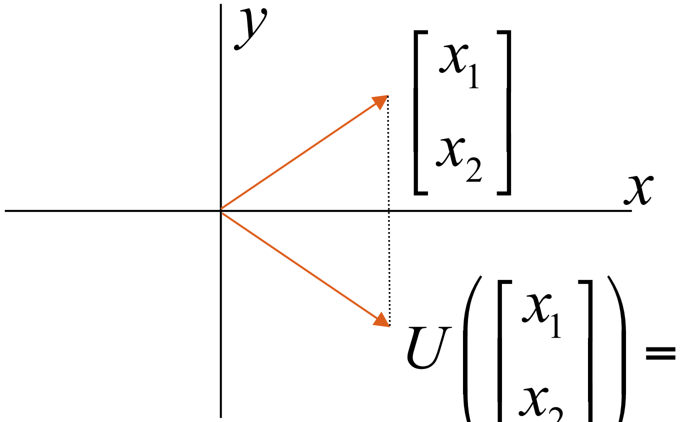
Note that, by Theorem 2.9, the standard matrix A of T has the property that $T(\mathbf{v}) = A\mathbf{v}$ for every $\mathbf{v}$ in R^n.

Example:

$$T : \mathcal{R}^3 \to \mathcal{R}^2 \text{ with } T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 - 4x_2 \\ 2x_1 + x_3 \end{bmatrix}$$

Find the standard matrix of T.

Example: Reflection of $\mathcal{R}^2$ about the x-axis.



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$U\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}$$

$U(\mathbf{e}_1) = \mathbf{e}_1$, and $U(\mathbf{e}_2) = -\mathbf{e}_2$
$\Rightarrow$ the standard matrix of U

is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Homework Set for Section 2.7

Section 2.7: Problems 1-6, 10-12, 21-23, 26, 29, 32-34