Chapter 4

Linear Transformations

Outlines

- **Definition and Example**
- **Matrix representation of Linear Transformation**
- **Similarity**

4.1 Definition and Examples
Definition

A mapping L from a vector space V into a vector Space W is said to be a **linear transformation** if

$$L(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2) \qquad (1)$$

for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and for all scalars α and β.

- If L is a linear transformation mapping a vector space V into W, from (1) we get

$$\boldsymbol{L}(\mathbf{v}_1 + \mathbf{v}_2) = \boldsymbol{L}(\mathbf{v}_1) + \boldsymbol{L}(\mathbf{v}_2) \qquad (\alpha = \beta = 1) \qquad (1)$$

$$\boldsymbol{L}(\alpha\mathbf{v}) = \alpha\boldsymbol{L}(\mathbf{v}) \qquad (\mathbf{v} = \mathbf{v}_1,\ \beta = 0) \qquad (2)$$

- Conversely, if L satisfies (2) and (3), then

$$\begin{aligned} L(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) &= L(\alpha\mathbf{v}_1) + L(\beta\mathbf{v}_2) && (\text{from}(2)) \\ &= \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2) && (\text{from}(3)) \end{aligned}$$

L is a linear transformation if and only if L satisfies (2) and (3).

Notation

- A mapping L from a vector space V into a vector space W will be denoted

$$L: V \rightarrow W$$

- If the vector space V and W are the same, the linear transformation

$$L: V \rightarrow V$$

  is referred to as a ***linear operator*** on V.

  A linear operator is a linear transformation that maps a vector space V into itself.

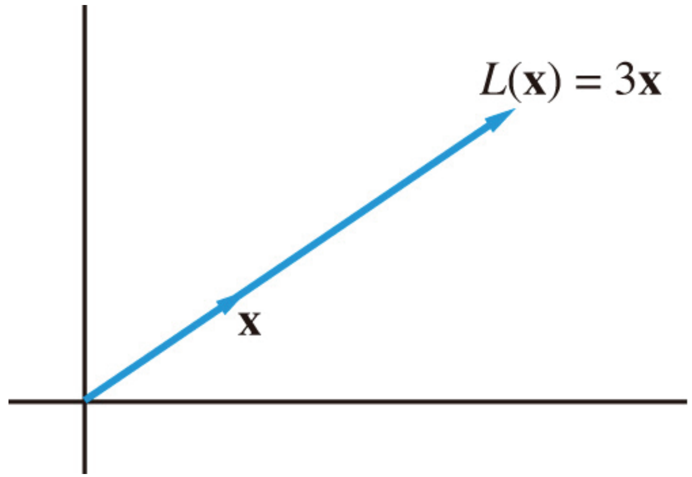
Linear Transformations on R^2 Example 1

- Let L be the operator defined by $L(\mathbf{x}) = 3\mathbf{x}$, for each $\mathbf{x} \in R^2$.
- *Sol:*

  Since $L(\alpha\mathbf{x}) = 3(\alpha\mathbf{x}) = \alpha(3\mathbf{x}) = \alpha L(\mathbf{x})$

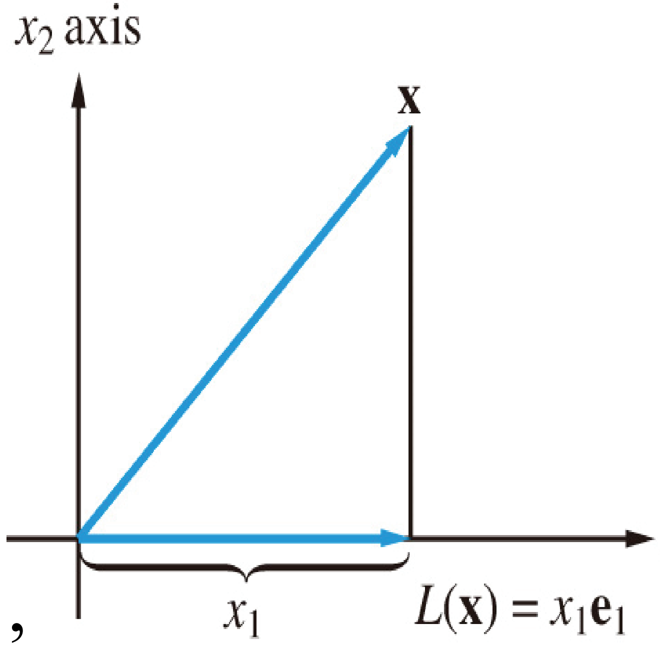
  and $L(\mathbf{x}+\mathbf{y}) = 3(\mathbf{x}+\mathbf{y}) = 3\mathbf{x} + 3\mathbf{y} = L(\mathbf{x}) + L(\mathbf{y})$

  $\Rightarrow L$ is a linear transformation



Example 2

- Consider the mapping L defined by $L(\mathbf{x}) = x_1\mathbf{e}_1$ for each $\mathbf{x} = (x_1, x_2)^T \in R^2$.



- *Sol:*

If $\mathbf{x} = (x_1, x_2)^T$, then $L(\mathbf{x}) = \mathbf{x}_1\mathrm{e}_1 = \mathbf{x}_1\ (1, 0)^T = (x_1, 0)^T$

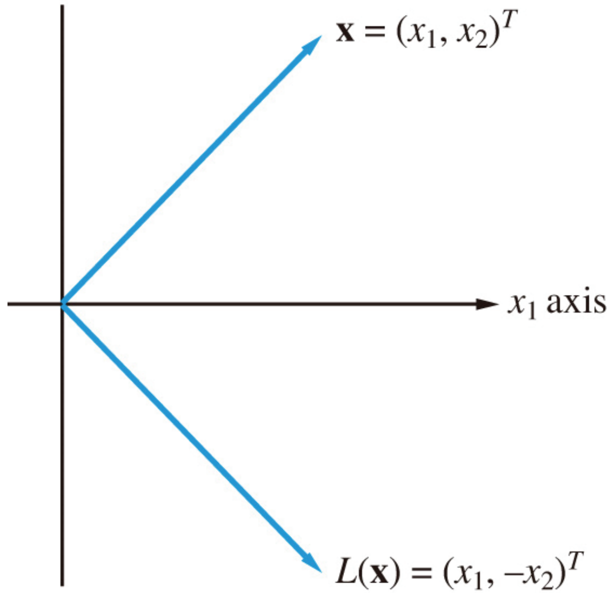
If $\mathbf{y} = (y_1, y_2)^T$, then

$$L(\alpha\mathbf{x}+\beta\mathbf{y}) = L\left(\alpha\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \beta\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \boldsymbol{L}\left(\begin{bmatrix} \alpha x_1+\beta y_1 \\ \alpha x_2+\beta y_2 \end{bmatrix}\right) = \begin{bmatrix} \alpha x_1+\beta y_1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \alpha x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} \beta y_1 \\ 0 \end{bmatrix} = \alpha\begin{bmatrix} x_1 \\ 0 \end{bmatrix} + \beta\begin{bmatrix} y_1 \\ 0 \end{bmatrix} = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$$

$\Rightarrow$ L is a linear transformation.

Example 3



- Let L be the operator defined by $L(\mathbf{x}) = (x_1, -x_2)^T$ for each $\mathbf{x} = (x_1, x_2)^T$ in R^2.
- *Sol:*

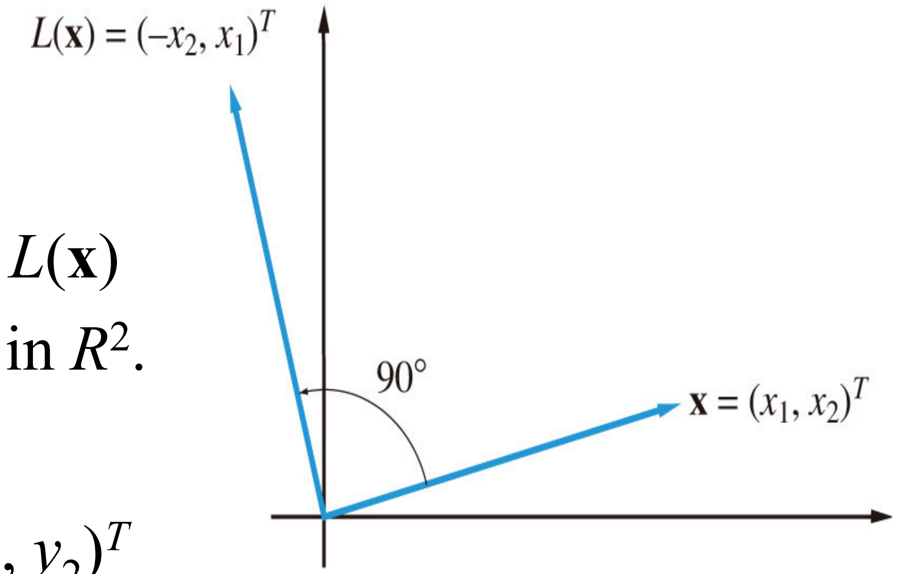
  For each $\mathbf{x} = (x_1, x_2)^T$ and $\mathbf{y} = (y_1, y_2)^T$

$$L(\alpha\mathbf{x} + \beta\mathbf{y}) = L\left(\alpha\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \beta\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \boldsymbol{L}\left(\begin{bmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \end{bmatrix}\right) = \begin{bmatrix} \alpha x_1 + \beta y_1 \\ -(\alpha x_2 + \beta y_2) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha x_1 \\ -\alpha x_2 \end{bmatrix} + \begin{bmatrix} \beta y_1 \\ -\beta y_2 \end{bmatrix} = \alpha\begin{bmatrix} x_1 \\ -x_2 \end{bmatrix} + \beta\begin{bmatrix} y_1 \\ -y_2 \end{bmatrix} = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$$

$\Rightarrow$ L is a linear operator.

Example 4



- Let L be the operator defined by $L(\mathbf{x}) = (-x_2, x_1)^T$ for each $\mathbf{x} = (x_1, x_2)^T$ in R^2.
- *Sol:*

  For each $\mathbf{x} = (x_1, x_2)^T$ and $\mathbf{y} = (y_1, y_2)^T$

$$L(\alpha\mathbf{x}+\beta\mathbf{y}) = L\left(\alpha\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \beta\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = L\left(\begin{bmatrix} \alpha x_1+\beta y_1 \\ \alpha x_2+\beta y_2 \end{bmatrix}\right) = \begin{bmatrix} -(\alpha x_2+\beta y_2) \\ \alpha x_1+\beta y_1 \end{bmatrix}$$

$$= \begin{bmatrix} -\alpha x_2 \\ \alpha x_1 \end{bmatrix} + \begin{bmatrix} -\beta y_2 \\ \beta y_1 \end{bmatrix} = \alpha\begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} + \beta\begin{bmatrix} -y_2 \\ y_1 \end{bmatrix} = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$$

$\Rightarrow$ L is a linear transformation.

Linear Transformations from R^n to R^m Example 5

- $L: R^2 \to R^1$ defined by $L(\mathbf{x}) = x_1 + x_2$
- *Sol:*

For each $\mathbf{x} = (x_1, x_2)^T$ and $\mathbf{y} = (y_1, y_2)^T$

$$L(\alpha\mathbf{x}+\beta\mathbf{y}) = L\left(\alpha\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \beta\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = L\left(\begin{bmatrix} \alpha x_1+\beta y_1 \\ \alpha x_2+\beta y_2 \end{bmatrix}\right)$$

$$= \alpha x_1+\beta y_1+\alpha x_2+\beta y_2$$

$$= \alpha(x_1+x_2)+\beta(y_1+y_2) = \alpha L(\mathbf{x}) \ + \ \beta L(\mathbf{y})$$

$\Rightarrow$ L is a linear transformation.

Example 6

- Consider the mapping M defined by $M(\mathbf{x}) = (x_1{}^2 + x_2{}^2)^{1/2}$.
- *Sol:*

Since: $M(\alpha\mathbf{x}) = M\left(\alpha\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = M\left(\begin{bmatrix} \alpha x_1 \\ \alpha x_2 \end{bmatrix}\right) = \left[(\alpha x_1)^2 + (\alpha x_2)^2\right]^{1/2}$

$= |\alpha|(x_1^2 + x_2^2)^{1/2}$

$\alpha M(\mathbf{x}) = \alpha(x_1^2 + x_2^2)^{1/2}$

$\Rightarrow M(\alpha\mathbf{x}) \neq \alpha M(\mathbf{x})$ whenever $\alpha < 0$

$\Rightarrow$ ***M*** is **not** a linear transformation.

Example 7

- L: $R^2 \to R^3$ defined by $L(\mathbf{x}) = (x_2, x_1, x_1+x_2)^T$.
- *Sol:*

$$L(\alpha\mathbf{x}) = L\left(\alpha\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = L\left(\begin{bmatrix} \alpha x_1 \\ \alpha x_2 \end{bmatrix}\right) = \begin{bmatrix} \alpha x_2 \\ \alpha x_1 \\ \alpha x_1+\alpha x_2 \end{bmatrix} = \alpha\begin{bmatrix} x_2 \\ x_1 \\ x_1+x_2 \end{bmatrix} = \alpha L(\mathbf{x})$$

$$L(\mathbf{x}+\mathbf{y}) = L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = L\left(\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \end{bmatrix}\right)$$

$$= \begin{bmatrix} x_2+y_2 \\ x_1+y_1 \\ x_1+y_1+x_2+y_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \\ x_1+x_2 \end{bmatrix} + \begin{bmatrix} y_2 \\ y_1 \\ y_1+y_2 \end{bmatrix} = L(\mathbf{x}) + L(\mathbf{y})$$

$\Rightarrow$ L is a linear transformation.

Example 7 (con.)

- Note that if we define the matrix A by $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$

- Then $L\mathbf{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \\ x_1 + x_2 \end{bmatrix} = A(\mathbf{x})$

  for each $\mathbf{x} \in R^2$.

- In general, if A is any $m \times n$ matrix, we can define a linear transformation L_A from R^n to R^m by

$$L_A(\mathbf{x}) = A\mathbf{x} \quad \text{for each } \mathbf{x} \in R^n$$

- The transformation LA is linear since

$$\begin{aligned} LA(\alpha\mathbf{x} + \beta\mathbf{y}) &= A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A\mathbf{x} + \beta A\mathbf{y} \\ &= \alpha LA(\mathbf{x}) + \beta LA(\mathbf{y}) \end{aligned}$$

- We can think of each $m \times n$ matrix as defining a linear transformation from R^n to R^m.

Linear Transformations from V to W

- If L is a linear transformation mapping a vector space V into a vector space W, then

(1) $L(\mathbf{0}_V) = \mathbf{0}_W$ (where $\mathbf{0}_V$ and $\mathbf{0}_W$ are zero vectors in V and W)

(2) $L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n)$
$= \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \ldots + \alpha_n L(\mathbf{v}_n)$

(3) $L(-\mathbf{v}) = -L(\mathbf{v})$ for all $\mathbf{v} \in V$.

Example 8

- If V is any vector space, then the identity operator $\mathcal{I}$ is define by

$$\mathcal{I}(\mathbf{v}) = \mathbf{v}, \qquad \text{for all } \mathbf{v} \in V$$

- *Sol:*

$$\mathcal{I}(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2 = \alpha\,\mathcal{I}(\mathbf{v}_1) + \beta\,\mathcal{I}(\mathbf{v}_2)$$

$\mathcal{I}$ is a linear transformation that maps V into itself.

Example 9

- Let L be the mapping from $C[a, b]$ to R^1 defined by

$$L(f) = \int_a^b f(x)dx$$

If f and g are any vectors in $C[a, b]$, then

$$L(\alpha f + \beta g) = \int_a^b (\alpha f + \beta g)(x)\,dx$$

$$= \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx = \alpha L(x) + \beta L(g)$$

Therefore, L is a linear transformation.

Example 10

- Let D be the linear transformation mapping $C^1[a, b]$ into $C[\mathrm{a}, \mathrm{b}]$ and defined by

  $$D(f) = f' \qquad \text{(the derivative of } f\text{)}$$

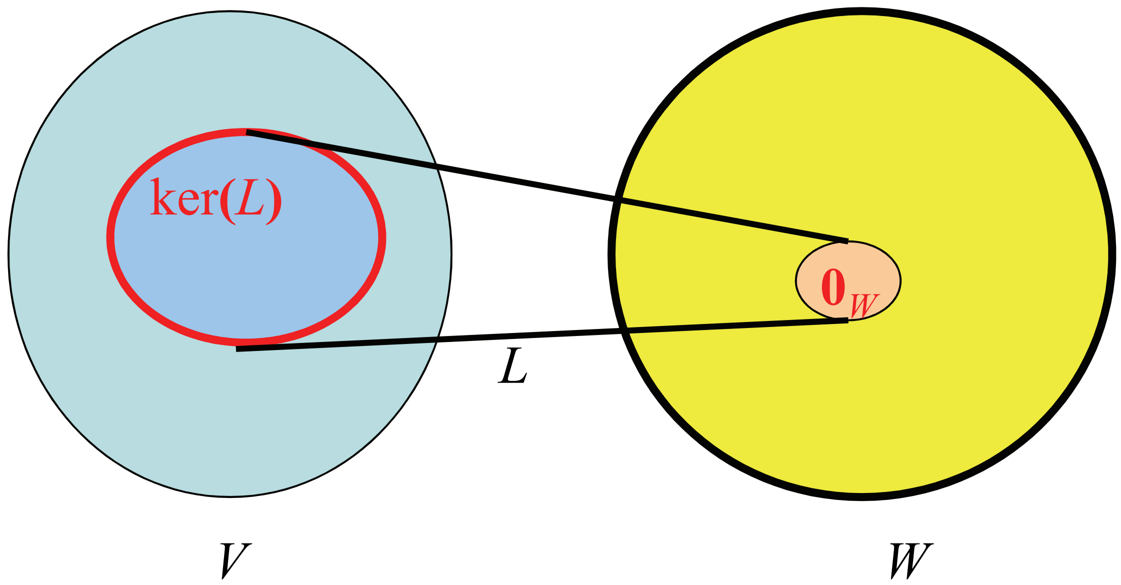
  D is a linear transformation, since

  $$D(\alpha f + \beta g) = \alpha f' + \beta g' = \alpha D(f) + \beta D(g)$$

The Image and Kernel Definition

Let $L: V \to W$ be a linear transformation. The **kernel** of L, denoted $\ker(L)$, is define by

$$\ker(L) = \{\mathbf{v} \in V \mid L(\mathbf{v}) = \mathbf{0}_W\}$$

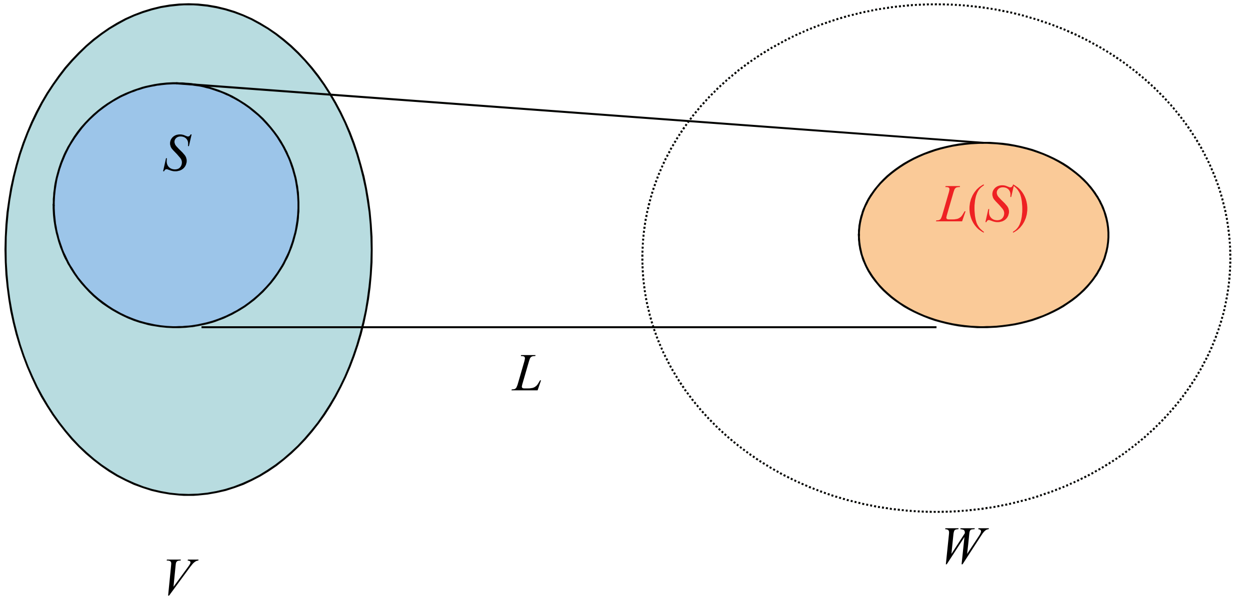


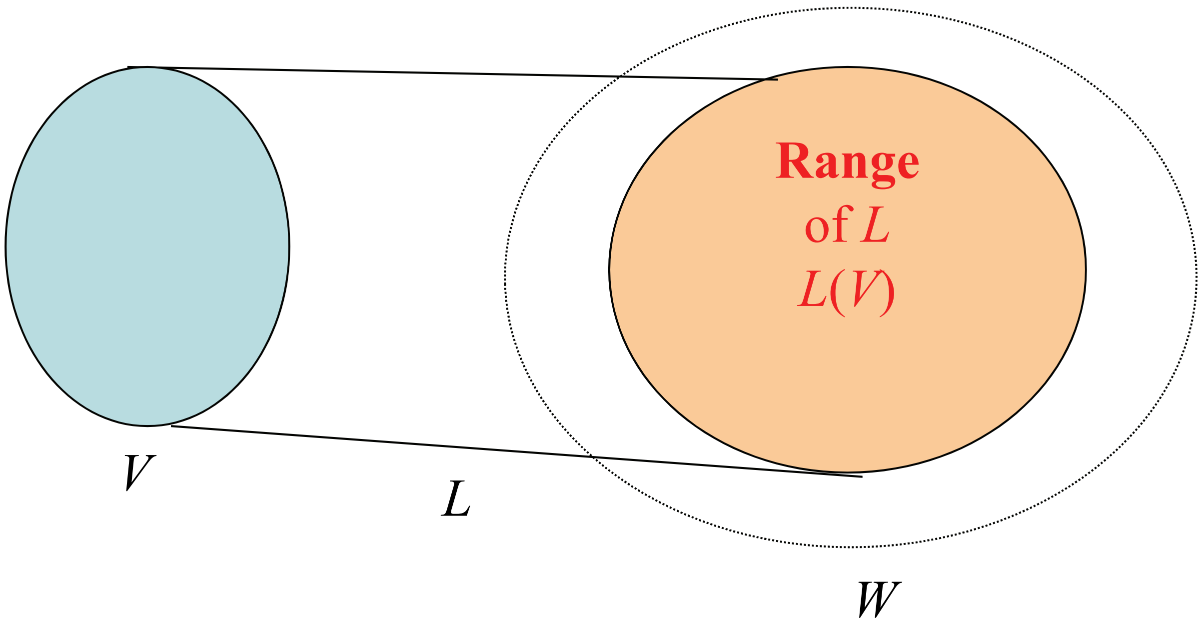
Definition

Let L: $V \to W$ be a linear transformation and let S be a subspace of V. The **image** of S, denoted $L(S)$, is defined by

$$L(S) = \{\mathbf{w} \in W \mid \mathbf{w} = L(\mathbf{v}) \text{ for some } \mathbf{v} \in S\}$$

The **image** of the entire vector space, $L(V)$, is called the **range** of L.


S
L(S)
L
W
V


Range
of L
L(V)
V
L
W

Theorem 4.1.1

If $L: V \to W$ is a linear transformation and S is a subspace of V, then

(1) $\ker(L)$ is a subspace of V

(2) $L(S)$ is a subspace of W

Theorem 4.1.1 *proof*

(1) by C_1: If $\mathbf{v} \in \ker(L)$ and α is a scalar $\Rightarrow L(\mathbf{v}) = \mathbf{0}_W$

$L(\alpha\mathbf{v}) = \alpha L(\mathbf{v}) = \alpha\mathbf{0}_W = \mathbf{0}_W$

$\Rightarrow \alpha\mathbf{v} \in \ker(L)$

by C_2: If $\mathbf{v}_1$ and $\mathbf{v}_2 \in \ker(L) \Rightarrow L(\mathbf{v}_1) = L(\mathbf{v}_2) = \mathbf{0}_W$,

then

$L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W$

$\Rightarrow \mathbf{v}_1 + \mathbf{v}_2 \in \ker(L)$

$\ker(L)$ is a subspace of V.

Theorem 4.1.1 *proof*

(2) by C_1: If $\mathbf{w} \in L(S)$, then $\mathbf{w} = L(\mathbf{v})$ for some $\mathbf{v} \in S$

$\alpha\mathbf{w} = \alpha L(\mathbf{v}) = L(\alpha\mathbf{v})$

Since S is a subspace $\Rightarrow \alpha\mathbf{v} \in S$

$\Rightarrow \alpha\mathbf{w} \in \boldsymbol{L}(S)$

by C_2: If $\mathbf{w}_1$ and $\mathbf{w}_2 \in L(S)$, then there exist $\mathbf{v}_1$ and $\mathbf{v}_2$ $\in \boldsymbol{S}$ such that

$L(\mathbf{v}_1) = \mathbf{w}_1$ and $L(\mathbf{v}_2) = \mathbf{w}_2$

$\Rightarrow \mathbf{w}_1+\mathbf{w}_2 = L(\mathbf{v}_1) + L(\mathbf{v}_2) = L(\mathbf{v}_1+\mathbf{v}_2)$

Since S is a subspace $\Rightarrow \mathbf{v}_1+\mathbf{v}_2 \in S$

$\Rightarrow \mathbf{w}_1+\mathbf{w}_2 \in L(S)$

$L(S)$ is a subspace of L.

Example 11

- $L: R^2 \rightarrow R^2$ defined by $L(\mathbf{x}) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$
- If $\mathbf{x} \in$ Ker(L), i.e.,

$$L(\mathbf{x}) = L\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- Thus, $x_1 = 0$,

$$\ker(\boldsymbol{L}) = \left\{\begin{bmatrix} 0 \\ x_2 \end{bmatrix}\right\} = x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_2 \mathbf{e}_2$$

- Thus, ker(L) is a one-dimensional subspace of R^2 spanned by $\mathbf{e}_2$. The range of L, $L(R^2)$, is a one-dimensional subspace of R^2 spanned by $\mathbf{e}_1$.

Example 12

- Thus, ker(L) is a one-dimensional subspace of R^2 spanned by $\mathbf{e}_2$. The range of L, $L(R^2)$, is a one-dimensional subspace of R^2 spanned by $\mathbf{e}_1$.
- *Sol:*

(1)If $\mathbf{x} \in$ ker(L), $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \xrightarrow{L} \begin{bmatrix} x_1+x_2 \\ x_2+x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0}$

Therefore, $x_1 + x_2 = 0$ and $x_2 + x_3 = 0$

let $x_3 = a \Rightarrow x_2 = -a$ and $x_1 = a$

$$\ker(L) = \{\begin{bmatrix} a \\ -a \\ a \end{bmatrix} \mid a \in R\} = \{a\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \mid a \in R\}$$

$\Rightarrow$ $\ker(L)$ is a one–dimension subspace of R^3.

Example 12 (con.)

(2)If $\mathbf{x} \in S$, $\mathbf{x} = a\,\mathbf{e}_1 + b\,\mathbf{e}_3 = a\begin{bmatrix}1\\0\\0\end{bmatrix} + b\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}a\\0\\b\end{bmatrix}$

Thus, $L(\mathrm{x}) = \begin{bmatrix}a+0\\0+b\end{bmatrix} = \begin{bmatrix}a\\b\end{bmatrix} = R^2$

$L(S) = R^2$, the image of S is R^2

$\Rightarrow L(R^3) = R^2$, the entire range of L must be R^2

Example 13

- Let D: $P_3 \rightarrow P_3$ be the differentiation operator, defined by

$$D(p(x)) = p'(x)$$

- The kernel of D consists of all polynomials of degree 0. Thus $\text{Ker}(D) = P_1$. The derivative of any polynomial in P_3 will be a polynomial of degree 1 or less.
- Conversely, ant polynomial in P_2 will have antiderivatives in P_3, so each polynomial in P_2 will be the image of polynomials in P_3 under the operator D. It then follows that $D(P_3) = P_2$.

4.2 Matrix Representations of Linear Transformations

- Each $m \times n$ matrix A defines a linear transformation L_A from R^n to R^m, where
$$L_A(\mathbf{x}) = A\mathbf{x}$$
for each $\mathbf{x} \in R^n$.
- For each linear transformation L mapping R^n into R^m there is an $m \times n$ matrix A such that
$$L(\mathbf{x}) = A\mathbf{x}$$

Theorem 4.2.1

If L is a linear transformation mapping R^n into R^m, there is an $m \times n$ matrix A such that

$$L(\mathbf{x}) = A\mathbf{x}$$

for each $\mathbf{x} \in R^n$. In fact, the jth column vector of A is given by

$$\mathbf{a}_j = L(\mathbf{e}_j), \quad j = 1, 2, \ldots, n$$

Theorem 4.2.1 *proof*

- Consider $A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$
- Define $\mathbf{a}_j = (a_{1j}, a_{2j}, \ldots, a_{mj})^T = L(\mathbf{e}_j), j = 1, 2, \ldots, n$
  Let $A = (a_{ij}) = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$
- If $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$ is any element of R^n, then

Theorem 4.2.1 *proof*

- $L(\mathbf{x}) = L(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n)$

$$= x_1L(\mathbf{e}_1) + x_2L(\mathbf{e}_2) + \ldots + x_nL(\mathbf{e}_n)$$

$$= x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n$$

$$= (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= A\mathbf{x}$$

Note

- $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n] = [L(\mathbf{e}_1) \quad L(\mathbf{e}_2) \quad \cdots \quad L(\mathbf{e}_n)]$
- A is referred to as the **standard matrix representation** of L

Example 1

- L: $R^3 \rightarrow R^2$ defined by $L(\mathbf{x})$ = for each $\mathbf{x} = (x_1, x_2, x_3)^T$ in R^3.
- *Sol:*
- Let $L(\mathbf{x}) = A\mathbf{x}$

$$\mathbf{a}_1 = L(\mathbf{e}_1) = L(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}) = \begin{bmatrix} 1+0 \\ 0+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\mathbf{a}_2 = L(\mathbf{e}_2) = L(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 0+1 \\ 1+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{a}_3 = L(\mathbf{e}_3) = L\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0+0 \\ 0+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow A = (\mathbf{a}_1,\ \mathbf{a}_2,\ \mathbf{a}_3) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

To check the result, compute $A\mathbf{x}$

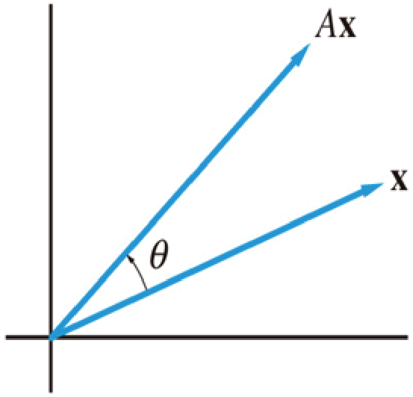
$$A\mathbf{x} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix}$$

Example 2

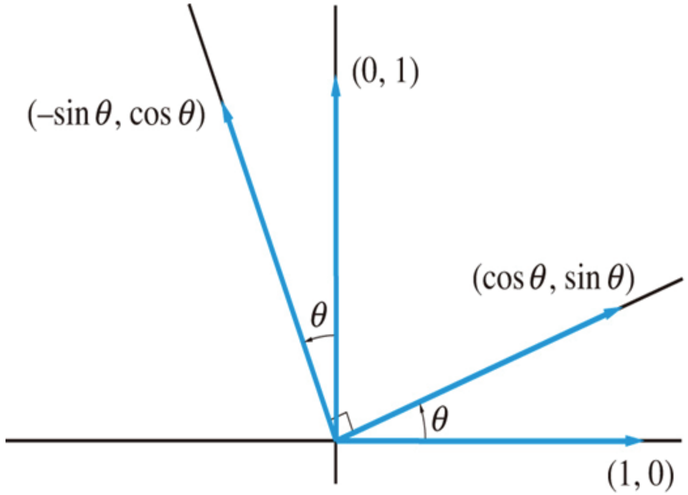
- Let L: $R^2 \rightarrow R^2$ which rotates each vector by an angle θ in the counterclockwise direction.
- *Sol:*
- Let $L(\mathbf{x}) = A\mathbf{x}$
  Since $\mathbf{a}_1 = L(\mathbf{e}_1) = (\cos\theta, \sin\theta)^T$, and
  $\mathbf{a}_2 = L(\mathbf{e}_2) = (-\sin\theta, \cos\theta)^T$

  $\Rightarrow$

$$A = (\mathbf{a}_1, \mathbf{a}_2) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$



(b)



(a)

Question

- *How to find a similar representation for linear transformations from an n-dimensional vector space V into an m-dimensional vector space W?*

- Let $E = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $F = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ be ordered bases for vector spaces V and W, and L be a linear transformation mapping V into W. If $\mathbf{v} \in V$, then $\mathbf{v}$ can be expressed in terms of the basis E:

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n$$

- There exists an $m \times n$ matrix A representing the linear transformation L:

$$A\mathbf{x} = \mathbf{y} \text{ iff } L(\mathbf{v}) = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \ldots + y_m\mathbf{w}_m$$

- If $\mathbf{x}$ is the coordinate vector of $\mathbf{v}$ w. r. t. (with respect to) E, then the coordinate vector of $L(\mathbf{v}) = \mathbf{y}$ w. r. t. F is given by:

$$\mathbf{y} = [L(\mathbf{v})]_F = A\mathbf{x}$$

How to determine the matrix representation A?

- Let $\mathbf{a}_j = (a_{1j}, a_{2j}, \ldots, a_{mj})^T$ be the <u>coordinate vector</u> of $L(\mathbf{v}_j)$ w. r. t. $F = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$, $j = 1, 2, \ldots, n$
$$L(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \ldots + a_{mj}\mathbf{w}_m \qquad 1 \le j \le n$$
- Let $A = (a_{ij}) = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$. If
$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n$$
then
$$\begin{aligned} L(\mathbf{v}) &= L(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n) \\ &= x_1L(\mathbf{v}_1) + x_2L(\mathbf{v}_2) + \ldots + x_nL(\mathbf{v}_n) \end{aligned}$$

$$= L\left(\sum_{j=1}^{n} x_j \mathbf{v}_j\right) = \sum_{j=1}^{n}\left(x_j\ L(\mathbf{v}_j)\right) = \sum_{j=1}^{n}\left(x_j (\sum_{i=1}^{m} a_{ij}\mathbf{w}_i)\right)$$

$$= \sum_{i=1}^{m}\left(\left(\sum_{j=1}^{n} a_{ij} x_j\right)\mathbf{w}_i\right)$$

- Let $y_i = \sum_{j=1}^{n} a_{ij} x_j = \mathrm{a}(i,:)^T \mathbf{x}, \text{ for } i = 1, 2, ..., m$

  $\mathbf{y} = (y_1, y_2, \ldots, y_m)^{\mathrm{T}} = A\mathbf{x}$ is the coordinate vector of $L(\mathbf{v})$ w. r. t. $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$.

Theorem 4.2.2

(Matrix Representation Theorem)

If $E = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $F = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ are ordered bases for vector spaces V and W, respectively, then corresponding to each linear transformation $L: V \rightarrow W$ there is an $m \times n$ matrix A such that

$$[L(\mathbf{v})]_F = A\,[\mathbf{v}]_E \qquad \text{for each } \mathbf{v} \in V$$

A is the matrix representing L relative to the ordered bases E and F. In fact,

$$\mathbf{a}_j = [L(\mathbf{v}_j)]_F \quad j = 1, 2, \ldots, n$$

Theorem 4.2.2

$$
\begin{tikzcd}
\mathbf{v} \in V \arrow[rr, "L = L_A"] \arrow[d, leftrightarrow] & & \mathbf{w} = L(\mathbf{v}) \in W \arrow[d, leftrightarrow] \\
\mathbf{x} = [\mathbf{v}]_E \in R^n \arrow[rr, "A"] & & A\mathbf{x} = [\mathbf{w}]_F \in R^m
\end{tikzcd}
$$

$\mathbf{x} = [\mathbf{v}]_E$: the coordinate vector of $\mathbf{v}$ with respect to E

$\mathbf{y} = [\mathbf{w}]_F$: the coordinate vector of $\mathbf{w}$ with respect to F

$\Rightarrow$ L maps $\mathbf{v}$ into $\mathbf{w}$ iff A maps $\mathbf{x}$ into $\mathbf{y}$

Example 3

- L: $R^3 \rightarrow R^2$ defined by

$$L(\mathbf{x}) = x_1\mathbf{b}_1 + (x_2 + x_3)\mathbf{b}_2$$

for each $\mathbf{x} \in R^3$, where

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Find the matrix A representing L w. r. t. the ordered bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{b}_1, \mathbf{b}_2\}$.

Example 3 (con.)

- *Sol:* $L(\mathbf{e}_1) = L(\begin{bmatrix}1\\0\\0\end{bmatrix}) = 1\cdot\mathbf{b}_1 + 0\cdot\mathbf{b}_2 \Rightarrow \mathbf{a}_1 = [L(\mathbf{e}_1)]_F = \begin{bmatrix}1\\0\end{bmatrix}$

$$L(\mathbf{e}_2) = L(\begin{bmatrix}0\\1\\0\end{bmatrix}) = 0\cdot\mathbf{b}_1 + 1\cdot\mathbf{b}_2 \Rightarrow \mathbf{a}_2 = [L(\mathbf{e}_2)]_F = \begin{bmatrix}0\\1\end{bmatrix}$$

$$L(\mathbf{e}_3) = L(\begin{bmatrix}0\\0\\1\end{bmatrix}) = 0\cdot\mathbf{b}_1 + 1\cdot\mathbf{b}_2 \Rightarrow \mathbf{a}_3 = [L(\mathbf{e}_3)]_F = \begin{bmatrix}0\\1\end{bmatrix}$$

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

Example 4

- L: $R^2 \rightarrow R^2$ defined by $L(\alpha\,\mathbf{b}_1 + \beta\,\mathbf{b}_2) = (\alpha + \beta)\mathbf{b}_1 + 2\beta\,\mathbf{b}_2$, where $\{\mathbf{b}_1, \mathbf{b}_2\}$ is the ordered basis defined in Example 3. Find the matrix A representing L w. r. t. $\{\mathbf{b}_1, \mathbf{b}_2\}$.
- *Sol:*
- $L(\mathbf{b}_1) = L(1\mathbf{b}_1 + 0\mathbf{b}_2) = (1+0)\mathbf{b}_1 + (2\times0)\mathbf{b}_2$
  $= 1\mathbf{b}_1 + 0\mathbf{b}_2 \Rightarrow \mathbf{a}_1 = [L(\mathbf{b}_1)]_{\boldsymbol{F}} = (1, 0)^T$
- $L(\mathbf{b}_2) = L(0\mathbf{b}_1 + 1\mathbf{b}_2) = (0+1)\mathbf{b}_1 + (2\times1)\mathbf{b}_2$
  $= 1\mathbf{b}_1 + 2\mathbf{b}_2 \Rightarrow \mathbf{a}_2 = [L(\mathbf{b}_2)]_{\boldsymbol{F}} = (1, 2)^T$

$$\Rightarrow A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

Example 5

- The linear transformation D defined by $D(p) = p'$ maps P_3 into P_2. Given the ordered bases $[x^2, x, 1]$ and $[x, 1]$ for P_3 and P_2, respectively, we wish to determine a matrix representation for D. To do this, we apply D to each of the basis elements of P_3:

$$D(x^2) = 2x + 0 \cdot 1$$

$$D(x) = 0x + 1 \cdot 1$$

$$D(1) = 0x + 0 \cdot 1$$

Example 5 (con.)

- In P_2, the coordinate vectors dor $D(x^2)$, $D(x)$, and $D(1)$ are $(2, 0)^T$, $(0, 1)^T$, and $(0, 0)^T$, respectively. The matrix A is formed with these vectors as its columns.

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

- If $p(x) = ax^2 + bx + c$, then the coordinate vector of p with respect to the ordered basis of P_3 is $(a, b, c)^T$.

Example 5 (con.)

- To find the coordinate vector of $D(p)$ with respect to the ordered basis of P_2, we simply multiply

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2a \\ b \end{bmatrix}$$

- Thus,

$$D(ax^2 + bx + c) = 2ax + b$$

Theorem 4.2.3

Let $E = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$ and $F = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m]$ be ordered bases for R^n and R^m, respectively. If $L: R^n \to R^m$ is a linear transformation and A is the matrix representing L with respect to E and F, then

$$\mathbf{a}_j = B^{-1}L(\mathbf{u}_j) \quad \text{for } j = 1, 2, \ldots, n$$

where $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m)$.

Theorem 4.2.3 *proof*

- If A is representing L with respect to E and F, then for $j = 1, 2, \ldots, n$

  $L(\mathbf{u}_j) = a_{1j}\mathbf{b}_1 + a_{2j}\mathbf{b}_2 + \ldots + a_{mj}\mathbf{b}_m$

  (**Note**: $\mathbf{a}_j = [L(\mathbf{u}_j)]_F$)

  $$= (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m)\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = B\mathbf{a}_j$$

- $\Rightarrow \mathbf{a}_j = B^{-1}L(\mathbf{u}_j)$ for $j = 1, 2, \ldots, n$

Corollary 4.2.4

If A is the matrix representing the linear transformation $L: R^n \to R^m$ with respect to the bases $E = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$ and $F = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m]$, then the reduced row echelon form of $(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m \mid L(\mathbf{u}_1), \ldots, L(\mathbf{u}_n))$ is $(I \mid A)$.

Corollary 4.2.4 *proof*

- Let $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m)$
  The matrix $(B|\ L(\mathbf{u}_1), \ldots , L(\mathbf{u}_n))$ is row equivalent to
  $$B^{-1}\ (B|\ L(\mathbf{u}_1), \ldots , L(\mathbf{u}_n))$$
  $$= (\ I\ |\ B^{-1}L(\mathbf{u}_1), \ldots , \mathrm{B}^{-1}L(\mathbf{u}_n))$$
  $$= (\ I\ |\ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$$
  $$= (\ I\ |\ A)$$

Example 6

- Let L: $R^2 \to R^3$ defined by $L(\mathbf{x}) = \begin{bmatrix} x_2 \\ x_1+x_2 \\ x_1-x_2 \end{bmatrix}$

  Find the matrix representation of L with respect to the ordered bases $[\mathbf{u}_1, \mathbf{u}_2]$ and $[\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$, where

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ and } \mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Example 6 (con.)

- *Sol:*

$$L(\mathbf{u}_1) = L(\begin{bmatrix} 1 \\ 2 \end{bmatrix}) = \begin{bmatrix} 2 \\ 1+2 \\ 1-2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$$

$$L(\mathbf{u}_2) = L(\begin{bmatrix} 3 \\ 1 \end{bmatrix}) = \begin{bmatrix} 1 \\ 3+1 \\ 3-1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$$

$$(B \mid L(\mathbf{u}_1)\ L(\mathbf{u}_2)) = \left[\begin{array}{ccc|cc} 1 & 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 3 & 4 \\ 0 & 0 & 1 & -1 & 2 \end{array}\right] \rightarrow \left[\begin{array}{ccc|cc} 1 & 0 & 0 & -1 & -3 \\ 0 & 1 & 0 & 4 & 2 \\ 0 & 0 & 1 & -1 & 2 \end{array}\right]$$

$$\Rightarrow A = \begin{bmatrix} -1 & -3 \\ 4 & 2 \\ -1 & 2 \end{bmatrix}$$

Verification

$$L(\mathbf{u}_1) = -\mathbf{b}_1 + 4\mathbf{b}_2 - \mathbf{b}_3 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 4 \\ 4 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$$

$$L(\mathbf{u}_2) = -3\mathbf{b}_1 + 2\mathbf{b}_2 + 2\mathbf{b}_3 = \begin{bmatrix} -3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$$

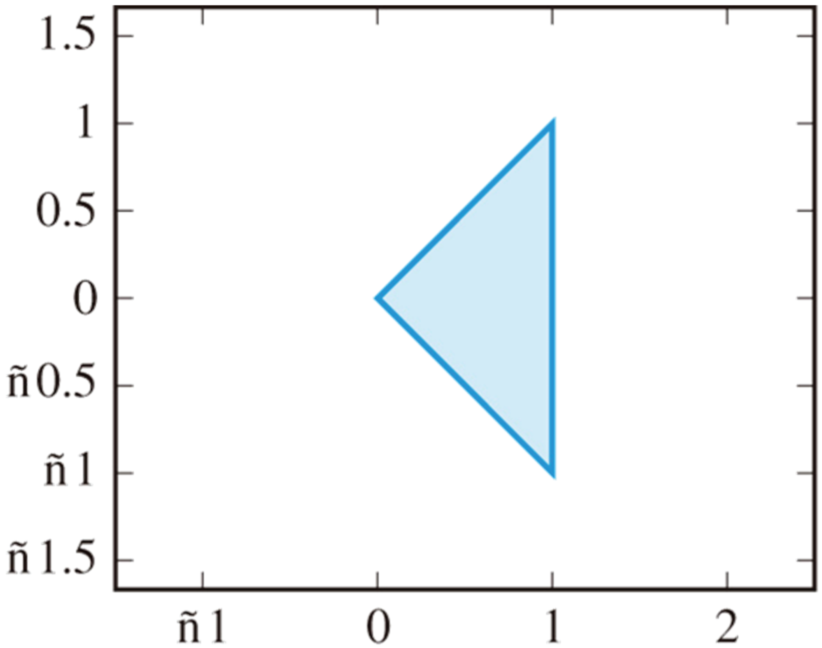
Application: Computer Graphics and Animation

- A picture with n vertices can be stored in a $2\times n$ matrix. The x coordinates of the vertices are stored in the first row and the y coordinates in the second. Each successive pair of points is connected by a straight line.

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n & x_1 \\ y_1 & y_2 & \cdots & y_n & y_1 \end{bmatrix}$$

Figure 4.2.3(a)

- An example: a triangle with 3 vertices (0, 0), (1, 1) and (1, -1) are stored in a matrix:



(a) Triangle defined by T

$$T = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

- Four primary geometric transformations that are used in computer graphics are: **dilations and contractions**, **reflection about an axis**, **rotation**, **translation**

- Four primary geometric transformations that are used in computer graphics are: **dilations and contractions**, **reflection about an axis**, **rotation**, **translation.**

1. **Dilations and Contractions:** a linear transformation of the form

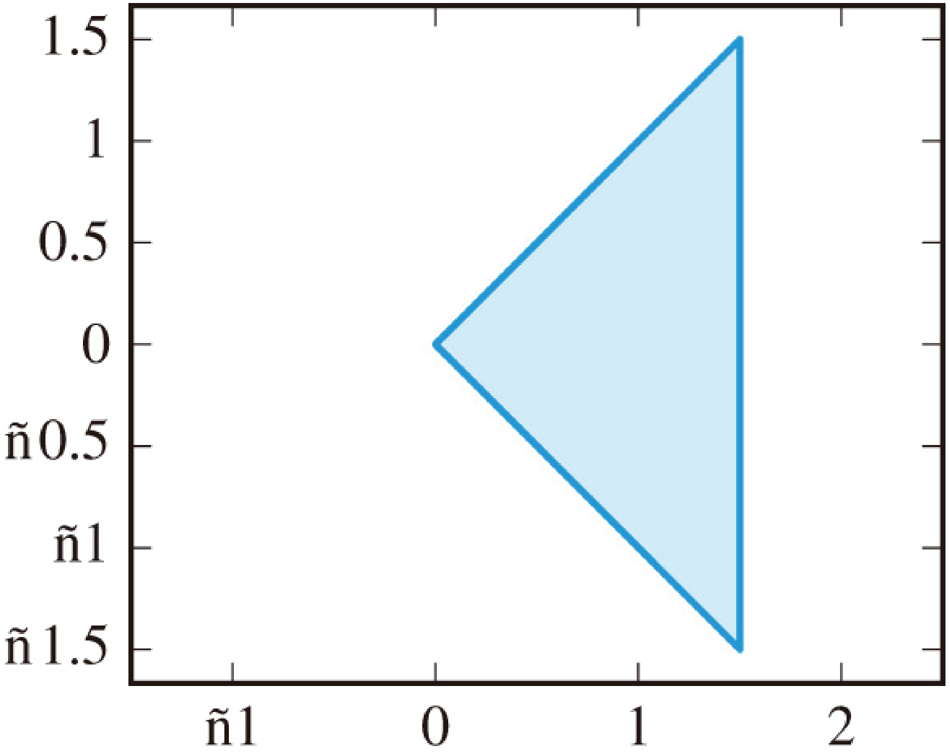
$$L(\mathbf{x}) = c\mathbf{x}$$

is a dilation if $c > 1$ and a contraction if $0 < c < 1$.

The transformation L is represented by the matrix $A = c\mathbf{I}$, where $\mathbf{I}$ is the 2×2 identity matrix.

$$AT = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & c & c & 0 \\ 0 & c & -c & 0 \end{bmatrix}$$

Figure 4.2.3(b)



(b) Dilation by factor of 1.5

- **Reflection about an axis:** If Lx is a transformation that reflects a vector about the x-axis, then $L\boldsymbol{x}$ is a linear operator and can be represented by a 2×2 matrix A.
- Since

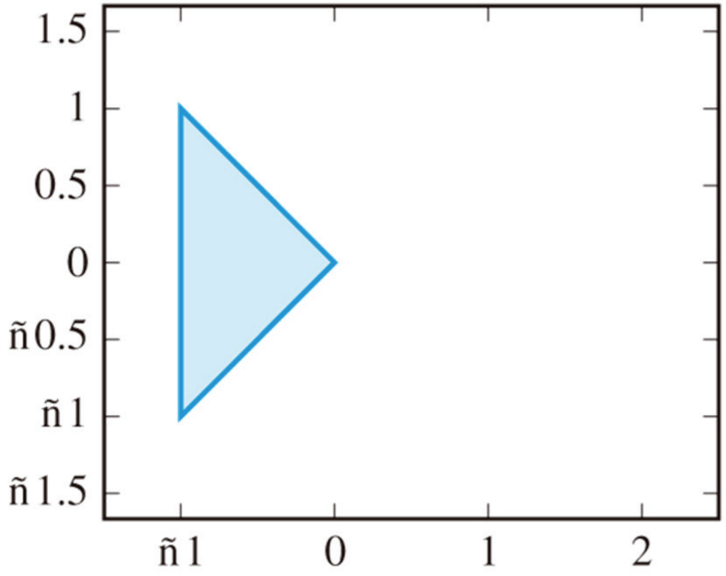
  $L_x(\mathbf{e}_1) = \mathbf{e}_1$ and $L_x(\mathbf{e}_2) = -\mathbf{e}_2$ **(Note: $L_x(\mathbf{x}) = (x_1, -x_2)^T$)**

  $\Rightarrow A = (\mathbf{a}_1, \mathbf{a}_2) = (L_x(\mathbf{e}_1), L_x(\mathbf{e}_2))$

  $= (\mathbf{e}_1, -\mathbf{e}_2) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Figure 4.2.3(c)

- Similarly, if L_y is a linear transformation that reflects a vector about the y-axis, then $\boldsymbol{L_y}$ is represented by a 2×2 matrix A.



(c) Reflection about y axis

- $\Rightarrow \quad A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

(Note: $L_y(\mathbf{x}) = (-x_1, x_2)^T$)

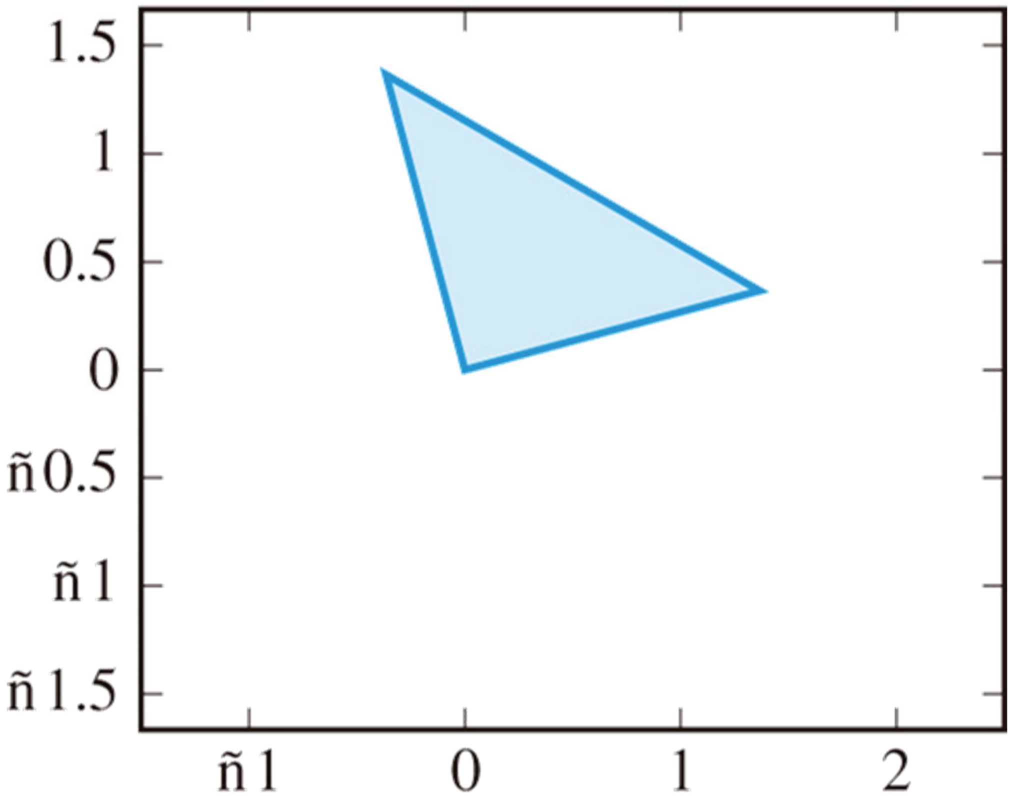
$$AT = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

- **Rotations:** Let L be a transformation that rotates a vector about the origin by an angle θ in the counterclockwise direction. In fact, L is linear transformation and that $L(\mathbf{x}) = A\mathbf{x}$ where

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

$$AT = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & \cos\theta - \sin\theta & \cos\theta + \sin\theta & 0 \\ 0 & \sin\theta + \cos\theta & \sin\theta - \cos\theta & 0 \end{bmatrix}$$

Figure 4.2.3(d)



(d) Rotation by 60°

- **Translations:** A translation by a vector $\mathbf{a}$ is a transformation of the form

$$L(\mathbf{x}) = \mathbf{x} + \mathbf{a}$$

If $\mathbf{a} \neq \mathbf{0}$, L is not a linear transformation and hence L cannot be represented by a 2×2 matrix.

Homogeneous Coordinates Systems

- **Homogeneous Coordinates Systems** – equating each vector in R^2 with a vector in R^3 having the same first two coordinates and 1 as its third coordinate:

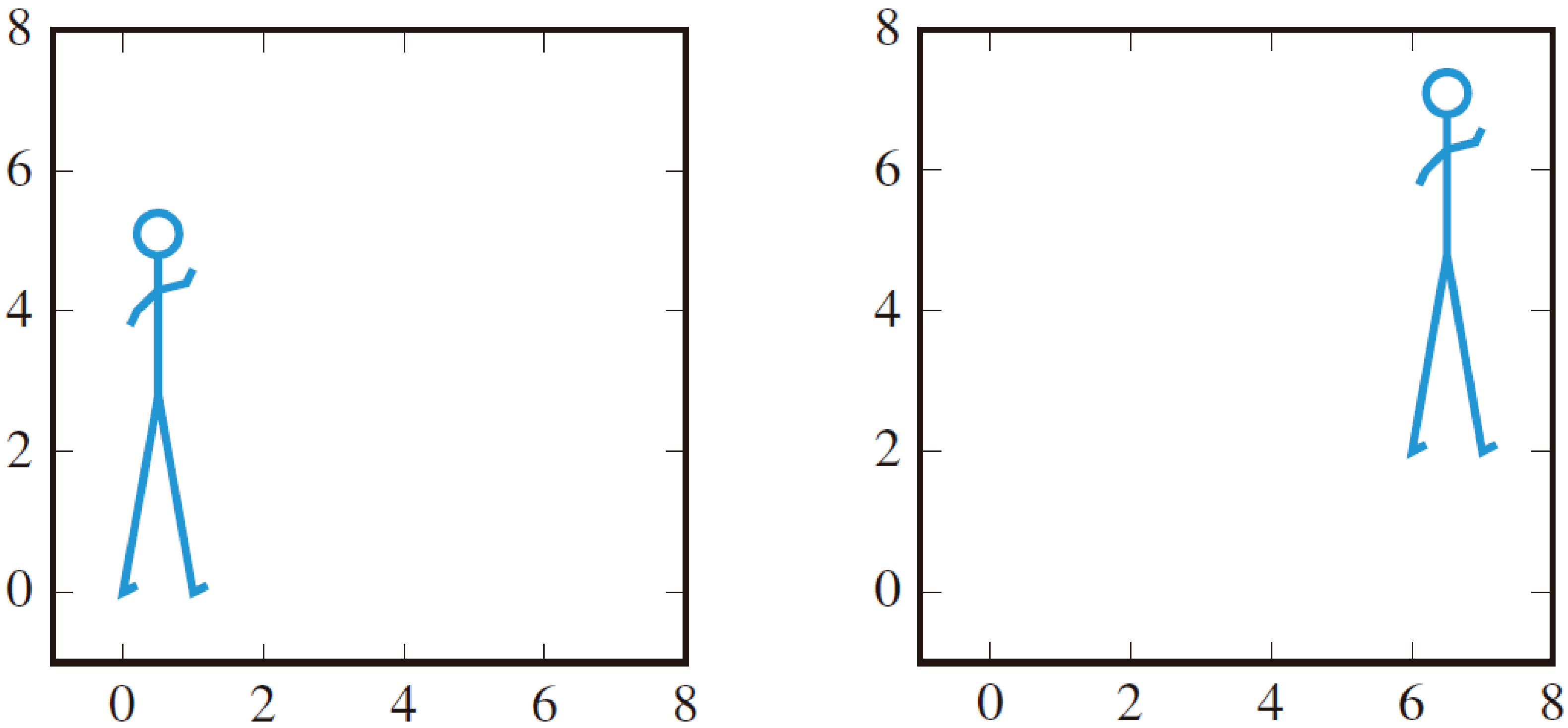
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

- The 2×2 matrix representation can be augmented into a 3×3 matrix by <u>attaching the third row and third column of the identity matrix</u> <u>***I***</u>:

$$\begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} \rightarrow \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ 1 \end{bmatrix}$$

- If L is a translation by a vector **a** in R^2, the matrix representation can be formed by replacing the first two entries in the last column of the identity matrix ***I*** with the entries of **a**:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & a_1 \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & a_1 \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 + a_1 \\ x_2 + a_2 \\ 1 \end{bmatrix}$$

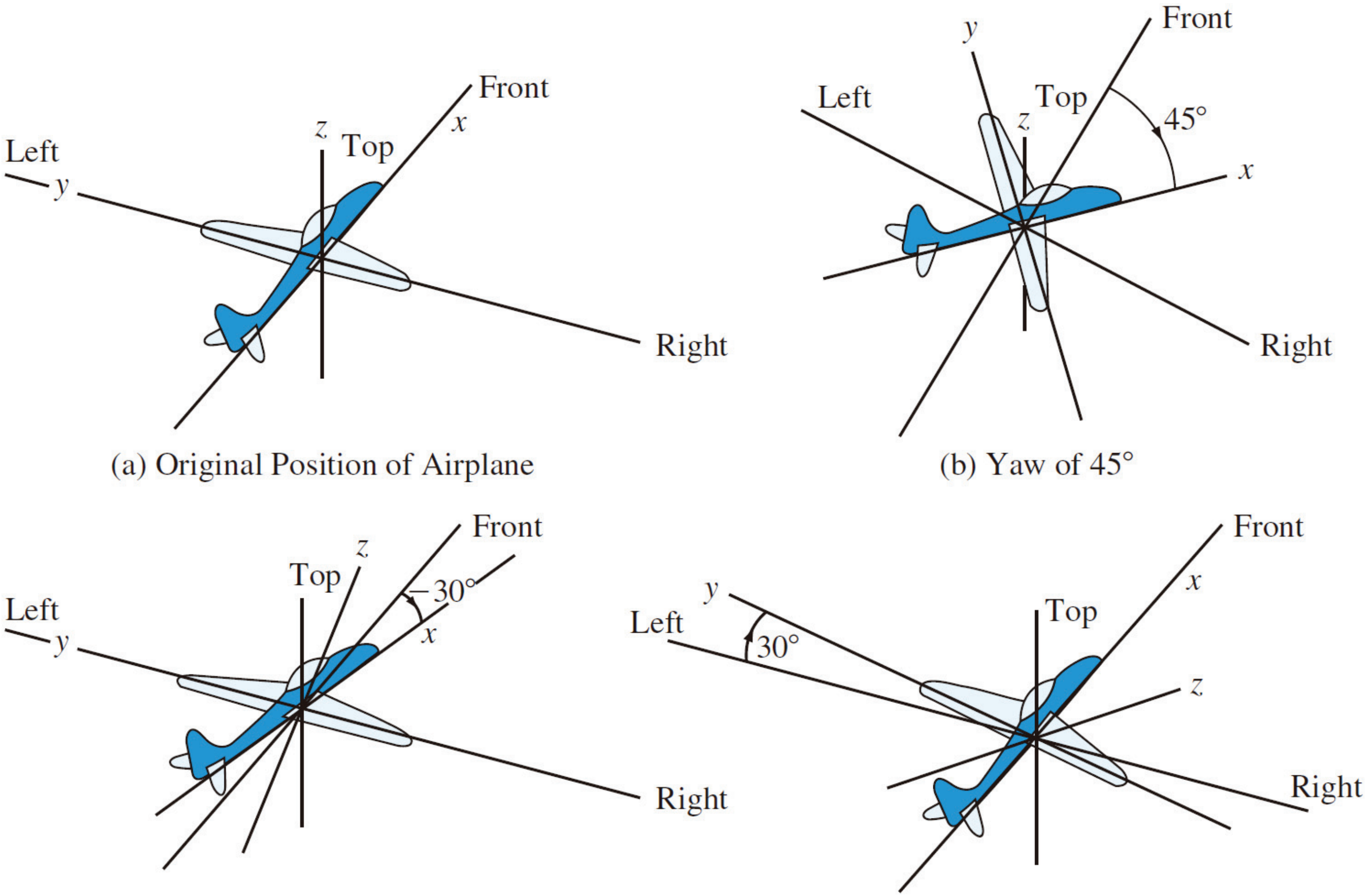


(a) Graph of 3×81 matrix S

(b) Graph of translated figure AS

Figure 4.2.4.

APPLICATION 2



(a) Original Position of Airplane

(b) Yaw of 45°

(c) Pitch of −30°

(d) Roll of 30°

Figure 4.2.5.

4.3 Similarity

- **Example:** Let L be the linear transformation mapping R^2 into itself defined by

$$L(\mathbf{x}) = (2x_1, x_1 + x_2)^T$$

Since $L(\mathbf{e}_1) = L(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $L(\mathbf{e}_2) = L(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Thus, the matrix representing L w. r. t. $[\mathbf{e}_1, \mathbf{e}_2]$ is

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$$

- If we use $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ as the basis for R^2, then

$$L(\mathbf{u}_1) = A\mathbf{u}_1 = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$L(\mathbf{u}_2) = A\mathbf{u}_2 = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

- Since the transition matrix from $[\mathbf{u}_1, \mathbf{u}_2]$ to $[\mathbf{e}_1, \mathbf{e}_2]$ is

$$U = (\mathbf{u}_1, \mathbf{u}_2) = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

- So the transition matrix from $[\mathbf{e}_1, \mathbf{e}_2]$ to $[\mathbf{u}_1, \mathbf{u}_2]$ is

$$U^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

- Then, the coordinates of $L(\mathbf{u}_1)$ and $L(\mathbf{u}_2)$ w. r. t. $[\mathbf{u}_1, \mathbf{u}_2]$ is

$$U^{-1}L(\mathbf{u}_1) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$U^{-1}L(\mathbf{u}_2) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$\Rightarrow$ $L(\mathbf{u}_1) = 2\mathbf{u}_1 + 0\mathbf{u}_2$

$L(\mathbf{u}_2) = -1\mathbf{u}_1 + 1\mathbf{u}_2$

$\Rightarrow$ The matrix representing L w. r. t. $[\mathbf{u}_1, \mathbf{u}_2]$ is $B = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$

How are A and B related?

- Since $\begin{bmatrix} 2 \\ 0 \end{bmatrix} = U^{-1}L(\mathbf{u}_1) = U^{-1}\mathrm{A}\mathbf{u}_1$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix} = U^{-1}L(\mathbf{u}_2) = U^{-1}\mathrm{A}\mathbf{u}_2$

Hence, $B = (U^{-1}A\mathbf{u}_1, \mathrm{U}^{-1}A\mathbf{u}_2) = \mathrm{U}^{-1}A(\mathbf{u}_1, \mathbf{u}_2) = \mathrm{U}^{-1}A\mathrm{U}$

Conclusion

- If (i) B is the matrix representing L w. r. t. $[\mathbf{u}_1, \mathbf{u}_2]$
  (ii) A is the matrix representing L w. r. t. $[\mathbf{e}_1, \mathbf{e}_2]$
  (iii) U is the transition matrix corresponding to the change of basis from

  $$[\mathbf{u}_1, \mathbf{u}_2] \text{ to } [\mathbf{e}_1, \mathbf{e}_2]$$

  then $B = U^{-1}AU$.

Theorem 4.3.1

Let $E = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ and $F = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n]$ be two ordered bases for a vector space V and let L be a linear operator on V. Let S be the transition matrix representing the change from F to E. If A is the matrix representing L w. r. t. E and B is the matrix representing L w. r. t. F, then $B = S^{-1}AS$.

Theorem 4.3.1 *proof*

Basis E: $[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$

$$
\begin{tikzcd}
\mathbf{y} \arrow[r, "A"] & \mathbf{t} \arrow[d, "S^{-1}"] \\
\mathbf{x} \arrow[u, "S"] \arrow[r, "B"'] & \mathbf{z}
\end{tikzcd}
$$

Basis F: $[\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n]$

- Let $\mathbf{x}$ be any vector in R^n and let $\mathbf{x} = [\mathbf{v}]_F$

$$\mathbf{v} = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \ldots + x_n\mathbf{w}_n$$

- Since S is the transition matrix representing the change from F to E

  Let $\mathbf{y} = S\mathbf{x}$, $\mathbf{t} = A\mathbf{y}$, $\mathbf{z} = B\mathbf{x}$

  Let $\mathbf{x}$ be any vector in R^n and let $\mathbf{y} = [\mathbf{v}]_E$

$$\Rightarrow \mathbf{v} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \ldots + y_n\mathbf{v}_n$$

- Since A represents L w. r. t. E and B represents L w. r. t. F, we have

$$\mathbf{t} = [L(\mathbf{v})]_E \text{ and } \mathbf{z} = [L(\mathbf{v})]_F$$

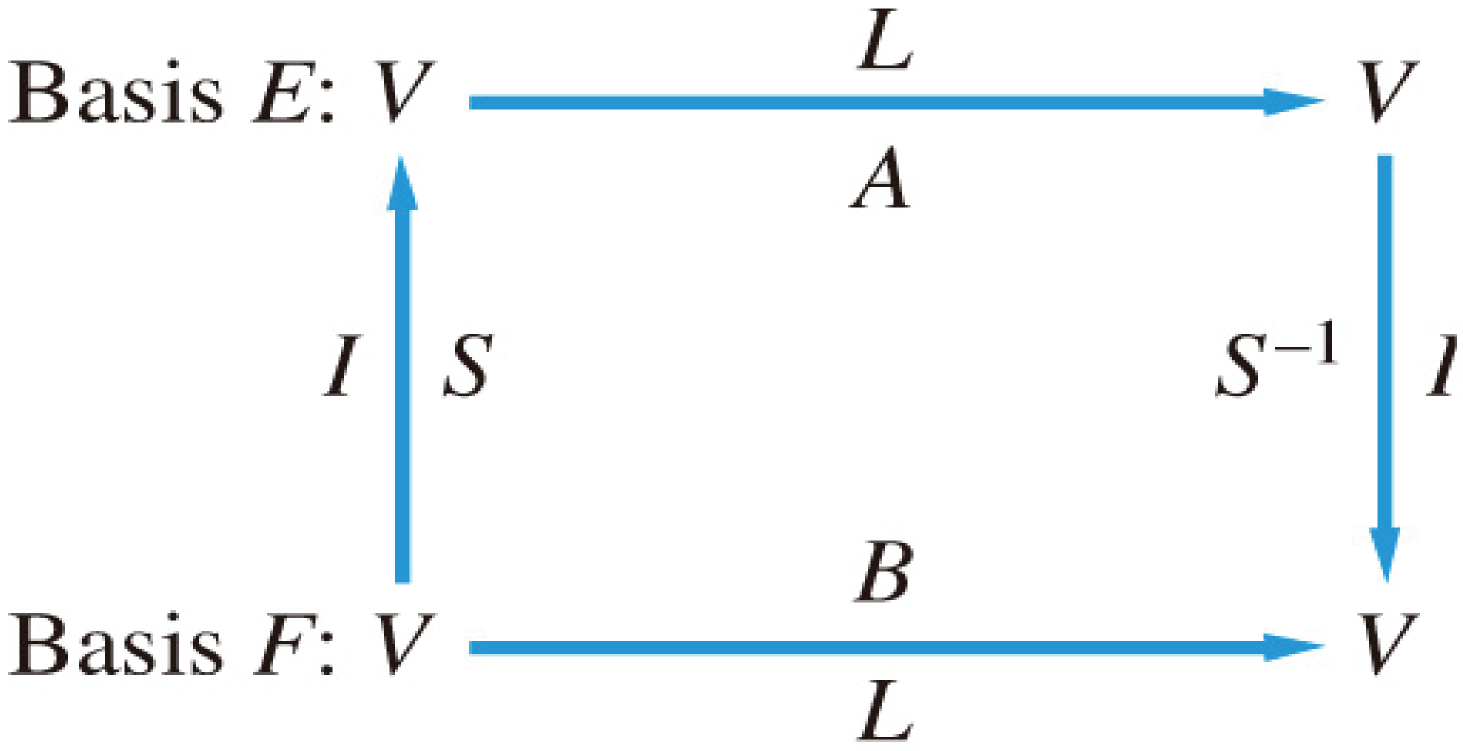
- Since the transition from E to F is S^{-1}, therefore

$$S^{-1}\mathbf{t} = \mathbf{z}$$

$$\Rightarrow S^{-1}\mathbf{t} = S^{-1}A\mathbf{y} = S^{-1}AS\mathbf{x} = \mathbf{z} = B\mathbf{x}$$

$$\Rightarrow S^{-1}AS = B$$

Figure 4.3.2



Definition

Let A and B be two $n \times n$ matrices. B is said to be **similar** to A if there exists a nonsingular matrix S such that $B = S^{-1}AS$.

Note

- $B = S^{-1}AS$

  $\Leftrightarrow SBS^{-1} = S\,(S^{-1}AS)\,S^{-1}$

  $\Leftrightarrow SBS^{-1} = A$

  $\Leftrightarrow A = (S^{-1})^{-1}B(S^{-1})$
- That is, A is similar to B, we may say that A and B are similar matrices.

Example 1

- Let D be the differentiation operator on P_3. Find the matrix B representing D with respect to $[1, x, x^2]$ and the matrix A representing D with respect to $[1, 2x, 4x^2 - 2]$.
- *Sol:*

$$D(1) = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2$$

$$D(x) = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2$$

$$D(x^2) = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2$$

- The matrix B is then given by $B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

- Apply D to 1, $2x$, and $4x^2 - 2$, we obtain

$$D(1) = 0 \cdot 1 + 0 \cdot 2x + 0 \cdot (4x^2 - 2)$$

$$D(2x) = 2 \cdot 1 + 0 \cdot 2x + 0 \cdot (4x^2 - 2)$$

$$D(4x^2 - 2) = 0 \cdot 1 + 4 \cdot 2x + 0 \cdot (4x^2 - 2)$$

- Thus

$$A = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

- The transition matrix S corresponding to the change of basis from $[1, 2x, 4x^2 - 2]$ to $[1, x, x^2]$ and its inverse are given by

$$S = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \quad \text{and} \quad S^{-1} = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$$

Example 2

- L: $R^3 \rightarrow R^3$ defined by $L(\mathbf{x}) = A\mathbf{x}$, where $A = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix}$
- Thus the matrix A represents L with respect to [$\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$], find the matrix representing L with respect to [$\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$] where

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{y}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \mathbf{y}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

- *Sol:*

$$L(\mathbf{y}_1) = A\mathbf{y}_1 = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0\mathbf{y}_1 + 0\mathbf{y}_2 + 0\mathbf{y}_3$$

$$L(\mathbf{y}_2) = A\mathbf{y}_2 = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = 0\mathbf{y}_1 + 1\mathbf{y}_2 + 0\mathbf{y}_3$$

$$L(\mathbf{y}_3) = A\mathbf{y}_3 = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = 0\mathbf{y}_1 + 0\mathbf{y}_2 + 4\mathbf{y}_3$$

- Thus the matrix representing L with respect to [$\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$] is

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

- Try to verify $D = S^{-1}AS$

S: the transition matrix from $[\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3]$ to $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$

$$\Rightarrow S = (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3) = \begin{bmatrix} 1 & -2 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\Rightarrow S^{-1}AS = \begin{bmatrix} 0 & -1 & 1 \\ \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} = D$$