Ch. 4 Linear Transformation

4.1 Definition & Examples

Def. A linear transformation (or a linear mapping) L is a mapping from a vector space V into a vector space W (we write $L: V \longrightarrow W$) that has the following property:

$$\boxed{L(\alpha v_1 + \beta v_2) = \alpha L(v_1) + \beta L(v_2)} \qquad \forall\, v_1, v_2 \in V \text{ and } \forall\, \alpha, \beta \in \mathbb{R}$$

we say that the mapping L preserves the operations of addition and scalar multiplication. In particular, we have:

(1) $L(v_1 + v_2) = L(v_1) + L(v_2)$ (put $\alpha = \beta = 1$)

(2) $L(\alpha v) = \alpha \cdot L(v)$ (put $v_1 = v$, $\beta = 0$)

We then have: L is a linear transformation iff L satisfies (1) and (2).

In the case when $V = W$, the linear transformation $L: V \rightarrow V$ is reffered to as a linear operator on V. In other words, a linear operator is a linear transformation that maps a vector space V into itself.

Ex. 1 Consider the mapping $L(\mathbf{x}) = 2\mathbf{x}$ $\quad \forall\, \mathbf{x} \in \mathbb{R}^2$.

L is a linear operator on $\mathbb{R}^2$, because:

1) L maps $V = \mathbb{R}^2$ into $W = \mathbb{R}^2$ $\quad (V = W = \mathbb{R}^2)$

2) L satisfies (1) and (2):

$L(\mathbf{x} + \mathbf{y}) = 2(\mathbf{x} + \mathbf{y}) = 2\mathbf{x} + 2\mathbf{y} = L(\mathbf{x}) + L(\mathbf{y})$ $\quad \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^2$

$L(\alpha \mathbf{x}) = 2(\alpha \mathbf{x}) = \alpha(2\mathbf{x}) = \alpha L(\mathbf{x})$ $\quad \forall\, \mathbf{x} \in \mathbb{R}^2, \forall\, \alpha \in \mathbb{R}$.

$\boxed{\text{EX. 2}}$ consider the mapping $L(\mathbf{x}) = (-x_2, x_1)^T \quad \forall\, \mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$.

L is a linear operator on $\mathbb{R}^2$, because:

1) L maps $V = \mathbb{R}^2$ into $W = \mathbb{R}^2$ $\quad (V = W = \mathbb{R}^2)$

2) L satisfies (1) and (2):

$$L(\mathbf{x}+\mathbf{y}) = \begin{bmatrix} -(x_2+y_2) \\ x_1+y_1 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} + \begin{bmatrix} -y_2 \\ y_1 \end{bmatrix} = L(\mathbf{x}) + L(\mathbf{y}) \qquad \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^2$$

$$L(\alpha \mathbf{x}) = \begin{bmatrix} -(\alpha x_2) \\ \alpha x_1 \end{bmatrix} = \alpha \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} = \alpha L(\mathbf{x}) \qquad \forall\, \mathbf{x} \in \mathbb{R}^2,\ \forall\, \alpha \in \mathbb{R}.$$

NOTE: In item 2 above, instead of showing properties (1) and (2), we can directly show $L(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$, as follows:

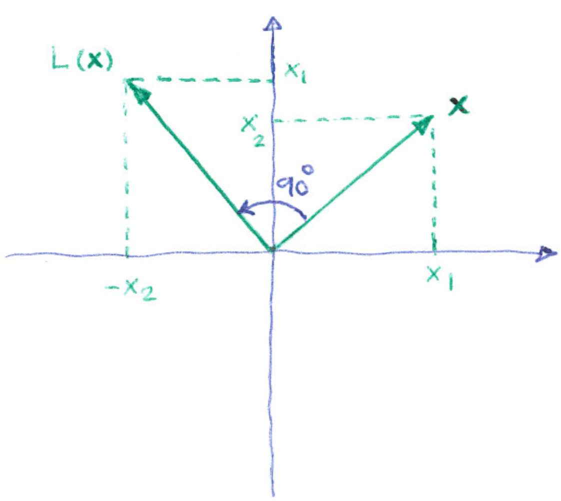
$$L(\alpha\mathbf{x} + \beta\mathbf{y}) = \begin{bmatrix} -(\alpha x_2 + \beta y_2) \\ \alpha x_1 + \beta y_1 \end{bmatrix} = \begin{bmatrix} -\alpha x_2 \\ \alpha x_1 \end{bmatrix} + \begin{bmatrix} -\beta y_2 \\ \beta y_1 \end{bmatrix} =$$

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

$\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^2,\ \forall\, \alpha, \beta \in \mathbb{R}$

Geometric interpretation of EX. 2:

The linear operator $L: \mathbb{R}^2 \to \mathbb{R}^2$ rotates vector $\mathbf{x} = (x_1, x_2)^T$ by 90° in the counterclockwise direction and generates vector $L(\mathbf{x}) = (-x_2, x_1)^T$.



$\boxed{\text{Ex. 3}}$ Consider the mapping $L(\mathbf{x}) = x_1 + x_2 \quad \forall\, \mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$.

L is a linear transformation from $V = \mathbb{R}^2$ into $W = \mathbb{R}$, because:

$$\begin{aligned} L(\alpha \mathbf{x} + \beta \mathbf{y}) &= (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \\ &= \alpha (x_1 + x_2) + \beta (y_1 + y_2) \\ &= \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \qquad \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^2, \forall\, \alpha, \beta \in \mathbb{R}. \end{aligned}$$

we write: $L : \mathbb{R}^2 \longrightarrow \mathbb{R}$.

$\boxed{\text{Ex. 4}}$ Consider the mapping $L(\mathbf{x}) = \sin(x_1 + x_2) \quad \forall\, \mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$.

Although L maps $\mathbb{R}^2$ into $\mathbb{R}$, it is <u>not</u> a linear transformation, because:

$$L(\alpha \mathbf{x}) = \sin(\alpha x_1 + \alpha x_2) \neq \alpha \cdot \sin(x_1 + x_2) = \alpha L(\mathbf{x}).$$

$$L(\mathbf{x} + \mathbf{y}) = \sin(x_1 + y_1 + x_2 + y_2) \neq \sin(x_1 + x_2) + \sin(y_1 + y_2) = L(\mathbf{x}) + L(\mathbf{y}).$$

$\boxed{\text{Ex. 5}}$ Consider the mapping $L(\mathbf{x}) = A\mathbf{x} \quad \forall\, \mathbf{x} \in \mathbb{R}^n$ wher $A \in \mathbb{R}^{m \times n}$.

L is a linear transformation from $V = \mathbb{R}^n$ into $W = \mathbb{R}^m$, because:

1) It takes a vector $\mathbf{x}$ in $\mathbb{R}^n$ and generates a vector $A\mathbf{x}$ in $\mathbb{R}^m$.

2) $L(\alpha \mathbf{x} + \beta \mathbf{y}) = A(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha A \mathbf{x} + \beta A \mathbf{y} = \alpha L(\mathbf{x}) + \beta L(\mathbf{y}) \quad \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \forall\, \alpha, \beta \in \mathbb{R}$.

We write $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

Sometimes, L is denoted by L_A.

<u>NOTE</u>: If $n = m$, then L will be an operator, since $V = W = \mathbb{R}^n$.

In <u>Ex. 5</u>, what happens when $\mathbf{x} = \mathbf{0} \in \mathbb{R}^n$?

$$L(\mathbf{x}) = L(\mathbf{0}) = A\mathbf{0} = \mathbf{0} \in \mathbb{R}^m.$$

Def. The identity operator $\mathcal{I}$ on a vector space is defined as

$$\mathcal{I}(v) = v \qquad \forall v \in V.$$

$\mathcal{I}$ takes v as input and generates v as output.

It is easy to show that $\mathcal{I}: V \longrightarrow V$ is a linear operator:

$$\mathcal{I}(\alpha v_1 + \beta v_2) = \alpha v_1 + \beta v_2 = \alpha \mathcal{I}(v_1) + \beta \mathcal{I}(v_2) \qquad \forall v_1, v_2 \in V.$$

$\boxed{\text{Ex. 6}}$ Consider the mapping $L(f) = \int_a^b f(x)\,dx \qquad \forall f \in C[a,b]$.

L is a linear transformation from $V = C[a,b]$ into $W = \mathbb{R}$, because:

$$\begin{aligned} L(\alpha f + \beta g) &= \int_a^b [\alpha f(x) + \beta g(x)]\,dx \\ &= \int_a^b \alpha f(x)\,dx + \int_a^b \beta g(x)\,dx \\ &= \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx \\ &= \alpha L(f) + \beta L(g) \qquad \forall f, g \in C[a,b],\ \forall \alpha, \beta \in \mathbb{R}. \end{aligned}$$

$\boxed{\text{Exercise}}$ What about $L(f) = \int_a^b |f(x)|^2\,dx \qquad \forall f \in C[a,b]$?

$\boxed{\text{Ex. 7}}$ Consider the mapping $D(f) = f' \qquad \forall f \in C^1[a,b]$.

L is a linear transformation from $V = C^1[a,b]$ into $W = C[a,b]$, because:

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

$$\forall f, g \in C^1[a,b],\ \forall \alpha, \beta \in \mathbb{R}.$$

Image and Kernel

Def. Let $L: V \to W$ be a linear transformation.

The kernel of L is $\ker(L) = \{ v \in V \mid L(v) = 0_W \}$

NOTE: 0_W means the zero vector in the vector space W.

For example, if $W = \mathbb{R}^2$, then $0_W = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, and if $W = \mathbb{R}^{2\times 2}$, then $0_W = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Def. Let $L: V \to W$ be a linear transformation, and let S be a subspace of V.

The image of S is $L(S) = \{ w \in W \mid w = L(v),\ v \in S \}$

The image of the entire vector space V is called the range of L.
We write $L(V) = \text{range}(L)$

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

1. $\ker(L)$ is a subspace of V.
2. $L(S)$ is a subspace of W.
3. $L(V)$ is a subspace of W.

proof for 1:

1) $\ker(L)$ is a subset of V.

2) $\ker(L)$ is non-empty because $0_V \in \ker(L)$.

3) $\ker(L)$ is closed under scalar multiplication:

Let $v \in \ker(L)$ and $\alpha \in \mathbb{R}$. Then $L(\alpha v) = \alpha L(v) = \alpha \cdot 0_W = 0_W \Rightarrow \alpha v \in \ker(L)$.

4) $\ker(L)$ is closed under addition:

Let $v_1, v_2 \in \ker(L)$. Then $L(v_1 + v_2) = L(v_1) + L(v_2) = 0_W + 0_W = 0_W \Rightarrow v_1 + v_2 \in \ker(L)$.

---

proof of 2 is similar.

3 is a special case of 2.

$\boxed{\text{Ex. 8}}$ Consider the mapping $L: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

$$L(x) = (x, 0)^T \qquad \forall\, x = (x_1, x_2)^T \in \mathbb{R}^2$$

1) is L a linear transformation ?

2) Find $\ker(L)$.

3) Find range of L.

---

1) L is a linear transformation from $V = \mathbb{R}^2$ into $W = \mathbb{R}^2$, because

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

$\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^2, \ \forall\, \alpha, \beta \in \mathbb{R}$.

2) $\ker(L) = \{ \mathbf{x} \in \mathbb{R}^2 \mid L(\mathbf{x}) = \mathbf{0} \}$

For $L(\mathbf{x}) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$ to be $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, we need $x_1 = 0$.

this means $\ker(L) = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;\middle|\; x_1 = 0 \right\}$.

Hence, $\ker(L)$ contains all vectors of form $\begin{pmatrix} 0 \\ x_2 \end{pmatrix}$.

In other words, $\ker(L)$ is a subspace of $\mathbb{R}^2$ spanned by $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

3) $\text{range}(L) = \{ \mathbf{y} \in W = \mathbb{R}^2 \mid \mathbf{y} = L(\mathbf{x}), \ \mathbf{x} \in V = \mathbb{R}^2 \}$

Hence, range(L) contains vectors of form $\begin{pmatrix} x_1 \\ 0 \end{pmatrix}$.

In other words, range(L) is a subspace of $W = \mathbb{R}^2$ spanned by $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

<u>Remark</u>. This is an example where range$(L) \neq W$.

$\boxed{\text{Ex. 9}}$ Let $L: \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation defined by

$$L(\mathbf{x}) = (x_1 + x_2,\ x_2 + x_3)^T \qquad \forall\ \mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3.$$

1) Find $\ker(L)$.

2) If S is a subspace of $\mathbb{R}^3$ spanned by $\mathbf{e}_1$ and $\mathbf{e}_3$, Find the image of S.

3) Find range (L).

---

1) $\ker(L) = \{ \mathbf{x} \in \mathbb{R}^3 \mid L(\mathbf{x}) = \mathbf{0} \}$.

Hence, if $\mathbf{x} \in \ker(L)$, then $\begin{cases} x_1 + x_2 = 0 \\ x_2 + x_3 = 0 \end{cases} \xrightarrow{\text{Free variable } x_3 = a \in \mathbb{R}} \mathbf{x} = \begin{pmatrix} a \\ -a \\ a \end{pmatrix} = a \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$

Therefore, $\ker(L)$ is a subspace of $\mathbb{R}^3$ spanned by $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

---

2) If $\mathbf{x} \in S$, then $\mathbf{x}$ must be of the form $\begin{pmatrix} a \\ 0 \\ b \end{pmatrix}$, and hence $L(\mathbf{x}) = \begin{pmatrix} a \\ b \end{pmatrix}$.

Hence, $L(S) = \{ \mathbf{y} \in W \mid \mathbf{y} = \begin{pmatrix} a \\ b \end{pmatrix} \} = \mathbb{R}^2$

3) Since the image of S is the whole $\mathbb{R}^2$, the range of L is also $\mathbb{R}^2$:

$L(\mathbb{R}^3) = \mathbb{R}^2 = W$.

<u>Remark</u>. whenever range$(L) = W$, we say the transformation L is <u>onto</u>.

$\boxed{\text{Exercise}}$ Consider the mapping $D: P_3 \to P_2$ with $D(p(x)) = p'(x) \quad \forall\ p \in P_3$.

1) Show that D is a linear transformation.

2) Find $\ker(D)$.

3) Find range (D).

Answers: 2) $\ker(D) = P_1$

3) range $(D) = P_2$

$\boxed{\text{EX. 10}}$ Let $L(v) = Av$ be a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^3$

with $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 9 \end{pmatrix}$.

1) Find a basis for $\ker(L)$.

2) Find a basis for range (L).

---

1) $\ker(L) = \{ \mathbf{x} \in \mathbb{R}^2 \mid L(\mathbf{x}) = \mathbf{0} \}$

This means that we need to find $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ such that $A\mathbf{x} = \mathbf{0}$.

This means we need to find the null space of matrix A.

$$A\mathbf{x} = \mathbf{0}: \quad \left(\begin{array}{cc|c} 1 & 3 & 0 \\ 2 & 6 & 0 \\ 3 & 9 & 0 \end{array}\right) \xrightarrow{\text{elimination}} \underbrace{\left(\begin{array}{cc|c} 1 & 3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)}_{\text{RE form}} \Rightarrow x_1 + 3x_2 = 0$$

free variable: $x_2 = \alpha \Rightarrow x_1 = -3\alpha \Rightarrow \mathbf{x} = \begin{pmatrix} -3\alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} -3 \\ 1 \end{pmatrix}$

$\Rightarrow \ker(L) = N(A) = \text{span}((-3, 1)^T)$.

2) $\text{range}(L) = \{ \mathbf{y} \in \mathbb{R}^3 \mid \mathbf{y} = A\mathbf{x},\ \mathbf{x} \in \mathbb{R}^2 \}$

Hence, rang(L) contains vectors of form $\mathbf{y} = \begin{pmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 9 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = a \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + b \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$.

This means that the columns of A span range(L).

It is clear that the second column is 3 times the first column. Hence, the two columns are linearly independent, and only one of them contributes. Hence, $\text{range}(L) = C(A) = \text{span}((1, 2, 3)^T)$.

Remark. Since $W = \mathbb{R}^3$, then this example shows that range(L) is not necessarily equal to W. Here, dimension of range(L) is one, while dimension of $W = \mathbb{R}^3$ is three. So, we have range$(L) \subset W$.

Hence, the transformation L in Ex. 10 is NOT onto.

4.2 Matrix representation of linear transformation

In preveious section, we saw that a matrix $A \in \mathbb{R}^{m\times n}$ defines a linear transformation L_A from $\mathbb{R}^n$ to $\mathbb{R}^m$, where $L_A(x) = Ax$, $\forall x \in \mathbb{R}^n$.

In this section, we will see that for each linear transformation $L: \mathbb{R}^n \to \mathbb{R}^m$ there is a matrix $A \in \mathbb{R}^{m\times n}$ such that $L(x) = Ax$.

Theorem If $L: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, there is a matrix $A \in \mathbb{R}^{m\times n}$ such that: $L(x) = Ax$, $\forall x \in \mathbb{R}^n$.

The jth column of A is $a_j = L(e_j)$ for $j = 1, \cdots, n$.

Let us show that $L(e_j) = a_j$ through an example:

Consider $A = \begin{pmatrix} 1 & -2 \\ 2 & 4 \\ 3 & -6 \end{pmatrix}$.

The 1st column of A is $(1,2,3)^T$.

We also have $L(e_1) = Ae_1 = \begin{pmatrix} 1 & -2 \\ 2 & 4 \\ 3 & -6 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

Hence, $a_1 = L(e_1)$.

Similarly we have $a_2 = L(e_2)$.

The above theorem tells us how to construct A from L.

$\boxed{\text{Ex. 11}}$ Let the linear transformation $L: \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$L(x) = (x_1 + x_2,\ x_2 + x_3)^T \quad \forall \bar{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3.$$

Find a matrix A such that $L(x) = Ax$, $\forall x \in \mathbb{R}^3$.

We first note that $A \in \mathbb{R}^{2\times 3}$. So, A has 3 columns: $A = [a_1\ a_2\ a_3]$.

$$\begin{cases} a_1 = L(e_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ a_2 = L(e_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ a_3 = L(e_3) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{cases} \Rightarrow A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

So far, we have seen that matrices can be used to represent linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Can we use matrices to represent linear transformation from any n-dimensional vector space V to any m-dimensional vector space W?

Matrix Representation Theorem

Assume that

$E = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis for V.

$F = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ is a basis for W.

$L : V \longrightarrow W$ is a linear transformation.

Then, corresponding to L, there is $A \in \mathbb{R}^{m\times n}$ such that

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

We also have: $\mathbf{a}_j = [L(\mathbf{v}_j)]_F$.

Recall:

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

this means: $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n$

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

This means: $\mathbf{w} = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \cdots + y_m\mathbf{w}_m$

If we let $\mathbf{w} = L(\mathbf{v})$, then the theorem states that $\mathbf{y} = A\mathbf{x}$.

In other words, if $\mathbf{x}$ is the coordinate vector of $\mathbf{v}$ with respect to E, then the coordinate vector of $L(\mathbf{v})$ with respect to F is $[L(\mathbf{v})]_F = \mathbf{y} = A\mathbf{x}$.

The formula $\mathbf{a}_j = [L(\mathbf{v}_j)]_F$ helps us to find the columns of matrix A.

To find $\mathbf{a}_j$, we need to represent $L(\mathbf{v}_j)$ as a linear combination of $\mathbf{w}_1, \dots, \mathbf{w}_m$:

$$L(\mathbf{v}_j) = \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \cdots + \alpha_m\mathbf{w}_m \Longrightarrow \mathbf{a}_j = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$$

$\boxed{\text{EX. 12}}$ Let $L : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$L(\mathbf{x}) = x_1 \mathbf{b}_1 + (x_2 + x_3)\mathbf{b}_2 \qquad \forall\, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \text{ with } \mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{b}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Find the matrix A representing L with respect to the bases $E = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $F = \{\mathbf{b}_1, \mathbf{b}_2\}$.

We first note that $A \in \mathbb{R}^{2\times3}$. Let $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3]$.

$\mathbf{a}_1 = [L(\mathbf{e}_1)]_F$ wher $\mathbf{e}_1 = (1,0,0)^T$

we have $L(\mathbf{e}_1) = 1 \times \mathbf{b}_1 + 0 \times \mathbf{b}_2 \implies \mathbf{a}_1 = [L(\mathbf{e}_1)]_F = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$\mathbf{a}_2 = [L(\mathbf{e}_2)]_F$ where $\mathbf{e}_2 = (0,1,0)^T$

we have $L(\mathbf{e}_2) = 0 \times \mathbf{b}_1 + 1 \times \mathbf{b}_2 \implies \mathbf{a}_2 = [L(\mathbf{e}_2)]_F = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$\mathbf{a}_3 = [L(\mathbf{e}_3)]_F$ where $\mathbf{e}_3 = (0,0,1)^T$

we have $L(\mathbf{e}_3) = 0 \times \mathbf{b}_1 + 1 \times \mathbf{b}_2 \implies \mathbf{a}_3 = [L(\mathbf{e}_3)]_F = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Hence, $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

$\boxed{\text{EX. 13}}$ Let $D : P_3 \to P_2$ be defined by $D(p) = p'$ $\forall\, p \in P_3$.

Find the matrix representation for D with respect to the bases $E = \{x^2, x, 1\}$ and $F = \{x, 1\}$

We know that $A \in \mathbb{R}^{2\times3}$. Let $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3]$.

$D(x^2) = 2x \implies D(x^2) = 2 \times x + 0 \times 1 \implies [D(x^2)]_F = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \mathbf{a}_1$

$D(x) = 1 \implies D(x) = 0 \times x + 1 \times 1 \implies [D(x)]_F = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{a}_2$

$D(1) = 0 \implies D(1) = 0 \times x + 0 \times 1 \implies [D(1)]_F = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{a}_3$

Hence, $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

Theorem Let
$E = \{u_1, \dots, u_n\}$ be basis for $\mathbb{R}^n$
$F = \{b_1, \dots, b_m\}$ be basis for $\mathbb{R}^m$
$L: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation
A be the representive matrix for L with respect to E and F.

Then,

$$a_j = B^{-1} L(u_j) \quad , \quad B = (b_1, \dots, b_m) \quad , \quad j = 1, 2, \dots, n.$$

---

proof: We have $L(u_j) = a_{1j} b_1 + a_{2j} b_2 + \dots + a_{mj} b_m = B a_j$

Since B is nonsingular, then $a_j = B^{-1} L(u_j)$.

Application to Computer Graphics & Animation

A two-dimensional picture is stored in computer as a set of vertices or nodes, connected by lines.

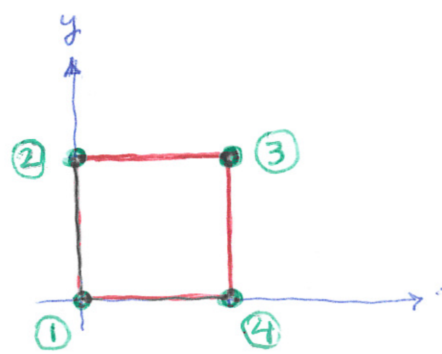
If there are N vertices, they are stored in a $2 \times N$ matrix, whose first row contains the x-coordinates of vertices and second row contains the y-coordinates. We will also need another matrix to determine which nodes are connected by straight lines.

For example, to generate a unit square, we need $N = 4$ vertices $(0,0)$, $(0,1)$, $(1,1)$, $(1,0)$ and we store them in a 2×4 matrix:

$$X = \begin{array}{c} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \\ \;\;\;1 \quad\; 2 \quad\; 3 \quad\; 4 \end{array}.$$

then, we need to connect the following nodes:

node 1 to node 2
node 2 to node 3
node 3 to node 4
node 4 to node 1



we can store this in a matrix like the following:

$$Z = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{bmatrix}.$$

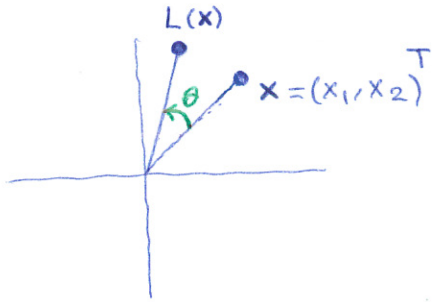
Animation

We can transform a figure by changing the positions of the vertices and then draw the transformed figure. A succession of such drawings will produce the effect of animation.

If the transformation is linear, it can be done by a matrix multiplication.

the main geometric transformations in computer graphics are:

① rotation



Let L be a transformation that rotates a vector about the origin by an angle θ in the counterclockwise direction.

Such transformation is linear and can be written as

$$L(\mathbf{x}) = A\mathbf{x} \quad \text{where} \quad A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

② dilation / contraction

A linear operator of the form $L(\mathbf{x}) = B\mathbf{x}$ where $B = \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix}$

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

A dilation increases the size of the figure by a factor $b > 1$.

A contraction shrinks the figure by a factor $b < 1$.

③ translation

A transformation of the form $L(\mathbf{x}) = \mathbf{x} + \mathbf{c}$ where $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ is a translation. Such transformation is not linear, because $L(\alpha\mathbf{x} + \beta\mathbf{y}) \neq \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$

Hence L cannot be represented by a 2x2 matrix. However, we can do a trick, called homogeneous coordinates as follows:

1) chang $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ into $\hat{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$

2) Find $\mathbf{y} = C\hat{\mathbf{x}}$ with the 3x3 matrix $C = \begin{pmatrix} 1 & 0 & c_1 \\ 0 & 1 & c_2 \\ 0 & 0 & 1 \end{pmatrix}$

3) Finally throw the 3rd entry of $\mathbf{y}$. The first two entries of $\mathbf{y}$ are $\mathbf{x} + \mathbf{c}$.

This can easily be seen:

$$Y = C\hat{x} = \begin{pmatrix} 1 & 0 & c_1 \\ 0 & 1 & c_2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 + c_1 \\ x_2 + c_2 \\ 1 \end{pmatrix}$$

If we would like to perform all three transformations at the same time, it is better to also change the 2x2 matrices A and B into 3x3 matrices, as follows:

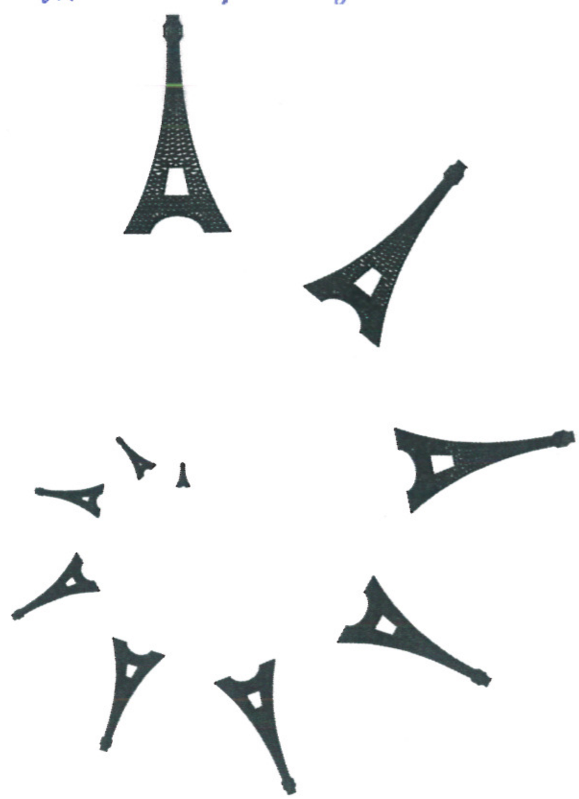
$$A = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & c_1 \\ 0 & 1 & c_2 \\ 0 & 0 & 1 \end{pmatrix}$$

Then, compute $y = CBA\hat{x}$ where $\hat{x} = \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$.

Finally pick the first two entries of y as the transformed point.

---

this is what we did in class, using the above geometric translations:



4.3 Similarity

Recall: (1) the linear transformation $L: \mathbb{R}^n \to \mathbb{R}^m$ can be defined by a matrix $A \in \mathbb{R}^{m\times n}$ such that $\boxed{L(x) = Ax}$, $\forall\, x \in \mathbb{R}^n$.

In this case, we have $A = [a_1\ a_2 \cdots a_m]$ with $a_j = L(e_j)$

(2) Consider a linear transformation $L: V \longrightarrow W$, where

$$\begin{cases} E = \{v_1, \dots, v_n\} \text{ is a basis for } V \\ F = \{w_1, \dots, w_m\} \text{ is a basis for } W \end{cases}$$

Then, there is a matrix $A \in \mathbb{R}^{m\times n}$ corresponding to L such that

$$\boxed{[L(v)]_F = A[v]_E}, \quad \forall\, v \in V.$$

In this case, we have $A = [a_1\ a_2 \cdots a_m]$ with $a_j = [L(v_j)]_F$.

Remark 1. Item (1) is a particular case of item (2).

In fact, if $V = \mathbb{R}^n$, we can always consider the standard basis $E = \{e_1, e_2, \dots, e_n\}$. Similarly, if $W = \mathbb{R}^m$, we can consider the standard basis $F = \{e_1, \dots, e_m\}$.

Hence for any vector $v = (v_1, v_2, \dots, v_n)^T \in \mathbb{R}^n$, we can write:

$$v = v_1 e_1 + v_2 e_2 + \dots + v_n e_n \implies [v]_E = v$$

Similarly, we will have $[L(v)]_F = L(v)$.

Hence $[L(v)]_F = A[v]_E$ reduces to $L(v) = Av$ when $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ with $E = \{e_1, \dots, e_n\}$ and $F = \{e_1, \dots, e_m\}$.

We say A is the representative matrix for the linear transformation L.

Remark 2. Since the basis for V and the basis for W are not unique, that is, since we may form V and W by many different bases, there is not a unique representative matrix for $L: V \to W$.

For the rest of this section, we only consider linear operators $L: V \to V$.

EX. 14 Consider $L: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $L(x) = \begin{pmatrix} 2x_1 \\ x_1 + x_2 \end{pmatrix}$ for $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

1) If we consider the basis $E_1 = \{e_1, e_2\}$ for $\mathbb{R}^2$, then the representative matrix will be: $A = [a_1 \ a_2]$ where $a_1 = L(e_1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $a_2 = L(e_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

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

2) If we use a different basis for $\mathbb{R}^2$, say $E_2 = \{u_1, u_2\}$ where $u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ then we need to find the representative matrix $B = [b_1 \ b_2]$ for L with respect to $\{u_1, u_2\}$. How to find matrix B?

We first notice that we have $V = W = \mathbb{R}^2$ with basis $E_2 = \{u_1, u_2\}$.

Hence, the representative matrix B satisfies: $\boxed{[L(v)]_{E_2} = B[v]_{E_2} \quad \forall v \in \mathbb{R}^2}$

This means that $b_j = [L(u_j)]_{E_2}$

In other words, b_j is the coordinate vector of $L(u_j)$ with respect to the basis $E_2 = \{u_1, u_2\}$. Hence, we first find $L(u_j)$ and then express it in terms of u_1 and u_2:

$$L(u_1) = Au_1 = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$L(u_2) = Au_2 = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

To express these vectors in terms of u_1 and u_2, we need to multiply them by $U^{-1} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}$:

$$b_1 = [L(u_1)]_{E_2} = \left[\begin{pmatrix} 2 \\ 2 \end{pmatrix}\right]_{E_2} = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$$b_2 = [L(u_2)]_{E_2} = \left[\begin{pmatrix} -2 \\ 0 \end{pmatrix}\right]_{E_2} = \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix}\begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\Rightarrow B = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$$

How are $A = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$ related?

Answer : $B = [\, U^{-1} A u_1, \; U^{-1} A u_2] = U^{-1} A [u_1 \; u_2] = U^{-1} A U$

To summarize:

If i) A is the matrix representing L w.r.t. $\{e_1, e_2\}$

ii) B is the matrix representing L w.r.t. $\{u_1, u_2\}$

iii) U is the transition matrix to change basis from $\{u_1, u_2\}$ to $\{e_1, e_2\}$

then $\boxed{B = U^{-1} A U}$

This result holds for any $L: V \to V$:

Theorem Consider a linear operator $L: V \to V$.

Let $E = \{v_1, \dots, v_n\}$ and $F = \{u_1, \dots, u_n\}$ be two bases for V.

Let S be the transition matrix to change basis 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

$$\boxed{B = S^{-1} A S}$$

Def. Let $A, B \in \mathbb{R}^{n \times n}$. We say B is similar to A if $\exists$ a non-singular matrix S s.t. $B = S^{-1} A S$.

Ex. 15 If $B = S^{-1} A S$, write A in terms of B.

$$B = S^{-1} A S \Rightarrow SB = S S^{-1} A S \Rightarrow SB = AS \Rightarrow SBS^{-1} = ASS^{-1}$$

$$\Rightarrow \boxed{A = SBS^{-1}}$$

This shows that if B is similar to A, then A is also similar to B.

It follows from the theorem that:

If A and B are $n\times n$ matrices representing the same operator (w.r.t. different bases), then A and B are similar.

$\boxed{\text{Ex. 16}}$ Let D be the differentiation operator on P_3:

$$D: P_3 \to P_3, \quad D(p) = p' \quad \forall p \in P_3.$$

1) Find the matrix A representing D w.r.t. $E = \{1, x, x^2\}$.

2) Find the matrix B representing D w.r.t. $F = \{1, 2x, 4x^2 - 3\}$.

3) Find the transition matrix S to change basis from F to E, and verify that $B = S^{-1}AS$.

1) We need to find A s.t. $[D(v)]_E = A[v]_E \quad \forall v \in P_3$.

This means that $A = [a_1 \; a_2 \; a_3]$, where
$$\begin{cases} a_1 = [D(1)]_E = [0\times 1 + 0\times x + 0\times x^2]_E = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ a_2 = [D(x)]_E = [1\times 1 + 0\times x + 0\times x^2]_E = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ a_3 = [D(x^2)]_E = [0\times 1 + 2\times x + 0\times x^2]_E = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \end{cases}$$

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

2) We need to find B s.t. $[D(v)]_F = B[v]_F \quad \forall v \in P_3$.

This means that $B = [b_1 \; b_2 \; b_3]$, where
$$\begin{cases} b_1 = [D(1)]_F = [0\times 1 + 0\times(2x) + 0\times(4x^2-3)]_F = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ b_2 = [D(2x)]_F = [2\times 1 + 0\times(2x) + 0\times(4x^2-3)]_F = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \\ b_3 = [D(4x^2-3)]_F = [0\times 1 + 4\times(2x) + 0\times(4x^2-3)]_F = \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} \end{cases}$$

$$\Rightarrow B = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}.$$

3) Given $p(x) = c_1 \cdot 1 + c_2 \cdot (2x) + c_3 \cdot (4x^2 - 3)$, find $p(x) = d_1 \cdot 1 + d_2 \cdot x + d_3 \cdot x^2$.

$$\Rightarrow \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \Rightarrow S = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Instead of verifying $B = S^{-1}AS$, we can equivalently verify $SB = AS$:

$$SB = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 8 \\ 0 & 0 & 0 \end{pmatrix}$$

$$AS = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 8 \\ 0 & 0 & 0 \end{pmatrix}$$

$\Rightarrow SB = AS \Rightarrow B = S^{-1}AS$. <u>O.K.</u>