Lecture 3: Linear Transformations

Jie Mei, Harbin Institute of Technology, Shenzhen

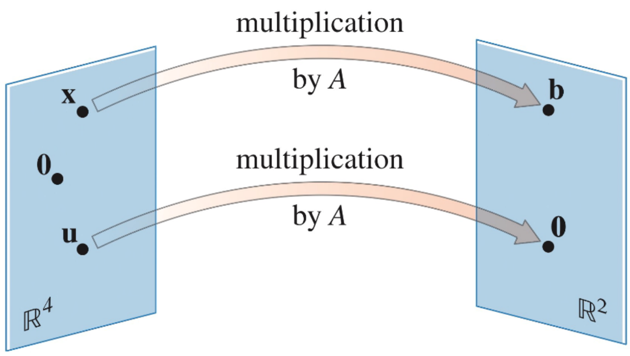
April 10, 2025

Review: For linear equations,
$Ax = b \Leftrightarrow$ linear combinations of column vectors: $x_1a_1 + \cdots + x_na_n = b$.

When a matrix A multiples a vector x, it "transforms" x into another vector Ax. For instance, the equations:

$$\underset{\uparrow\atop A}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}\underset{\uparrow\atop \mathbf{x}}{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}} = \underset{\uparrow\atop \mathbf{b}}{\begin{bmatrix} 5 \\ 8 \end{bmatrix}} \quad \text{and} \quad \underset{\uparrow\atop A}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}\underset{\uparrow\atop \mathbf{u}}{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}} = \underset{\uparrow\atop \mathbf{0}}{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}$$

say that multiplication by A transforms x into b and transforms u into the zero vector.



A transformation(mapping or function) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector in $\mathbb{R}^n$ a vector $T(x)$ in $\mathbb{R}^m$.

3.1 Definition of Linear Transformation

Definition 3.1.1: *A mapping T from a vector space V to a vector space W is said to be a* **linear transformation** *if*

(1) $T(v_1 + v_2) = T(v_1) + T(v_2)$ *(Additivity property)*

(2) $T(\alpha v) = \alpha T(v)$ *(Homogeneity property)*

for all $v_1, v_2, v \in V$ and any scalar α.

In the case that the vector spaces V and W are the same, we will refer to a linear transformation $T : V \to V$ as a **linear operator** *on V.*

Notation $\mathcal{L}(V, W)$**:** The set of all linear transformations from V to W.

Example 3.1.1:

1. $T(x) = 3x$ *for each* $x \in \mathbb{R}^2$.

2. $T(x) = x_1 e_1$ *for each* $x \in \mathbb{R}^2$. $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $T(x) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$.

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$
$$T(x+y) = \begin{bmatrix} x_1 + y_1 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} y_1 \\ 0 \end{bmatrix} = T(x) + T(y)$$
$$T(\alpha x) = \begin{bmatrix} \alpha x_1 \\ 0 \end{bmatrix} = \alpha \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \alpha T(x)$$

3. $T(x) = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}$ *for each* $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ *in* $\mathbb{R}^2$.

4. $T : \mathbb{R}^2 \to \mathbb{R}^1$. $T(x) = x_1 + x_2$ *for each* $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

5. $T : \mathbb{R}^2 \to \mathbb{R}^1$. $T(x) = (x_1^2 + x_2^2)^{1/2}$.

6. $T : \mathbb{R}^2 \to \mathbb{R}^3$. $T(x) = \begin{bmatrix} x_2 \\ x_1 \\ x_1 + x_2 \end{bmatrix}$ *for* $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Note that

$$
\begin{aligned}
T(x) &= x_1 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\end{aligned}
$$

■

In general, if A is any $m \times n$ matrix, $T_A : \mathbb{R}^n \to \mathbb{R}^m$, $T_A(x) = Ax$.

$$T_A(\alpha x + \beta y) = A(\alpha x + \beta y) = \alpha Ax + \beta Ay = \alpha T_A(x) + \beta T_A(y).$$

Thus, we can think of each $m \times n$ matrix A as defining a linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$.

Note 3.1.1: *If $T \in \mathcal{L}(V, W)$,*

(i) $T(0_V) = 0_W$.

(ii) If $v_1, \ldots, v_n$ are elements of V and $\alpha_1, \ldots, \alpha_n$ are scalars, then

$$T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n).$$

(iii) $T(-v) = -T(v)$.

■

Example 3.1.2:

1. $T : C[a, b] \to \mathbb{R}$. $\quad T(f) = \int_a^b f(x)dx$.
2. $D : C^1(a, b] \to C[a, b]$. $\quad D(f) = f'$.
3. $T : \mathbb{R}^2 \to \mathbb{R}^2$. $\quad T(x) = x + x_0$ *for a fixed nonzero vector* x_0 *in* $\mathbb{R}^2$.

■

3.2 Null Spaces and Ranges

Definition 3.2.1: *For $T \in \mathcal{L}(V, W)$, the* **null space** *of T, denoted $\operatorname{null} T$, is defined by*

$$\operatorname{null} T=\left\{v \in V: T(v)=0_W\right\}.$$

Definition 3.2.2: *Let $T \in \mathcal{L}(V, W)$ and let S be a subspace of V. The* **image** *of S, denoted $T(S)$, is defined by*

$$T(S)=\{w \in W \mid w=T(v) \quad \text{for some} \quad v \in S\}.$$

The image of the entire vector space, $T(V)$, denoted $\operatorname{range} T$, is called the **range** *of T.*

Theorem 3.2.1: *If $T \in \mathcal{L}(V, W)$ and S is a subspace of V, then*

(i) $\operatorname{null} T$ *is a subspace of V.*

(ii) $\operatorname{range} T$ *is a subspace of W.*

Proof: (i) $\operatorname{null} T \subset V$ and $\quad 0_V \in \operatorname{null} T$.

Let $v \in \operatorname{null} T$, then $T(v)=0_W$. For a scalar $\alpha \in \mathbb{F}$,

$$T(\alpha v)=\alpha T(v)=\alpha 0_W=0_W.$$

Hence, $\alpha v \in \operatorname{null} T$.

Let $v_1, v_2 \in \operatorname{null} T$. Then $T\left(v_1\right)=T\left(v_2\right)=0_W$.

$$T\left(v_1+v_2\right)=T\left(v_1\right)+T\left(v_2\right)=0_W+0_W=0_W.$$

Hence, $v_1+v_2 \in \operatorname{null} T$.

(ii) Note that $0_W=T\left(0_V\right) \in \operatorname{range} T$.

If $w \in \operatorname{range} T$, then $w=T(v)$ for some $v \in V$. Thus for a scalar $\alpha \in \mathbb{F}$,

$$\alpha w=\alpha T(v)=T(\alpha v).$$

Since $\alpha v \in V$, it follows that $\alpha w \in \operatorname{range} T$.

If $w_1, w_2 \in \operatorname{range} T$, then $w_1=T\left(v_1\right)$ and $w_2=T\left(v_2\right)$ for some $v_1, v_2 \in V$. Thus

$$w_1+\omega_2=T\left(v_1\right)+T\left(v_2\right)=T\left(v_1+v_2\right).$$

Since $v_1+v_2 \in V$, it follows that $w_1+w_2 \in \operatorname{range} T$. ■

Example 3.2.1:

1. *Zero linear transformation.*

   $\mathbf{0}v = 0_W$, *for any* $v \in V$. $\quad \operatorname{null} \mathbf{0} = V$, $\operatorname{range} \mathbf{0} = \{0_W\}$.

   Identity transformation.

   $\mathbf{I}v = v$, *for any* $v \in V$. $\quad V = W$, $\operatorname{null} \mathbf{I} = \{O_v\}$, $\operatorname{range} \mathbf{I} = W$.

2. *Let* $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ *be defined by*

$$T(x) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

   A vector $x \in \operatorname{null} T$ *if and only if* $x_1 = 0$.

   A basis of $\operatorname{null} T$ *is* $\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$, $\quad \dim \operatorname{null} T = 1$.

   A vector $y \in \operatorname{range} T$ *if and only if* y *is a multiple of* $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

   A basis of $\operatorname{range} T$ *is* $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$, $\quad \dim \operatorname{range} T = 1$.

   Let $T \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$ *be defined by*

$$T(x) = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

   If $x \in \operatorname{null} T$,

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_2 + x_3 &= 0 \\ \Rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= 0 \\ \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} & \end{aligned}$$

   A basis of $\operatorname{null} T$ *is* $\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$, $\quad \dim \operatorname{null} T = 1$.

Let $S = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)$.

If $x \in S$*, then* x *must be of the form* $\begin{bmatrix} x_1 \\ 0 \\ x_3 \end{bmatrix}$.

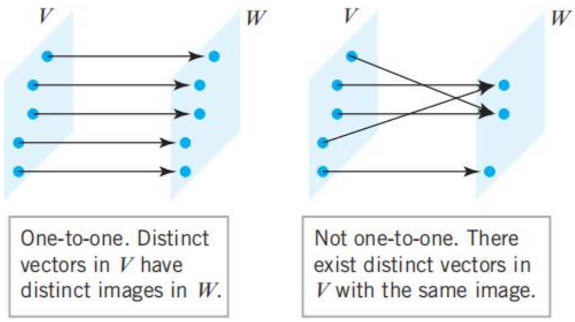
$$T(x) = \begin{pmatrix} x_1 \\ x_3 \end{pmatrix}.$$

$$T(S) = \mathbb{R}^2. \quad T\left(\mathbb{R}^3\right) = \mathbb{R}^2.$$

■

3.3 Injective and Surjective

Definition 3.3.1: *A function $T : V \to W$ is called* **injective** *if $T(u) = T(v)$ implies $u = v$.*



Theorem 3.3.1: *Let $T \in \mathcal{L}(V, W)$. Then T is injective if and only if $\operatorname{null} T = \{0_V\}$.*

Proof: ($\Longrightarrow$) Suppose T is injective. For any $v \in \operatorname{null} T$,

$$T(v) = 0_W = T(0_V)$$

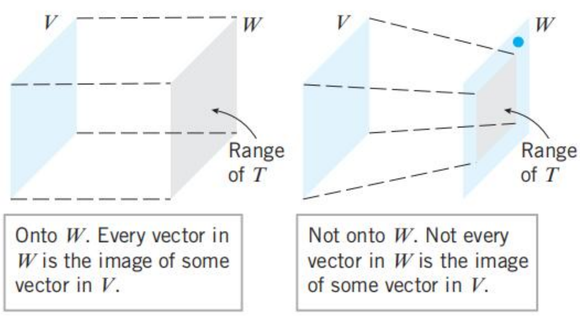
Since T is injective, $v = 0_V$. Therefore, $\operatorname{null} T = \{0_V\}$.

($\Longleftarrow$) Let $u, v \in V$ and $T(u) = T(v)$. Thus

$$0_W = T(u) - T(v) = T(u - v).$$

Therefore, $u - v \in \operatorname{null} T = \{0_V\}$. Therefore, $u - v = 0_V \Rightarrow u = v$. Hence, T is injective. ■

Definition 3.3.2: *A function $T : V \to W$ is called* **surjective** *if $\operatorname{range} T = W$.*



Example 3.3.1:

1. *Zero linear transformation.*
2. *Identity linear transformation.*
3. *The differentiation transformation:*
$$D \in \mathcal{L}\left(P_5(\mathbb{R}), P_5(\mathbb{R})\right)$$
$$D \in \mathcal{L}\left(P_5(\mathbb{R}), P_4(\mathbb{R})\right)$$
4. *The shifting operators:*
Let $V = \mathbb{R}^\infty$ be the vector space of infinite sequences of real numbers.
$$T_1(u_1, u_2, \ldots) = (0, u_1, u_2, \ldots)$$
$$T_2(u_1, u_2, \ldots) = (u_2, u_3, \ldots)$$

3.4 Fundamental Theorem of Linear Transformations

Theorem 3.4.1: **Fundamental Theorem of Linear Transformations:**
Suppose V is a finite-dimensional vector space and $T \in \mathcal{L}(V, W)$. Then range T *is finite-dimensional and*

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T. \tag{3.1}$$

Proof: Assume that $\dim V = n, \dim \operatorname{null} T = m$. Since $\operatorname{null} T$ is a subspace of V, $m \leq n$. We need to prove that $\dim \operatorname{range} T = n - m$.

Let $u_1, \cdots, u_m$ be a basis of $\operatorname{null} T$. It can be extended to a basis of V:

$$u_1, \cdots, u_m, v_1, \cdots, v_{n-m}.$$

Note that $T(u_i) = 0, i = 1, \ldots, m$ and $T(v_i) \neq 0, i = 1, , \ldots, n - m$. For any $v \in V$, v is a linear combination of $u_1, \cdots, u_m, v_1, \ldots, v_{n-m}$, $i.e.$,

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \ldots + b_{n-m} v_{n-m}$$

for some scalars a_i and b_j.

Since $T(u_i) = 0$ and $T \in \mathcal{L}(V, W)$,

$$\begin{aligned} T(v) &= a_1 T(u_1) + \cdots + a_m T(u_m) + b_1 T(v_1) + \cdots + b_{n-m} T(v_{n-m}) \\ &= b_1 T(v_1) + \cdots + b_{n-m} T(v_{n-m}) \end{aligned}$$

We can get $\operatorname{Span}(T(v_1), \cdots, T(v_{n-m})) = \operatorname{range} T$. Thus $\operatorname{range} T$ is finite-dimensional.

To show that $T(v_1), \ldots, T(v_{n-m})$ is linearly independant, consider

$$c_1 T(v_1) + \cdots + c_{n-m} T(v_{n-m}) = 0$$

Since $T \in \mathcal{L}(V, W)$, we have

$$T(c_1 v_1 + \cdots + c_{n-m} v_{n-m}) = 0$$

Hence $c_1 v_1 + \cdots + c_{n-m} v_{n-m} \in \operatorname{null} T$. Therefore, it can be written as a linear combination of $u_1, \ldots, u_m$. Assume that

$$\begin{aligned} & c_1 v_1 + \cdots + c_{n-m} v_{n-m} = d_1 u_1 + \cdots + d_m u_m \\ \Leftrightarrow & -d_1 u_1 - \cdots - d_m u_m + c_1 v_1 + \cdots + c_{n-m} v_{n-m} = 0 \end{aligned}$$

Since $u_1, \ldots, u_m, v_1, \ldots, v_{n-m}$ is a basis of V and thus linearly independent. We have $d_1 = \cdots = d_m = c_1 = \cdots = c_{n-m} = 0$. Thus $T(v_1), \ldots, T(v_{n-m})$ is linearly independent. Also note that $\operatorname{Span}(T(v_1), \ldots, T(v_{n-m})) = \operatorname{range} T$. Therefore, $T(v_1), \ldots, T(v_{n-m})$ is a basis of $\operatorname{range} T$ and thus $\dim \operatorname{range} T = n - m$. ■

Note 3.4.1:

1. *Consider $T = Ax$ with $A \in \mathbb{R}^{m \times n}$.*

   $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, $\operatorname{null} T = N(A)$, $\operatorname{range} T = C(A)$.

$$\begin{array}{ccccc} \dim V & = \dim \operatorname{null} T & + & \dim \operatorname{range} T \\ \downarrow & \downarrow & & \downarrow \\ \dim \mathbb{R}^n & = \dim N(A) & + & \dim C(A) \\ \downarrow & \downarrow & & \downarrow \\ n & = \quad n - r & + & r \end{array}$$

2. *Let V and W be finite-dimensional vector spaces.*

   A transformation to a smaller dimensional space is not injective.

   ($\dim V > \dim W$)

   Injective : $\operatorname{null} T = \{0\}$ *and* $\dim \operatorname{null} T = 0$.

$$\begin{aligned} \dim \operatorname{null} T &= \dim V - \dim \operatorname{range} T \quad (\operatorname{range} T \subset W) \\ &\geqslant \dim V - \dim W > 0 \end{aligned}$$

   A homogeneous system of linear equations with more variables than equations has nonzero solutions. $(Ax = 0, A \in \mathbb{R}^{m \times n}, m < n)$**.**

   A transformation to a larger dimensional space is not surjective.

   ($\dim V < \dim W$)

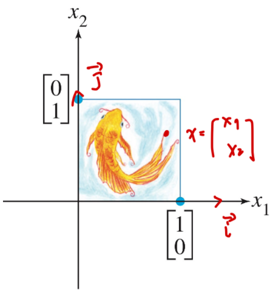
   Surjective: $\operatorname{range} T = W \Rightarrow \dim \operatorname{range} T = \dim W$.

$$\begin{aligned} \dim \operatorname{range} T &= \dim V - \dim \operatorname{null} T \\ &\leq \dim V < \dim W. \end{aligned}$$

   An inhomogeneous system of linear equations with more equations than variables has no solution for some choice of b.

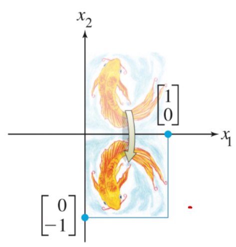
3.5 Matrix Representation of Linear Transformations

Motivation Examples: Geometric Transformations of $\mathbb{R}^2$.



Transformation T

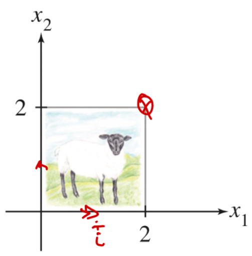
Reflection through the x_1-axis



$$x \xrightarrow{T} y = T(x)$$

$$\begin{aligned}
x &= x_1\vec{i} + x_2\vec{j} = \begin{bmatrix} \vec{i} & \vec{j} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
y &= T(x) \\
&= T\left(x_1\vec{i} + x_2\vec{j}\right) \\
&= x_1 T(\vec{i}) + x_2 T(\vec{j}) \\
&= [T(\vec{i}) \quad T(\vec{j})] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\end{aligned}$$

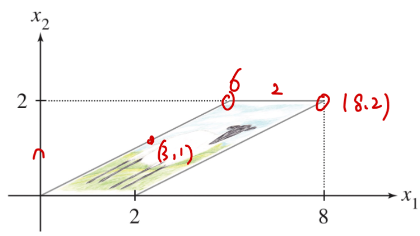
$$\vec{i} \xrightarrow{T} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \vec{j} \xrightarrow{T} \begin{bmatrix} 0 \\ -1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}$$



T

Shear Transformation

剪切变换



$$x = \begin{bmatrix} \vec{i} & \vec{j} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \xrightarrow{T} y = \begin{bmatrix} T(\vec{i}) & T(\vec{j}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

check $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$.

Theorem 3.5.1: *Let $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Then there is an $m \times n$ matrix A such that*

$$T(x) = Ax$$

for each $x \in \mathbb{R}^n$. In fact, A is the $m \times n$ matrix whose jth column is the vector $T(e_j)$:

$$A = [T(e_1) \cdots T(e_n)].$$

Proof: For any $x \in \mathbb{R}^n$,

$$x = x_1 e_1 + \cdots + x_n e_n = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Then

$$\begin{aligned} T(x) &= T(x_1 e_1 + \cdots + x_n e_n) \\ &= x_1 T(e_1) + \cdots + x_n T(e_n) \\ &= [T(e_1) \cdots T(e_n)] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &\triangleq Ax. \end{aligned}$$

■

The matrix A is called the standard matrix for the linear transformation T.

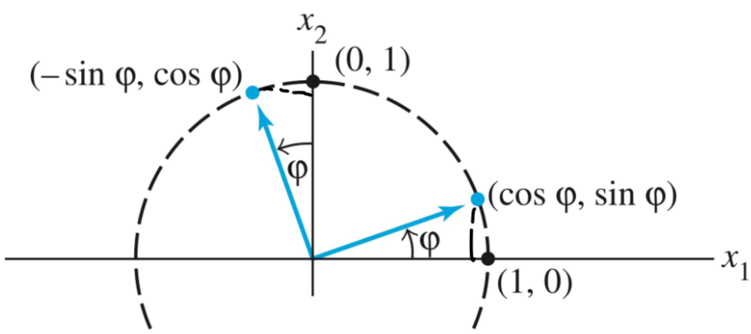
Example 3.5.1:

1. *Define $T \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$ by $T(x) = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix}$ with $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.*

$$\begin{aligned} A &= \begin{bmatrix} T(e_1) & T(e_2) & T(e_3) \end{bmatrix} \\ &= \left[T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) \quad T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) \quad T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) \right] \\ &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}. \end{aligned}$$

$$Ax = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix}.$$

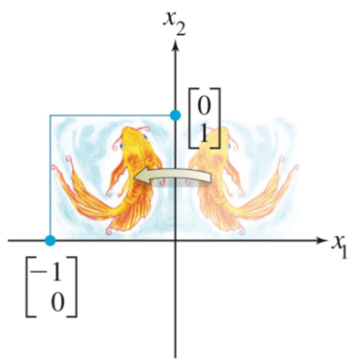
2. *Let $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ be the transformation that rotates each point in $\mathbb{R}^2$ about the origin through an angle φ, in the counterclockwise direction.*



$$T(e_1) = \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}, \quad T(e_2) = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix}$$

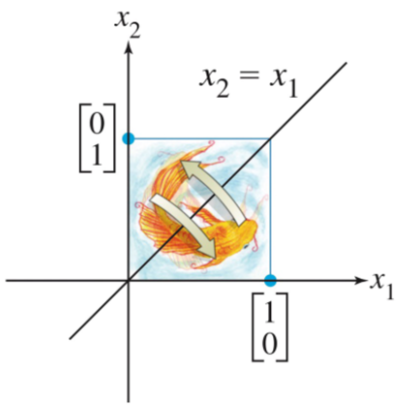
$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

3. *Reflection through the x_2-axis:*



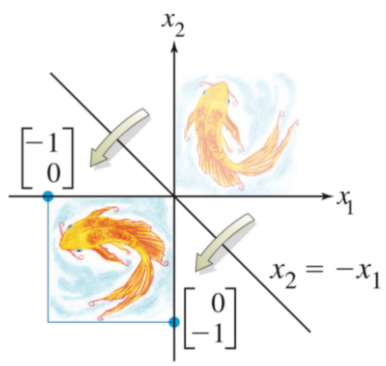
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

4. Reflection through the line $x_2 = x_1$:



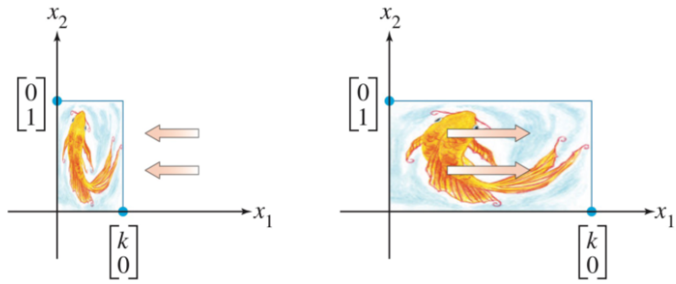
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

5. Reflection through the line $x_2 = -x_1$:



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} -x_2 \\ -x_1 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

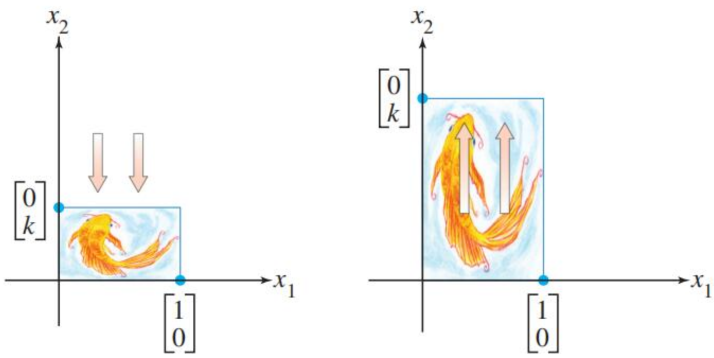
6. Horizontal contraction and expansion:



$0 < k < 1$ $\quad$ $k > 1$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} kx_1 \\ x_2 \end{bmatrix}, \quad \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$$

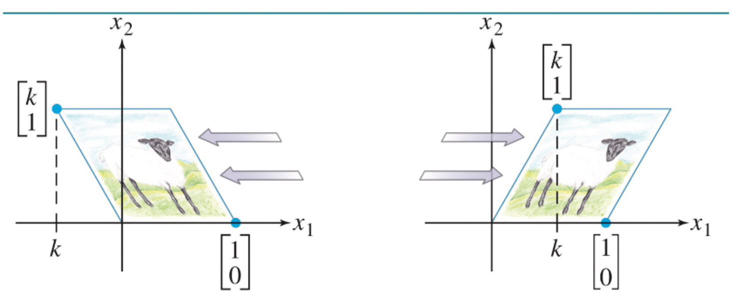
7. *Vertical contraction and expansion:*



$0 < k < 1$ $\quad$ $k > 1$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} x_1 \\ kx_2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$$

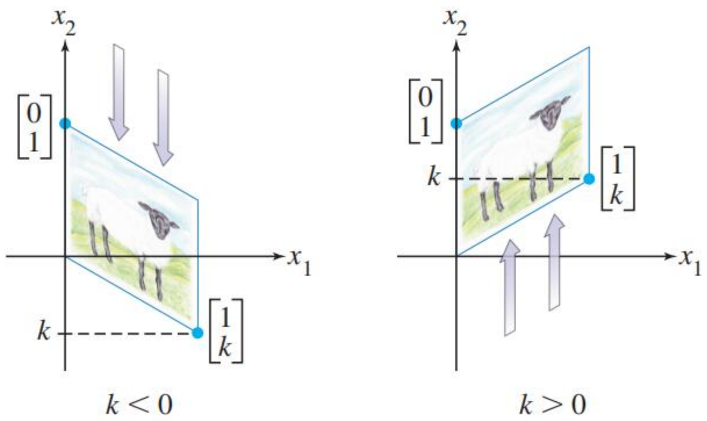
8. *Horizontal shear:*



$k < 0$ $\quad$ $k > 0$

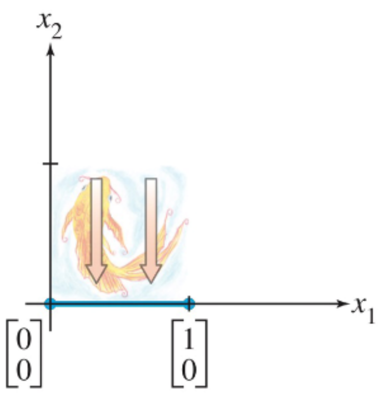
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} x_1 + kx_2 \\ x_2 \end{bmatrix}, \quad \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

9. Vertical shear:



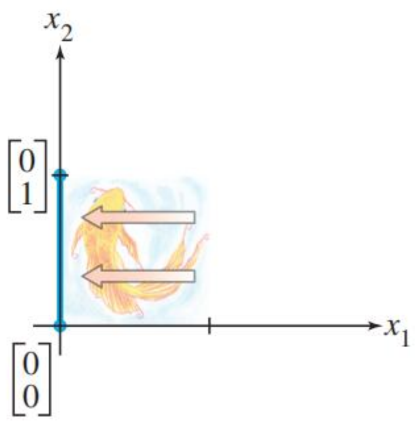
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} x_1 \\ kx_1 + x_2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$

10. Projection onto the x_1-axis:



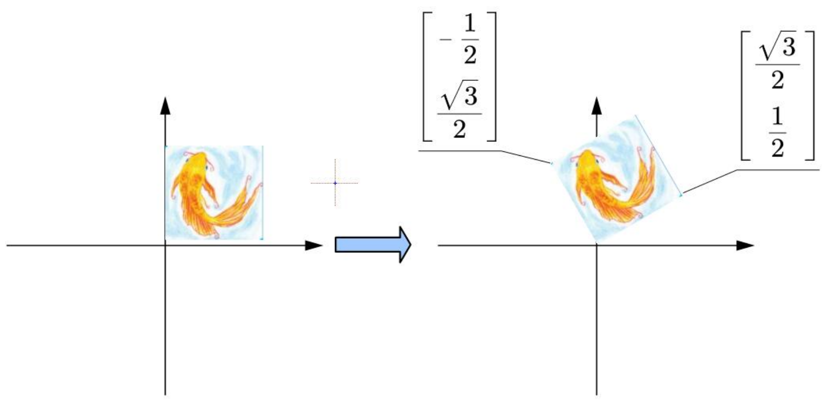
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} x_1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

11. Projection onto the x_2-axis:



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \to \begin{bmatrix} 0 \\ x_2 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

12. Rotation about the origin through a positive angle $\theta = \frac{\pi}{6}$:



$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

■

We have seen how matrices are used to represent linear transformations from $\mathbb{R}^n \to \mathbb{R}^m$. We may ask whether it is possible to find a similar representation for linear transformations from V into W, where V and W are vector spaces of dimension n and m, respectively.

Consider $T \in \mathcal{L}(V, W)$.
Let $E = \{v_1, \ldots, v_n\}$ be an ordered basis for V.
Let $F = \{w_1 \ldots, w_m\}$ be an ordered basis for W.

$$v \in V \quad \xrightarrow{T} \quad w = T(v) \in W$$

For any $v \in V$ and $w = T(v) \in W$,

$$\begin{aligned} v =& x_1 v_1 + \cdots + x_n v_n = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ w =& y_1 w_1 + \cdots + y_m w_m = [w_1 \cdots w_m] \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \end{aligned}$$

$$\begin{aligned} w =& \ T(v) \\ &= T\left(x_1 v_1 + \cdots + x_n v_n\right) \\ &= x_1 T\left(v_1\right) + \cdots + x_n T\left(v_n\right) \\ &= \begin{bmatrix} T\left(v_1\right) & \cdots & T\left(v_n\right) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \end{aligned}$$

Since $T\left(v_j\right) \in W, \forall j = 1, \cdots, n$,

$$T\left(v_1\right) = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$$

$$\vdots$$

$$T\left(v_j\right) = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

Then

$$
\begin{aligned}
w &= T(v) \\
&= \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \underbrace{\begin{bmatrix} a_{11} & a_{22} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{x}} \\
&= \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{\mathbf{y}}
\end{aligned}
$$

$$\Longrightarrow \quad y = Ax.$$

Theorem 3.5.2: **Matrix Representation Theorem**
If $E = \{v_1, \ldots, v_n\}$ and $F = \{w_1, \ldots, w_m\}$ are ordered bases for vector spaces V and W, respectively, then, corresponding to each linear transformation $T : V \to W$, there is an $m \times n$ matrix such that

$$[T(v)]_F = A[v]_E \quad \textit{for each} \quad v \in V.$$

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

$$A = \begin{bmatrix} [T(v_1)]_F & \cdots & [T(v_n)]_F \end{bmatrix}.$$

Note 3.5.1: *For $A(T, E, F)$, if $[v]_E = x$, $[w]_F = y$, then T maps v into w if and only if A maps x into y ($y = Ax$).*

$$
\begin{tikzcd}
v \in V \arrow[r, "T"] \arrow[d, leftrightarrow] & w = T(v) \in W \arrow[d, leftrightarrow] \\
x = [v]_E \in R^n \arrow[r, "A"] & y = Ax = [w]_F \in R^m
\end{tikzcd}
$$

Example 3.5.2:

1. $T \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$,

$$T(x) = x_1 b_1 + (x_2 + x_3)\, b_2 = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 + x_3 \end{bmatrix}$$

for each $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$, *where* $b_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ *is a basis of* $\mathbb{R}^2$,
$\left\{ e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ *is a basis of* $\mathbb{R}^3$.

$$A = \begin{bmatrix} T(e_1) & T(e_2) & T(e_3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Check $Ax = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 + x_3 \end{bmatrix}$.

2. $T \in \mathcal{L}(\mathbb{R}^2)$

$$T(\alpha b_1 + \beta b_2) = (\alpha + \beta) b_1 + 2\beta b_2 = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} \alpha + \beta \\ 2\beta \end{bmatrix}$$

where $b_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $b_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

$$\begin{aligned} T(b_1) &\underset{\alpha=1,\beta=0}{=} \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ T(b_2) &\underset{\alpha=0,\beta=1}{=} \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ A &= \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \end{aligned}$$

3. $T \in \mathcal{L}(\mathbb{R}^2)$. $T(v) = v$ *is the identity transformation.*

One might expect its matrix to be I*, but that only happens when the input basis is the same as the output basis.*

Let $E = \left\{ \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ 8 \end{bmatrix} \right\}$*(input basis),* $F = \left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\}$*(output basis).*

$$v = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

$$\begin{aligned} \Rightarrow \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} &= \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 3 & 6 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \end{aligned}$$

$$[[T(v_1)]_F \quad [(T(v_2)]_F] = [[v_1]_F \quad [v_2]_F]$$

$$\begin{aligned} v_1 &= 1 \cdot w_1 + 1 \cdot w_2 \\ v_2 &= 2 \cdot w_1 + 3 \cdot w_2 \end{aligned}$$

4. *$D \in \mathcal{L}(\mathbb{P}_2(\mathbb{R}), \mathbb{P}_1(\mathbb{R}))$ with the linear transformation D defined by $D(p) = p'$ mapping $\mathbb{P}_2(\mathbb{R}) = \{a_2x^2 + a_1x + a_0, a_2, a_1, a_0 \in \mathbb{R}\}$ into $\mathbb{P}_1(\mathbb{R}) = \{b_1x + b_0, b_1, b_0 \in \mathbb{R}\}$. A basis of $\mathbb{P}_2$ is $\{x^2, x, 1\}$ and a basis of $\mathbb{P}_1$ is $\{x, 1\}$.*

$$\begin{aligned} D\left(x^2\right) =& 2x + 0 \cdot 1 = \begin{bmatrix} x & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} \\ D(x) = 1 =& \begin{bmatrix} x & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ D(1) = 0 =& \begin{bmatrix} x & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ A =& \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{aligned}$$

For $p(x) = a_2x^2 + a_1x + a_0, D(p) = 2a_2x + a_1$.

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 2a_2 \\ a_1 \end{bmatrix}$$

5. *For $T \in \mathcal{L}(V, W)$, sometimes, $T(V)$ is just a subspace of W, denoted U. Is the matrix represention of $T \in \mathcal{L}(V, W)$ the same as that of $T \in \mathcal{L}(V, U)$?*

For $D(\mathbb{P}_2(\mathbb{R}), \mathbb{P}_2(\mathbb{R}))$, *where a basis of* $\mathbb{P}_2$ *is* $\{x^2, x, 1\}$. *Then*

$$
\begin{aligned}
D\left(x^2\right) =& 2x + 0 \cdot 1 = \left[\begin{array}{lll} x^2 & x & 1 \end{array}\right] \left[\begin{array}{l} 0 \\ 2 \\ 0 \end{array}\right] \\
D(x) = 1 = & \left[\begin{array}{lll} x^2 & x & 1 \end{array}\right] \left[\begin{array}{l} 0 \\ 0 \\ 1 \end{array}\right] \\
D(1) = 0 = & \left[\begin{array}{lll} x^2 & x & 1 \end{array}\right] \left[\begin{array}{l} 0 \\ 0 \\ 0 \end{array}\right] \\
A = & \left[\begin{array}{lll} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right] \in \mathbb{R}^{3 \times 3}.
\end{aligned}
$$

Note 3.5.2: *To find the matrix representation* A *for a linear transformation* T *with respect to the ordered bases* $E = \{v_1, \ldots, v_n\}$ *and* $F = \{w_1, \ldots, w_m\}$, *the key is to represent each vector* $T(v_j)$ *as a linear combination of* $w_1, \cdots, w_m$.

For $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, the above process equivalents to solving the linear equation

$$
\left[\begin{array}{lll} w_1 & \cdots & w_m \end{array}\right] \left[\begin{array}{c} x_1 \\ \vdots \\ x_m \end{array}\right] = T(v_j)
$$

Denote $B = \left[\begin{array}{lll} w_1 & \cdots & w_m \end{array}\right]$, $a_j = \left[\begin{array}{c} x_1 \\ \vdots \\ x_m \end{array}\right]$, then $a_j = B^{-1} T(v_j)$.

The augmented matrix

$$
\begin{aligned}
& B^{-1} \left[\begin{array}{c:ccc} B & T(v_1) & \cdots & T(v_n) \end{array}\right] \\
= & \left[\begin{array}{c:ccc} I & B^{-1} T(v_1) & \cdots & B^{-1} T(v_n) \end{array}\right] \\
= & \left(\begin{array}{c:c} I & A \end{array}\right)
\end{aligned}
$$

Example 3.5.3: $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$

$$
T(x) = \left[\begin{array}{c} x_2 \\ x_1 + x_2 \\ x_1 - x_2 \end{array}\right] \textit{ for } x \in \left[\begin{array}{l} x_1 \\ x_2 \end{array}\right].
$$

Find the matrix representation of T with respect to the ordered bases $\{v_1, v_2\}$ and $\{w_1, w_2, w_3\}$ where

$$v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad w_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, w_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

The augmented matrix

$$\begin{bmatrix} 1 & 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 3 & 4 \\ 0 & 0 & 1 & -1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -1 & -3 \\ 0 & 1 & 1 & 3 & 4 \\ 0 & 0 & 1 & -1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -1 & -3 \\ 0 & 1 & 0 & 4 & 2 \\ 0 & 0 & 1 & -1 & 2 \end{bmatrix}.$$

$$A = \begin{bmatrix} -1 & -3 \\ 4 & 2 \\ -1 & 2 \end{bmatrix} \Rightarrow T(v_1) = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix},$$

$$T(v_2) = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} -3 \\ 2 \\ 2 \end{bmatrix}$$

Question:

$$A(T, \{v_1, v_2\}, \{w_1, w_2, w_3\}) = \begin{bmatrix} -1 & -3 \\ 4 & 2 \\ -1 & 2 \end{bmatrix}$$

find $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$?

$$T(v) = w:$$

$$v = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \triangleq \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

$$w = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

From $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = A \cdot \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

$$\Longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = A \cdot \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cdot A \cdot \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

■

3.6 Equivalent Matrix and Similar Matrix

Let $T \in \mathcal{L}(V, W)$. $\dim V = n, \dim W = m$.
Let $E = \{v_1, \cdots, v_n\}$ be an ordered basis of V and $F = \{w_1, \cdots, w_m\}$ an ordered basis of W.

$$\begin{tikzcd} v \arrow[r, "T"] \arrow[d, leftrightarrow] & T(v) \in W. \\ {[v]_E} \arrow[r, "A"] & {[T(v)]_F} \end{tikzcd}$$

$$[T(v)]_F = A \cdot [v]_E. \quad A \in R^{m \times n}$$

Let $E' = \{v_1', \cdots, v_n'\}$ be a new basis of V and $F' = \{w_1', \cdots, w_m'\}$ a new basis of W.

Question: $[T(v)]_{F'} = B[v]_{E'}$, **what is B?**
(1) There is a transition matrix $P \in \mathbb{R}^{n \times n}$, such that

$$\begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} P.$$

(2) There is a transition matrix $Q \in \mathbb{R}^{m \times m}$, such that

$$\begin{bmatrix} w_1' & \cdots & w_m' \end{bmatrix} = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} Q.$$

Then for any $v \in V$, $T(v) \in W$,

$$\begin{aligned} v &= \begin{bmatrix} v_1' & \cdots & v_n' \end{bmatrix} [v]_{E'} \\ &= \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} P[v]_{E'} \\ &= \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} [v]_E \\ \implies & \quad [v]_E = P[v]_{E'}. \end{aligned}$$

$$\begin{aligned} T(v) &= \begin{bmatrix} w_1' & \cdots & w_n' \end{bmatrix} [T(v)]_{F'} \\ &= \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} Q[T(v)]_{F'} \\ &= \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} [T(v)]_F \\ \implies & \quad [T(v)]_F = Q[T(v)]_{F'}. \end{aligned}$$

From $[T(v)]_F = A[v]_E$, we have $Q[T(v)]_{F'} = AP[v]_{E'}$. And thus

$$[T(v)]_{F'} = \underbrace{Q^{-1}AP}_{B}[v]_{E'}.$$

Definition 3.6.1: **Equivalent Matrices**
Let A and B be $m \times n$ matrices. B is said to be **equivalent** *to A if there exist a nonsingular $n \times n$ matrix P and a nonsingular $m \times m$ matrix Q, such that*

$$B = Q^{-1}AP.$$

If $V = W, T \in \mathcal{L}(V)$ is a linear operator. Let $E = F$ and $E' = F'$. We have $P = Q$. And thus $B = Q^{-1}AQ$.

Definition 3.6.2: **Similar Matrices**
Let A and B be $n \times n$ matrices. B is said to be **similar** *to A if there exists a nonsingular $n \times n$ matrix S, such that*

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

Note 3.6.1: *Let $E = \{v_1, \ldots, v_n\}$ and $E' = \{v'_1, \ldots, v'_n\}$ be two ordered bases for a vector space V, and let $T \in \mathcal{L}(V)$. If A is the matrix representing T with respect to E, and B is the matrix representing T with respect to E', then $B = S^{-1}AS$, where S is the transition matrix from E' to E.*

3.7 Warps and Morphs

Among the more interesting image-manipulation techniques available for computer graphics are warps and morphs. In this section we introduce the mathematics (mainly linear transformations) behind image warping and morphing. And we will show how linear transformations can be used to distort a single picture to produce a warp, or to distort and blend two pictures to produce a morph.

Motivation:

Begin-picture End-picture

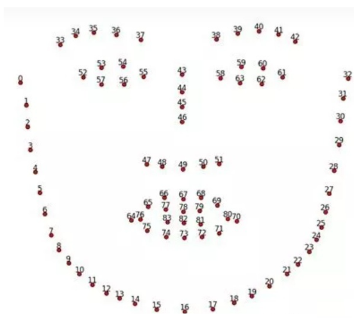
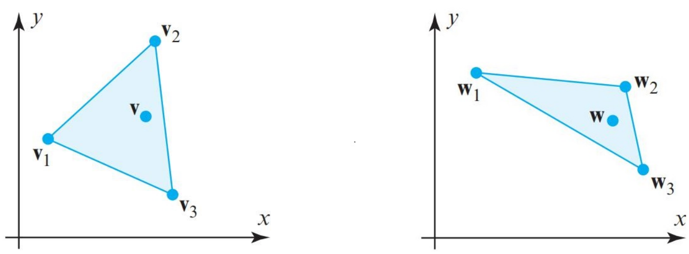


Figure 1: Facial feature points.



Consider the mapping that maps v_1 to w_1, v_2 to w_2, and v_3 to w_3. This transformation is not a linear transformation since it may not map 0 to 0. Actually, it combines a linear transformation with a translation, called the **affine transformation**.

Let us first consider the linear transformation T in this affine transformation, by regrading v_1 and w_1 as the 'zero vectors'. Then, T maps $v_2 - v_1$ to $w_2 - w_1$ and maps $v_3 - v_1$ to $w_3 - w_1$, which is a linear operator with respect to the standard basis on $\mathbb{R}^2$. And we have

$$T(v_2 - v_1) = w_2 - w_1, \quad T(v_3 - v_1) = w_3 - w_1$$

and its matrix represention is

$$A = \begin{bmatrix} T(e_1) & T(e_2) \end{bmatrix}.$$

Let $v_2 - v_1 = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}$ and $v_3 - v_1 = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix}$.
Let $w_2 - w_1 = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix}$ and $w_3 - w_1 = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} c_{12} \\ c_{22} \end{bmatrix}$.
We have

$$\begin{bmatrix} v_2 - v_1 & v_3 - v_1 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \underbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}}_{B}$$

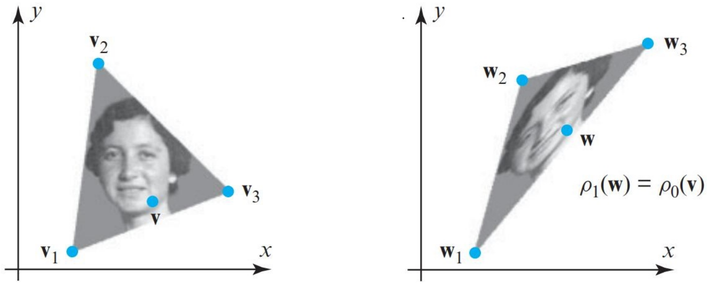
$$\begin{bmatrix} w_2 - w_1 & w_3 - w_1 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \underbrace{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}}_{C}$$

Therefore,

$$\begin{aligned} A &= \begin{bmatrix} T(e_1) & T(e_2) \end{bmatrix} \\ &= \begin{bmatrix} T(v_2 - v_1) & T(v_3 - v_1) \end{bmatrix} B^{-1} \\ &= \begin{bmatrix} w_2 - w_1 & w_3 - w_1 \end{bmatrix} B^{-1} \\ &= \begin{bmatrix} e_1 & e_2 \end{bmatrix} C B^{-1} \\ &= C B^{-1} \end{aligned}$$

We then combines the above linear transformation with a translation from v_1 to w_1, and finally we have

$$w = Av + b$$

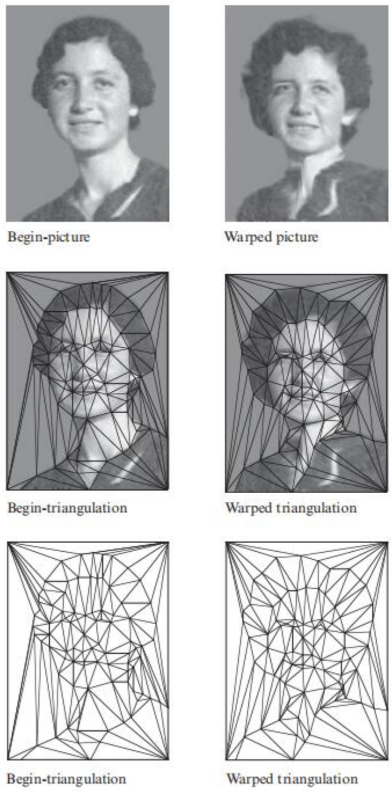


with $b = w_1 - v_1$.

At each pixel in the triangle be given by the three noncollinear point w_1, w_2, and w_3, its pixel-density is determined by

$$\rho_1(w) = \rho_0(v)$$

where ρ_0 and ρ_1 are the pixel-density of the begin-triangle and warped-triangle, repectively.



A time-varying warp is the set of warps generated when the vertex points of the begin-picture are moved continually in time from their original positions to

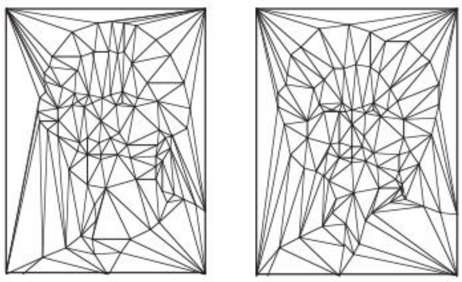
$t = 0.00$ $t = 0.25$ $t = 0.50$ $t = 0.75$ $t = 1.00$

specified final positions. This gives us a motion picture in which the begin-picture is continually warpedvto a final warp. Let us choose time units so that t = 0 corresponds to our begin-picture and $t = 1$ corresponds to our final warp. The simplest way of moving the vertex points from time 0 to time 1 is with constant velocity along straight-line paths from their initial positions to their final positions.

To describe such a motion, let $u_i(t)$ denote the position of the ith vertex point at any time t between 0 and 1. Thus $u_i(0) = v_i$ (its given position in the begin-picture) and $u_i(1) = w_i$ (its given position in the final warp). In between, we determine its position by

$$u_i(t) = (1 - t)v_i + tw_i$$

Begin-picture End-picture



Begin-triangulation End-triangulation

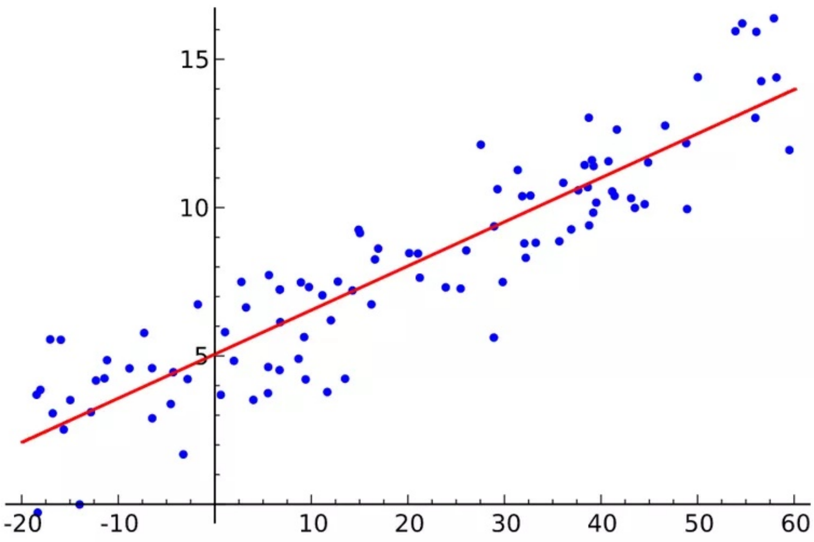
A time-varying morph can be described as a blending of two time-varying warps of two different pictures using two triangulations that match corresponding features in the two pictures. One of the two pictures is designated as the begin-picture and the other as the end-picture. First, a time-varying warp from $t = 0$ to $t = 1$ is generated in which the begin-picture is warped into the shape of the end-picture. Then a time-varying warp from $t = 1$ to $t = 0$ is generated in which the end-picture is warped into the shape of the begin-picture. Finally, a weighted average of picture-density of the two warps at each time t is produced to generate the morph of the two images at time t, as

$$\rho_t(u) = (1 - t)\rho_0(v) + t\rho_1(w).$$

$t = 0.00$ $t = 0.25$ $t = 0.50$ $t = 0.75$ $t = 1.00$

3.8 Projections and Least Squares Problems

Motivation:



Note 3.8.1: *A least squares problem can generally be formulated as an overdetermined linear system of equations involving more equations than unknowns.*

Given a system of equations $Ax = b$, where A is an $m \times n$ matrix with $m > n$, we cannot expect in general to find a vector $x \in \mathbb{R}^n$ for which Ax equals b. Instead, we can look for a vector x for which Ax is closest to b.

For each $x \in \mathbb{R}^n$, we can form a **residual**

$$r(x) = b - Ax.$$

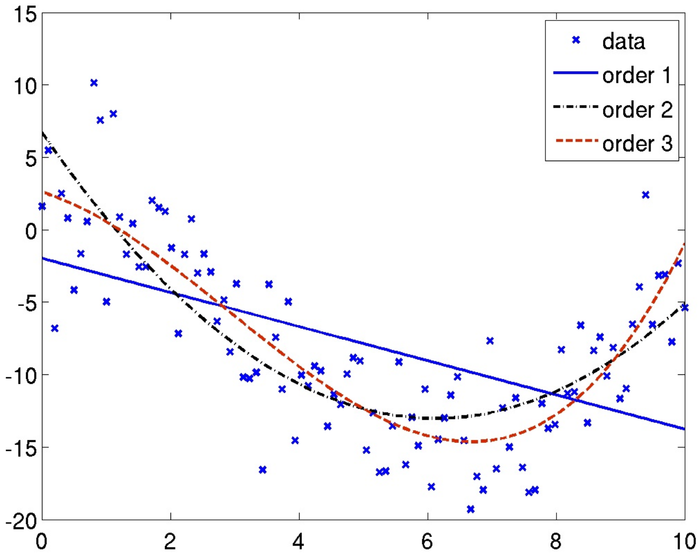
The distance between b and Ax is given by $\|b - Ax\| = \|r(x)\|$. We wish to find a vector $x \in \mathbb{R}^n$ for which $\|r(x)\|$ will be a minimum. Minimizing $\|r(x)\|$ is equivalent to minimizing $\|r(x)\|^2$. A vector $\hat{x}$ that accomplishes this is said to be a **least squares solution** of the system $Ax = b$.

If $\hat{x}$ is a least squares solution of the system $Ax = b$ and $p = A\hat{x}$, then p is a vector in the column space of A ($C(A)$) that is closest to b.

Definition 3.8.1: **Projection onto a Subspace**
Let V be a finite-dimensional vector space and S is a subspace of V. For any direct sum $V = S \oplus W$ and any $v \in V$, the **projection of v onto S along W** *is*

$$\mathrm{Proj}_{S,W}(v) = p$$

where $v = p + w$ with $p \in S$ and $w \in W$.



Theorem 3.8.1: **Projection is a Linear Transformation (Operator)**
The projection defined in Definition 3.8.1 is a linear transformation (operator) on V.

Proof: Since $V = S \oplus W$, for any $v \in V$, v can be expressed uniquely as a sum

$$v = p + w$$

where $p \in S$ and $w \in W$. Therefore, $\text{Proj}_{S,W}$ defines a mapping on V.

For $v_1, v_2 \in V$, they can be expressed uniquely as the following sums

$$v_1 = p_1 + w_1, \qquad v_2 = p_2 + w_2,$$

where $p_1, p_2 \in S$ and $w_1, w_2 \in W$. Then $\text{Proj}_{S,W}(v_1) = p_1$ and $\text{Proj}_{S,W}(v_2) = p_2$.

Since S and W are subspaces, $p_1 + p_2 \in S$ and $w_1 + w_2 \in W$. And the sum

$$v_1 + v_2 = (p_1 + w_1) + (p_2 + w_2) = (p_1 + p_2) + (w_1 + w_2)$$

is unique. Then we have

$$\text{Proj}_{S,W}(v_1 + v_2) = p_1 + p_2 = \text{Proj}_{S,W}(v_1) + \text{Proj}_{S,W}(v_2).$$

For any scalar α,

$$\text{Proj}_{S,W}(\alpha v) = \text{Proj}_{S,W}(\alpha p + \alpha w) = \alpha p = \alpha \, \text{Proj}_{S,W}(v).$$

Note 3.8.2:

$$\text{null}\,\text{Proj}_{\text{S,W}} = W, \quad \text{range}\,\text{Proj}_{\text{S,W}} = S.$$

Definition 3.8.2: **Orthogonal Projection onto a Subspace**

Let S be a subspace of $\mathbb{R}^m$ and $S^\perp$ be its orthogonal complement. Then $\mathbb{R}^m = S \oplus S^\perp$. The **orthogonal projection** *of any vector $b \in \mathbb{R}^m$ onto S along $S^\perp$ is*

$$\text{Proj}_S(b) = p$$

where $b = p + w$ with $p \in S$ and $w \in S^\perp$.

Theorem 3.8.2:

Let S be a subspace of $\mathbb{R}^m$, and let $b \in \mathbb{R}^m$. Then there is a unique $p \in S$ satisfies

$$\|b - p\| \leq \|b - y\|, \qquad y \in S$$

if and only if $b - p \in S^\perp$. Specifically, $p = \text{Proj}_S(b) \in S$ is the unique closest vector in S to $b \in \mathbb{R}^m$.

Proof: Since $\mathbb{R}^m = S \oplus S^\perp$, b can be expressed uniquely as a sum

$$b = p + w$$

where $p = \text{Proj}_S(b) \in S$ and $w = b - p \in S^\perp$.

For any $y \neq p \in S$,

$$\|b - y\|^2 = \|(b - p) + (p - y)\|.$$

Since $b - p \in S^\perp$ and $p - y \in S$, it follows

$$\begin{aligned}
\|b - y\|^2 &= ((b - p) + (p - y))^T ((b - p) + (p - y)) \\
&= (b - p)^T (b - p) + 2(b - p)^T (p - y) + (p - y)^T (p - y) \\
&= \|b - p\|^2 + \|p - y\|^2 \\
&> \|b - p\|^2.
\end{aligned} \tag{3.2}$$

Thus if $p \in S$ and $b - p \in S^\perp$, then p is the unique vector in S that is closest to b. Conversely, if $q \in S$ and $b - q \notin S^\perp$, then $q \neq p$, and it follows from (3.2) that $\|b - q\| > \|b - p\|$, which means that q is not the closest vector in S to b. ■

A vector $\hat{x}$ will be a solution of the least squares problem $Ax = b$ if and only if $p = A\hat{x}$ is the vector in $C(A)$ that is closest to b, *i.e.*, $p = \mathrm{Proj}_{C(A)}(b)$. From Theorem 3.8.2, we can get

$$b - p = b - A\hat{x} \in C(A)^{\perp} = N(A^T).$$

Thus,

$$A^T(b - A\hat{x}) = 0 \Longleftrightarrow A^T A\hat{x} = A^T b.$$

Theorem 3.8.3: *If A is an $m \times n$ matrix of rank n, the normal equations*

$$A^T A\hat{x} = A^T b$$

have a unique solution

$$\hat{x} = (A^T A)^{-1} A^T b$$

and $\hat{x}$ is the unique least squares solution of the system $Ax = b$.

Proof: It is sufficient to show that $A^T A$ is nonsingular. Since $\operatorname{rank} A = n$, $\dim C(A) = n$ and $\dim N(A) = 0$. Thus $N(A) = \{0\}$.

Let $z \in N(A^T A)$. We have $A^T Az = 0$. Then $z^T A^T Az = \|Az\|^2 = 0$, and thus $Az = 0$. Clearly, $z \in N(A)$. On the other hand, for any $z \in N(A)$, $Az = 0$. We have $A^T Az = 0$ and thus $z \in N(A^T A)$. Therefore, $N(A^T A) = N(A) = \{0\}$ and $\operatorname{rank}(A^T A) = \dim C(A^T A) = n$. We can conclude that $A^T A$ is nonsingular. ■

Note 3.8.3: *$p = \mathrm{Proj}_{C(A)}(b) = A\hat{x} = A(A^T A)^{-1} A^T b$ is the vector in $C(A)$ that is closest to b in the least squares sense. The matrix represention the projection onto $C(A)$ along $C(A)^{\perp} = N(A)$ is*

$$A(A^T A)^{-1} A^T$$

is called the **projection matrix.**

3.9 Algebraic Operations on $\mathcal{L}(V, W)$

Definition 3.9.1: **Addition and scalar multiplication on** $\mathcal{L}(V, W)$
Suppose $S, T \in \mathcal{L}(V, W)$ and α is a scalar. The sum $S + T$ and the product λT are linear transformations V to W defined by

$$(S + T)(v) = S(v) + T(v) \quad (\lambda T)(v) = \lambda(T(v)).$$

Note 3.9.1: *With the operations of addition and scalar multiplication as defined above, $\mathcal{L}(V, W)$ is a vector space.*

Given a linear transformation $T \in \mathcal{L}(V, W)$, let $M(T)$ be its matrix representation.

Question: matrix addition, $M(S + T) = M(S) + M(T)$?

$S, T, S + T \in \mathcal{L}(V, W)$, $\dim V = n$, $\dim W = m$.
$\{v_1, \cdots, v_n\}$ is a basis of V, $\quad \{w_1, \ldots, w_m\}$ is a basis of W.

Let

$$\begin{aligned} M(S) &\triangleq A = [a_{i,j}] \in \mathbb{R}^{m \times n}, \\ M(T) &\triangleq B = [b_{i,j}] \in \mathbb{R}^{m \times n}, \\ M(S + T) &\triangleq C = [c_{i,j}] \in \mathbb{R}^{m \times n}. \end{aligned}$$

$$\begin{aligned} \Longrightarrow S(v_j) &= \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{bmatrix} = \sum_{i=1}^{m} a_{i,j} w_i, \\ T(v_j) &= \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} b_{1,j} \\ \vdots \\ b_{m,j} \end{bmatrix} = \sum_{i=1}^{m} b_{i,j} w_i, \\ (S + T)(v_j) &= \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} c_{1,j} \\ \vdots \\ c_{m,j} \end{bmatrix} = \sum_{i=1}^{m} c_{i,j} w_i. \end{aligned}$$

From Definition 3.9.1, $(S + T)(v_j) = S(v_j) + T(v_j)$. Then we have

$$\begin{aligned} \sum_{i=1}^{m} c_{i,j} w_i &= \sum_{i=1}^{m} a_{i,j} w_i + \sum_{i=1}^{m} b_{i,j} w_i \\ &= \sum_{i=1}^{m} (a_{i,j} + b_{i,j}) w_i. \end{aligned}$$

$$\Longrightarrow c_{i,j} = a_{i,j} + b_{i,j} \quad (w_1, \ldots, w_m \quad \text{are linearly independent.})$$

Question: scalar multiplication, $M(\lambda T) = \lambda M(T)$?

Let

$$M(T) \triangleq B = [b_{i,j}] \in \mathbb{R}^{m \times n},$$
$$M(\lambda T) \triangleq C = [c_{i,j}] \in \mathbb{R}^{m \times n}.$$

$$\Longrightarrow T(v_j) = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} b_{1,j} \\ \vdots \\ b_{m,j} \end{bmatrix} = \sum_{i=1}^{m} b_{i,j} w_i,$$
$$(\lambda T)(v_j) = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} c_{1,j} \\ \vdots \\ c_{m,j} \end{bmatrix} = \sum_{i=1}^{m} c_{i,j} w_i.$$

From Definition 3.9.1, $(\lambda T)(v_j) = \lambda(T(v_j))$. Then we have

$$\sum_{i=1}^{m} c_{i,j} w_i = \lambda \sum_{i=1}^{m} b_{i,j} w_i$$
$$= \sum_{i=1}^{m} (\lambda b_{i,j}) w_i.$$

$$\Longrightarrow c_{i,j} = \lambda b_{i,j} \quad (w_1, \ldots, w_m \quad \text{are linearly independent.})$$

Definition 3.9.2: **Product of Linear Transformations**
Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$. The product $ST \in \mathcal{L}(U, W)$ is defined by

$$(ST)(u) = S(T(u))$$

for any $u \in U$.

Question: $(ST)(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 (ST)(u_1) + \alpha_2 (ST)(u_2)$???

$$
\begin{aligned}
&(ST)\left(\alpha_1 u_1+\alpha_2 u_2\right) \\
&\stackrel{Def.}{=} S\left(T\left(\alpha_1 u_1+\alpha_2 u_2\right)\right) \\
&\stackrel{T\in\mathcal{L}(U,V)}{=} S\left(\alpha_1 T\left(u_1\right)+\alpha_2 T\left(u_2\right)\right) \\
&\stackrel{S\in\mathcal{L}(V,W)}{=} \alpha_1 S\left(T\left(u_1\right)\right)+\alpha_2 S\left(T\left(u_2\right)\right) \\
&\stackrel{Def.}{=} \alpha_1(ST)\left(u_1\right)+\alpha_2(ST)\left(u_2\right).
\end{aligned}
$$

Theorem 3.9.1: **Algebraic Properties of Products of Linear Transformations**
Associativity: $(T_1T_2)T_3 = T_1(T_2T_3)$.
Identity: $TI_V = I_WT = T$.
Distributive Properties: $(S_1+S_2)T = S_1T+S_2T$ *and* $S(T_1+T_2) = ST_1+ST_2$.

Question: matrix multiplication, $M(ST) = M(S) * M(T)$??

For $T \in \mathcal{L}(U,V), S \in \mathcal{L}(V,W), ST \in \mathcal{L}(U,W)$, $\dim U = p$, $\dim V = n$, $\dim W = m$, $\{u_1,\dots,u_p\}$ is a basis of U, $\{v_1,\dots,v_n\}$ is a basis of V, $\{w_1,\dots,w_m\}$ is a basis of W.

Let

$$
\begin{aligned}
M(S) &\triangleq A = [a_{i,j}] \in \mathbb{R}^{m\times n}, \\
M(T) &\triangleq B = [b_{i,j}] \in \mathbb{R}^{n\times p}, \\
M(ST) &\triangleq C = [c_{i,j}] \in \mathbb{R}^{m\times p}.
\end{aligned}
$$

$$
\Longrightarrow \text{for } v_k \in V, S\left(v_k\right) = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} a_{1,k} \\ \vdots \\ a_{m,k} \end{bmatrix} = \sum_{i=1}^{m} a_{i,k} w_i,
$$

$$
\text{for } u_j \in U, T\left(u_j\right) = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} b_{1,j} \\ \vdots \\ b_{n,j} \end{bmatrix} = \sum_{k=1}^{n} b_{k,j} v_k,
$$

$$
\text{for } u_j \in U, (ST)\left(u_j\right) = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} c_{1,j} \\ \vdots \\ c_{m,j} \end{bmatrix} = \sum_{i=1}^{m} c_{i,j} w_i.
$$

From Definition 3.9.2, $(ST)(u_j) = S(T(u_j))$. Then we have

$$\begin{aligned}
\sum_{i=1}^{m} c_{i,j} w_i =& S(T(u_j)) \\
=& S(\sum_{k=1}^{n} b_{k,j} v_k) \\
=& \sum_{k=1}^{n} b_{k,j} S(v_k) \quad \text{(Since } S \text{ is a linear transformation.)} \\
=& \sum_{k=1}^{n} b_{k,j} \sum_{i=1}^{m} a_{i,k} w_i \\
=& \sum_{k=1}^{n} \sum_{i=1}^{m} b_{k,j} a_{i,k} w_i \\
=& \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{i,k} b_{k,j} \right) w_i
\end{aligned}$$

$$\implies c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j} \quad (w_1, \ldots, w_m \quad \text{are linearly independent.})$$

3.10 Invertibility and Isomorphic Vector Spaces

Definition 3.10.1: *Let $T \in \mathcal{L}(V, W)$. We say that $S \in \mathcal{L}(W, V)$ is the inverse of T if $ST = I_V$ and $TS = I_W$, where I_V and I_W are the identity transformations (operators) on V and W, respectively. We say that T is invertible if it has an inverse, and call the inverse T^{-1}; thus $TT^{-1} = I_W$ and $T^{-1}T = I_V$.*

Note 3.10.1: *The definition is symmetric: if S is the inverse ef T, then T is the inverse of S.*

Theorem 3.10.1: *An invertible linear transformation has a unique inverse.*

Proof: Suppose $T \in \mathcal{L}(V, W)$ is invertible and S_1, S_2 are inverses of T.

$$S_1 = S_1 I_W = S_1 (TS_2) = (S_1 T) S_2 = I_V S_2 = S_2.$$

■

Theorem 3.10.2: *Let $T \in \mathcal{L}(V, W)$. T is invertible if and only if T is injective and surjective.*

Proof: ($\Longrightarrow$) Suppose T is invertible. Let $u, v \in V$ and $T(u) = T(v) \in W$. Then $u = I_V u = (T^{-1}T)(u) = T^{-1}(T(u)) = T^{-1}(T(v)) = v$. Hence, T is injective.

For any $w \in W, T^{-1}(w) \in V$. $T\left(T^{-1}(w)\right) = w \in \operatorname{range} T$. Since $\operatorname{range} T$ is a subspace of W, then $\operatorname{range} T = W$. Hence T is surjective.

($\Longleftarrow$) Suppose T is injective and surjective. For any $w \in W$, since T is surjective, we have $w = T(v)$ for some $v \in V$; since T is injective, this v is unique.

Define $S(w) = v$, with $S : W \to V$. For any w, $T(S(w)) = T(v) = w$. Thus $TS = I_W$.

Next we show that $ST = I_V$, *i.e.*, $(ST)(v) = v$ for any $v \in V$. Since $(TS)(w) = w$ for any $w \in W$. In particular, $TS(T(v)) = T(v)$. Since T is injective, $(ST)(v) = v \Rightarrow ST = I_V$.

Let $w_1, w_2 \in W$; we need to show that $S(w_1 + w_2) = S(w_1) + S(w_2)$.

$$\begin{aligned}
T\left(S(w_1 + w_2)\right) =& (TS)\left(w_1 + w_2\right) \\
=& I_W\left(w_1 + w_2\right) \\
=& I_W(w_1) + I_W(w_2) \\
=& (TS)\left(w_1\right) + (TS)\left(w_2\right) \\
=& T\left(S\left(w_1\right)\right) + T\left(S\left(w_2\right)\right) \\
=& T\left(S\left(w_1\right) + S\left(w_2\right)\right)
\end{aligned}$$

Sine T is injective, $S(\omega_1 + w_2) = S(w_1) + S(\omega_2)$.
Let $w \in W$ and λ be a scalar.

$$\begin{aligned} T(S(\lambda w)) = (TS)(\lambda w) = I_W(\lambda w) = \lambda I_W(w) &= \lambda(TS)(w) \\ &= \lambda T(S(w)) \\ &= T(\lambda S(w)). \end{aligned}$$

Since T is injective, $S(\lambda w) = \lambda S(w)$. ■

Definition 3.10.2: **Isomorphic Vector Spaces**

(1) An **isomorphism** *is an invertible linear transformation.*

(2) Two vector spaces V and W are said to be **isomorphic** *if there is an invertible linear transformation from one space to another.*

Example 3.10.1: *$T : \mathbb{R}^3 \to \mathbb{P}_2(\mathbb{R})$ defined by*

$$T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = ax^2 + bx + c.$$

■

Theorem 3.10.3: *Two finite-dimensional spaces V and W are isomorphic of and only if $\dim V = \dim W$.*

Proof: ($\Longrightarrow$) If V and W are isomorphic, there is an invertible $T \in \mathcal{L}(V, W)$, which is injective and surjective. Thus, $\operatorname{null} T = \{0\}$, $\operatorname{range} T = W$. From the fundamental theorem of linear transformations, we have

$$\dim V = \dim \operatorname{null}(T) + \dim \operatorname{range} T = 0 + \dim W = \dim W.$$

($\Longleftarrow$) Suppose that $\dim V = \dim W = n$. Then V has a basis $\{v_1, ..., v_n\}$ and W has a basis $\{w_1, ..., w_n\}$. Define a linear transformation $T : V \to W$ by $T(v_j) = w_j$, $j = 1, \ldots, n$. Clearly, for any $v = a_1v_1 + ... + a_nv_n \in V$, $T(v) = a_1w_1 + ... + a_nw_n \in W$.

If $T(v) = 0$, we have $a_1w_1 + \cdots + a_nw_n = 0$. Since $\{w_1, \cdots w_n\}$ is a basis of W and thus linearly independent, $a_1 = \cdots = a_n = 0$. Therefore, $v = a_1v_1 + \cdots + a_nv_n = 0$. Thus, $\operatorname{null} T = \{0\}$, *i.e.*, T is injective.

Note that $\operatorname{range} T = \operatorname{Span}(w_1, \cdots, w_n) = W$, since $\{w_1, \cdots, w_n\}$ is a basis of W. Thus, T is also surjective. Therefore, T is invertible. Clearly, V and W are isomorphic.

■

Theorem 3.10.4: *Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(i) T is invertible;

(ii) T is injective;

(iii) T is surjective;

Proof: (i) $\Longrightarrow$ (ii): it is easily obtained from Theorem 3.10.2.

(ii) $\Longrightarrow$ (iii): If T is injective, $\operatorname{null} T = 0$. From the fundamental theorem of linear transformations, for a linear operator $T \in \mathcal{L}(V)$,

$$\dim V = \operatorname{null} T + \dim \operatorname{range} T = \dim \operatorname{range} T$$

Thus, $\operatorname{range} T = V$. Therefore, T is surjective.

(iii) $\Longrightarrow$ (ii): If T is surjective, $\operatorname{range} T = V$.

$$\operatorname{null} T = \dim V - \dim \operatorname{range} T = 0.$$

Thus, $\operatorname{null} T = \{0\}$. Therefore, T is injective.

■

Theorem 3.10.5: **Linear Transformation Lemma:**

If $v_1, \ldots, v_n$ is a basis for vector space V and $w_1, \ldots, w_n$ is a basis for vector space W, then there is a unique linear transformation $T : V \to W$ with $Tv_i = w_i$.

Proof: For any $v \in V$, there exist scalars $c_1, \cdots, c_n$ such that

$$v = c_1 v_1 + \cdots + c_n v_n.$$

Define $T : V \to W$ by

$$T(v) = T(c_1 v_1 + \cdots + c_n v_n) = c_1 w_1 + \cdots + c_n w_n.$$

For each i, let $c_i = 1$, $c_j = 0$, $j \neq i$. We have $T(v_i) = w_i$.

Next, we show that T is a linear transformation. For $u \in V$, there exist scalars $b_1, \cdots, b_n$ such that

$$u = b_1 v_1 + \cdots + b_n v_n,$$

and thus

$$T(u) = b_1 w_1 + \cdots + b_n w_n.$$

For scalars α and β,

$$\begin{aligned}
T(\alpha u + \beta v) &= T\left(\alpha(b_1 v_1 + \cdots + b_n v_n) + \beta(c_1 v_1 + \cdots + c_n v_n)\right) \\
&= T\left((\alpha b_1 + \beta c_1) v_1 + \cdots + (\alpha b_n + \beta c_n) v_n\right) \\
&= (\alpha b_1 + \beta c_1) w_1 + \cdots + (\alpha b_n + \beta c_n) w_n \\
&= \alpha(b_1 w_1 + \cdots + b_n w_n) + \beta(c_1 w_1 + \cdots + c_n w_n) \\
&= \alpha T(u) + \beta T(v).
\end{aligned}$$

Thus, $T \in \mathcal{L}(V, W)$.

Note that $T(c_1 v_1 + \cdots + c_n v_n) = c_1 w_1 + \cdots + c_n w_n$. Since $\{v_1, \cdots, v_n\}$ is a basis of V, for any $v \in V$, $c_1, \cdots, c_n$ are uniquely determined. Therefore, T is uniquely determined on V. ■

Theorem 3.10.6: *Suppose $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_m$ is a basis of W. Then the matrix representation of linear transformation $\mathcal{M}$ is an isomorphism between $\mathcal{L}(V, W)$ and $\mathbb{F}^{m,n}$. And $\dim \mathcal{L}(V, M) = \dim \mathbb{F}^{m,n} = mn$.*

Proof: For $T_1, T_2 \in \mathcal{L}(V, W)$. Let $\mathcal{M}(T_1), \mathcal{M}(T_2)$ be their matrix representations, respectively. Under the algebraic operations of linear transformations, we have

$$\mathcal{M}(T_1 + T_2) = \mathcal{M}(T_1) + \mathcal{M}(T_2), \quad \mathcal{M}(\lambda T_1) = \lambda \mathcal{M}(T_1).$$

Thus, $\mathcal{M} : \mathcal{L}(V, W) \to \mathbb{F}^{m,n}$ is linear. We need to show that $\mathcal{M}$ is injective and surjective.

Suppose $T \in \mathcal{L}(V, W)$ and $\mathcal{M}(T) = 0$. Then $T(v_i) = 0$, $i = 1, \cdots, n$. Since $v_1, \cdots, v_n$ is a basis of V. Then $T(v) = 0$ for all $v \in V$. $T = 0$. Thus $\mathcal{M}$ is injective.

To prove that $\mathcal{M}$ is surjective, suppose $A \in \mathbb{F}^{m,n}$. Construct the following map from V to W:

$$T(c_1 v_1 + \cdots + c_n v_n) = c_1 \sum_{i=1}^{m} a_{i,1} w_i + \cdots + c_n \sum_{i=1}^{m} a_{i,n} w_i$$

Thus,

$$T(v_j) = \sum_{i=1}^{m} a_{i,j} w_i = \begin{bmatrix} w_1 & \cdots & w_m \end{bmatrix} \begin{bmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{bmatrix}.$$

We next show that $T \in \mathcal{L}(V, W)$. For any $u, v \in V$, there exist scalars $b_1, \ldots, b_n$ and $c_1, \ldots, c_n$ such that

$$u = b_1 v_1 + \cdots + b_n v_n, \quad v = c_1 v_1 + \cdots + c_n v_n$$

$$
\begin{aligned}
T(\alpha u + \beta v) &= T\left(\alpha(b_1 v_1 + \cdots + b_n v_n) + \beta(c_1 v_1 + \cdots + c_n v_n)\right) \\
&= T\left((\alpha b_1 + \beta c_1) v_1 + \cdots + (\alpha b_n + \beta c_n) v_n\right) \\
&= (\alpha b_1 + \beta c_1) \sum_{i=1}^{m} a_{i,1} w_i + \cdots + (\alpha b_n + \beta c_n) \sum_{i=1}^{m} a_{i,n} w_i \\
&= \alpha T(u) + \beta T(v).
\end{aligned}
$$

Clearly, $T \in \mathcal{L}(V, W)$ and $\mathcal{M}(T)$ is its matrix representation. And $\dim \mathcal{L}(V, W) = \dim \mathbb{F}^{m,n} = mn$. ■

3.11 Products and Quotients of Vector Spaces

Definition 3.11.1: **Product of Vector Spaces** *Suppose $V_1, \ldots, V_m$ are vector spaces over* $\mathbf{F}$. *The product $V_1 \times \cdots \times V_m$ is defined by*

$$V_1 \times \cdots \times V_m = \{(v_1, \ldots, v_m) : v_1 \in V_1, \ldots, v_m \in V_m\}.$$

Addition and scalar multiplication on $V_1 \times \cdots \times V_m$ are defined by

$$(u_1, \ldots, u_m) + (v_1, \ldots, v_m) = (u_1 + v_1, \ldots, u_m + v_m)$$

and

$$\lambda(v_1, \ldots, v_m) = (\lambda v_1, \ldots, \lambda v_m),$$

where $\lambda \in \mathbb{F}$.

Theorem 3.11.1: *Suppose $V_1, \ldots, V_m$ are vector spaces over $\mathbb{F}$. Then $V_1 \times \cdots \times V_m$ is a vector space over $\mathbb{F}$.*

Example 3.11.1: *Product of the vector spaces $\mathbb{P}_5(\mathbb{R})$ and $\mathbb{R}^3$.*

Elements of $\mathbb{P}_5(\mathbb{R}) \times \mathbb{R}^3$ are lists of length two, with the first item in the list an element of $\mathbb{P}_5(\mathbb{R})$ and the second item in the list an element of $\mathbb{R}^3$.

For example, $(5 - 6x + 4x^2, (3, 8, 7))$ and $(x + 9x^5, (2, 2, 2))$ are elements of $\mathbb{P}_5(\mathbb{R}) \times \mathbb{R}^3$. Their sum is defined by

$$\begin{aligned} &(5 - 6x + 4x^2, (3, 8, 7)) + (x + 9x^5, (2, 2, 2)) \\ =&(5 - 5x + 4x^2 + 9x^5, (5, 10, 9)). \end{aligned}$$

Also, $2(5 - 6x + 4x^2, (3, 8, 7)) = (10 - 12x + 8x^2, (6, 16, 14))$.

■

Example 3.11.2: *Elements of the vector space $\mathbb{R}^2 \times \mathbb{R}^3$ are lists*

$$((x_1, x_2), (x_3, x_4, x_5)),$$

where $x_1, x_2, x_3, x_4, x_5 \in \mathbb{R}$. Elements of $\mathbb{R}^5$ are lists

$$(x_1, x_2, x_3, x_4, x_5),$$

where $x_1, x_2, x_3, x_4, x_5 \in \mathbb{R}$.

■

Theorem 3.11.2: *Suppose $V_1, \dots, V_m$ are finite-dimensional vector spaces. Then $V_1 \times \cdots \times V_m$ is finite-dimensional and*

$$\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m.$$

Proof: Choose a basis of each V_i, $i = 1, \cdots, m$, denoted by $\{v_{i,1}, \cdots, v_{i,\dim V_i}\}$. Consider the following vectors of $V_1 \times \cdots \times V_m$.

$$\begin{aligned}
&(v_{1,1}, 0, \cdots, 0), \cdots, (v_{1,\dim V_1}, 0, \cdots, 0) &&\to \dim V_1 \text{ vectors}\\
&(0, v_{2,1}, 0, \cdots, 0), \cdots, (0, v_{2,\dim V_2}, 0, \cdots, 0) &&\to \dim V_2 \text{ vectors}\\
&\vdots\\
&(0, \cdots, 0, v_{m,1}), \cdots, (0, \cdots, 0, v_{m,\dim V_m}) &&\to \dim V_m \text{ vectors}
\end{aligned}$$

Let

$$\begin{aligned}
&a_{1,1}(v_{1,1}, 0, \cdots, 0) + \cdots + a_{1,\dim V_1}(v_{1,\dim V_1}, 0, \cdots, 0) + \cdots\\
&+a_{m,1}(0, \cdots, 0, v_{m,1}) + \cdots + a_{m,\dim V_m}(0, \cdots, 0, v_{m,\dim V_m}) = 0.
\end{aligned}$$

$$\Rightarrow \left(\underbrace{a_{1,1}v_{1,1} + \cdots + a_{1,\dim V_1}v_{1,\dim V_1}}_{0_{V_1}}, \cdots, \underbrace{a_{m,1}v_{m,1} + \cdots + a_{m,\dim V_m}v_{m,\dim V_m}}_{0_{V_m}}\right) = 0.$$

$\Rightarrow a_{1,1} = \cdots = a_{1,\dim V_1} = \cdots = a_{m,1} = \cdots = a_{m,\dim V_m} = 0.$

Clearly, the list of these vectors is linearly independent and spans $V_1 \times \cdots \times V_m$, thus is a basis of $V_1 \times \cdots \times V_m$. The length of this basis is $\dim V_1 + \cdots + \dim V_m$, as desired. ■

Theorem 3.11.3: *Suppose that $V_1, \dots, V_m$ are subspaces of V. Define a linear transformation $\Gamma : V_1 \times \cdots \times V_m \to V_1 + \cdots + V_m$ by*

$$\Gamma(v_1, \dots, v_m) = v_1 + \cdots + v_m.$$

Then $V_1 + \cdots + V_m$ is a direct sum if and only if Γ is injective.

Proof: ($\Rightarrow$). For any $v = (v_1, \cdots, v_m) \in$ null T. $T(v) = v_1 + \cdots + v_m = 0$. Since $V_1 + \cdots + V_m$ is a direct sum, $v_1 = \cdots = v_m = 0$. Thus $v = (0, \cdots, 0)$ and null $T = \{(0, \cdots, 0)\}$.

($\Leftarrow$). Since T is injective, if $T(v_1, \cdots, v_m) = 0$, then $v_1 = \cdots = v_m = 0$. Thus, $V_1 + \cdots + V_m$ is a direct sum. ■

Theorem 3.11.4: *Suppose V is finite-dimensional and $V_1, \dots, V_m$ are subspaces of V. Then $V_1 + \cdots + V_m$ is a direct sum if and only if*

$$\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m.$$

Proof: Note that T in Theorem 3.11.3 is injective, and also surjective. $V_1+\cdots+V_m$ is a direct sum if and only if

$$\dim(V_1 \times \cdots \times V_m) = \dim(V_1 + \cdots + V_m)$$

From Theorem 3.11.2, $\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m$. We have $\dim V_1 + \cdots + \dim V_m = \dim(V_1 + \cdots + V_m)$. ■

Let V be a vector space over a field $\mathbb{F}$, and let W be a subspace of V. There is a sense in which we can "divide" V by W to get a new vector space. Of course, the word "divide" is in quotation marks because we can't really divide vector spaces in the usual sense of division, but there is still an analog of division we can construct. This leads the notion of what's called a *quotient vector space*.

Definition 3.11.2: *A* ***partition*** *of a set S is a collection $\pi := \{B_1, \ldots, B_k\}$ consisting of pairwise disjoint nonempty subsets of S such that $S = \bigcup_{j=1}^{k} B_j$.*

Definition 3.11.3: *Let U be a set. An* ***equivalence relation*** *on U is a binary relation $\sim$ satisfying three properties:*

1) $x \sim x$ for all $x \in U$ (reflexivity)
2) if $x \sim y$ then $y \sim x$ (symmetry)
3) if $x \sim y$ and $y \sim z$ then $x \sim z$ (transitivity)

Definition 3.11.4: *Let V be a vector space and U a subspace. We can define an equivalence relation on V by $v \sim v'$ if $v' - v \in U$.*

(1) $v - v = 0 \in U \Rightarrow v \sim v$.

(2) $v \sim v' \Rightarrow v' - v \in U \Rightarrow v - v' \in U \Rightarrow v' \sim v$.

(3) $v_1 \sim v_2, v_2 \sim v_3 \Rightarrow v_2 - v_1 \in U, v_3 - v_2 \in U$.

$$v_3 - v_1 = (v_3 - v_2) + (v_2 - v_1) \in U \Rightarrow v_1 \sim v_3$$

$$\begin{aligned}\overline{v} :=& \{u \in V : v \sim u\} \\ =& \{u \in V : u - v \in U\} \\ =& \{u \in V : u - v = w \in U\} \\ =& \{u \in V : u = v + w, w \in U\} \\ =& \{v + w : w \in U\} \triangleq v + U\end{aligned}$$

Definition 3.11.5: *Let V be a vector space and U a subspace. Then the* ***equivalence class*** *of a vector v under the relation $\sim$ above is the set*

$$v + U = \{v + w : w \in U\};$$

this is called an ***affine subspace****, a* ***coset****, or a* ***translate*** *for v and U.*

Example 3.11.3:

1. *If U is the line in $\mathbb{R}^2$ defined by $U = \{(x, 2x) \in \mathbb{R}^2 : x \in \mathbb{R}\}$, then all lines in $\mathbb{R}^2$ with slope 2 are translates of U.*
2. *More generally, if U is a line in $\mathbb{R}^2$, then the set of all translates of U is the set of all lines in $\mathbb{R}^2$ that are parallel to U.*
3. *If $U = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\}$, then the translates of U are the planes in $\mathbb{R}^3$ that are parallel to the xy-plane U.*
4. *More generally, if U is a plane in $\mathbb{R}^3$, then the set of all translates of U is the set of all planes in $\mathbb{R}^3$ that are parallel to U.*

■

Theorem 3.11.5: *Suppose U is a subspace of V and $v, w \in V$. Then*

$$v - w \in U \iff v + U = w + U \iff (v + U) \cap (w + U) \neq \varnothing.$$

Proof: Suppose $v - w \in U$. For $u \in U$,

$$\left.\begin{array}{l} v + u = w + (v - w + u) \in w + U \Rightarrow v + U \subseteq w + U \\ w + u = v + (w - v + u) \in v + U \Rightarrow w + U \subseteq v + U \end{array}\right\} \Rightarrow v + U = w + U$$

Suppose $v + U = w + U$. $(v + U) \cap (w + U) = v + U \neq \varnothing$.
Suppose $(v + U) \cap (w + U) \neq \varnothing$. There exist $u_1, u_2 \in U$ such that

$$v + u_1 = w + u_2 \Rightarrow v - w = u_2 - u_1 \in U$$

■

Definition 3.11.6: *Suppose U is a subspace of V. Then the* ***quotient space*** *V/U is the set of all translates of U. Thus*

$$V/U = \{v + U : v \in V\}.$$

Theorem 3.11.6: *Suppose U is a subspace of V. V/U is a vector space itself, with addition and scalar multiplication defined by*

$$(v+U)+(w+U)=(v+w)+U \qquad a\cdot(v+U)=(a\cdot v)+U.$$

Proof: Suppose that $v_1, v_2, w_1, w_2 \in V$ and

$$v_1+U=v_2+U, \quad w_1+U=w_2+U.$$

We need to show that $v_1+w_1+U=v_2+w_2+U$.

Since $v_1+U=v_2+U$ and $w_1+U=w_2+U$, we have $v_1-v_2\in U$ and $w_1-w_2\in U$.

$$(v_1+w_1)-(v_2+w_2)=(v_1-v_2)+(w_1-w_2)\in U \Rightarrow v_1+w_1+U=v_2+w_2+U.$$

Thus, the definition of addition on V/U makes sense.

Suppose $v_1+U=v_2+U$ with $v_1, v_2\in V$.

$$v_1-v_2\in U\Rightarrow \lambda(v_1-v_2)\in U, \quad \lambda v_1+U=\lambda v_2+U.$$

Thus, the definition of scalar multiplication on V/U also makes sense.

Zero: $0+U=U$. Additive inverse of $v+U$: $-v+U$. ∎

Definition 3.11.7: *Suppose U is a subspace of V. The **quotient map** $\pi: V\to V/U$ is the linear transformation defined by*

$$\pi(v)=v+U$$

for each $v\in V$.

$$\begin{aligned}\pi(\alpha v_1+\beta v_2)&=\alpha v_1+\beta v_2+U\\&=\alpha v_1+U+\beta v_2+U\\&=\alpha(v_1+U)+\beta(v_2+U)\\&=\alpha\pi(v_1)+\beta\pi(v_2)\end{aligned}$$

Theorem 3.11.7: *Suppose V is finite-dimensional and U is a subspace of V. Then*

$$\dim V/U=\dim V-\dim U.$$

Proof: Consider the linear transformation π in Def. 3.11.7. Note that

$$v + U = 0 + U \Leftrightarrow v \in U.$$

Then null $\pi = U$. Clearly range $\pi = V/U$.

$$\begin{aligned} \dim V &= \dim \text{null}\ \pi + \dim \text{range}\ \pi \\ &= \dim U + \dim V/U \\ \Rightarrow \dim V/U &= \dim V - \dim U. \end{aligned}$$

■

Find a basis of $\dim V/U$:

Suppose $\dim U = m$, $\dim V = n$. Let $\{v_1, \cdots, v_m\}$ be a basis of U, which can be extended to a basis of V, $\{v_1, \cdots, v_m, v_{m+1}, \cdots, v_n\}$.

Consider $v_{m+1} + U, \cdots, v_n + U$. Since $v_1, \cdots, v_m \in U$, $v_1 + U = \cdots = v_m + U = U$.

Let

$$\begin{aligned} &\beta_{m+1}(v_{m+1} + U) + \cdots + \beta_n(v_n + U) = 0 \\ \Rightarrow &(\beta_{m+1}v_{m+1} + \cdots + \beta_n v_n) + U = 0 \\ \Rightarrow &\beta_{m+1}v_{m+1} + \cdots + \beta_n v_n \in U. \end{aligned}$$

Since $v_{m+1}, \cdots, v_n$ are linearly independent, $\beta_{m+1} = \cdots = \beta_n = 0$. Therefore, $v_{m+1} + U, \cdots, v_n + U$ are linearly independent.

For any $v + U \in V/U$, $v = c_1v_1 + \cdots + c_nv_n$.

$$\begin{aligned} v + U &= (c_1v_1 + U) + \cdots + (c_mv_m + U) + (c_{m+1}v_{m+1} + U) + \cdots + (c_nv_n + U) \\ &= c_{m+1}(v_{m+1} + U) + \cdots + c_n(v_n + U) \in \text{Span}(v_{m+1} + U, \cdots, v_n + U) \end{aligned}$$

Clearly, $\text{Span}(v_{m+1} + U, \cdots, v_n + U) \subseteq V/U$. Thus

$$V/U = \text{Span}(v_{m+1} + U, \cdots, v_n + U).$$

Therefore, $v_{m+1} + U, \cdots, v_n + U$ is a basis of V/U and $\dim V/U = n - m = \dim V - \dim U$.

Theorem 3.11.8: *Suppose V is finite-dimensional and U is a subspace of V. If V/U has a basis, $\beta_1 + U, \cdots, \beta_t + U$. Let $W = \text{span}(\beta_1, \ldots, \beta_t)$. Then $V = W \oplus U$, and $\beta_1, \ldots, \beta_t$ is a basis of W.*

Proof: For any $v \in V$, $v + U \in V/U$.

$$\begin{aligned} v + U &= c_1(\beta_1 + U) + \cdots + c_t(\beta_t + U) \\ &= (c_1\beta_1 + \cdots + c_t\beta_t) + U \end{aligned}$$

$$\Rightarrow v - (c_1\beta_1 + \cdots + c_t\beta_t) \in U$$

Let $\beta = c_1\beta_1 + \cdots + c_t\beta_t \in W$. Then $\eta = v - \beta \in U$ and $v = \beta + \eta$, so $V = W + U$. For any $r \in W \cap U$, $r \in W$, $r = c_1\beta_1 + \cdots + c_t\beta_t$. Since $r \in U$, $r + U = U$.

$$\begin{aligned} U &= (c_1\beta_1 + \cdots + c_t\beta_t) + U \\ &= c_1(\beta_1 + U) + \cdots + c_t(\beta_t + U) \end{aligned}$$

Since $\beta_1 + U, \cdots, \beta_t + U$ is a basis of V/U, $c_1 = \cdots = c_t = 0$, so $r = 0_V$. Then $W \cap U = \{0_V\}$, and $V = W \oplus U$.

Let $k_1\beta_1 + \cdots + k_t\beta_t = 0$.

$$U = 0 + U = (k_1\beta_1 + \cdots + k_t\beta_t) + U \Rightarrow k_1 = \cdots = k_t = 0$$

■

Each linear map T on induces a linear map $\widetilde{T}$ on $V/(\text{null})$, as defined below.

Definition 3.11.8: *Suppose $T \in \mathcal{L}(V, W)$. Define $\widetilde{T} : V/(\text{null } T) \to W$ by*

$$\widetilde{T}(v + \text{null } T) = Tv.$$

Theorem 3.11.9: *Suppose $T \in \mathcal{L}(V, W)$. Then*

1. *$\widetilde{T} \circ \pi = T$, where π is the quotient map of V onto $V/(\text{null } T)$;*
2. *$\widetilde{T}$ is injective;*
3. *range $\widetilde{T}$ = range T;*
4. *$V/(\text{null } T)$ and range T are isomorphic vector spaces.*