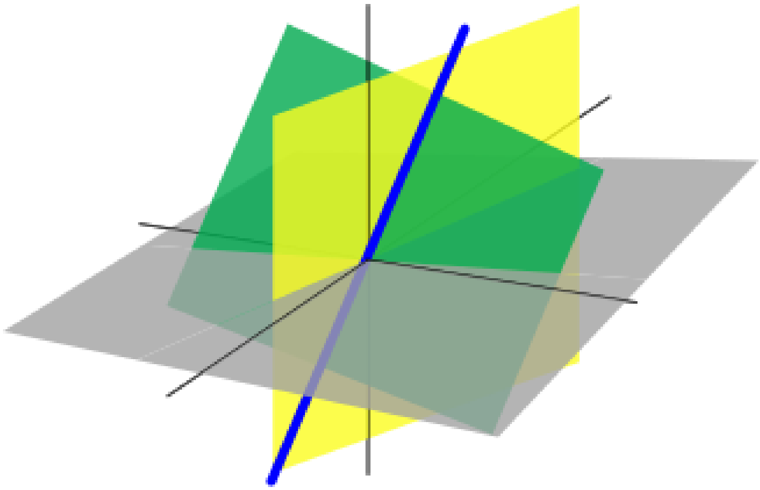
2 Vector Spaces (向量空间)

2.6

LINEAR TRANSFORMATION (线性变换)

- Definition & Examples
- Matrix representations
- Kernel (核)



I. Linear Transformation: Definition & Examples

A **function** (函数) f from a set A to a set B is a rule that assigns to each element of A a *single* element of B. We often write

$$f : A \to B$$
$$a \mapsto f(a)$$

where $f(a)$ is often defined by some equation, with

$\text{range}(f) = \{f(a) \mid a \in A\} \subseteq B.$ (range: 值域; domain: 定义域)

For example,

- $f(x) =$ sin$x$ is a function from **R** to $[-1, 1]$.
- $f : (x, y) \mapsto (2x, 3y)$ maps $\mathbf{R}^2$ to $\mathbf{R}^2$.
- $f : (x, y) \mapsto (x, y, x + y)$ maps $\mathbf{R}^2$ to $\mathbf{R}^3$.
- $f : (x, y, z) \mapsto (x, z)$ maps $\mathbf{R}^3$ to $\mathbf{R}^2$.

Definition 1 A function f from a vector space V to a vector space W is called a **linear transformation (线性变换)** if

(1) $f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$ for all vectors $\boldsymbol{u}, \boldsymbol{v} \in V$;

(2) $f(c\boldsymbol{v}) = cf(\boldsymbol{v})$ for all vectors $\boldsymbol{v} \in V$ and all $c \in \mathbf{R}$.

(1)&(2) $\Leftrightarrow$ $f(c\boldsymbol{u} + d\boldsymbol{v}) = cf(\boldsymbol{u}) + df(\boldsymbol{v})$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$ and all $c, d \in \mathbf{R}$.

Linear transformations *preserve the operations of vector addition and scalar multiplication*. **(线性变换保持加法和数乘运算)**

Examples

- $f: (x, y) \mapsto (2x, 3y)$ maps $\mathbf{R}^2$ to $\mathbf{R}^2$.
- $f: (x, y) \mapsto (x, y, x + y)$ maps $\mathbf{R}^2$ to $\mathbf{R}^3$.
- $f: (x, y, z) \mapsto (x, z)$ maps $\mathbf{R}^3$ to $\mathbf{R}^2$.

Examples

- $f(x) = x^2$ is not a linear transformation from $\mathbf{R}$ to $\mathbf{R}$, since

$$f(x + y) = (x + y)^2 \neq x^2 + y^2 = f(x) + f(y), \text{ except } xy=0.$$

- $f(x) = \sin x$ is not a linear transformation from $\mathbf{R}$ to $\mathbf{R}$, since $f(x + y) = \sin(x + y) \neq \sin x + \sin y = f(x) + f(y)$ does not always hold.

Other examples

We take as examples the spaces $\mathbf{P}_n$, in which the vectors are polynomials $p(t)$ of degree n.

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n,$$

and the dimension of the vector space is $n + 1$.

- The operation of ***differentiation*** is linear

$$\frac{d}{dt} p(t) = a_1 + 2a_2 t + \cdots + n a_n t^{n-1}.$$

- ***Integration*** from 0 to t is also linear (*it takes* $\mathbf{P}_n$ *to* $\mathbf{P}_{n+1}$)

$$\int_0^t p(s)ds = a_0 t + \frac{1}{2} a_1 t^2 + \cdots + \frac{a_n}{n+1} t^{n+1}.$$

- ***Multiplication*** by a fixed polynomial like $2 + 3t$ is linear (*it also takes* $\mathbf{P}_n$ *to* $\mathbf{P}_{n+1}$) :

$$(2 + 3t)p(t) = 2a_0 + \cdots + 3a_n t^{n+1}.$$

II. Transformations Represented by Matrices

Let
$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

For any vector $\boldsymbol{x} \in \mathbf{R}^n$, the product $\boldsymbol{Ax}$ is a vector in $\mathbf{R}^m$:

$$\boldsymbol{Ax} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \in \mathbf{R}^m$$

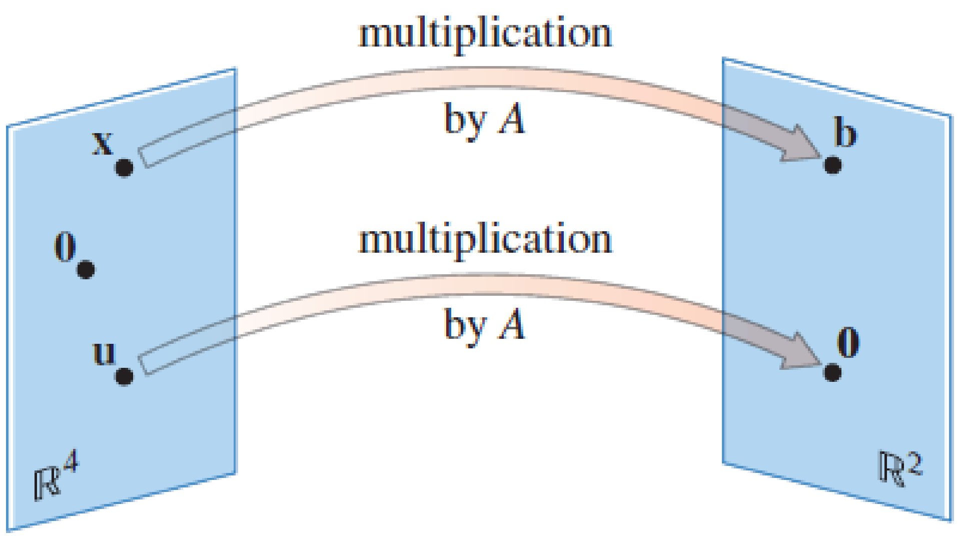
This defines a function f from $\mathbf{R}^n$ to $\mathbf{R}^m$: $\quad f : \boldsymbol{x} \mapsto \boldsymbol{Ax}.$

Suppose $\boldsymbol{x}$ is an n-dimensional vector.

When $\boldsymbol{A}$ multiplies $\boldsymbol{x}$, it *transforms* that vector into a new vector $\boldsymbol{A}\boldsymbol{x}$, which is an m-dimensional vector.

For instance,

$$\underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{bmatrix} 5 \\ 8 \end{bmatrix}}_{\boldsymbol{b}} \qquad \underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}}_{\boldsymbol{u}} = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\boldsymbol{0}}$$



It is a *linear transformation* as, for all $\mathbf{v}$, $\mathbf{w} \in \mathbf{R}^n$ and $c \in \mathbf{R}$,

$$f(\mathbf{v}+\mathbf{w}) = \boldsymbol{A}(\mathbf{v}+\mathbf{w}) = \boldsymbol{A}\mathbf{v} + \boldsymbol{A}\mathbf{w} = f(\mathbf{v}) + f(\mathbf{w}),$$

$$f(c\mathbf{v}) = \boldsymbol{A}(c\mathbf{v}) = c\boldsymbol{A}\mathbf{v} = cf(\mathbf{v}).$$

That is to say, matrix multiplication satisfies ***the rule of linearity* (线性性).**

Remark: If $\boldsymbol{A}$ is square (n by n):

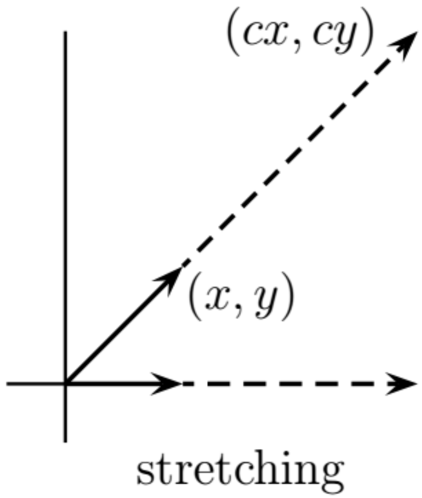
Suppose $\boldsymbol{x}$ is an n-dimensional vector, then $\boldsymbol{Ax}$ is also an n-dimensional vector.

This happens at every point $\boldsymbol{x}$ of the n-dimensional space $\mathbf{R}^n$.

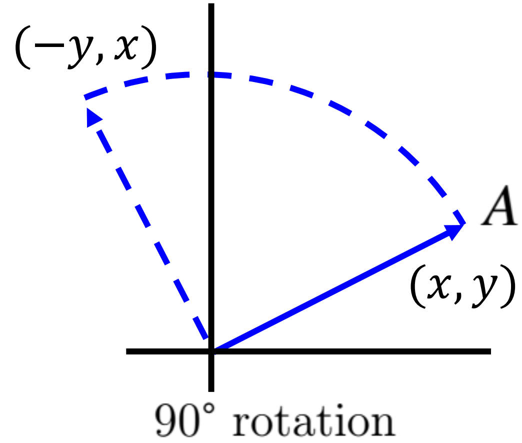
The whole space is transformed, or “mapped into itself,” by the matrix $\boldsymbol{A}$. (整个空间$\mathbf{R}^n$在方阵$\boldsymbol{A}$的作用下, 变换/映射到自身: $\mathbf{R}^n$)

https://www.geogebra.org/m/avrasb6d

Matrix representation - Examples in Geometry

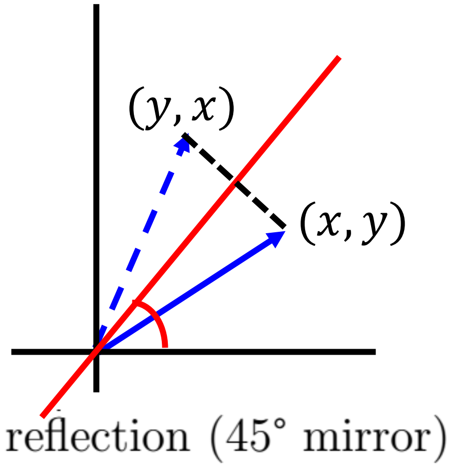


$$A = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}$$

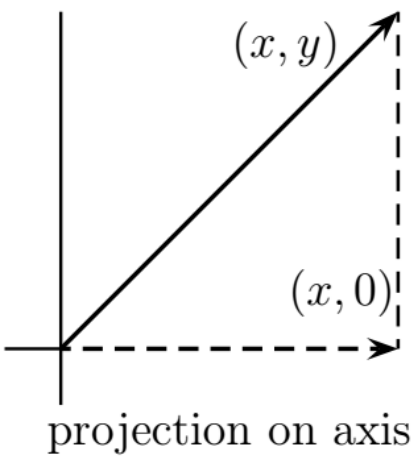


$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

What if -- rotations through other angles?



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



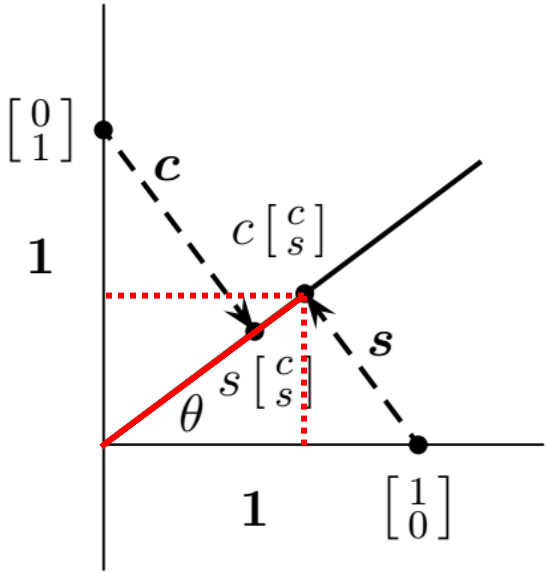
$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

What if -- reflections in other mirrors? *What if -- projections onto other lines?*

Can you find the matrix representation?

https://www.geogebra.org/m/guhhgudj

Projection (投影)



Projection onto the θ-line

In general, we may find the projection to the θ-line (the line at the angle θ from the x-axis). Thus the linear transformation p is such that

$$p: \begin{cases} \begin{bmatrix}1\\0\end{bmatrix} \to c\begin{bmatrix}c\\s\end{bmatrix} \\ \begin{bmatrix}0\\1\end{bmatrix} \to s\begin{bmatrix}c\\s\end{bmatrix} \end{cases} \qquad \begin{aligned} c &= \cos\theta \\ s &= \sin\theta \end{aligned}$$

The point $\begin{bmatrix}x\\y\end{bmatrix} = x\begin{bmatrix}1\\0\end{bmatrix} + y\begin{bmatrix}0\\1\end{bmatrix}$ is

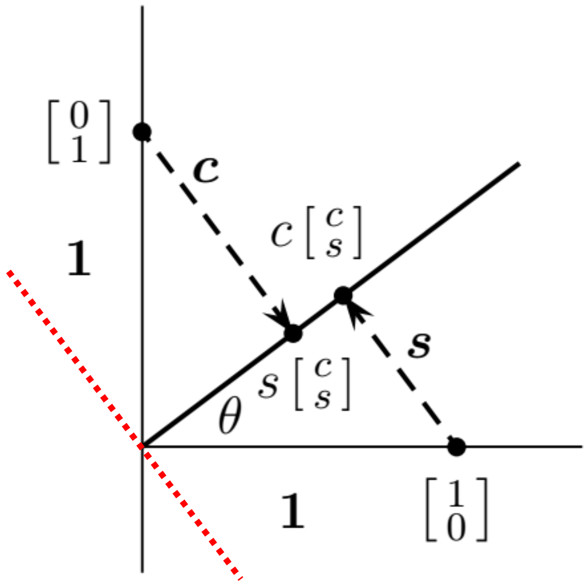
projected to: $x \cdot p\left(\begin{bmatrix}1\\0\end{bmatrix}\right) + y \cdot p\left(\begin{bmatrix}0\\1\end{bmatrix}\right)$

$$= x\begin{bmatrix}c^2\\cs\end{bmatrix} + y\begin{bmatrix}cs\\s^2\end{bmatrix}$$

$$= \boxed{\begin{bmatrix}c^2 & cs\\cs & s^2\end{bmatrix}}\begin{bmatrix}x\\y\end{bmatrix}$$

P : projection matrix

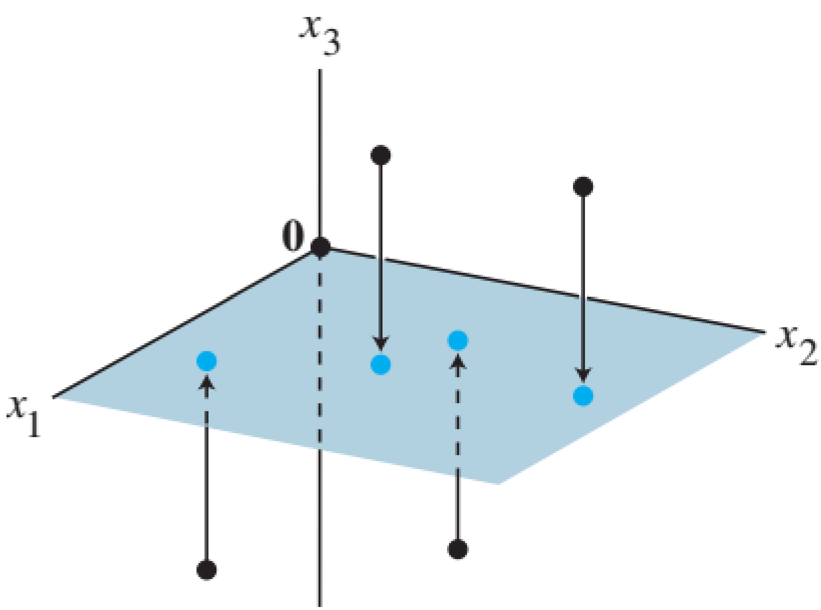
$\boldsymbol{P} = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}$ has some natural properties.



Projection onto the θ-line

- This matrix has no inverse (det($\boldsymbol{P}$)=0), because the transformation has no inverse.
- Points on the perpendicular line are projected onto the origin; that line is the nullspace of $\boldsymbol{P}$.
- Points on the θ -line are projected to themselves!
- Projecting twice is the same as projecting once, and $\boldsymbol{P}^2 = \boldsymbol{P}$ (幂等矩阵, idempotent matrix), i.e., *a projection matrix equals its own square*.(投影矩阵等于自身的平方)

What 3 by 3 matrices represent the transformations that project every vector onto the x_1-x_2 plane?



If $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$,

then the transformation

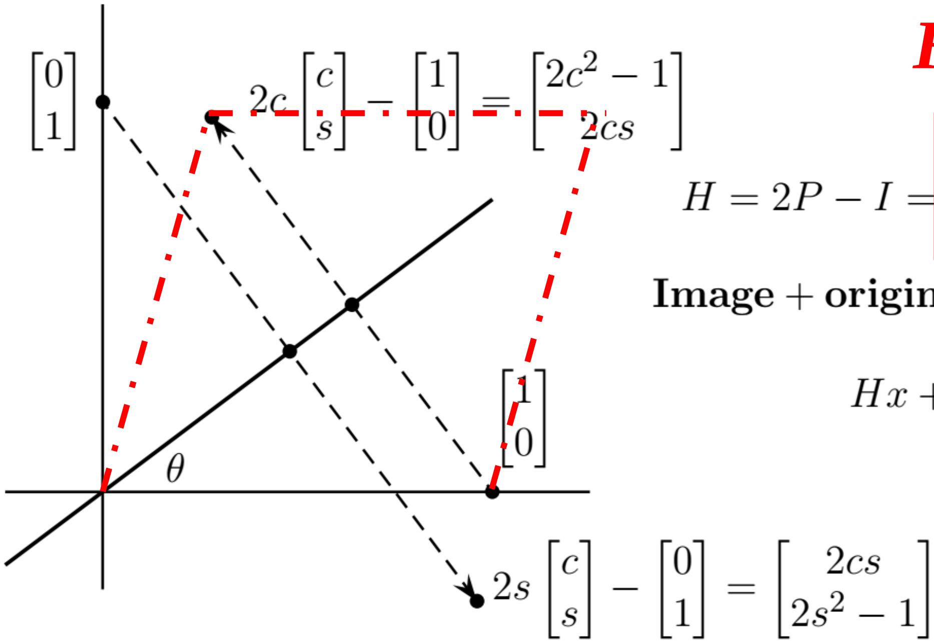
$$\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$$

projects points in $\mathbf{R}^3$ onto the x_1-x_2 plane because

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

Reflection (反射)

In general, we may find reflection along θ-line in a similar way:



H : reflection matrix

$$H = 2P - I = \boxed{\begin{bmatrix}2c^2-1 & 2cs\\ 2cs & 2s^2-1\end{bmatrix}}$$

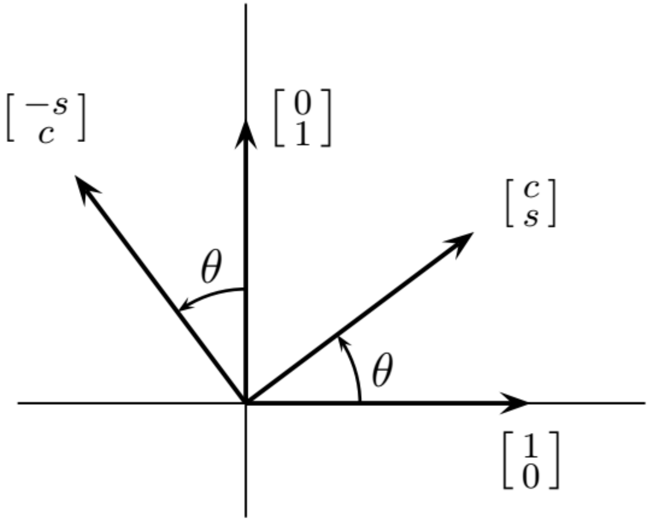
$$\textbf{Image} + \textbf{original} = 2 \times \textbf{projection}$$

$$Hx + x = 2Px$$

H has some remarkable properties:

- $\boldsymbol{H}^2 = \boldsymbol{I}$. *Two reflections bring back the original.*
  $(\boldsymbol{H}^2 = (2\boldsymbol{P} - \boldsymbol{I})^2 = 4\boldsymbol{P}^2 - 4\boldsymbol{P} + \boldsymbol{I} = \boldsymbol{I}$, since $\boldsymbol{P}^2 = \boldsymbol{P}$.)
- $\boldsymbol{H}^{-1} = \boldsymbol{H}$. *A reflection is its own inverse.*

Rotation (旋转)



Rotation through θ

$\boldsymbol{Q}_\theta = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$: rotation matrix

is a perfect example showing the correspondence between transformations and matrices:

- $\boldsymbol{Q}_\theta \boldsymbol{Q}_{-\theta} = \boldsymbol{I}$

  The inverse of $\boldsymbol{Q}_\theta$ equals $\boldsymbol{Q}_{-\theta}$ (rotation backward through θ)

- $\boldsymbol{Q}_\theta^2 = \boldsymbol{Q}_{2\theta}$

  The square of $\boldsymbol{Q}_\theta$ equals $\boldsymbol{Q}_{2\theta}$ (rotation through a double angle)

- $\boldsymbol{Q}_\theta \boldsymbol{Q}_\phi = \boldsymbol{Q}_{\theta+\phi}$

  The product of $\boldsymbol{Q}_\theta$ and $\boldsymbol{Q}_\phi$ equals $\boldsymbol{Q}_{\theta+\phi}$ (rotation through ϕ then θ)

Matrix Representations of Linear Transformations (线性变换的矩阵表示： *general case*)

Example 1 Let f be a linear transformation from $\mathbf{R}^2$ to $\mathbf{R}^3$ such that

$$f\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}1\\2\\3\end{bmatrix},\quad f\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}0\\-1\\2\end{bmatrix},$$

Determine $f\left(\begin{bmatrix}x\\y\end{bmatrix}\right)$.

Solution

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

$$=xf\left(\begin{bmatrix}1\\0\end{bmatrix}\right)+yf\left(\begin{bmatrix}0\\1\end{bmatrix}\right)$$

$$=x\begin{bmatrix}1\\2\\3\end{bmatrix}+y\begin{bmatrix}0\\-1\\2\end{bmatrix}=\begin{bmatrix}x\\2x-y\\3x+2y\end{bmatrix}$$

$$=\begin{bmatrix}1&0\\2&-1\\3&2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}.$$

Theorem 1 Let T: $\mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Then there exists a unique matrix $\boldsymbol{A}$ such that

$$T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} \quad \text{for all } \boldsymbol{x} \text{ in } \mathbf{R}^n.$$

In fact, $\boldsymbol{A}$ is the $m \times n$ matrix whose j-th column is the vector $T(\boldsymbol{e}_j)$, where $\boldsymbol{e}_j$ is the j-th column of the identity matrix in $\mathbf{R}^n$:

$$\boldsymbol{A} = [T(\boldsymbol{e}_1) \quad \ldots \ T(\boldsymbol{e}_n)].$$

Proof Write $\boldsymbol{x} = \boldsymbol{I}_n\boldsymbol{x} = [\boldsymbol{e}_1 \ \ldots \boldsymbol{e}_n]\boldsymbol{x} = x_1\boldsymbol{e}_1 + \ \ldots + x_n\boldsymbol{e}_n$, and use the linearity of T to compute

$$T(\boldsymbol{x}) = T(x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n) = x_1T(\boldsymbol{e}_1) + \cdots + x_nT(\boldsymbol{e}_n)$$

$$= \boxed{[T(\boldsymbol{e}_1) \quad \ldots \quad T(\boldsymbol{e}_n)]} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \boldsymbol{A}\boldsymbol{x}.$$

The matrix $\boldsymbol{A}$ is called the **standard matrix for the linear transformation** T.

(寻找矩阵A的关键, 是知道线性变换T对于标准基各列的作用)

Example 2 Define the linear transformation $L\colon \mathbf{R}^3 \to \mathbf{R}^2$ by

$$L(\boldsymbol{x}) = (x_1 + x_2, x_2 + x_3)^T.$$

for each $\boldsymbol{x} = (x_1, x_2, x_3)^T$ in $\mathbf{R}^3$.

We wish to find a matrix $\boldsymbol{A}$ such that $L(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ for each $\boldsymbol{x} \in \mathbf{R}^3$.

Solution: To do this, we calculate

$$L(\boldsymbol{e}_1) = L\big((1,0,0)^T\big) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$L(\boldsymbol{e}_2) = L\big((0,1,0)^T\big) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$L(\boldsymbol{e}_3) = L\big((0,0,1)^T\big) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

We choose these vectors to be the columns of the matrix

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

To check the result, we compute $\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \end{bmatrix}.$

What if – general case: 非自然基

To transform a space to itself, one basis is enough.

A transformation from one space to another requires a basis for each. They can be bases other than the standard bases.

Theorem 2 Suppose the vectors $E = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ are a basis for the space V, and vectors $F = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_m\}$ are a basis for W.

Each linear transformation T from V to W is represented by a matrix $\boldsymbol{A}$. The jth column of $\boldsymbol{A}$ is found by applying T to the jth basis vector $\boldsymbol{v}_j$, and writing $T(\boldsymbol{v}_j)$ as a combination of the $\boldsymbol{w}$'s:

> **Column j of A** is the coordinate vector of $T(\boldsymbol{v}_j)$ with respect to $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_m\}$, which means

$$T(\boldsymbol{v}_j) = \boldsymbol{A}\boldsymbol{v}_j = a_{1j}\boldsymbol{w}_1 + a_{2j}\boldsymbol{w}_2 + \cdots + a_{mj}\boldsymbol{w}_m.$$

$E = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ are a basis for V,

$F = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_m\}$ are a basis for W.

$$T(\boldsymbol{v}_j) = \boldsymbol{A}\boldsymbol{v}_j = a_{1j}\boldsymbol{w}_1 + a_{2j}\boldsymbol{w}_2 + \cdots + a_{mj}\boldsymbol{w}_m$$

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$[T(\boldsymbol{v}_1)]_F \quad [T(\boldsymbol{v}_2)]_F \quad [T(\boldsymbol{v}_n)]_F$$

where

$$T(\boldsymbol{v}_1) = a_{11}\boldsymbol{w}_1 + a_{21}\boldsymbol{w}_2 + \cdots + a_{m1}\boldsymbol{w}_m$$

$$T(\boldsymbol{v}_2) = a_{12}\boldsymbol{w}_1 + a_{22}\boldsymbol{w}_2 + \cdots + a_{m2}\boldsymbol{w}_m$$

$$\vdots$$

$$T(\boldsymbol{v}_n) = a_{1n}\boldsymbol{w}_1 + a_{2n}\boldsymbol{w}_2 + \cdots + a_{mn}\boldsymbol{w}_m$$

Example 3 Define the linear transformation L: $\mathbf{R}^3 \to \mathbf{R}^2$ by

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

for each $\boldsymbol{x} = (x_1, x_2, x_3)^T$ in $\mathbf{R}^3$, where $\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\boldsymbol{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Find the matrix $\boldsymbol{A}$ representing L with respect to the ordered bases $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ and $\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$.

Solution:

$$L(\boldsymbol{e}_1) = 1\boldsymbol{b}_1 + 0\boldsymbol{b}_2$$
$$L(\boldsymbol{e}_2) = 0\boldsymbol{b}_1 + 1\boldsymbol{b}_2$$
$$L(\boldsymbol{e}_3) = 0\boldsymbol{b}_1 + 1\boldsymbol{b}_2$$

The jth column of $\boldsymbol{A}$ is determined by the coordinates of $L(\boldsymbol{e}_j)$ with respect to $\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ for $i = 1, 2, 3$. Thus

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

Example 4 Let L be a linear transformation mapping $\mathbf{R}^2$ into itself defined by

$$L(\alpha \boldsymbol{b}_1 + \beta \boldsymbol{b}_2) = (\alpha + \beta)\boldsymbol{b}_1 + 2\beta \boldsymbol{b}_2.$$

where $\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\boldsymbol{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Find the matrix $\boldsymbol{A}$ representing L with respect to $\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$.

Solution:

$$L(\boldsymbol{b}_1) = 1\boldsymbol{b}_1 + 0\boldsymbol{b}_2$$
$$L(\boldsymbol{b}_2) = 1\boldsymbol{b}_1 + 2\boldsymbol{b}_2$$

Thus

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

Example 5 Let $L:\mathbf{R}^2 \to \mathbf{R}^3$ be the linear transformation defined by

$$L(\boldsymbol{x}) = (x_2, x_1 + x_2, x_1 - x_2)^T.$$

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

$$\boldsymbol{u}_1 = (1,2)^T, \boldsymbol{u}_2 = (3,1)^T$$

and

$$\boldsymbol{b}_1 = (1,0,0)^T, \boldsymbol{b}_2 = (1,1,0)^T, \boldsymbol{b}_3 = (1,1,1)^T.$$

Solution 1

$$L(\boldsymbol{u}_1) = (2, 3, -1)^T, \qquad L\,(\boldsymbol{u}_2) = (1, 4, 2)^T.$$

We need to write them as combinations of $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$:

$$\begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} = a_{11}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + a_{21}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + a_{31}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} = a_{12}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + a_{22}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + a_{32}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Thus $\quad \boldsymbol{A} = \begin{bmatrix} -1 & -3 \\ 4 & 2 \\ -1 & 2 \end{bmatrix}.$

Theorem 3 (An equivalent way to find the matrix A)

Let $E = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ and $F = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m\}$ be ordered bases for $\mathbf{R}^n$ and $\mathbf{R}^m$, respectively.

If $L: \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation and $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n]$ is the matrix representing L with respect to E and F, then

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

where $\boldsymbol{B} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m]$.

Proof. If $\boldsymbol{A}$ is representing L with respect to E and F, then, for $j = 1, \ldots, n$,

$$L(\boldsymbol{u}_j) = a_{1j}\boldsymbol{b}_1 + a_{2j}\boldsymbol{b}_2 + \cdots + a_{mj}\boldsymbol{b}_m = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m]\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

$$= \boldsymbol{B}\boldsymbol{a}_j$$

The matrix $\boldsymbol{B}$ is nonsingular since its column vectors form a basis for $\mathbf{R}^m$. Hence

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

The way to find the matrix representation of the transformation is :

by computing the reduced row echelon form of an augmented matrix.

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

$$[\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m \mid L(\boldsymbol{u}_1), \ldots, L(\boldsymbol{u}_n)]$$

is

$$[\boldsymbol{I} \mid \boldsymbol{A}] .$$

Example 5 (continued)

Let $L: \mathbf{R}^2 \to \mathbf{R}^3$ be the linear transformation defined by

$$L(\boldsymbol{x}) = (x_2, x_1 + x_2, x_1 - x_2)^T.$$

Find the matrix representations of L with respect to the ordered bases $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ and $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$, where $\boldsymbol{u}_1 = (1,2)^T, \boldsymbol{u}_2 = (3,1)^T$ and $\boldsymbol{b}_1 = (1,0,0)^T, \boldsymbol{b}_2 = (1,1,0)^T, \boldsymbol{b}_3 = (1,1,1)^T$.

Solution 2 We must compute $L(\boldsymbol{u}_1)$ and $L(\boldsymbol{u}_2)$ and then transform the augmented matrix $[\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3 \mid L(\boldsymbol{u}_1), L(\boldsymbol{u}_2)]$ to reduced row echelon form:

$$L(\boldsymbol{u}_1) = (2, 3, -1)^T, \qquad L(\boldsymbol{u}_2) = (1, 4, 2)^T.$$

$$\left[\begin{array}{ccc|cc} 1 & 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 3 & 4 \\ 0 & 0 & 1 & -1 & 2 \end{array}\right] \to \left[\begin{array}{ccc|cc} 1 & 0 & 0 & -1 & -3 \\ 0 & 1 & 0 & 4 & 2 \\ 0 & 0 & 1 & -1 & 2 \end{array}\right].$$

The matrix representing L with respect to the given ordered bases is

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

Next we find matrices that represent **differentiation**.

Basis for $P_3 : E = \{p_1, p_2, p_3, p_4\}$. $\quad p_1 = 1, p_2 = t, p_3 = t^2, p_4 = t^3$.

The derivatives of those four basis vectors

Action of $\frac{d}{dt} : \frac{d}{dt}p_1 = 0, \frac{d}{dt}p_2 = p_1, \frac{d}{dt}p_3 = 2p_2, \frac{d}{dt}p_4 = 3p_3$.

Coordinate vectors: $[p_1]_E = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, [p_2]_E = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, [p_3]_E = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, [p_4]_E = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

Differentiation matrix

$$\boldsymbol{A}_{\text{diff}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{array}{cccc} \downarrow & \downarrow & \downarrow & \downarrow \\ Ap_1 & Ap_2 & Ap_3 & Ap_4 \end{array}$$

$$p_1 = 1, \qquad p_2 = t, \qquad p_3 = t^2, \qquad p_4 = t^3$$

For the matrix $\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$,

$N(\boldsymbol{A}) = \text{Span}\{p_1\}$, $\text{nullity}(\boldsymbol{A}) = 1$;

$C(\boldsymbol{A}) = \text{Span}\{p_1, p_2, p_3\}$, $\text{rank}(\boldsymbol{A}) = 3$.

For any vector in the vector space $\mathbf{P}_3$, the derivative is decided by linearity.

For example, for $p = 2 + t - t^2 - t^3$,

$$\frac{dp}{dt} = \boldsymbol{A}p \rightarrow \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -3 \\ 0 \end{bmatrix},$$

so the derivative for p is $1 - 2t - 3t^2$.

Example: Integration

$$\textbf{Integration}: V(=P_3) \longrightarrow W(=P_4)$$

$$\textbf{Basis for } P_3\text{:}\ \ \boldsymbol{x}_1 = 1, \boldsymbol{x}_2 = t, \boldsymbol{x}_3 = t^2, \boldsymbol{x}_4 = t^3$$

$$\textbf{Basis for } P_4\text{:}\ \ \boldsymbol{y}_1 = 1, \boldsymbol{y}_2 = t, \boldsymbol{y}_3 = t^2, \boldsymbol{y}_4 = t^3, \boldsymbol{y}_5 = t^4$$

$$\int_0^t 1ds = t \ \text{ or } \ \boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{y}_2, \boldsymbol{A}\boldsymbol{x}_2 = \frac{1}{2}\boldsymbol{y}_3, \int_0^t s^3 ds = \frac{t^4}{4} \ \text{ or } \ \boldsymbol{A}\boldsymbol{x}_4 = \frac{1}{4}\boldsymbol{y}_5$$

$$\boldsymbol{A}_{\text{int}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/4 \end{bmatrix}$$

$$\boldsymbol{A}\boldsymbol{x}_3 = \frac{1}{3}\boldsymbol{y}_4$$

Theorem 4 Suppose $\boldsymbol{A}$ and $\boldsymbol{B}$ are linear transformations from V to W and from U to V. Their product $\boldsymbol{AB}$ starts with a vector $\boldsymbol{u}$ in U, goes to $\boldsymbol{Bu}$ in V, and finishes with $\boldsymbol{ABu}$ in W. This "composition" $\boldsymbol{AB}$ is again a linear transformation (from U to W). Its matrix is the product of the individual matrices representing $\boldsymbol{A}$ and $\boldsymbol{B}$.

$$
\begin{array}{ccccc}
U & \xrightarrow{\boldsymbol{B}} & V & \xrightarrow{\boldsymbol{A}} & W \\
\boldsymbol{u} & \mapsto & \boldsymbol{Bu} & \mapsto & \boldsymbol{ABu}
\end{array}
$$

Example 6 Let T be a linear transformation of $\mathbf{R}^2$ such that

$$T\left(\begin{bmatrix}1\\2\end{bmatrix}\right)=\begin{bmatrix}3\\4\end{bmatrix},\quad T\left(\begin{bmatrix}2\\1\end{bmatrix}\right)=\begin{bmatrix}5\\6\end{bmatrix}.$$

Find the matrix for T.

Solution The matrix $\begin{bmatrix}1 & 2\\2 & 1\end{bmatrix}^{-1}$ is such that $\begin{bmatrix}1\\2\end{bmatrix}\to\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}2\\1\end{bmatrix}\to\begin{bmatrix}0\\1\end{bmatrix}$.

The matrix $\begin{bmatrix}3 & 5\\4 & 6\end{bmatrix}$ is such that $\begin{bmatrix}1\\0\end{bmatrix}\to\begin{bmatrix}3\\4\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}\to\begin{bmatrix}5\\6\end{bmatrix}$.

Thus the matrix

$$\boldsymbol{A}=\begin{bmatrix}3 & 5\\4 & 6\end{bmatrix}\begin{bmatrix}1 & 2\\2 & 1\end{bmatrix}^{-1}=\frac{1}{3}\begin{bmatrix}7 & 1\\8 & 2\end{bmatrix}$$

is such that $\begin{bmatrix}1\\2\end{bmatrix}\to\begin{bmatrix}3\\4\end{bmatrix}$ and $\begin{bmatrix}2\\1\end{bmatrix}\to\begin{bmatrix}5\\6\end{bmatrix}$, and $T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\frac{1}{3}\begin{bmatrix}7 & 1\\8 & 2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$.

III. Ranges (值域)

We notice that, if f is a linear transformation, then

$$f(c\boldsymbol{u} + d\boldsymbol{v}) = f(c\boldsymbol{u}) + f(d\boldsymbol{v}) = cf(\boldsymbol{u}) + df(\boldsymbol{v}).$$

Theorem 5 Let $f: \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Then the range of f is a subspace of $\mathbf{R}^m$.

Proof Let $S = \text{range}(f)$. Then $f(\boldsymbol{0}) = \boldsymbol{0} \in S$,

because $f(\boldsymbol{0}) = f(\boldsymbol{0} - \boldsymbol{0}) = f(\boldsymbol{0}) - f(\boldsymbol{0}) = \boldsymbol{0}$.

Let $\boldsymbol{x}, \boldsymbol{y} \in S$, i.e., $\boldsymbol{x} = f(\boldsymbol{u})$ and $\boldsymbol{y} = f(\boldsymbol{v})$ for some $\boldsymbol{u}, \boldsymbol{v} \in \mathbf{R}^n$.

Thus $\boldsymbol{x} + \boldsymbol{y} = f(\boldsymbol{u}) + f(\boldsymbol{v}) = f(\boldsymbol{u} + \boldsymbol{v}) \in S$.

Let $a \in \mathbf{R}$ and $\boldsymbol{x} \in S$, i.e., $\boldsymbol{x} = f(\boldsymbol{u})$ for some $\boldsymbol{u} \in \mathbf{R}^n$.

Thus $a\boldsymbol{x} = af(\boldsymbol{u}) = f(a\boldsymbol{u}) \in S$.

Therefore, S is a subspace of $\mathbf{R}^m$.

Note: ***A linear transformation is determined by the effect it has on a basis.***

Rank-nullity theorem

Let f be a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$.

Definition 2 The **kernel (核)** ker(f) is the set $\{\boldsymbol{u} \in \mathbf{R}^n \mid f(\boldsymbol{u}) = \boldsymbol{0}\}$.

Example 7 Let f be a linear transformation from $\mathbf{R}^2$ to $\mathbf{R}^2$ defined by

$$f(x, y) = (x + y, x + y).$$

Then the range

$$\mathrm{range}(f) = \{(a, a) \mid a \in \mathbf{R}\},$$

which is a line. The kernel is also a line:

$$\ker(f) = \{(a, -a) \mid a \in \mathbf{R}\}.$$

Theorem 6 (Rank-nullity theorem)

Let f be a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$.

Then the range of f is a subspace of $\mathbf{R}^m$, and the kernel of f is a subspace of $\mathbf{R}^n$. Moreover,

$$\dim(\ker(f)) + \dim(\mathrm{range}(f)) = n.$$

Key words:

Linear transformation: definition and examples;
matrix of linear transformation;
Range, kernel

Homework

See Blackboard