Week 4
Linear Space

Christopher Ting

http://cting.x10host.com/

Hiroshima University

✉: cting@hiroshima-u.ac.jp

☎: +81 082-424-6451

A1棟 131-1

Table of Contents

Inspiring Mystery

The eternal silence of these infinite spaces frightens me.

—Blaise Pascal

Learning Outcomes

- Build a strong case for studying linear space in the context of informatics and data science.
- Demonstrate a deeper understanding of vectors and a plane in the 3-dimensional space from the geometrical standpoint.
- Compute the inner product of two vectors, and the angle formed by two location vectors.
- Illustrate the concepts of space, linear space and subspace by providing appropriate examples.
- Use inner product to formulate linear classifiers of machine learning

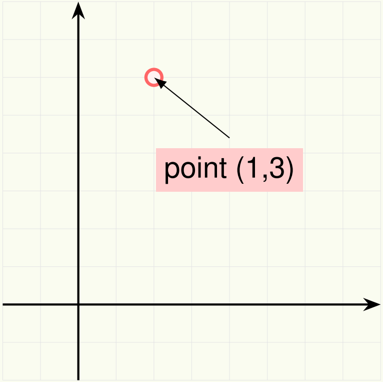
Learning Outcomes (つづき)

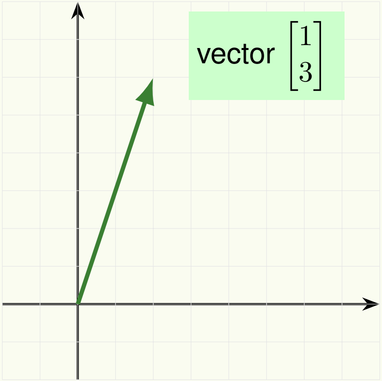
- Make a connection between the system of first-order equations to the concept of subspace.
- Analyze the relations among the concepts of linear independence, linear dependence, linear combination, and the rank of a matrix.
- Define and derive the basis transformation matrix.
- Create the links among span, basis, dimension, rank, system of linear equations, and concepts that appear in the discussion of linear space.

Preamble

- In Part 1, a vector is defined abstractly as a set of numbers arranged as an array vertically.
- The term “dimension” refers to the number of rows or entries a vector has.
- Likewise, a matrix is simply $p \times q$ numbers arranged as a rectangle. It is also defined as a vector of row vectors, or a row vector of vectors.
- The rank of a matrix is abstractly defined as the sum of the main components in the context of matrix simplification.
- But geometrically, what are the meanings of vectors and matrices?

Point and Vector



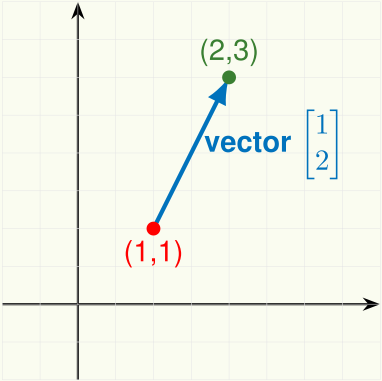


What is the difference between a point and a vector in $\Re^2$?

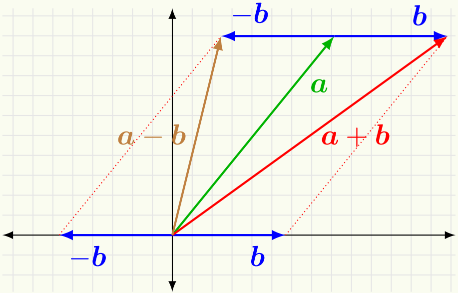
Interpretation of a Vector

- The arrow from one point to another point gives rise to **direction**.
- The **length** between two points gives rise to **magnitude**.
- Let there be the Euclidean co-ordinate system. Suppose $\boldsymbol{a}$ is an n-dimensional vector. Its length is defined as

$$\|\boldsymbol{a}\| := \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

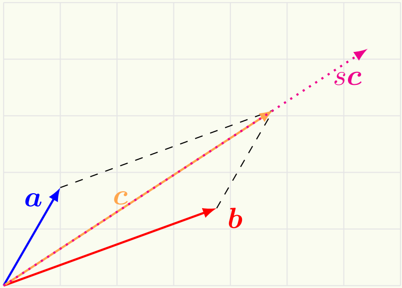


Addition and Scalar Multiplication of Vectors

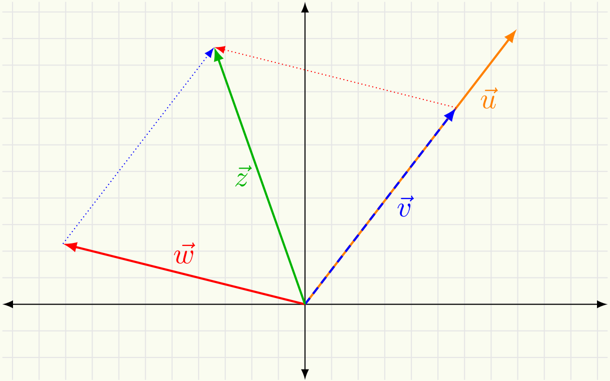


Vectors addition and subtraction

Vector multiplication by a scalar s

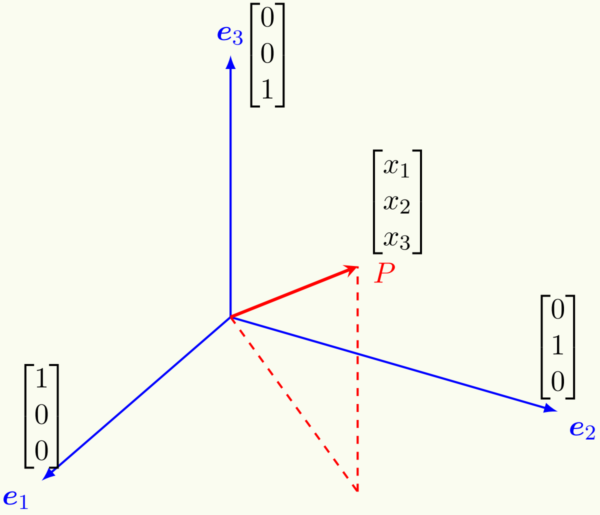


Geometric Vectors: A Summary

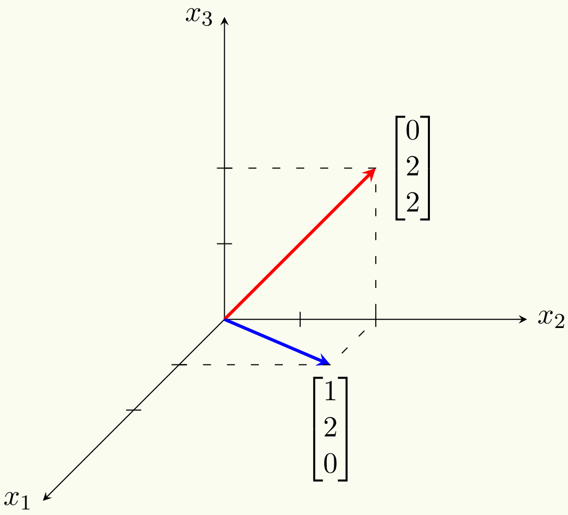


- magnitude and direction
- addition and scalar multiplication

Basis Vectors e_i and Location Vector



Example of 3-Dimensional Location Vectors



Inner Product

Definition 2.1 (Inner Product).

For two n-dimensional location vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ which are not null vectors, the **inner product** $(\boldsymbol{a}, \boldsymbol{b})$ is defined as

$$\boldsymbol{a} \cdot \boldsymbol{b} \equiv (\boldsymbol{a}, \boldsymbol{b}) := \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| \cos(\theta),$$

where θ is the **angle** formed by $\boldsymbol{a}$ and $\boldsymbol{b}$, and $\| \cdot \|$ is the **length** of the vector defined on $\mathfrak{R}^n$.

Properties of vector's inner product

1. $(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a})$
2. $(\boldsymbol{a} + \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a}, \boldsymbol{c}) + (\boldsymbol{b}, \boldsymbol{c})$
3. $(s\boldsymbol{a}, \boldsymbol{b}) = s(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a}, s\boldsymbol{b})$, where $s \in \mathfrak{R}$ is the scalar.
4. $(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij}$, where $i, j = 1, 2, \ldots, n$ for the basis vectors $\boldsymbol{e}_i$.

Example of Inner Product

Example 2.2.

Compute the inner product of two 3-dimensional location vectors $\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

Find the angle θ formed by them.

- We first express $\boldsymbol{a}$ and $\boldsymbol{b}$ in terms of the basis vectors as follows:

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3, \quad \boldsymbol{b} = b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2 + b_3\boldsymbol{e}_3.$$

- Then from Property (4) of the basis vectors with $n = 3$,

$$(\boldsymbol{a}, \boldsymbol{b}) = \sum_{i,j}^{3} a_i b_j (\boldsymbol{e}_i, \boldsymbol{e}_j) = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

- $\cos\theta = \dfrac{(\boldsymbol{a}, \boldsymbol{b})}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|} = \dfrac{a_1 b_1 + a_2 b_2 + a_3 b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}\sqrt{b_1^2 + b_2^2 + b_3^2}}.$

Example of a Plane

Example 2.3.

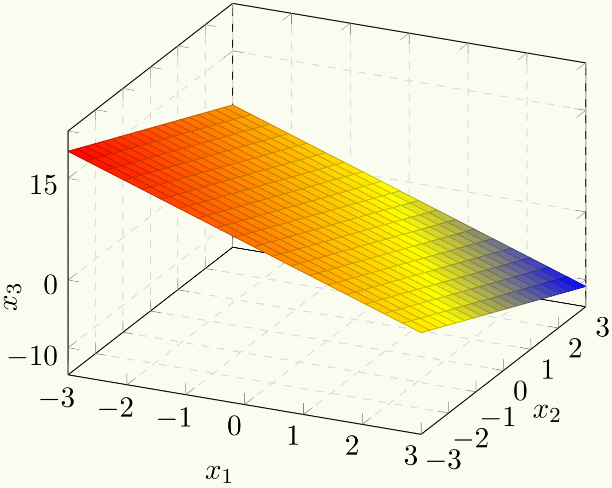
Let the location vector on a point P be $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $ax_1 + bx_2 + cx_3 = d$, where a, b, c and d are constants, with $(a, b, c) \neq (0, 0, 0)$. Show that the set of all points is a plane.

- Without loss of generality, suppose $c \neq 0$. Then, for any x_1 and x_2, we can express x_3 in terms of x_1 and x_2. Moreover,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \dfrac{d - ax_1 - bx_2}{c} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dfrac{d}{c} \end{bmatrix} + x_1 \begin{bmatrix} 1 \\ 0 \\ \dfrac{-a}{c} \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \dfrac{-b}{c} \end{bmatrix} =: \boldsymbol{p}_0 + x_1 \boldsymbol{b}_1 + x_2 \boldsymbol{b}_2.$$

- For illustration, let $c = 1$, $d = 4$, $a = 3$, $b = 2 \implies x_3 = 4 - 3x_1 - 2x_2$.

Plot of $x_3 = 4 - 3x_1 - 2x_2$



Two Types of Boundary 二種類の境界

Definition 3.1 (Decision Boundary 決定境界).

The decision boundary is the set of points $\boldsymbol{x}$ that satisfy

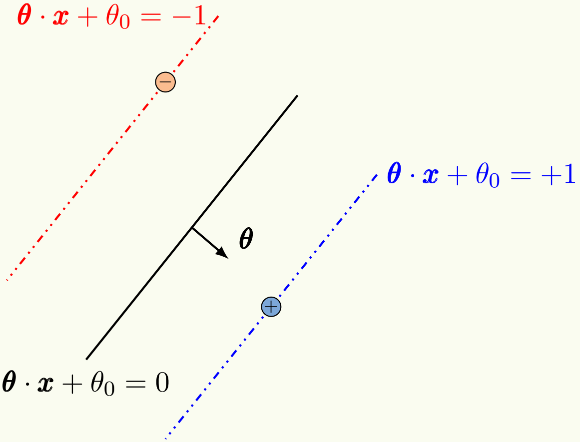
$$\boldsymbol{\theta} \cdot \boldsymbol{x} + \theta_0 = 0.$$

Definition 3.2 (Margin Boundary マージン(周縁部)境界).

The margin boundary is the set of points $\boldsymbol{x}$ that satisfy

$$\boldsymbol{\theta} \cdot \boldsymbol{x} + \theta_0 = \pm 1.$$

Decision Boundary and Margin 決定境界とマージン



Shortest Distance

+ Recall the definition of inner product, also called **dot product**:

$$\boldsymbol{\theta} \cdot \boldsymbol{p} := (\boldsymbol{\theta}, \boldsymbol{p}) = ||\boldsymbol{\theta}|| \, ||\boldsymbol{p}|| \cos(\alpha)$$

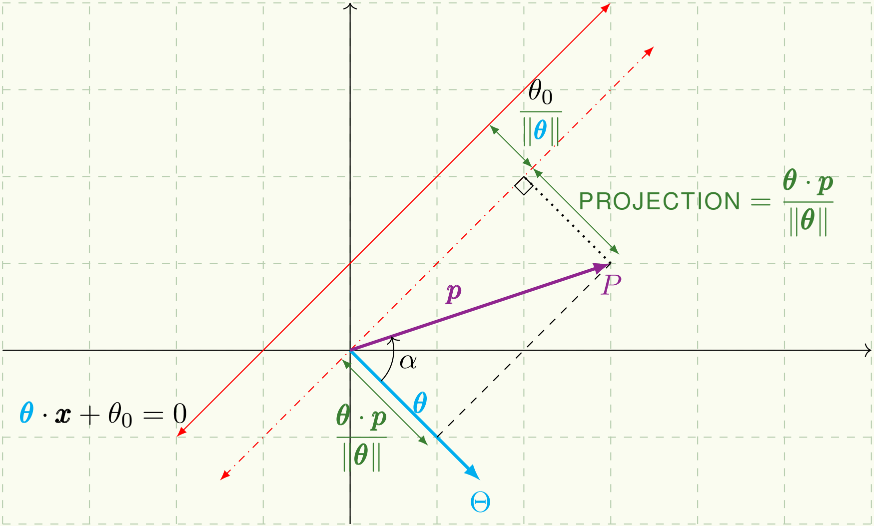
+ If there is no **offset**, i.e., $\theta_0 = 0$, the **projection** of $\boldsymbol{p}$ on $\boldsymbol{\theta}$ is by definition given by

$$||\boldsymbol{p}|| \cos(\alpha) = \frac{\boldsymbol{\theta} \cdot \boldsymbol{p}}{||\boldsymbol{\theta}||}.$$

+ With the offset θ_0 added, the shortest distance d from a point P to the line L defined by $\boldsymbol{\theta} \cdot \boldsymbol{x} + \theta_0$ is

$$d = \frac{\boldsymbol{\theta} \cdot \boldsymbol{p} + \theta_0}{||\boldsymbol{\theta}||}. \tag{1}$$

Projection of Feature Vector p on θ



Linear Space

Definition 4.1 (Linear Space).

If a nonempty set V satisfies the following two conditions, then V is called the **linear space** or **vector space**.

(I) For any elements $\boldsymbol{u}$ and $\boldsymbol{v}$ of V, and any scalar $a \in \mathfrak{R}$, vector addition $\boldsymbol{u} + \boldsymbol{v}$ and scalar multiplication $a\boldsymbol{u}$ are defined, and the resulting outcomes of these operations also belong to V.

(II) The 8 laws concerning addition and scalar multiplication are valid.

1. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$
2. $1\boldsymbol{u} = \boldsymbol{u}$
3. $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$
4. $a(b\boldsymbol{u}) = (ab)\boldsymbol{u}$
5. $a(\boldsymbol{u} + \boldsymbol{v}) = a\boldsymbol{u} + a\boldsymbol{v}$
6. $(a + b)\boldsymbol{u} = a\boldsymbol{u} + b\boldsymbol{u}$
7. $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$, where $\boldsymbol{0} \in V$ is the zero vector.
8. $\boldsymbol{u} + \widetilde{\boldsymbol{u}} = \boldsymbol{0}$, where the **inverse vector** $\widetilde{\boldsymbol{u}} = -\boldsymbol{u}$.

Examples of Linear Space

- $$\mathbb{R}^2 = \left\{ \boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \middle| a_1, a_2 \in \mathfrak{R} \right\}$$
- In the same way, $\mathbb{R}^n$—a set of all n-dimensional vectors—is a linear space.
- The set of all $m \times n$ real-valued matrices $M_{m,n}(\mathfrak{R})$ is a linear space under matrix addition and scalar multiplication.
- Consider the set of all polynomial functions whose coefficients are real numbers:
  $$\mathbb{R}[x]_n = \left\{ a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \middle| a_0, a_1, \ldots, a_n \in \mathfrak{R} \right\}$$
  under addition and scalar multiplication, $\mathbb{R}[x]_n$ is a linear space.
- Consider the set V consisting the sine and cosine functions.
  $$V = \left\{ a_0 + a_1 \cos(x) + a_2 \sin(x) + a_3 \cos(2x) + a_4 \sin(2x) \middle| a_0, a_1, \ldots, a_4 \in \mathfrak{R} \right\}$$
  under the usual addition and scalar multiplication, V is a linear space.

Another Example

- Let $C(I)$ be the space of all real-valued continuous functions on an interval $I = [a, b]$.
- For any elements $f(x)$ and $g(x)$ in $C(I)$, define function addition and scalar multiplication as

$$(f + g)(x) = f(x) + g(x),$$
$$(cf)(x) = c(f(x)),$$

  where $c \in \Re$.
- Given this definition, $C(I)$ becomes a linear space.

Subspace

Definition 4.2 (Subspace of a Linear Space).

When a nonempty subset W of the linear space V satisfies the conditions (I) and (II) of Definition 4.1, it is called the **subspace** of V.

Theorem 4.3 (Closure under Addition and Scalar Multiplication).

The statement "W is a subspace of the linear space V" is equivalent to the following set of 3 statements:

(1) $\mathbf{0} \in W$

(2) *If* $\boldsymbol{u}, \boldsymbol{v} \in W$, *then* $\boldsymbol{u} + \boldsymbol{v} \in W$.

(3) *If* $a \in \mathfrak{R}$ *is a scalar and* $\boldsymbol{u} \in W$, *then* $a\boldsymbol{u}$ *also belongs to* W.

Proof of Theorem 4.3

- If W is a subspace of V, by definition, it satisfies conditions (I) and (II) of Definition 4.1, which include the set of 3 statements.
- Next, we start with the set of 3 statements to show that W is a subspace of V.
- The statement (1) that $\mathbf{0} \in W$ satisfies condition (II) ⑦ of Definition 4.1.
- The statement (2) that $\boldsymbol{u} + \boldsymbol{v} \in W$ satisfies condition (I) of Definition 4.1. So is statement (3) that $a\boldsymbol{u} \in W$.
- With $a = -1$, we have $-\boldsymbol{u}$ as the inverse of $\boldsymbol{u}$ in W, which satisfies condition (II) ⑧.
- The remaining conditions II ① to ⑥ in Definition 4.1 are also satisfied by statements (2) and (3).
- Hence, the subset W is a subspace of V. $\square$

Example of Subspace: A Plane in $\Re^3$

- Consider the set of all location vectors on the plane $x_1 - x_2 + x_3 = 0$:

$$W = \left\{ \boldsymbol{x} \in \Re^3 \middle| x_1 - x_2 + x_3 = 0 \right\}.$$

- First, note that the origin ($x_1 = x_2 = x_3 = 0$) satisfies the equation that describes the plane. So the point $\mathbf{0} = (0, 0, 0) \in W$.
- For two vectors $\boldsymbol{u}' = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$ and $\boldsymbol{v}' = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$ in W, their sum is $\begin{bmatrix} u_1 + v_1 & u_2 + v_2 & u_3 + v_3 \end{bmatrix}$. This sum satisfies the equation of the plane:

$$(u_1 + v_1) - (u_2 + v_2) + (u_3 + v_3) = (u_1 - u_2 + u_3) + (v_1 - v_2 + v_3) = 0 + 0 = 0.$$

  Hence, $\boldsymbol{u} + \boldsymbol{v} \in W$.
- It is easy to see that the scalar multiple of $\boldsymbol{w} = c\boldsymbol{u}$ also satisfies the plane equation, implying that $\boldsymbol{w} \in W$.
- It follows from Theorem 4.3 that W is a subspace. $\square$

Other Examples of Subspace

- If W_1 and W_2 are subspaces of V, then their sum

$$W_1 + W_2 := \{\boldsymbol{w}_1 + \boldsymbol{w}_2 \big| \boldsymbol{w}_1 \in W_1 \text{ and } \boldsymbol{w}_2 \in W_2\}$$

  is a also a subspace of V.

- Likewise, the intersection of W_1 and W_2, i.e.,

$$W_1 \cap W_2 = \{\boldsymbol{w} \big| \boldsymbol{w} \in W_1 \text{ and } \boldsymbol{w} \in W_2\},$$

  is also a subspace of V.

An Example of not a Subspace

- However, the union of W_1 and W_2 is not necessarily a subspace of V, except only when $W_1 \subset W_2$ or $W_2 \subset W_1$.
- To show that this statement is true, suppose $V = \Re^2$, and the subspaces are

$$W_1 = \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \middle| x \in \Re \right\}, \qquad W_2 = \left\{ \begin{bmatrix} 0 \\ y \end{bmatrix} \middle| y \in \Re \right\},$$

- The union is $W_1 \cup W_2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \middle| x = 0 \text{ or } y = 0 \right\}$.
- But with $\boldsymbol{w}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{w}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, both being in the union $W_1 \cup W_2$, their sum $\boldsymbol{w}_1 + \boldsymbol{w}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin W_1 \cup W_2$.

System of First-Order Equations

Theorem 4.4 (Set of All Solutions).

Let $\boldsymbol{A}$ be a $m \times n$ matrix. The set of all solutions of the system of first-order equations $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$, i.e.,

$$W := \left\{ \boldsymbol{x} \in \Re^n \middle| \boldsymbol{A}\boldsymbol{x} = \mathbf{0} \right\},$$

is a subspace of $\Re^n$.

Proof.

By Theorem 4.3, We just need to show that the 3 conditions of subspace are satisfied.
(1) Since $\boldsymbol{A}\mathbf{0} = \mathbf{0}$, the element $\mathbf{0} \in W$.
(2) Suppose $\boldsymbol{x}, \boldsymbol{y} \in W$. Then since $\boldsymbol{A}(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}\boldsymbol{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}$, $\boldsymbol{x} + \boldsymbol{y} \in W$.
(3) For any scalar c, $\boldsymbol{A}(c\boldsymbol{x}) = c\boldsymbol{A}\boldsymbol{x} = c\mathbf{0} = \mathbf{0}$. It follows that $c\boldsymbol{x} \in W$. □

Examples

Example 4.5.

Is $W = \left\{ \boldsymbol{x} \in \Re^3 \middle| \begin{array}{l} x_1 - x_2 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{array} \right\}$ a subspace of $\Re^3$?

Answer: Let $\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. We can write the system of first-order equations as $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$.

From Theorem 4.4, W is a subspace.

Example 4.6.

Is $W = \left\{ \boldsymbol{x} \in \Re^3 \middle| \begin{array}{l} x_1 - x_2 + x_3 = 1 \\ x_1 + x_2 + x_3 = 5 \end{array} \right\}$ a subspace of $\Re^3$?

Answer: Now, $W = \left\{ \boldsymbol{x} \in \Re^3 \middle| \boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} 1 \\ 5 \end{bmatrix} \right\}$. Since $\boldsymbol{A}\boldsymbol{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 5 \end{bmatrix}$, condition (1) in Theorem 4.3 of a subspace is not satisfied. Hence, W is not a subspace of $\Re^3$.

Linear Combination

Definition 4.7 (Linear Combination).

Let the vectors of the linear space be $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$. Suppose $c_1, c_2, \ldots, c_n$ are scalars. The linear combination of these n vectors is defined as

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n.$$

Definition 4.8 (Set of All Linear Combinations).

The set of all linear combinations of the vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ belonging to the linear space V is expressed as

$$< \boldsymbol{u}_1, \boldsymbol{u_2}, \ldots, \boldsymbol{u}_n >:= \left\{ c_1\boldsymbol{u}_1 + \cdots + c_n\boldsymbol{u}_n \middle| c_i \in \Re, \boldsymbol{u}_i \in V \ \ (i = 1, 2, \ldots, n) \right\}.$$

The set $< \boldsymbol{u}_1, \boldsymbol{u_2}, \ldots, \boldsymbol{u}_n >$ is called the **span** of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$.

Theorem 4.9.

$W =< \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n >$ *is a subspace of* V.

Theorem 4.10

Theorem 4.10 (Link to the System of Linear Equations).

Let $W =< \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n >$ be the subspace generated by the n vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ of $\mathfrak{R}^n$. Also, let the $m \times n$ matrix be $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}$. For the m-dimensional vector $\boldsymbol{b}$ to belong to W, the necessary and sufficient condition is that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has solutions.

Proof.

- For $\boldsymbol{b}$ to belong to W, it is equivalent to the existence of a vector $\boldsymbol{c} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}'$ that satisfies the linear combination $c_1\boldsymbol{a}_1 + \cdots + c_n\boldsymbol{a}_n = \boldsymbol{b}$, which is the system of linear equations.
- In the vector-matrix form, we have $\boldsymbol{A}\boldsymbol{c} = \boldsymbol{b}$.
- Thus, we see that the necessary and sufficient condition for $\boldsymbol{b} \in W$ is that the solutions $\boldsymbol{x} = \boldsymbol{c}$ exist.

$\square$

Linear Relation

- The linear relation of n vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ in V is defined by

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n = \mathbf{0}, \quad \text{where } c_1, c_2, \ldots, c_n \in \mathfrak{R}.$$

  This is called the **first-order relation** or **linear relation**.

- The trivial linear relation is defined as the case where $c_1 = c_2 = \cdots = c_n = 0$. If this is the only solution for the linear relation, then these n vectors are said to be **linearly independent**.

- When the n vectors are not independent, they are said to be **linearly dependent**.

Examples of Linear Relation

- The basis vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ of the linear space $\mathfrak{R}^n$ are linearly independent. In fact,

$$c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2 + \cdots + c_n\boldsymbol{e}_n = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{0} \quad \Longleftrightarrow \quad c_1 = c_2 = \cdots = c_n = 0.$$

- Consider the linear space $\mathbb{R}[x]_n$ of n-order polynomials. The $n+1$ vectors of $\mathbb{R}[x]_n$,

$$1, x, x^2, \ldots, x^n,$$

are linearly independent.
Proof: Now we need to examine, for any $x \in \mathfrak{R}$, the full-blown polynomial:

$$c_0 1 + c_1 x + c_2 x^2 + \cdots + c_n x^n = 0. \tag{2}$$

Examples of Linear Relation (つづき)

- Let $x = 0$ and (2) suggests that $c_0 = 0$.
- Differentiate both sides of (2), we obtain $c_1 + 2c_2x + \cdots + nc_nx^{n-1} = 0$. Upon letting $x = 0$, we obtain $c_1 = 0$.
- In the same fashion, differentiate (2) twice, and after setting $x = 0$, we obtain $c_2 = 0$.
- Proceeding in the same way, we can conclude that $c_0 = c_1 = c_2 = \cdots = c_n = 0$.
- Hence, $1, x, x^2, \ldots, x^n$ are linearly independent.

Linear Independence and Rank

Theorem 5.1 (Rank $\equiv$ Linear Independence).

Suppose $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}$, where each vector $\boldsymbol{a}_i$ in $\Re^m$ has m rows. Then, $\boldsymbol{A}$ is an $m \times n$ matrix and the n m-dimensional vectors $\boldsymbol{a}_i$ are linearly independent if and only if

$$\operatorname{rank} \boldsymbol{A} = n.$$

In particular, when $m = n$, it is equivalent to the condition that $\boldsymbol{A}$ is a regular matrix.

Theorem 5.2 (Linear Combination $\equiv$ Linear Dependence).

For the n vectors in the linear space V to be linearly dependent, among $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$, at least one of them can be written as a linear combination of the other $(n-1)$ vectors, and vice versa.

Linear Combination Theorem and Notation

Theorem 5.3.

Suppose n vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ are linearly independent. A vector $\boldsymbol{v}$ is added to these n vectors. If the $(n+1)$ vectors $\boldsymbol{v}, \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ are linearly dependent, then $\boldsymbol{v}$ can be written as a linear combination of the n vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$.

- The notation for a general linear combination is

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n =: (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n)\boldsymbol{c}, \qquad \boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \in \mathfrak{R}^n.$$

- For m linear combinations of $\boldsymbol{v}_i$, as an extension, the notation is $\boldsymbol{v}_i = (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n)\boldsymbol{c}_i$, $i = 1, 2, \ldots, m$, and

$$(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m) := (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n)\begin{bmatrix}\boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_m\end{bmatrix} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n)\boldsymbol{C}.$$

Example

Example 5.4.

Consider two vectors $g_1(x)$ and $g_2(x)$ of $\mathbb{R}[x]_n$, which are written as linear combinations of $f_1(x)$ and $f_2(x)$ as follows:

$$g_1(x) = 3f_1(x) + f_2(x), \qquad g_2(x) = 2f_1(x) - f_2(x).$$

Then,

$$g_1(x) = (f_1(x), f_2(x)) \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \qquad g_2(x) = (f_1(x), f_2(x)) \begin{bmatrix} 2 \\ -1 \end{bmatrix},$$

Moreover,

$$(g_1(x), g_2(x)) = (f_1(x), f_2(x)) \begin{bmatrix} 3 & 2 \\ 1 & -1 \end{bmatrix}.$$

When Vectors Are Linearly Independent

Theorem 5.5.

When the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ of a linear space are linearly independent, then (1) and (2) below hold.

(1) *With respect to $\boldsymbol{x} \in \Re^m$,*

$$(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)\boldsymbol{x} = \boldsymbol{0} \quad \Longleftrightarrow \quad \boldsymbol{x} = \boldsymbol{0}.$$

(2) *With respect to $m \times n$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$,*

$$(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)\boldsymbol{A} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)\boldsymbol{B} \quad \Longleftrightarrow \quad \boldsymbol{A} = \boldsymbol{B}.$$

Number of Linearly Independent Vectors

Definition 5.6.

The largest number r of linearly independent vectors in a linear space is such that $r+1$ vectors become linearly dependent.

Theorem 5.7.

Given two groups of vectors $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ *and* $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *of a linear space* V, *if the following two conditions are satisfied, then* $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ *are linearly dependent.*

(1) *Every vector in the group* $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *can be written as a linear combination of* $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$.

(2) $n > m$.

Theorem 5.8.

The rank of a matrix $\boldsymbol{A}$ *equals the largest number* m *of linearly independent vectors that make up the* m *columns of* $\boldsymbol{A}$.

Example

Example 5.9.

Consider three vectors: $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \quad \boldsymbol{a}_2 = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix}, \quad \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. Find the largest number of vectors that are linearly independent.

- Let $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix}$.
- Theorem 5.8 indicates that we just need to find the rank of $\boldsymbol{A}$.
- We perform simplification of $\boldsymbol{A}$ by applying the basic matrix operations:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 1 \\ -1 & -2 & 0 \\ 2 & 5 & 1 \end{bmatrix} \longrightarrow \boldsymbol{B} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

- Hence the largest number of independent vectors is 2.

Basis

Definition 6.1.

When the n vectors $\boldsymbol{u}_i$ $(i = 1, 2, \ldots, n)$ of a linear space V satisfy the following two conditions, they are said to be the **basis** of V:

1. $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ are linearly independent.
2. $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ **generate** V.

- The set of basic vectors $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ of $\Re^n$ is the basis of $\Re^n$, and it is called the **standard basis**.

- What is a basis for $\Re^{2\times 2}$?

  Answer: There are many possible answers. A possible basis is

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

An Example of Basis

- In $\mathfrak{R}^3$, every vector has form $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$, where a, b, c are real numbers.
- Note that $\mathfrak{R}^3$ is spanned by the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \quad \text{since} \quad a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

- Clearly, $a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ if and only if $a = b = c = 0$.
- Hence, the set consists of linearly independent vectors that spans $\mathfrak{R}^3$ and is therefore a basis for $\mathfrak{R}^3$.

An Example

Example 6.2.

Show that the set of $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$, and $\boldsymbol{a}_3 = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}$ forms the basis of $\mathfrak{R}^3$.

- Define the matrix $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix}$.
- Since the determinant $|\boldsymbol{A}| = \begin{vmatrix} 1 & -2 & 0 \\ 0 & 3 & -1 \\ -2 & 1 & 2 \end{vmatrix} = 3 \neq 0$, we conclude that $\boldsymbol{A}$ is regular.
- Theorem 5.1 indicates that $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ are linearly independent.

An Example (つづき)

- Consider an arbitrary vector $\boldsymbol{x} \in \Re^3$, and express it as a linear combination with coefficients c_1, c_2 and c_3 as follows:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + c_3\boldsymbol{a}_3 = \boldsymbol{A}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \boldsymbol{A}\boldsymbol{c}.$$

- Since $\boldsymbol{A}^{-1}$ exists, we can solve for $\boldsymbol{c}$

$$\boldsymbol{c} = \boldsymbol{A}^{-1}\boldsymbol{x}.$$

- Having found $\boldsymbol{c}$, $\boldsymbol{x}$ indeed can be written as a linear combination of $\boldsymbol{a}_1, \boldsymbol{a}_2$, and $\boldsymbol{a}_3$, the set $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ indeed is a basis of $\Re^3$.

Basis Transformation Matrix

Definition 6.3.

For two sets of basis, $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ and $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ of an n-dimensional linear space V, the regular matrix $\boldsymbol{P}$ such that

$$\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n \end{bmatrix} \boldsymbol{P}$$

is called the **basis transformation matrix**.

Example of Basis Transformation Matrix

Example 6.4.

Consider two sets of basis of $\Re^2$:

$$\left\{ \boldsymbol{u}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}, \quad \left\{ \boldsymbol{v}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}.$$

Find the basis transformation matrix.

- Let $\boldsymbol{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$. We need to find the unknown entries p_{11}, p_{12}, p_{21}, and p_{22}.
- Now $$\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} \boldsymbol{P}.$$
- In short, $\boldsymbol{V} = \boldsymbol{U}\boldsymbol{P}$, where $\boldsymbol{U} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix}$, and $\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix}$. Multiplying from the left $\boldsymbol{U}^{-1}$, we obtain

$$\boldsymbol{P} = \boldsymbol{U}^{-1}\boldsymbol{V} = \begin{bmatrix} 3 & -2 \\ -4 & 3 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} -9 & 5 \\ 13 & -7 \end{bmatrix}.$$

Dimension

Definition 6.5.

If the basis formed by a finite number of vectors in the linear space V exists, then V is said to be a **linear space of finite dimension**. The number of vectors that form the basis is said to be the **dimension** of V, and it is expressed as $\dim V$, with the convention that if $V = \{\mathbf{0}\}$, then $\dim V = 0$.

- $\dim \Re^n = n$, since the number of basis vectors e_i is n.
- For $\mathbb{R}[x]_n$, the vectors $1, x, x^2, \ldots, x^n$ are linearly independent. Thus $\dim \mathbb{R}[x]_n = n + 1$.

Simple Theorems

Theorem 6.6.

The dimension of a finite linear space V is equal to the largest number of linearly independent vectors in V.

Theorem 6.7.

When W is a subspace of V, if $\dim W = \dim V$, then $W = V$.

Other Theorems

- Let $\boldsymbol{A}$ be a $m \times n$ matrix. The space of solutions of the first-order system of equations $\boldsymbol{Ax} = \mathbf{0}$ is expressed as
$$W = \left\{\boldsymbol{x} \in \Re^n | \boldsymbol{Ax} = \mathbf{0}\right\}.$$
Then $\dim W = n - \operatorname{rank} \boldsymbol{A}$.

- The following 3 conditions are equivalent for n-dimensional linear space V:
  1. $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ of V are linearly independent.
  2. $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ generate V.
  3. $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ are the basis of V.

- Let the basis of an n-dimensional linear space be $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$. Suppose $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_\ell$ with $\ell \leqq n$ are independent. By selecting $(n - \ell)$ vectors $\boldsymbol{v}_{i_1}, \ldots, \boldsymbol{v}_{i_{n-\ell}}$ $\left(1 \leqq i_1, \cdots, i_{n-\ell} \leqq n\right)$, the set of vectors $\boldsymbol{w}_1, \ldots \boldsymbol{w}_\ell, \boldsymbol{v}_{i_1}, \ldots, \boldsymbol{v}_{i_{n-\ell}}$ forms the basis of V.

Takeaways

- Space is a mathematical construct to contain objects called elements.
- Inner product is important for geometrical vectors but not for the linear space of (abstract) vectors.
- The concept of subspace of a linear space marks a distinction between $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ vis-à-vis $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.
- Span is the set of all linear combinations of n vectors.
- A vector $\boldsymbol{b}$'s membership of a subspace is linked to whether the solutions for $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ exist (The columns of $\boldsymbol{A}$ are the vectors that span the subspace.).

Takeaways (つづき)

- For n vectors to be linearly independent, the matrix formed by them must have the rank equal to n, which corresponds to the largest number of linearly independent vectors for a liner space.

- The dimension of a linear space is equal to the largest number of linearly independent vectors.

- Basis$\longleftrightarrow$ Span$\longleftrightarrow$Linear Independence

Keywords