

# Mathematics for Machine Learning

— Linear Algebra: Basis, Rank, Linear Mappings & Affine Spaces

Joseph Chuang-Chieh Lin

Department of Computer Science & Engineering,  
National Taiwan Ocean University

Spring 2025

## Credits for the resource

- The slides are based on the textbooks:
  - *Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong: Mathematics for Machine Learning. Cambridge University Press. 2020.*
  - *Howard Anton, Chris Rorres, Anton Kaul: Elementary Linear Algebra. Wiley. 2019.*
- We could partially refer to the monograph:  
*Francesco Orabona: A Modern Introduction to Online Learning.*  
<https://arxiv.org/abs/1912.13213>

# Outline

- 1 Why linear algebra?
- 2 Vector Space
- 3 Basis & Dimension & Rank
- 4 Linear Mappings
- 5 Affine Spaces

# Outline

- 1 Why linear algebra?
- 2 Vector Space
- 3 Basis & Dimension & Rank
- 4 Linear Mappings
- 5 Affine Spaces

# Why linear algebra?

- Crucial in the graduate school entrance examination.

# Why linear algebra?

- Crucial in the graduate school entrance examination.
- Matrix operations.

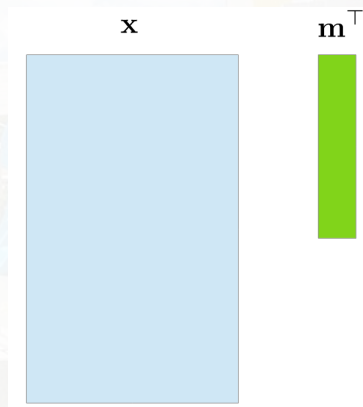
# Why linear algebra?

- Crucial in the graduate school entrance examination.
- Matrix operations.
- Vectorization.

# Vectorization Example (1/3)

$$\begin{aligned}y_i &= \langle \mathbf{m}, \mathbf{x}_i \rangle \\ &= m_1 x_{i,1} + m_2 x_{i,2} + \dots + m_k x_{i,k}.\end{aligned}$$

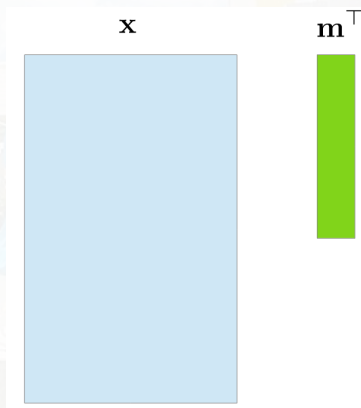
```
m = np.random.rand(1,5)
x = np.random.rand(5000000,5)
#assume k=5
```





## Vectorization Example (2/3)

```
start = time.time()
zer = []
for i in range(0,5000000):
    total = 0
    for j in range(0,5):
        total = total + x[i][j]*m[0][j]
    zer.append(total)
zer = np.array(zer)
end = time.time()
```



## Vectorization Example (2/3)

```
start = time.time()
zer = []
for i in range(0,5000000):
    total = 0
    for j in range(0,5):
        total = total + x[i][j]*m[0][j]
    zer.append(total)
zer = np.array(zer)
end = time.time()
```

```
In [8]: runfile('C:/Users/josep/_Project/
vectorization_matrix.py', wdir='C:/Users/josep/_Project')
Computation time = 13.515385389328003 seconds
```

## Vectorization Example (3/3)

```
start = time.time()  
zer = np.matmul(x, m.T)  
end = time.time()
```

```
In [13]: runfile('C:/Users/josep/_Project/  
vectorization_matrix.py', wdir='C:/Users/josep/_Project')  
Computation time = 0.010425329208374023 seconds
```

# Outline

- 1 Why linear algebra?
- 2 Vector Space**
- 3 Basis & Dimension & Rank
- 4 Linear Mappings
- 5 Affine Spaces

# Group

## Group

Consider a set  $\mathcal{G}$  and an operation  $\otimes : \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$  defined on  $\mathcal{G}$ . Then  $G : (\mathcal{G}, \otimes)$  is called a **group** if the following conditions hold:

# Group

## Group

Consider a set  $\mathcal{G}$  and an operation  $\otimes : \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$  defined on  $\mathcal{G}$ . Then  $G : (\mathcal{G}, \otimes)$  is called a **group** if the following conditions hold:

①  $\forall x, y \in \mathcal{G}, x \otimes y \in \mathcal{G}.$

# Group

## Group

Consider a set  $\mathcal{G}$  and an operation  $\otimes : \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$  defined on  $\mathcal{G}$ . Then  $G : (\mathcal{G}, \otimes)$  is called a **group** if the following conditions hold:

- ①  $\forall x, y \in \mathcal{G}, x \otimes y \in \mathcal{G}$ .
- ②  $\forall x, y, z \in \mathcal{G}, (x \otimes y) \otimes z = x \otimes (y \otimes z)$ .

# Group

## Group

Consider a set  $\mathcal{G}$  and an operation  $\otimes : \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$  defined on  $\mathcal{G}$ . Then  $G : (\mathcal{G}, \otimes)$  is called a **group** if the following conditions hold:

- ①  $\forall x, y \in \mathcal{G}, x \otimes y \in \mathcal{G}$ .
- ②  $\forall x, y, z \in \mathcal{G}, (x \otimes y) \otimes z = x \otimes (y \otimes z)$ .
- ③  $\exists e \in \mathcal{G}$  such that  $\forall x \in \mathcal{G}, x \otimes e = e \otimes x = x$ .



# Group

## Group

Consider a set  $\mathcal{G}$  and an operation  $\otimes : \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$  defined on  $\mathcal{G}$ . Then  $G : (\mathcal{G}, \otimes)$  is called a **group** if the following conditions hold:

- ①  $\forall x, y \in \mathcal{G}, x \otimes y \in \mathcal{G}$ .
- ②  $\forall x, y, z \in \mathcal{G}, (x \otimes y) \otimes z = x \otimes (y \otimes z)$ .
- ③  $\exists e \in \mathcal{G}$  such that  $\forall x \in \mathcal{G}, x \otimes e = e \otimes x = x$ .
- ④  $\forall x \in \mathcal{G}, \exists y \in \mathcal{G}$  such that  $x \otimes y = y \otimes x = e$ . We denote by  $x^{-1}$  the inverse element of  $x$ .

# Group

## Group

Consider a set  $\mathcal{G}$  and an operation  $\otimes : \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$  defined on  $\mathcal{G}$ . Then  $G : (\mathcal{G}, \otimes)$  is called a **group** if the following conditions hold:

- ①  $\forall x, y \in \mathcal{G}, x \otimes y \in \mathcal{G}$ .
  - ②  $\forall x, y, z \in \mathcal{G}, (x \otimes y) \otimes z = x \otimes (y \otimes z)$ .
  - ③  $\exists e \in \mathcal{G}$  such that  $\forall x \in \mathcal{G}, x \otimes e = e \otimes x = x$ .
  - ④  $\forall x \in \mathcal{G}, \exists y \in \mathcal{G}$  such that  $x \otimes y = y \otimes x = e$ . We denote by  $x^{-1}$  the inverse element of  $x$ .
- If  $G$  is a group and  $\forall x, y \in \mathcal{G}$  we have  $x \otimes y = y \otimes x$ , then  $G$  is an **Abelian** group.

# Examples

- $(\mathbb{Z}, +)$ : an Abelian group.
- $(\mathbb{N} \cup \{0\}, +)$  is NOT a group.
- $(\mathbb{Z}, \cdot)$  is NOT a group.
- $(\mathbb{R}, \cdot)$  is NOT a group.
- $(\mathbb{R} \setminus \{0\}, \cdot)$  is an Abelian group.
- $(\mathbb{R}^{m \times n}, +)$  is an Abelian group.

# Vector Space

## Vector Space

A real-valued vector space  $V = (\mathcal{V}, +, \cdot)$  is a set  $\mathcal{V}$  with two operations:

$$+ : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$$

$$\cdot : \mathbb{R} \times \mathcal{V} \rightarrow \mathcal{V}$$

where

- $(\mathcal{V}, +)$  is an Abelian group.
- Distributivity holds:
  - $\forall \lambda \in \mathbb{R}, \mathbf{x}, \mathbf{y} \in \mathcal{V}: \lambda \cdot (\mathbf{x} + \mathbf{y}) = \lambda \cdot \mathbf{x} + \lambda \cdot \mathbf{y}.$
  - $\forall \lambda, \psi \in \mathbb{R}, \mathbf{x} \in \mathcal{V}: (\lambda + \psi) \cdot \mathbf{x} = \lambda \cdot \mathbf{x} + \psi \cdot \mathbf{x}.$
- $\forall \lambda, \psi \in \mathbb{R}, \mathbf{x} \in \mathcal{V}: \lambda \cdot (\psi \cdot \mathbf{x}) = (\lambda\psi) \cdot \mathbf{x}.$
- $\forall \mathbf{x} \in \mathcal{V}: 1 \cdot \mathbf{x} = \mathbf{x}.$

★ Note: A vector multiplication is not defined.

# Vector Subspaces

## Vector Subspace

Let  $V = (\mathcal{V}, +, \cdot)$  be a vector space and  $\mathcal{U} \subset \mathcal{V}$  and  $\mathcal{U} \neq \emptyset$ . Then  $U = (\mathcal{U}, +, \cdot)$  is called a vector **subspace** of  $V$  if  $U$  is a vector space with the operations  $+$  and  $\cdot$  restricted to  $\mathcal{U} \times \mathcal{U}$  and  $\mathbb{R} \times \mathcal{U}$  respectively.

- Denote by  $U \subseteq V$  a subspace  $U$  of  $V$ .

# Vector Subspaces

## Vector Subspace

Let  $V = (\mathcal{V}, +, \cdot)$  be a vector space and  $\mathcal{U} \subset \mathcal{V}$  and  $\mathcal{U} \neq \emptyset$ . Then  $U = (\mathcal{U}, +, \cdot)$  is called a vector **subspace** of  $V$  if  $U$  is a **vector space** with the operations  $+$  and  $\cdot$  **restricted to  $\mathcal{U} \times \mathcal{U}$  and  $\mathbb{R} \times \mathcal{U}$  respectively**.

- Denote by  $U \subseteq V$  a subspace  $U$  of  $V$ .

# Examples

- The trivial subspace of a vector space  $V$ :  $\{\mathbf{0}\}$  and  $V$ .

# Examples

- The trivial subspace of a vector space  $V$ :  $\{\mathbf{0}\}$  and  $V$ .
- The solution set of a homogeneous system of linear equations  $A\mathbf{x} = \mathbf{0}$  with  $n$  unknowns (i.e.,  $\mathbf{x} = [x_1, \dots, x_n]^\top$ ) is a subspace of  $\mathbb{R}^n$ .



# Examples

- The trivial subspace of a vector space  $V$ :  $\{\mathbf{0}\}$  and  $V$ .
- The solution set of a homogeneous system of linear equations  $A\mathbf{x} = \mathbf{0}$  with  $n$  unknowns (i.e.,  $\mathbf{x} = [x_1, \dots, x_n]^\top$ ) is a subspace of  $\mathbb{R}^n$ .
- The intersection of arbitrarily many subspaces is a subspace.

# Examples

- The trivial subspace of a vector space  $V$ :  $\{\mathbf{0}\}$  and  $V$ .
- The solution set of a homogeneous system of linear equations  $A\mathbf{x} = \mathbf{0}$  with  $n$  unknowns (i.e.,  $\mathbf{x} = [x_1, \dots, x_n]^\top$ ) is a subspace of  $\mathbb{R}^n$ .
- The intersection of arbitrarily many subspaces is a subspace.
- The solution of an **inhomogeneous system** of linear equations  $A\mathbf{x} = \mathbf{b}$  for  $\mathbf{b} \neq \mathbf{0}$  is NOT a subspace of  $\mathbb{R}^n$ .

# Linear Combination

## Linear Combination

Consider a vector space  $V$  and a finite number of vectors  $\mathbf{x}_1, \dots, \mathbf{x}_k \in V$ . Then, every  $\mathbf{v} \in V$  of the form

$$\mathbf{v} = \lambda_1 \mathbf{x}_1 + \dots + \lambda_k \mathbf{x}_k = \sum_{i=1}^k \lambda_i \mathbf{x}_i \in V$$

with  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$  is a **linear combination** of the vectors  $\mathbf{x}_1, \dots, \mathbf{x}_k$ .

- **Question:** How to represent  $\mathbf{0}$  as a linear combination of  $\mathbf{x}_1, \dots, \mathbf{x}_k$ ?

# Linearly Independent

## Linear (In)dependence

Consider a vector space  $V$  with  $k > 0$  vectors  $\mathbf{x}_1, \dots, \mathbf{x}_k \in V$ .

- If there is a nontrivial linear combination such that  $\mathbf{0} = \sum_{i=1}^k \lambda_i \mathbf{x}_i$  with at least one  $\lambda_i \neq 0$ , then we say  $\mathbf{x}_1, \dots, \mathbf{x}_k$  are **linearly dependent**.
- If only the trivial solution exists (i.e.,  $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ ), then we say  $\mathbf{x}_1, \dots, \mathbf{x}_k$  are **linearly independent**.

# Recall some facts

- If at least one of vectors  $\mathbf{x}_1, \dots, \mathbf{x}_k$  is  $\mathbf{0}$  then they are linearly dependent.

# Recall some facts

- If at least one of vectors  $\mathbf{x}_1, \dots, \mathbf{x}_k$  is  $\mathbf{0}$  then they are linearly dependent.
- Two identical vectors are linearly dependent.

# Recall some facts

- If at least one of vectors  $\mathbf{x}_1, \dots, \mathbf{x}_k$  is  $\mathbf{0}$  then they are linearly dependent.
- Two identical vectors are linearly dependent.
- Write all vectors as rows (or columns) of a matrix and perform Gaussian elimination until the matrix is in row echelon form.

## Remark (1/2)

Consider a vector space  $V$  with  $k$  linearly independent vectors  $\mathbf{b}_1, \dots, \mathbf{b}_k$  and  $m$  linear combinations

$$\mathbf{x}_1 = \sum_{i=1}^k \lambda_{i,1} \mathbf{b}_i$$

$$\vdots$$

$$\mathbf{x}_m = \sum_{i=1}^k \lambda_{i,m} \mathbf{b}_i$$



## Remark (1/2)

Consider a vector space  $V$  with  $k$  linearly independent vectors  $\mathbf{b}_1, \dots, \mathbf{b}_k$  and  $m$  linear combinations

$$\mathbf{x}_1 = \sum_{i=1}^k \lambda_{i,1} \mathbf{b}_i$$

$$\vdots$$

$$\mathbf{x}_m = \sum_{i=1}^k \lambda_{i,m} \mathbf{b}_i$$

- Define  $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_k]$  (i.e., a matrix), then

$$\mathbf{x}_j = \mathbf{B} \boldsymbol{\lambda}_j, \text{ for } \boldsymbol{\lambda}_j = \begin{bmatrix} \lambda_{1j} \\ \vdots \\ \lambda_{kj} \end{bmatrix}, j = 1, \dots, m.$$

## Remark (2/2)

We want to test whether  $\mathbf{x}_1, \dots, \mathbf{x}_m$  are linearly independent.

- $\sum_{j=1}^m \psi_j \mathbf{x}_j = \mathbf{0}.$

## Remark (2/2)

We want to test whether  $\mathbf{x}_1, \dots, \mathbf{x}_m$  are linearly independent.

- $\sum_{j=1}^m \psi_j \mathbf{x}_j = \mathbf{0}$ .
- So,

$$\sum_{j=1}^m \psi_j \mathbf{x}_j = \sum_{j=1}^m \psi_j \mathbf{B} \boldsymbol{\lambda}_j = \mathbf{B} \sum_{j=1}^m \psi_j \boldsymbol{\lambda}_j.$$

## Remark (2/2)

We want to test whether  $\mathbf{x}_1, \dots, \mathbf{x}_m$  are linearly independent.

- $\sum_{j=1}^m \psi_j \mathbf{x}_j = \mathbf{0}.$

- So,

$$\sum_{j=1}^m \psi_j \mathbf{x}_j = \sum_{j=1}^m \psi_j \mathbf{B} \boldsymbol{\lambda}_j = \mathbf{B} \sum_{j=1}^m \psi_j \boldsymbol{\lambda}_j.$$

- Why does the last equality hold?

## Remark (2/2)

We want to test whether  $\mathbf{x}_1, \dots, \mathbf{x}_m$  are linearly independent.

- $\sum_{j=1}^m \psi_j \mathbf{x}_j = \mathbf{0}.$

- So,

$$\sum_{j=1}^m \psi_j \mathbf{x}_j = \sum_{j=1}^m \psi_j \mathbf{B} \boldsymbol{\lambda}_j = \mathbf{B} \sum_{j=1}^m \psi_j \boldsymbol{\lambda}_j.$$

- Why does the last equality hold?
- $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$  are linearly independent iff  $\{\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_m\}$  are linearly independent.

## Remark (2/2)

We want to test whether  $\mathbf{x}_1, \dots, \mathbf{x}_m$  are linearly independent.

- $\sum_{j=1}^m \psi_j \mathbf{x}_j = \mathbf{0}.$

- So,

$$\sum_{j=1}^m \psi_j \mathbf{x}_j = \sum_{j=1}^m \psi_j \mathbf{B} \boldsymbol{\lambda}_j = \mathbf{B} \sum_{j=1}^m \psi_j \boldsymbol{\lambda}_j.$$

- Why does the last equality hold?
- $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$  are linearly independent iff  $\{\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_m\}$  are linearly independent.
- **Note:**  $m$  linear combinations of  $k$  vectors  $\mathbf{x}_1, \dots, \mathbf{x}_k$  are linearly *dependent* if  $m > k$ .

## Example

Consider a set of linearly independent vectors  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4 \in \mathbb{R}^n$  and

$$\begin{aligned}\mathbf{x}_1 &= \mathbf{b}_1 - 2\mathbf{b}_2 + \mathbf{b}_3 - \mathbf{b}_4 \\ \mathbf{x}_2 &= -4\mathbf{b}_1 - 2\mathbf{b}_2 + 4\mathbf{b}_4 \\ \mathbf{x}_3 &= 2\mathbf{b}_1 + 3\mathbf{b}_2 - \mathbf{b}_3 - 3\mathbf{b}_4 \\ \mathbf{x}_4 &= 17\mathbf{b}_1 - 10\mathbf{b}_2 + 11\mathbf{b}_3 + \mathbf{b}_4\end{aligned}$$

**Question:** Is  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$  linearly independent?

## Example

Consider a set of linearly independent vectors  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4 \in \mathbb{R}^n$  and

$$\begin{aligned}\mathbf{x}_1 &= \mathbf{b}_1 - 2\mathbf{b}_2 + \mathbf{b}_3 - \mathbf{b}_4 \\ \mathbf{x}_2 &= -4\mathbf{b}_1 - 2\mathbf{b}_2 + 4\mathbf{b}_4 \\ \mathbf{x}_3 &= 2\mathbf{b}_1 + 3\mathbf{b}_2 - \mathbf{b}_3 - 3\mathbf{b}_4 \\ \mathbf{x}_4 &= 17\mathbf{b}_1 - 10\mathbf{b}_2 + 11\mathbf{b}_3 + \mathbf{b}_4\end{aligned}$$

**Question:** Is  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$  linearly independent?

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 & -1 \\ -4 & -2 & 0 & 4 \\ 2 & 3 & -1 & 3 \\ 17 & -10 & 11 & 1 \end{bmatrix}$$



## Example

Consider a set of linearly independent vectors  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4 \in \mathbb{R}^n$  and

$$\begin{aligned}\mathbf{x}_1 &= \mathbf{b}_1 - 2\mathbf{b}_2 + \mathbf{b}_3 - \mathbf{b}_4 \\ \mathbf{x}_2 &= -4\mathbf{b}_1 - 2\mathbf{b}_2 + 4\mathbf{b}_4 \\ \mathbf{x}_3 &= 2\mathbf{b}_1 + 3\mathbf{b}_2 - \mathbf{b}_3 - 3\mathbf{b}_4 \\ \mathbf{x}_4 &= 17\mathbf{b}_1 - 10\mathbf{b}_2 + 11\mathbf{b}_3 + \mathbf{b}_4\end{aligned}$$

**Question:** Is  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$  linearly independent?

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 & -1 \\ -4 & -2 & 0 & 4 \\ 2 & 3 & -1 & 3 \\ 17 & -10 & 11 & 1 \end{bmatrix} \longrightarrow$$

## Example

Consider a set of linearly independent vectors  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4 \in \mathbb{R}^n$  and

$$\begin{aligned}\mathbf{x}_1 &= \mathbf{b}_1 - 2\mathbf{b}_2 + \mathbf{b}_3 - \mathbf{b}_4 \\ \mathbf{x}_2 &= -4\mathbf{b}_1 - 2\mathbf{b}_2 + 4\mathbf{b}_4 \\ \mathbf{x}_3 &= 2\mathbf{b}_1 + 3\mathbf{b}_2 - \mathbf{b}_3 - 3\mathbf{b}_4 \\ \mathbf{x}_4 &= 17\mathbf{b}_1 - 10\mathbf{b}_2 + 11\mathbf{b}_3 + \mathbf{b}_4\end{aligned}$$

**Question:** Is  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$  linearly independent?

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 & -1 \\ -4 & -2 & 0 & 4 \\ 2 & 3 & -1 & 3 \\ 17 & -10 & 11 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -2 & 1 & -1 \\ 0 & 1 & -\frac{2}{5} & 0 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

# Outline

- 1 Why linear algebra?
- 2 Vector Space
- 3 Basis & Dimension & Rank**
- 4 Linear Mappings
- 5 Affine Spaces

# Basis

## Spanning/Generating

Consider a vector space  $V = (\mathcal{V}, +, \cdot)$  and a set  $\mathcal{A} = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subseteq \mathcal{V}$ .

If every vector  $\mathbf{v} \in \mathcal{V}$  can be expressed as a linear combination of vectors in  $\mathcal{A}$ , then  $\mathcal{A}$  is called a **spanning set (or generating set)** of  $V$ .

- $\mathcal{A}$  spans  $V$ ;  $\text{span}(\mathcal{A}) = V$ .

# Basis

## Spanning/Generating

Consider a vector space  $V = (\mathcal{V}, +, \cdot)$  and a set  $\mathcal{A} = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subseteq \mathcal{V}$ .

If every vector  $\mathbf{v} \in \mathcal{V}$  can be expressed as a linear combination of vectors in  $\mathcal{A}$ , then  $\mathcal{A}$  is called a **spanning set (or generating set)** of  $V$ .

- $\mathcal{A}$  spans  $V$ ;  $\text{span}(\mathcal{A}) = V$ .

## Basis

Consider a vector space  $V = (\mathcal{V}, +, \cdot)$  and a set  $\mathcal{A} \subseteq \mathcal{V}$ . Then if one of the following condition holds, we say that  $\mathcal{A}$  is a **basis** of  $V$ .

- $\mathcal{A}$  is a minimal generating set of  $V$ .

# Basis

## Spanning/Generating

Consider a vector space  $V = (\mathcal{V}, +, \cdot)$  and a set  $\mathcal{A} = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subseteq \mathcal{V}$ .

If every vector  $\mathbf{v} \in \mathcal{V}$  can be expressed as a linear combination of vectors in  $\mathcal{A}$ , then  $\mathcal{A}$  is called a **spanning set (or generating set)** of  $V$ .

- $\mathcal{A}$  spans  $V$ ;  $\text{span}(\mathcal{A}) = V$ .

## Basis

Consider a vector space  $V = (\mathcal{V}, +, \cdot)$  and a set  $\mathcal{A} \subseteq \mathcal{V}$ . Then if one of the following condition holds, we say that  $\mathcal{A}$  is a **basis** of  $V$ .

- $\mathcal{A}$  is a minimal generating set of  $V$ .  
No smaller set  $\mathcal{A}' \subsetneq \mathcal{A} \subseteq \mathcal{V}$  that spans  $V$ .
- $\mathcal{A}$  spans  $V$  and is also linearly independent.

# Dimension

## Dimension

The number of basis vectors of a vector space  $V$  is the *dimension* of  $V$  and denoted by  $\dim(V)$ .

- For  $U \subset V$  a subspace of  $V$ ,  $\dim(U) \leq \dim(V)$

## Example

Let  $V = \mathbb{R}[x]$  be the vector space of all real-coefficient polynomials. Define

$$U = x\mathbb{R}[x] = \{xp(x) : p(x) \in \mathbb{R}[x]\},$$

the set of all polynomials whose constant term is 0.

**Claim.**  $U \subsetneq V$  and  $\dim(U) = \dim(V)$ .

- $U$  is a subspace of  $V$ : it is closed under addition and scalar multiplication by construction.
- $U$  is proper:  $1 \in V$  but  $1 \notin U$  (no polynomial  $p$  satisfies  $xp(x) = 1$ ).
- A standard basis of  $V$  is  $\mathcal{B}_V = \{1, x, x^2, x^3, \dots\}$ . Hence  $\dim(V) = |\mathcal{B}_V| = \aleph_0$  (countably infinite).
- A basis of  $U$  is  $\mathcal{B}_U = \{x, x^2, x^3, \dots\}$ , so  $\dim(U) = |\mathcal{B}_U| = \aleph_0$ .

Therefore  $U \subsetneq V$  but  $\dim(U) = \dim(V) = \aleph_0$ .



## Exercise

$$\text{Given } \mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 2 \\ -2 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 3 \\ -4 \\ 3 \\ 5 \\ -3 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} -1 \\ 8 \\ -5 \\ -6 \\ 1 \end{bmatrix}.$$

Find a basis of  $\text{span}(\{\mathbf{x}_1, \dots, \mathbf{x}_4\})$ .

# Rank

## Rank

The number of linearly independent columns of a matrix  $\mathbf{A} = \mathbb{R}^{m \times n}$ .

# Rank

## Rank

The number of linearly independent columns of a matrix  $\mathbf{A} = \mathbb{R}^{m \times n}$ .

- This equals the number of linearly independent rows of  $\mathbf{A}$ .
- Denote by  $\text{rank}(\mathbf{A})$  the rank of  $\mathbf{A}$ .

# Important Properties

- $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^\top)$ .
- For all  $\mathbf{A} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{A}$  is invertible if and only if  $\text{rank}(\mathbf{A}) = n$ .
- $\text{nullity}(\mathbf{A}) = \dim(\text{null}(\mathbf{A})) = n - \text{rank}(\mathbf{A})$ , where  $\text{null}(\mathbf{A})$  is the subspace of  $\mathbb{R}^n$  which solutions for  $\mathbf{A}\mathbf{x} = \mathbf{0}$ .
- If  $\text{rank}(\mathbf{A}) = \min\{m, n\}$  for a matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , then we say  $\mathbf{A}$  has **full rank**.

# Outline

- 1 Why linear algebra?
- 2 Vector Space
- 3 Basis & Dimension & Rank
- 4 Linear Mappings
- 5 Affine Spaces

# Linear Mappings/Linear Transformation

A mapping  $\Phi : V \rightarrow W$  preserves the structure of the vector space if

- $\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y})$
- $\Phi(\lambda \mathbf{x}) = \lambda \Phi(\mathbf{x})$

for all  $\mathbf{x}, \mathbf{y} \in V$  and  $\lambda \in \mathbb{R}$ .

# Linear Mappings/Linear Transformation

A mapping  $\Phi : V \rightarrow W$  preserves the structure of the vector space if

- $\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y})$
- $\Phi(\lambda \mathbf{x}) = \lambda \Phi(\mathbf{x})$

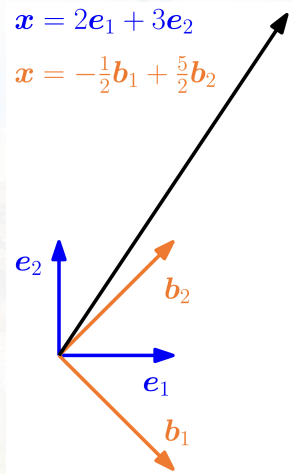
for all  $\mathbf{x}, \mathbf{y} \in V$  and  $\lambda \in \mathbb{R}$ .

## Linear Mapping

For two vector spaces  $V, W$ , a mapping  $\Phi : V \rightarrow W$  is a **linear mapping** if

$$\forall \mathbf{x}, \mathbf{y} \in V, \forall \lambda, \psi \in \mathbb{R} : \Phi(\lambda \mathbf{x} + \psi \mathbf{y}) = \lambda \Phi(\mathbf{x}) + \psi \Phi(\mathbf{y}).$$

# Different coordinate representation





# Transformation Matrix

## Transformation Matrix

Given vector spaces  $V, W$  with corresponding bases  $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$  and  $C = (\mathbf{c}_1, \dots, \mathbf{c}_m)$ . Consider a linear mapping  $\Phi : V \rightarrow W$ . For  $1 \leq j \leq n$ ,

$$\Phi(\mathbf{b}_j) = \alpha_{1,j}\mathbf{c}_1 + \dots + \alpha_{m,j}\mathbf{c}_m = \sum_{i=1}^m \alpha_{ij}\mathbf{c}_i$$

is the unique representation of  $\Phi(\mathbf{b}_j)$  w.r.t.  $C$  (i.e., coordinate). Then, we call the  $m \times n$  matrix  $\mathbf{A}_\Phi$ , whose elements are  $A_\Phi(i, j) = \alpha_{ij}$ , the **transformation matrix** of  $\Phi$ .

- If  $\hat{\mathbf{x}}$  is the coordinate of  $\mathbf{x} \in V$  w.r.t.  $B$  and  $\hat{\mathbf{y}} = \Phi(\mathbf{x}) \in W$  w.r.t.  $C$ , then

$$\hat{\mathbf{y}} = \mathbf{A}_\Phi(\hat{\mathbf{x}}).$$

## Example

Consider a linear mapping  $\Phi : V \rightarrow W$  and ordered bases  $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$  of  $V$  and  $C = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4)$  of  $W$ . Assume that

$$\Phi(\mathbf{b}_1) = \mathbf{c}_1 - \mathbf{c}_2 + 3\mathbf{c}_3 - \mathbf{c}_4$$

$$\Phi(\mathbf{b}_2) = 2\mathbf{c}_1 + \mathbf{c}_2 + 7\mathbf{c}_3 + 2\mathbf{c}_4$$

$$\Phi(\mathbf{b}_3) = 3\mathbf{c}_2 + \mathbf{c}_3 + 4\mathbf{c}_4.$$

The transformation matrix  $\mathbf{A}_\Phi$  w.r.t.  $B$  and  $C$  satisfying  $\Phi(\mathbf{b}_k) = \sum_{i=1}^4 \alpha_{ik} \mathbf{c}_i$  for  $k = 1, 2, 3$  is

## Example

Consider a linear mapping  $\Phi : V \rightarrow W$  and ordered bases  $B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$  of  $V$  and  $C = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4)$  of  $W$ . Assume that

$$\Phi(\mathbf{b}_1) = \mathbf{c}_1 - \mathbf{c}_2 + 3\mathbf{c}_3 - \mathbf{c}_4$$

$$\Phi(\mathbf{b}_2) = 2\mathbf{c}_1 + \mathbf{c}_2 + 7\mathbf{c}_3 + 2\mathbf{c}_4$$

$$\Phi(\mathbf{b}_3) = 3\mathbf{c}_2 + \mathbf{c}_3 + 4\mathbf{c}_4.$$

The transformation matrix  $\mathbf{A}_\Phi$  w.r.t.  $B$  and  $C$  satisfying  $\Phi(\mathbf{b}_k) = \sum_{i=1}^4 \alpha_{ik} \mathbf{c}_i$  for  $k = 1, 2, 3$  is

$$\mathbf{A}_\Phi = [\alpha_1, \alpha_2, \alpha_3] = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}.$$

# Basis Change (1/4)

- $[I]_B^{B'}$ : a transformation matrix that maps coordinates w.r.t.  $B$  onto coordinates w.r.t.  $B'$ .

# Basis Change (1/4)

- $[I]_B^{B'}$ : a transformation matrix that maps coordinates w.r.t.  $B$  onto coordinates w.r.t.  $B'$ .
- For example, let  $B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

# Basis Change (1/4)

- $[I]_B^{B'}$ : a transformation matrix that maps coordinates w.r.t.  $B$  onto coordinates w.r.t.  $B'$ .
- For example, let  $B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ ,  $B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ 
  - $[I]_{B'}^B =$

# Basis Change (1/4)

- $[I]_B^{B'}$ : a transformation matrix that maps coordinates w.r.t.  $B$  onto coordinates w.r.t.  $B'$ .
- For example, let  $B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ ,  $B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ 
  - $[I]_{B'}^B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ .

# Basis Change (1/4)

- $[I]_B^{B'}$ : a transformation matrix that maps coordinates w.r.t.  $B$  onto coordinates w.r.t.  $B'$ .
- For example, let  $B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ ,  $B' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ 
  - $[I]_{B'}^B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ .
  - What about  $[I]_B^{B'}$ ?



## Basis Change (2/4)

### Basis Change

Consider a transformation matrix

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

w.r.t. the standard basis (canonical basis) in  $\mathbb{R}^2$ .

## Basis Change (2/4)

### Basis Change

Consider a transformation matrix

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

w.r.t. the standard basis (canonical basis) in  $\mathbb{R}^2$ . Define a new basis

$$B = \left( \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)$$

Then, what about the transformation matrix  $\tilde{\mathbf{A}}$  w.r.t.  $B$ ?

# Basis Change (3/4)

## Basis Change

Given

- a linear mapping  $\Phi : V \rightarrow W$ , ordered bases

$$B = (\mathbf{b}_1, \dots, \mathbf{b}_n), \tilde{B} = (\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_n) \text{ of } V$$

$$C = (\mathbf{c}_1, \dots, \mathbf{c}_m), \tilde{C} = (\tilde{\mathbf{c}}_1, \dots, \tilde{\mathbf{c}}_m) \text{ of } W.$$

- a transformation matrix  $\mathbf{A}_\Phi$  of  $\Phi$  w.r.t.  $B$  and  $C$ .

Then, the corresponding transformation matrix  $\tilde{\mathbf{A}}_\Phi$  w.r.t.  $\tilde{B}$  and  $\tilde{C}$  is

$$\tilde{\mathbf{A}}_\Phi = \mathbf{T}^{-1} \mathbf{A}_\Phi \mathbf{S}.$$

where  $\mathbf{S} = [I]_{\tilde{B}}^B \in \mathbb{R}^{n \times n}$  and  $\mathbf{T} = [I]_{\tilde{C}}^C \in \mathbb{R}^{m \times m}$ .

# Preview of the whole picture

$$\begin{array}{ccc}
 V & \xrightarrow{\Phi} & W \\
 \\ 
 B & \xrightarrow[\textcolor{red}{A}_\Phi]{\Phi_{CB}} & C \\
 \uparrow \textcolor{red}{S} [I]_{\tilde{B}}^B & & \uparrow \textcolor{red}{T} [I]_{\tilde{C}}^C \\
 \tilde{B} & \xrightarrow[\Phi_{\tilde{C}\tilde{B}}]{\tilde{A}_\Phi} & \tilde{C}
 \end{array}$$

# Proof (1/2)

$$\tilde{\mathbf{b}}_j = s_{1j}\mathbf{b}_1 + \cdots s_{nj}\mathbf{b}_n = \sum_{i=1}^n s_{ij}\mathbf{b}_i, \quad j = 1, \dots, n.$$

$$\tilde{\mathbf{c}}_k = t_{1k}\mathbf{c}_1 + \cdots t_{m,k}\mathbf{c}_m = \sum_{\ell=1}^m t_{\ell k}\mathbf{c}_\ell, \quad k = 1, \dots, m.$$

Let  $\mathbf{S} = ((s_{ij})) = [\mathbf{I}]_{\tilde{\mathbf{B}}}^{\mathbf{B}} \in \mathbb{R}^{n \times n}$  and  $\mathbf{T} = ((t_{\ell k})) = [\mathbf{I}]_{\tilde{\mathbf{C}}}^{\mathbf{C}} \in \mathbb{R}^{m \times m}$ .

# Proof (1/2)

$$\tilde{\mathbf{b}}_j = s_{1j}\mathbf{b}_1 + \cdots s_{nj}\mathbf{b}_n = \sum_{i=1}^n s_{ij}\mathbf{b}_i, \quad j = 1, \dots, n.$$

$$\tilde{\mathbf{c}}_k = t_{1k}\mathbf{c}_1 + \cdots t_{m,k}\mathbf{c}_m = \sum_{\ell=1}^m t_{\ell k}\mathbf{c}_\ell, \quad k = 1, \dots, m.$$

Let  $\mathbf{S} = ((s_{ij})) = [I]_{\tilde{\mathbf{B}}}^{\mathbf{B}} \in \mathbb{R}^{n \times n}$  and  $\mathbf{T} = ((t_{\ell k})) = [I]_{\tilde{\mathbf{C}}}^{\mathbf{C}} \in \mathbb{R}^{m \times m}$ .

- Applying the mapping  $\Phi$ , we get that for all  $j = 1, \dots, n$ ,

$$\Phi(\tilde{\mathbf{b}}_j) = \sum_{k=1}^m \underbrace{\tilde{a}_{kj}\tilde{\mathbf{c}}_k}_{\in W}$$

# Proof (1/2)

$$\tilde{\mathbf{b}}_j = s_{1j}\mathbf{b}_1 + \cdots s_{nj}\mathbf{b}_n = \sum_{i=1}^n s_{ij}\mathbf{b}_i, \quad j = 1, \dots, n.$$

$$\tilde{\mathbf{c}}_k = t_{1k}\mathbf{c}_1 + \cdots t_{m,k}\mathbf{c}_m = \sum_{\ell=1}^m t_{\ell k}\mathbf{c}_\ell, \quad k = 1, \dots, m.$$

Let  $\mathbf{S} = ((s_{ij})) = [I]_{\tilde{\mathbf{B}}}^{\mathbf{B}} \in \mathbb{R}^{n \times n}$  and  $\mathbf{T} = ((t_{\ell k})) = [I]_{\tilde{\mathbf{C}}}^{\mathbf{C}} \in \mathbb{R}^{m \times m}$ .

- Applying the mapping  $\Phi$ , we get that for all  $j = 1, \dots, n$ ,

$$\Phi(\tilde{\mathbf{b}}_j) = \sum_{k=1}^m \underbrace{\tilde{a}_{kj}}_{\in W} \tilde{\mathbf{c}}_k = \sum_{k=1}^m \tilde{a}_{kj} \sum_{\ell=1}^m t_{\ell k} \mathbf{c}_\ell$$

# Proof (1/2)

$$\tilde{\mathbf{b}}_j = s_{1j}\mathbf{b}_1 + \cdots s_{nj}\mathbf{b}_n = \sum_{i=1}^n s_{ij}\mathbf{b}_i, \quad j = 1, \dots, n.$$

$$\tilde{\mathbf{c}}_k = t_{1k}\mathbf{c}_1 + \cdots t_{m,k}\mathbf{c}_m = \sum_{\ell=1}^m t_{\ell k}\mathbf{c}_\ell, \quad k = 1, \dots, m.$$

Let  $\mathbf{S} = ((s_{ij})) = [I]_{\tilde{\mathbf{B}}}^{\mathbf{B}} \in \mathbb{R}^{n \times n}$  and  $\mathbf{T} = ((t_{\ell k})) = [I]_{\tilde{\mathbf{C}}}^{\mathbf{C}} \in \mathbb{R}^{m \times m}$ .

- Applying the mapping  $\Phi$ , we get that for all  $j = 1, \dots, n$ ,

$$\Phi(\tilde{\mathbf{b}}_j) = \sum_{k=1}^m \underbrace{\tilde{a}_{kj}\tilde{\mathbf{c}}_k}_{\in W} = \sum_{k=1}^m \tilde{a}_{kj} \sum_{\ell=1}^m t_{\ell k}\mathbf{c}_\ell = \sum_{\ell=1}^m \left( \sum_{k=1}^m t_{\ell k}\tilde{a}_{kj} \right) \mathbf{c}_\ell.$$



## Proof (1/2)

$$\tilde{\mathbf{b}}_j = s_{1j}\mathbf{b}_1 + \cdots s_{nj}\mathbf{b}_n = \sum_{i=1}^n s_{ij}\mathbf{b}_i, \quad j = 1, \dots, n.$$

$$\tilde{\mathbf{c}}_k = t_{1k}\mathbf{c}_1 + \cdots t_{m,k}\mathbf{c}_m = \sum_{\ell=1}^m t_{\ell k}\mathbf{c}_\ell, \quad k = 1, \dots, m.$$

Let  $\mathbf{S} = ((s_{ij})) = [I]_{\tilde{\mathbf{B}}}^{\mathbf{B}} \in \mathbb{R}^{n \times n}$  and  $\mathbf{T} = ((t_{\ell k})) = [I]_{\tilde{\mathbf{C}}}^{\mathbf{C}} \in \mathbb{R}^{m \times m}$ .

- Applying the mapping  $\Phi$ , we get that for all  $j = 1, \dots, n$ ,

$$\Phi(\tilde{\mathbf{b}}_j) = \sum_{k=1}^m \underbrace{\tilde{a}_{kj}}_{\in W} \tilde{\mathbf{c}}_k = \sum_{k=1}^m \tilde{a}_{kj} \sum_{\ell=1}^m t_{\ell k} \mathbf{c}_\ell = \sum_{\ell=1}^m \left( \sum_{k=1}^m t_{\ell k} \tilde{a}_{kj} \right) \mathbf{c}_\ell.$$

- Alternatively,

$$\begin{aligned} \Phi(\tilde{\mathbf{b}}_j) &= \Phi \left( \sum_{i=1}^n s_{ij} \mathbf{b}_i \right) = \sum_{i=1}^n s_{ij} \Phi(\mathbf{b}_i) = \sum_{i=1}^n s_{ij} \sum_{\ell=1}^m a_{\ell i} \mathbf{c}_\ell \\ &= \sum_{\ell=1}^m \left( \sum_{i=1}^n a_{\ell i} s_{ij} \right) \mathbf{c}_\ell \end{aligned}$$

# Proof (2/2)

Hence,

$$\sum_{k=1}^m t_{\ell k} \tilde{a}_{kj} = \sum_{i=1}^n a_{\ell i} s_{ij}, \text{ for each } j$$

and it means that

## Proof (2/2)

Hence,

$$\sum_{k=1}^m t_{\ell k} \tilde{a}_{kj} = \sum_{i=1}^n a_{\ell i} s_{ij}, \text{ for each } j$$

and it means that

$$\mathbf{T} \tilde{\mathbf{A}}_{\Phi} = \mathbf{A}_{\Phi} \mathbf{S} \in \mathbb{R}^{m \times n},$$

such that

$$\tilde{\mathbf{A}}_{\Phi} = \mathbf{T}^{-1} \mathbf{A}_{\Phi} \mathbf{S}.$$

# Basis Change (4/4)

- The theorem tells us that

# Basis Change (4/4)

- The theorem tells us that

With

- a basis change in  $V$  (i.e.,  $B \rightarrow \tilde{B}$ ) and
- a basis change in  $W$  (i.e.,  $C \rightarrow \tilde{C}$ ),

the transformation matrix  $\mathbf{A}_\Phi$  of a linear mapping  $\Phi : V \rightarrow W$  is replaced by an equivalent matrix  $\tilde{\mathbf{A}}_\Phi$  with

$$\tilde{\mathbf{A}}_\Phi = \mathbf{T}^{-1} \mathbf{A}_\Phi \mathbf{S}.$$

$$\begin{array}{ccc}
 V & \xrightarrow{\Phi} & W \\
 B & \xrightarrow[\mathbf{A}_\Phi]{\Phi_{CB}} & C \\
 \uparrow [I]_{\tilde{B}}^B \mathbf{S} & & \uparrow \mathbf{T} [I]_{\tilde{C}}^C \\
 \tilde{B} & \xrightarrow[\Phi_{\tilde{C}\tilde{B}}]{\tilde{\mathbf{A}}_\Phi} & \tilde{C}
 \end{array}$$

$$\begin{array}{ccc}
 V & \xrightarrow{\Phi} & W \\
 \\ 
 B & \xrightarrow[\mathbf{A}_\Phi]{\Phi_{CB}} & C \\
 \uparrow [I]_{\tilde{B}}^B \quad \mathbf{S} & & \downarrow T^{-1} \quad [I]_{\tilde{C}}^{\tilde{C}} = [I]_C^{-1} \\
 \tilde{B} & \xrightarrow[\Phi_{\tilde{C}\tilde{B}}]{\tilde{\mathbf{A}}_\Phi} & \tilde{C}
 \end{array}$$

## Example

Consider a linear mapping  $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}^4$  with transformation matrix

$$\mathbf{A}_\Phi = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}$$

w.r.t. the standard bases

$$B = \left( \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right), \quad C = \left( \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right).$$

We seek the transformation matrix  $\tilde{\mathbf{A}}_\Phi$  of  $\Phi$  w.r.t. the new bases

$$\tilde{B} = \left( \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right), \quad \tilde{C} = \left( \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right).$$



$$S =$$

$$T =$$

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then,

$$\tilde{\mathbf{A}}_{\Phi} = \mathbf{T}^{-1} \mathbf{A}_{\Phi} \mathbf{S} = \dots$$

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then,

$$\tilde{\mathbf{A}}_{\Phi} = \mathbf{T}^{-1} \mathbf{A}_{\Phi} \mathbf{S} = \cdots = \begin{bmatrix} -4 & -4 & -2 \\ 6 & 0 & 0 \\ 4 & 8 & 4 \\ 1 & 6 & 3 \end{bmatrix}.$$

# Image and Kernel

## Image & Kernel

For  $\Phi : V \rightarrow W$ , we define

$$\ker(\Phi) := \Phi^{-1}(\mathbf{0}_W) = \{\mathbf{v} \in V \mid \Phi(\mathbf{v}) = \mathbf{0}_W\}$$

and

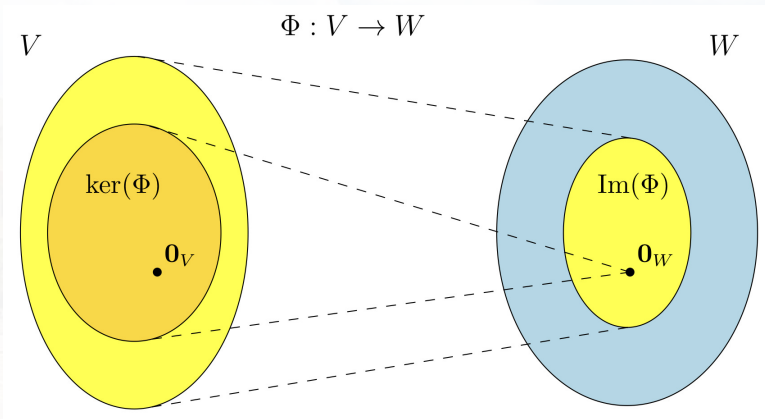
$$\text{Image}(\Phi) := \Phi(V) = \{\mathbf{w} \in W \mid \exists \mathbf{v} \in V \text{ s.t. } \Phi(\mathbf{v}) = \mathbf{w}\}.$$

- $V$ : domain of  $\Phi$
- $W$ : codomain of  $\Phi$

## Remark

For vector spaces  $V$  and  $W$  and a linear mapping  $\Phi : V \rightarrow W$ :

- $\Phi(\mathbf{0}_V) = \mathbf{0}_W$  so  $\mathbf{0} \in \ker(\Phi)$ .
- $\text{Image}(\Phi) \subseteq W$  is a subspace of  $W$
- $\ker(\Phi) \subseteq V$  is a subspace of  $V$ .
- $\Phi$  is injective (i.e., one-to-one) if and only if  $\ker(\Phi) = \{\mathbf{0}\}$ .
- $\text{Image}(\Phi) = \{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n\} = \{\sum_{i=1}^n x_i \mathbf{a}_i \mid x_1, \dots, x_n \in \mathbb{R}\} = \text{span}(\mathbf{a}_1, \dots, \mathbf{a}_n) \subseteq \mathbb{R}^m$ .
- $\text{rank}(\Phi) = \dim(\text{Image}(\Phi))$ .
- ★  $\dim(\ker(\Phi)) + \dim(\text{Image}(\Phi)) = \dim(V)$ .
  - $\text{null}(\mathbf{A}) + \text{rank}(\mathbf{A}) = \text{number of columns of } \mathbf{A}$ .
- If  $\dim(V) = \dim(W)$ , then  $\Phi$  is injective, surjective and bijective ( $\because \text{Image}(\Phi) \subseteq W$ ).



## Example

Consider the mapping  $\Phi : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ ,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_1 + x_4 \end{bmatrix}$$

## Example

Consider the mapping  $\Phi : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ ,

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} &\mapsto \begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_1 + x_4 \end{bmatrix} \\ &= x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

Image( $\Phi$ ) =



## Example

Consider the mapping  $\Phi : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ ,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_1 + x_4 \end{bmatrix}$$

$$= x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\text{Image}(\Phi) = \text{span} \left( \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \right)$$

## Example (contd.)

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

### Example (contd.)

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \cdots \longrightarrow$$

## Example (contd.)

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \cdots \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

## Example (contd.)

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \cdots \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Thus,

$$\ker(\Phi) =$$

## Example (contd.)

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \cdots \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Thus,

$$\ker(\Phi) = \text{span} \left\{ \begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ \frac{1}{2} \\ 0 \\ 1 \end{bmatrix} \right\}.$$

# Outline

- 1 Why linear algebra?
- 2 Vector Space
- 3 Basis & Dimension & Rank
- 4 Linear Mappings
- 5 Affine Spaces**

# Affine Spaces

- Spaces that are offset from the origin.
- They are NO LONGER vector (sub)spaces.



# Affine Spaces

- Spaces that are offset from the origin.
- They are NO LONGER vector (sub)spaces.

## Affine Subspace

Let  $V$  be a vector space,  $\mathbf{x}_0 \in V$ , and  $U \subseteq V$  be a subspace. Then,

$$\begin{aligned} L &= \mathbf{x}_0 + U := \{\mathbf{x}_0 + \mathbf{u} \mid \mathbf{u} \in U\} \\ &= \{\mathbf{v} \in V \mid \exists \mathbf{u} \in U : \mathbf{v} = \mathbf{x}_0 + \mathbf{u}\} \subseteq V \end{aligned}$$

is called **affine subspace** (or linear manifold) of  $V$ .

- $U$ : **direction space**.
- $\mathbf{x}_0$ : **support point**.

## Remark

- An affine subspace excludes  $\mathbf{0}$  if  $\mathbf{x}_0 \notin U$ .
- Examples: points, lines, and planes in  $\mathbb{R}^3$  which do not go through the origin.

## Remark

- An affine subspace excludes  $\mathbf{0}$  if  $\mathbf{x}_0 \notin U$ .
- Examples: points, lines, and planes in  $\mathbb{R}^3$  which do not go through the origin.
- One-dimensional affine subspaces:

$$\mathbf{y} = \mathbf{x}_0 + \lambda \mathbf{b}_1$$

for  $\lambda \in \mathbb{R}$  and  $U = \text{span}(\mathbf{b}_1)$  is a one-dimensional subspace of  $\mathbb{R}^n$ .

- Two-dimensional affine subspaces:

$$\mathbf{y} = \mathbf{x}_0 + \lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2$$

for  $\lambda_1, \lambda_2 \in \mathbb{R}$  and  $U = \text{span}(\{\mathbf{b}_1, \mathbf{b}_2\})$  is a two-dimensional subspace of  $\mathbb{R}^n$ .

•  
⋮

# Affine Mappings

## Affine Mappings

Given two vector spaces  $V, W$ , a linear mapping  $\Phi : V \rightarrow W$ , and  $\mathbf{a} \in W$ , the mapping  $\phi : V \rightarrow W$  with

$$\phi(\mathbf{x}) = \mathbf{a} + \Phi(\mathbf{x})$$

is called an **affine mapping** from  $V$  to  $W$ . The vector  $\mathbf{a}$  is called the **translation vector** of  $\phi$ .

# Discussions