Linear Algebra

Liang Zheng
Australian National University
liang.zheng@anu.edu.au

2.4.1 Groups

- Consider a set $\mathcal{G}$ and an operation $\otimes: \mathcal{G} \otimes \mathcal{G} \rightarrow \mathcal{G}$ defined on $\mathcal{G}$. Then $G := (\mathcal{G}, \otimes)$ is called a group if the following holds
  - Closure of $\mathcal{G}$ under $\otimes$: $\forall x, y \in \mathcal{G}: x \otimes y \in \mathcal{G}$
  - Associativity: $\forall x, y, z \in \mathcal{G}: (x \otimes y) \otimes z = x \otimes (y \otimes z)$
  - Neutral element: $\exists e \in \mathcal{G}\ \forall x \in \mathcal{G}: x \otimes e = x$ and $e \otimes x = x$
  - Inverse element: $\forall x \in \mathcal{G}\ \exists y \in \mathcal{G}: x \otimes y = e$ and $y \otimes x = e$. We often write x^{-1} to denote the inverse element of x
- Additionally, If $\forall x, y \in \mathcal{G}, x \otimes y = y \otimes x$ (commutative), then $G := (\mathcal{G}, \otimes)$ is an Abelian group.
- Examples
- $(\mathbb{Z}, +)$ is a group and an Abelian group
  - …,-5, -4, -3, -2, -1, 0, 1, 2, 3,4, …

Closure: √
Associativity: $(x + y) + z = x + (y + z)$ √
Neutral element: 0 √
Inverse element: $\forall x \in \mathbb{Z}, y = -x \in \mathbb{Z}$ √

- $(\mathbb{Z}, -)$ is not a group: it does not satisfy associativity, has no neutral element or inverse element

~~Associativity~~: $(x - y) - z \neq x - (y - z)$

- Examples
- $(\mathbb{R}^{m\times n}, +)$, the set of $m\times n$-matrices is Abelian (component-wise addition).
  - Closure: addition of any two matrices in $\mathbb{R}^{m\times n}$ is a matrix in $\mathbb{R}^{m\times n}$
  - Associativity: $\forall \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{m\times n}$, $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C} = \boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$
  - Neutral element: $\boldsymbol{0}$
  - Inverse element: $\forall \boldsymbol{A} \in \mathbb{R}^{m\times n}$, there exists its inverse element $-\boldsymbol{A}$
  - Commutative: $\forall \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m\times n}$, $\boldsymbol{A}+\boldsymbol{B} = \boldsymbol{B}+\boldsymbol{A}$

2.4.2 Vector spaces

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

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

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

- where
  - $(\mathcal{V}, +)$ is an Abelian group
  - Distributivity:

$$\forall \lambda \in \mathbb{R}, \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}: \qquad \lambda \cdot (\boldsymbol{x} + \boldsymbol{y}) = \lambda \cdot \boldsymbol{x} + \lambda \cdot \boldsymbol{y}$$
$$\forall \lambda, \varphi \in \mathbb{R}, \boldsymbol{x} \in \mathcal{V}: \qquad (\lambda + \varphi) \cdot \boldsymbol{x} = \lambda \cdot \boldsymbol{x} + \varphi \cdot \boldsymbol{x}$$

  - Associativity (outer operation $\cdot$):

$$\forall \lambda, \varphi \in \mathbb{R}, \boldsymbol{x} \in \mathcal{V}: \qquad \lambda \cdot (\varphi \cdot \boldsymbol{x}) = (\lambda\varphi) \cdot \boldsymbol{x}$$

  - Neutral element (w.r.t to outer operation $\cdot$):

$$\forall \boldsymbol{x} \in \mathcal{V}: \qquad 1 \cdot \boldsymbol{x} = \boldsymbol{x}$$

2.4.2 Vector spaces

- Elements $\boldsymbol{x} \in \mathcal{V}$ are called vectors
- The neutral element of $(\mathcal{V}, +)$ is the zero vector $\mathbf{0} = [0, \cdots, 0]^{\mathrm{T}}$
- $+$ is called vector addition
- Elements $\lambda \in \mathbb{R}$ are called scalars
- Outer operation $\cdot$ is a multiplication by scalars

- Example
- $\mathcal{V} = \mathbb{R}^n, n \in \mathbb{N}$ is a vector space. Its operations are defined as
  - Addition: $\boldsymbol{x} + \boldsymbol{y} = (x_1, \cdots, x_n)^{\mathrm{T}} + (y_1, \cdots, y_n)^{\mathrm{T}} = (x_1 + y_1, \cdots, x_n + y_n)^{\mathrm{T}}$, for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$
  - Multiplication by scalars: $\lambda \boldsymbol{x} = \lambda(x_1, \cdots, x_n)^{\mathrm{T}} = (\lambda x_1, \cdots, \lambda x_n)^{\mathrm{T}}$, for $\boldsymbol{x} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$
- Custom
- We usually write $\boldsymbol{x} \in \mathbb{R}^n$ in a column vector $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

Vector spaces - example

- $\mathcal{V} = \mathbb{R}^n, n \in \mathbb{N}$ is a vector space. Its operations are defined as

  - Addition: for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

  $\boldsymbol{x} + \boldsymbol{y} = (x_1, \cdots, x_n)^{\mathrm{T}} + (y_1, \cdots, y_n)^{\mathrm{T}} = (x_1 + y_1, \cdots, x_n + y_n)^{\mathrm{T}}$, for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

  - Multiplication by scalars: for $\boldsymbol{x} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$

  $\lambda \boldsymbol{x} = \lambda (x_1, \cdots, x_n)^{\mathrm{T}} = (\lambda x_1, \cdots, \lambda x_n)^{\mathrm{T}}$

- We usually write $\boldsymbol{x} \in \mathbb{R}^n$ in a column vector $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

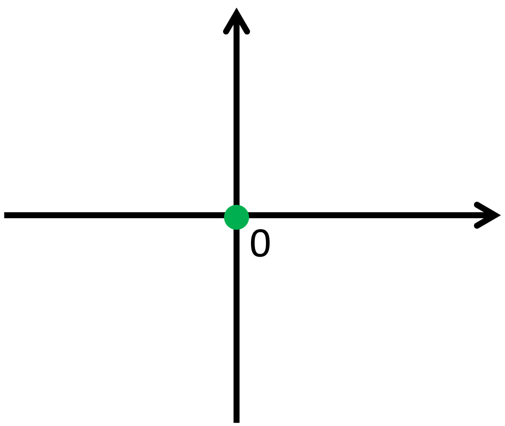
2.4.3 Vector Subspaces

- Sets contained in the original vector space
- "closed"
- When we perform vector space operations on elements within this subspace, we will never leave it

- $U = (\mathcal{U}, +, \cdot)$ is called vector subspace of $V = (\mathcal{V}, +, \cdot)$, if
- $\mathcal{U} \subseteq \mathcal{V}$,
- $\mathcal{U} \neq \emptyset$, in particular $\mathbf{0} \in \mathcal{U}$
- Closure of U
  - $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{U}, \boldsymbol{x} + \boldsymbol{y} \in \mathcal{U}$
  - $\forall \boldsymbol{x} \in \mathcal{U}, \lambda \in \mathbb{R}, \lambda \boldsymbol{x} \in \mathcal{U}$

2.4.3 Vector Subspaces

- Examples
- For every vector space V, the trivial subspaces are V itself and $\{\mathbf{0}\}$
- Is it a subspace of $\mathbb{R}^2$?



Is it a subset of $\mathbb{R}^2$? Yes

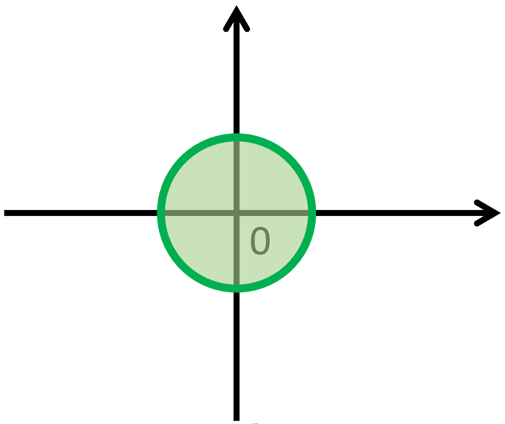
Does it satisfy $\mathcal{U} \neq \emptyset$, in particular $\mathbf{0} \in \mathcal{U}$ Yes

Does it satisfy closure? Yes

$$\boxed{\begin{array}{l} x + y \in \{0\} \\ \lambda x \in \{0\} \end{array}}$$

2.4.3 Vector Subspaces

- Examples
- Is it a subspace of $\mathbb{R}^2$?



Is it a subset of $\mathbb{R}^2$? Yes

Does it satisfy $\mathcal{U} \neq \emptyset$, in particular $\mathbf{0} \in \mathcal{U}$ Yes

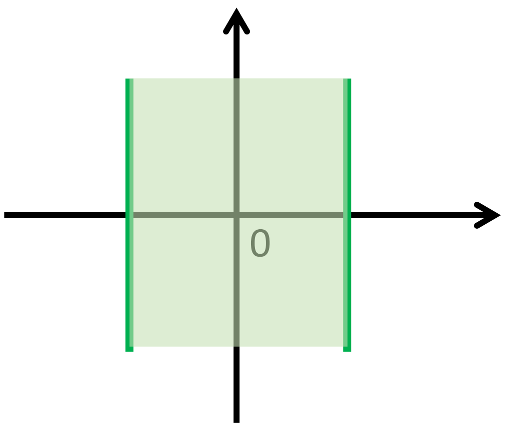
Does it satisfy closure? No

- Type equation here.

$$(0.8, 0) + (0.9, 0) = (1.7,0) \notin \mathcal{U}$$

2.4.3 Vector Subspaces

- Examples
- Is it a subspace of $\mathbb{R}^2$?



Is it a subset of $\mathbb{R}^2$? Yes

Does it satisfy $\mathcal{U} \neq \emptyset$, in particular $\mathbf{0} \in \mathcal{U}$ Yes

Does it satisfy closure? No

2.4.3 Vector Subspaces

- Examples

- The solution set of a homogeneous system of linear equations $\boldsymbol{Ax} = \boldsymbol{0}$ with n unknowns $\boldsymbol{x} = [x_1, \cdots, x_n]^{\mathrm{T}}$. Is it a subspace of $\mathbb{R}^n$?

Is it a subset of $\mathbb{R}^n$? Yes

Does it satisfy $\mathcal{U} \neq \emptyset$, in particular $\boldsymbol{0} \in \mathcal{U}$ Yes

Does it satisfy closure? Yes

$\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{U}$, we have $\boldsymbol{Ax} = \boldsymbol{0}, \boldsymbol{Ay} = \boldsymbol{0}$

1) We investigate whether $\boldsymbol{x} + \boldsymbol{y} \in \mathcal{U}$.

Because $\boldsymbol{A}(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{Ax} + \boldsymbol{Ay} = \boldsymbol{0}$,

We know $\boldsymbol{x} + \boldsymbol{y}$ is a solution, thus belonging to$\mathcal{U}$

2) We investigate whether $\lambda\boldsymbol{x} \in \mathcal{U}$.

Because $\boldsymbol{A}(\lambda\boldsymbol{x}) = \lambda(\boldsymbol{Ax}) = \boldsymbol{0}$,

We know$\lambda\boldsymbol{x}$ is a solution, thus belonging to$\mathcal{U}$

2.4.3 Vector Subspaces

- Examples

- The solution set of an inhomogeneous system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}, \boldsymbol{b} \neq \boldsymbol{0}$. Is it a subspace of $\mathbb{R}^n$?

Is it a subset of $\mathbb{R}^2$? Yes

Does it satisfy $\mathcal{U} \neq \emptyset$, in particular $\boldsymbol{0} \in \mathcal{U}$ No

Does it satisfy closure? No

Linear combination

- Consider a vector space V and k vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k \in V$. For $\lambda_1, \cdots, \lambda_k \in \mathbb{R}$, $\boldsymbol{v} \in V$ is called a linear combination of vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$, if

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

2.5 Linear Independence

- Consider a system of linear functions $\lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k = \mathbf{0}$
- If there is a non-trivial solution, $\lambda_1, \ldots, \lambda_k$, with at least one $\lambda_i \neq 0$, the vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ are linearly dependent

- If only the trivial solution exists, i.e., $\lambda_1 = \cdots = \lambda_k = 0$, then vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ are linearly independent

- Intuitively, a set of linearly independent vectors consists of vectors that have no redundancy, *i.e.,* if we remove any of those vectors from the set, we will lose something.

How to determine linear (in)dependence

- Write all vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ as columns of a matrix $\boldsymbol{A}$
- Perform Gaussian elimination until the matrix is in row echelon form
- The pivot columns correspond to independent vectors

$$\begin{array}{c} \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \\ \begin{matrix} x_1 & x_2 & x_3 \end{matrix} \end{array}$$

$$\boxed{x_2 = 2x_1}$$

- All column vectors are linearly independent if and only if all columns are pivot columns.
- If there is at least one non-pivot column, the columns are linearly dependent.

Determine linear (in)dependence

- Consider three vectors in $\mathbb{R}^3$

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \boldsymbol{x}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \boldsymbol{x}_3 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{\text{R1+R2->R2}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 4 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{\text{Swap R2 and R3}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{bmatrix}$$

$$\xrightarrow{\text{R3-2R2->R3}} \begin{array}{c} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \\ \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \end{array}$$

$$x_3 = x_1 + 2x_2$$

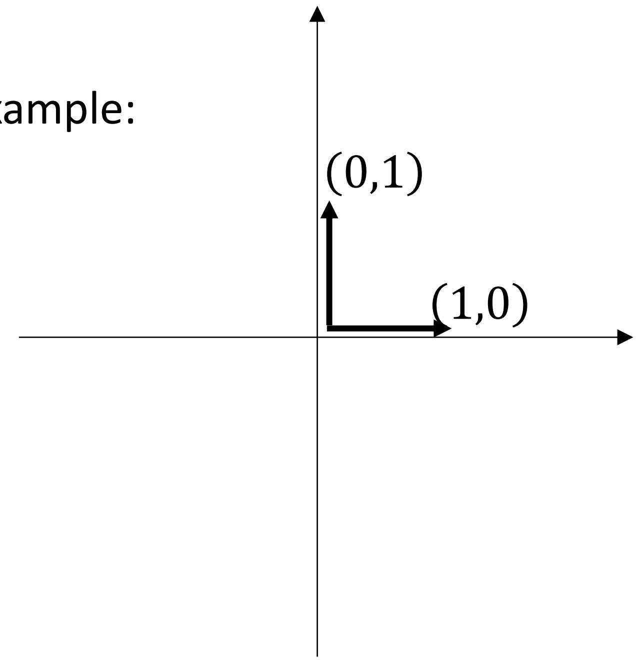
The Basis of a vector space

- A set of vectors $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\}$ is said to form a basis for a vector space if

(1) The vectors $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\}$ span the vector space: every vector in this space can be represented by a linear combination of $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\}$

(2) The vectors $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\}$ are linearly independent.

Example:



- Example
- In $\mathbb{R}^3$, the canonical/standard basis is

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

- Different bases in $\mathbb{R}^3$ are $\mathcal{B}_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

First, this REF has three pivots, so the three bases are linearly independent.

Second, do the three bases span $\mathbb{R}^3$?

Specifically, $\forall\ [a, b, c]^T \in \mathbb{R}^3$, we examine whether it can be obtained by a linear combination by the three bases.

$$\lambda_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

We can obtain the solution

$$\begin{cases} \lambda_3 = c \\ \lambda_2 = b - c \\ \lambda_1 = a - b \end{cases}$$

- Another different basis in $\mathbb{R}^3$ is

$$\mathcal{B}_2 = \left\{ \begin{bmatrix} 0.5 \\ 0.8 \\ 0.4 \end{bmatrix}, \begin{bmatrix} 1.8 \\ 0.3 \\ 0.3 \end{bmatrix}, \begin{bmatrix} -2.2 \\ -1.3 \\ 3.5 \end{bmatrix} \right\}$$

- Another example

$$\mathcal{A} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ -4 \end{bmatrix} \right\}$$

is linearly independent, but not a basis of $\mathbb{R}^4$: For instance, the $[1,0,0,0]^\top$ cannot be obtained by a linear combination of elements in $\mathcal{A}$.

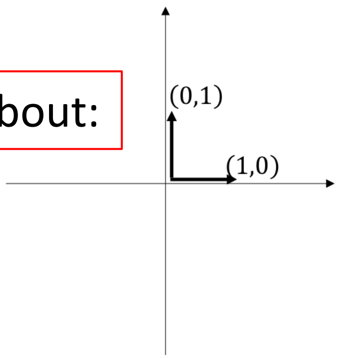
So, a couple of things about basis

- Let $V = (\mathcal{V}, +, \cdot)$ be a vector space and $\mathcal{B} \subseteq \mathcal{V}, \mathcal{B} \neq \emptyset$ be a basis of V.
- $\mathcal{B}$ is a maximal linearly independent set of vectors in V, i.e., adding any other vector to this set will make it linearly dependent.

- Every vector $\boldsymbol{x} \in V$ is a linear combination of vectors from $\mathcal{B}$, and every linear combination is unique, i.e., with

$$x = \sum_{i=1}^{k} \lambda_i \boldsymbol{b}_i = \sum_{i=1}^{k} \psi_i \boldsymbol{b}_i$$

Think about:



and $\lambda_i, \psi_i \in \mathbb{R}, \boldsymbol{b}_i \in B$ it follows that $\boldsymbol{\lambda}_i = \psi_i, i = 1, \cdots, k.$

- Every vector space V possesses a basis $\mathcal{B}$.
- There can be many bases of a vector space.
- All bases possess the same number of elements, called the basis vectors

$\mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$, then $\dim(\mathcal{B}) = 3$

- Dimension of (V): number of basis vectors of V. We write $\dim(V)$
- If $U \subseteq V$ is a subspace of V , then $\dim(U) \leq \dim(V)$
- $\dim(U) = \dim(V)$ if and only if $U = V$

Determining a Basis

- Write the spanning vectors as columns of a matrix $\boldsymbol{A}$
- Determine the row-echelon form of $\boldsymbol{A}$.
- The spanning vectors associated with the pivot columns are a basis of U.

- Example
- For a vector subspace $U \subseteq \mathbb{R}^5$, spanned by the vectors

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 2 \\ -2 \end{bmatrix}, \quad \boldsymbol{x}_3 = \begin{bmatrix} 3 \\ -4 \\ 3 \\ 5 \\ -3 \end{bmatrix}, \quad \boldsymbol{x}_4 = \begin{bmatrix} -1 \\ 8 \\ -5 \\ -6 \\ 1 \end{bmatrix} \in \mathbb{R}^5$$

Determining a Basis - Example

- Which vectors of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_4$ are a basis for U?
- Check whether $\boldsymbol{x}_1, \dots, \boldsymbol{x}_4$ are linearly independent.

$$\boxed{\sum_{i=1}^{4} \lambda_i \boldsymbol{x}_i = \boldsymbol{0}}$$

- A homogeneous system of equations with matrix

$$[\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4] = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & -4 & 8 \\ -1 & 1 & 3 & -5 \\ -1 & 2 & 5 & -6 \\ -1 & -2 & -3 & 1 \end{bmatrix}$$

- Through Gaussian Elimination, we obtain the row-echelon form

$$\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & -4 & 8 \\ -1 & 1 & 3 & -5 \\ -1 & 2 & 5 & -6 \\ -1 & -2 & -3 & 1 \end{bmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{array}{c} \begin{array}{cccc} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 & \boldsymbol{x}_4 \end{array} \\ \begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

$\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_4$ are linearly independent. Therefore, $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_4\}$ is a basis of U

2.6.2 Rank

- The number of linearly independent columns of a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is called the rank of $\boldsymbol{A}$, denoted by $\mathrm{rk}(\boldsymbol{A})$

- $\mathrm{rk}(\boldsymbol{A})$ also equals the number of linearly independent rows

- Rank gives us an idea of how much information a matrix contains

Important properties

- $\mathrm{rk}(\boldsymbol{A}) = \mathrm{rk}(\mathbf{A}^{\mathrm{T}})$
- Columns and rows of $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ can both span subspaces of the same dimension $\mathrm{rk}(\boldsymbol{A})$
- The basis of the subspace spanned by columns (rows) can be found by Gaussian elimination to $\boldsymbol{A}$ $(\mathbf{A}^{\mathrm{T}})$ to identify the pivot columns.
- For all $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ it holds that $\boldsymbol{A}$ is regular (invertible) if and only if $\mathrm{rk}(\boldsymbol{A}) = n$.

$$\begin{bmatrix} * & & & & \\ & * & & & \\ & & * & & \\ & & & \ddots & \\ & & & & * \end{bmatrix}_{n\times n}$$

- Example

- We use Gaussian elimination to determine the rank

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix} \rightsquigarrow \cdots \rightsquigarrow \begin{array}{c} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \\ \begin{matrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 \end{matrix} \end{array}$$

- 2 pivot columns. So $\mathrm{rk}(\boldsymbol{A}) = 2$

More properties

- For all $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and all $\boldsymbol{b} \in \mathbb{R}^{m}$ it holds that the linear equation system $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ can be solved if and only if $\mathrm{rk}(\boldsymbol{A}) = \mathrm{rk}(\boldsymbol{A}|\boldsymbol{b})$, where $\boldsymbol{A}|\boldsymbol{b}$ denotes the augmented matrix

- For $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ the subspace of solutions for $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{0}$ possesses dimension $n - \mathrm{rk}(\boldsymbol{A})$.

Let's look at a simpler case where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\mathrm{rk}(\boldsymbol{A}) = n$.
In this scenario, the dimension of the solution space is $n - \mathrm{rk}(\boldsymbol{A}) = \boldsymbol{0}$.
The only solution is $\boldsymbol{x} = \boldsymbol{0}$.

More properties

- A matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ has full rank if its rank equals the largest possible rank for a matrix of the same dimensions.

- The rank of a full-rank matrix is the lesser of the number of rows and columns, i.e., $\mathrm{rk}(\boldsymbol{A}) = \min(m, n)$.

For example, for $\boldsymbol{A} \in \mathbb{R}^{5 \times 3}$, $\mathrm{rk}(\boldsymbol{A})$ does not exceed 3.

- A matrix is said to be rank deficient if it does not have full rank.

2.7 Linear Mappings

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

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

- It implies the following

$$\Phi(\boldsymbol{x} + \boldsymbol{y}) = \Phi(\boldsymbol{x}) + \Phi(\boldsymbol{y}) \qquad \Phi(\lambda \boldsymbol{x}) = \lambda\, \Phi(\boldsymbol{x})$$

Example

- The mapping $\Phi\colon\ \mathbb{R}^2 \rightarrow \mathbb{C}, \Phi(\boldsymbol{x}) = x_1 + ix_2$, is a linear mapping:

$$\Phi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = (x_1 + y_1) + i(x_2 + y_2) = x_1 + ix_2 + y_1 + iy_2$$

$$= \Phi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) + \Phi\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right)$$

$$\Phi\left(\lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \lambda x_1 + \lambda i x_2 = \lambda(x_1 + ix_2) = \lambda\Phi\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)$$

2.7 Linear Mappings

- For linear mappings $\Phi: V \to W$ and $\Psi: W \to X$, the mapping $\Phi \circ \Psi: V \to X$ is also linear

- If $\Phi: V \to W$ and $\Psi: V \to W$ are both linear mappings, then $\Phi + \Psi$ and $\lambda\Phi, \lambda \in \mathbb{R}$ are also linear.

Coordinates of a vector

- Consider a vector space V and an ordered basis $B = (\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n)$ of V. For any $\boldsymbol{x} \in V$ we obtain a unique representation

$$\boldsymbol{x} = a_1\boldsymbol{b}_1 + \ldots + a_n\boldsymbol{b}_n$$

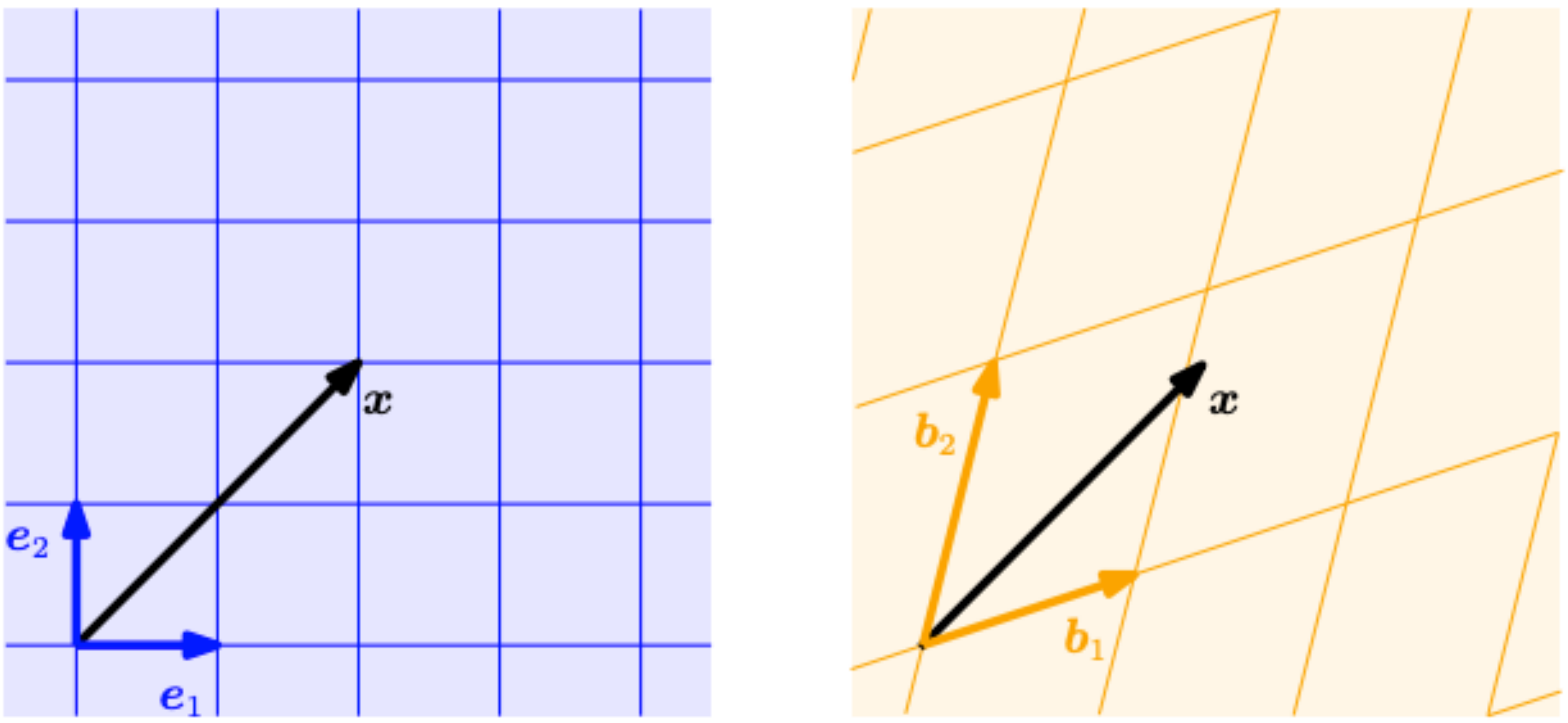
of $\boldsymbol{x}$ with respect to B. Then $\alpha_1, \cdots, \alpha_n$ are the coordinates of $\boldsymbol{x}$ with respect to B, and the vector

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{R}^n$$

is the coordinate vector/coordinate representation of $\boldsymbol{x}$ with respect to the ordered basis B.

Coordinates of a vector

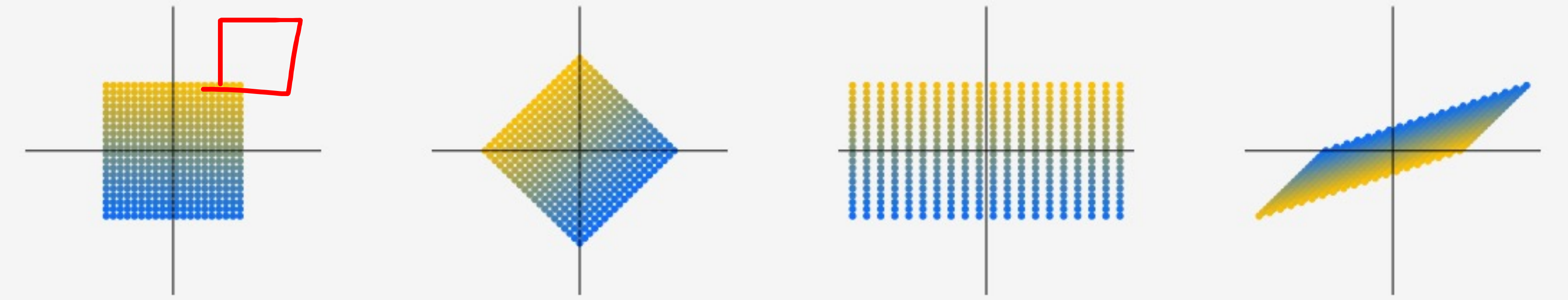
- [Left] A Cartesian coordinate system in two dimensions, which is spanned by the canonical basis vectors $\boldsymbol{e}_1, \boldsymbol{e}_2$.



- The same vector $\boldsymbol{x}$ may have different coordinates under different basis.

2.7.1 Matrix Representation of Linear Mappings

- Example - Linear Transformations of Vectors



(a) Original data. (b) Rotation by 45°. (c) Stretch along the horizontal axis. (d) General linear mapping.

- The following three linear transformations are used

$$A_1 = \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix} \quad A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad A_3 = \frac{1}{2}\begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix}$$

- Consider vector spaces V, W with corresponding bases $B = (\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n)$ and $C = (\boldsymbol{c}_1, \cdots, \boldsymbol{c}_m)$. We consider a linear mapping $\Phi: V \to W$. For $j \in \{1, \cdots, n\}$,

$$\Phi(\boldsymbol{b}_j) = \alpha_{1j}\boldsymbol{c}_1 + \cdots + \alpha_{mj}\boldsymbol{c}_m = \sum_{i=1}^{m} \alpha_{ij}\boldsymbol{c}_i$$

is the unique representation of $\Phi(\boldsymbol{b}_j)$ with respect to C. Then, we call the $m \times n$-matrix $\boldsymbol{A}_\Phi$ the transformation matrix of Φ, whose elements are given by

$$A_\Phi(i, j) = \alpha_{ij}$$

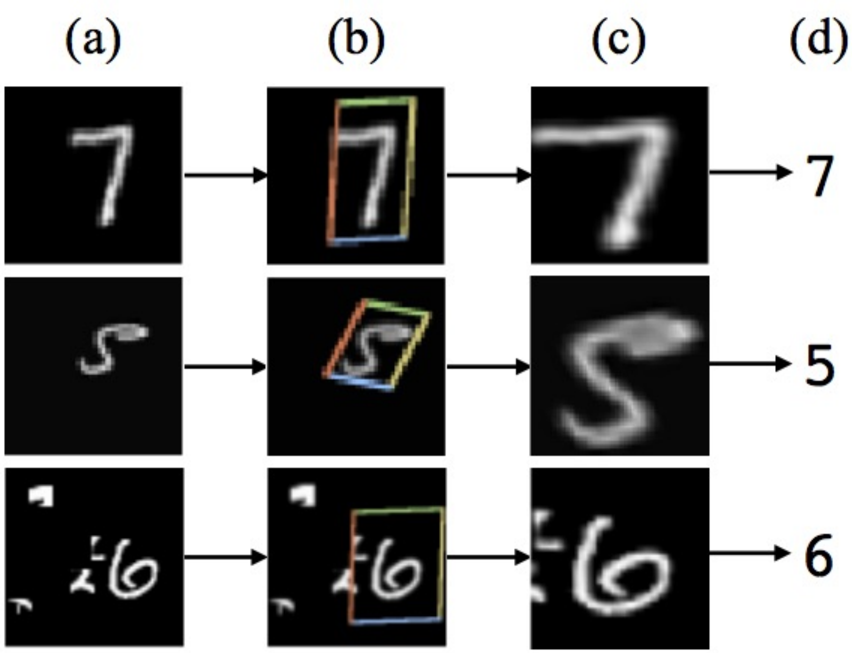
- If $\hat{\boldsymbol{x}}$ is the coordinate vector of $\boldsymbol{x} \in V$ with respect to B, and $\hat{\boldsymbol{y}}$ the coordinate vector of $\boldsymbol{y} = \Phi(\boldsymbol{x}) \in W$ with respect to C, then

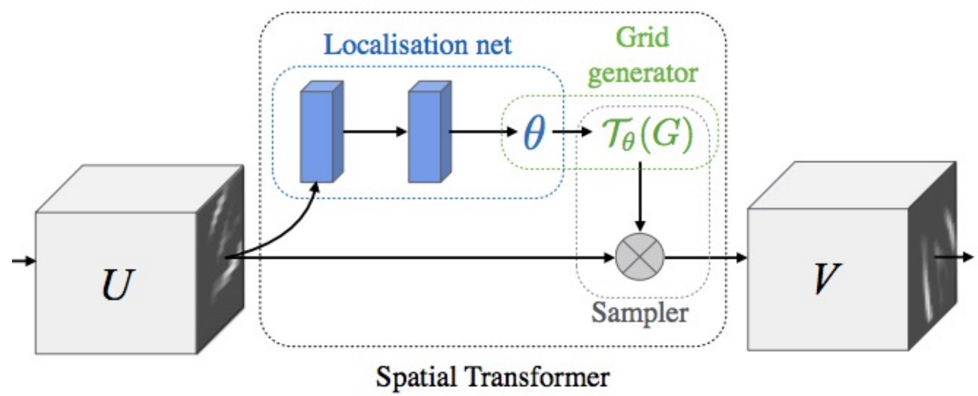
$$\hat{\boldsymbol{y}} = \boldsymbol{A}_\Phi \, \hat{\boldsymbol{x}}$$

Spatial Transformer Networks (Jaderberg et al., NIPS 2015)

$$\begin{pmatrix} x_i^s \\ y_i^s \end{pmatrix} = \mathcal{T}_\theta(G_i) = \mathtt{A}_\theta \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{bmatrix} \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix}$$

Affine transformation





Check your understanding

- Which of the following statements are correct?

(A) In a vector space, any vector can be represented as a linear combination of a certain set of vectors in this space

(B) The dimension of a vector equals the dimension of the space itis in.

(C) U is a vector subspace of V. Then vectors in U have lower dimension than vectors in V

(D) The set $\left\{\begin{bmatrix}2\\0\\0\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}6\\2\\-2\end{bmatrix}\right\}$ forms a basis for $\mathbb{R}^3$

(E) $U = \{(x, y): x = y, x \in \mathbb{R}, y \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^2$

(F) The vector $\mathbf{0}$ is linearly dependent with any vector in the same vector space