Chapter 1

Introduction to Vectors

Po-Ning Chen, Professor

Department of Electrical and Computer Engineering

National Chiao Tung University

Hsin Chu, Taiwan 30010, R.O.C.

Notes for this course

A few notes are given before the lectures begin.

1. In my slides,

   A red box usually indicates **theorems**, **lemmas**, etc.

   A blue box usually contains **definitions**, **important notions**, etc.

   A green box usually gives **tips**, **supplement stuff**, etc.

2. These slides are prepared mainly for lecturing, and are not intended to be replacements of the textbook. You are suggested to study the textbook and also to self-practice the problems at the end of each section.

3. The slides as well as the textbook are basically in **English**, and so will the exam. This will help and force you to be familiar with the **linear-algebraic terminologies** in **English**.

1.1 Vectors and their addition/subtraction

- What is a vector?

Answer: An ordered list of numbers. Usually, put in column (hence, in particular, it is called **column vector**.)

By dictionary, vector = a quantity with direction.

Example.

$$\boldsymbol{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

□

- How to define addition (or subtraction) of two vectors?

Answer: Component-wise addition/subtraction.

Example. Let $\boldsymbol{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Then,

$$\begin{aligned} \boldsymbol{w} + \boldsymbol{v} &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1+2 \\ 1+3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \\ \boldsymbol{w} - \boldsymbol{v} &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1-2 \\ 1-3 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} \end{aligned}$$

□

1.1 Scalar multiplication

- How to define scalar multiplication of a vector?

Answer: Component-wise multiplication.

Example. Let vector $\boldsymbol{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and scalar $c = 2$. Then,

$$c\,\boldsymbol{v} = 2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \times 2 \\ 2 \times 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

$\square$

By dictionary, scalar = a quantity without direction.

Scalar multiplication can be viewed (and should be viewed) as an extension of vector addition. For example, $2\boldsymbol{v} = \boldsymbol{v} + \boldsymbol{v}$ and $3\boldsymbol{v} = \boldsymbol{v} + \boldsymbol{v} + \boldsymbol{v}$.

Here, c is referred to as the *scalar*.
In contrast, c (alone) is a scalar and $[c]$ is a one-dimensional vector.

1.1 Linear combination

- What is **linear combination** of vectors?

Answer: To **combine** vectors in a **linear** fashion.

Example (of linear conception). Let $\boldsymbol{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $c_1 = 1$, $c_2 = 2$ and $c_3 = 3$.

Then,

$$\begin{aligned} c_1\,\boldsymbol{v} &= 1\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \times 2 \\ 1 \times 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ c_2\,\boldsymbol{v} &= 2\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \times 2 \\ 2 \times 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} \\ c_3\,\boldsymbol{v} &= 3\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \times 2 \\ 3 \times 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 9 \end{bmatrix} \end{aligned}$$

These three vectors lie along a **line** on a two-dimensional plane, hence they are **linear** to each others. □

Example (of linear combination). Now we wish to combine two vectors in a linear fashion.

$$c\,\boldsymbol{v} + d\,\boldsymbol{w} = c\begin{bmatrix} 2 \\ 3 \end{bmatrix} + d\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2c + d \\ 3c + d \end{bmatrix}$$

is a **linear combination** of vectors $\boldsymbol{v}$ and $\boldsymbol{w}$. □

1.1 Linear combination

- Why is **linear combination** important?

No answer but a tip:

- $c\,\boldsymbol{v}$ can characterize a line, if $\boldsymbol{v}$ is a non-zero vector (or equivalently, $\boldsymbol{v} \neq \mathbf{0}$).
- $c\,\boldsymbol{v} + d\,\boldsymbol{w}$ can characterize a 2-dimensional plane, if $\boldsymbol{v}$ and $\boldsymbol{w}$ are non-zero vectors and are not aligned.
- $c\,\boldsymbol{v} + d\,\boldsymbol{w} + e\,\boldsymbol{u}$ can characterize a 3-dimensional volumn (or 3-dimensional space), if $\boldsymbol{v}$, $\boldsymbol{w}$ and $\boldsymbol{u}$ are non-zero vectors and are not aligned.
- $\cdots$

Note: A **zero vector 0** can be viewed (and should be viewed or defined) as

$$\mathbf{0} = \boldsymbol{v} - \boldsymbol{v}.$$

For example, a two-dimensional zero vector is

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

1.1 Linear algebra

After the introduction of **linear combination**, it is perhaps time to talk about what do we mean by **linear algebra**.

- What is this course **Linear Algebra** about?

Algebra (代數)

> *The part of mathematics in which letters and other general symbols are used to represent numbers and quantities in formulas and equations.*

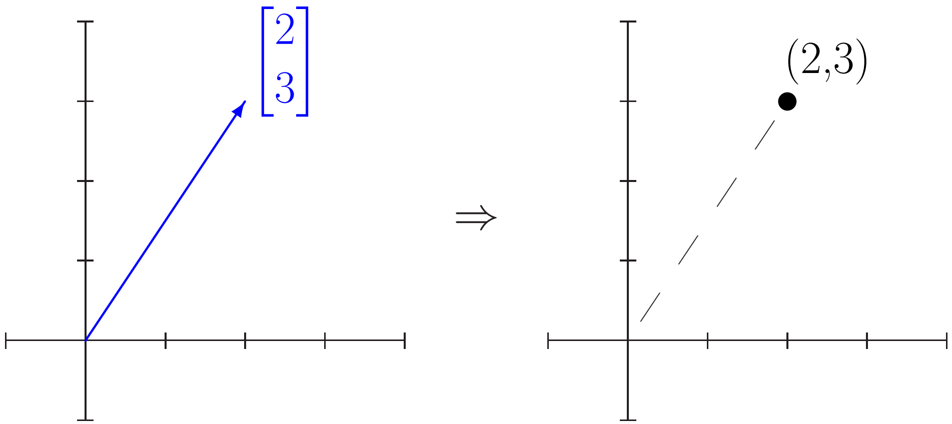
Linear Algebra (線性代數)

> *To combine these algebraic symbols (e.g., vectors) in a linear fashion.*

> So, we will **not** combine these algebraic symbols in a **nonlinear** fashion in this course!

1.1 Graphical representation of vectors

- Vector $\boldsymbol{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, as its name reveals (a quantity with direction), can be graphically represented by an **arrow** (pointing out its direction) from $(0,0)$ to $(2,3)$.

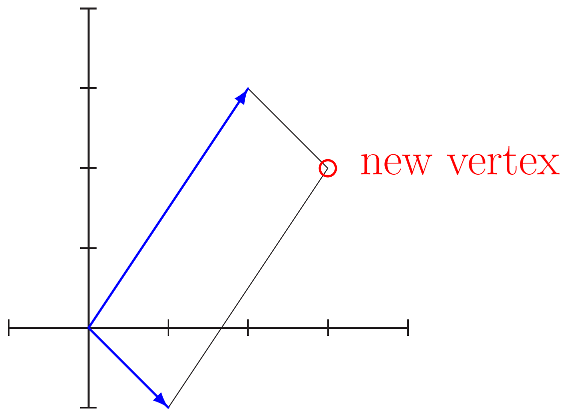


- Sometimes, we will only mark the **end point** of the directional arrow since a vector can be uniquely identified by its end point.

In notations, $(2,3)$ is the coordinate (or in Chinese, 坐標) in a 2-dimensional plane, and $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is a vector. As an end point can be used to identify a vector, these two representations will be used interchangeably.

1.1 Graphical representation of vector sum

- Graphically, the vector sum $\boldsymbol{v} + \boldsymbol{w}$ is the **new vertex** of the parallelogram formed by $\boldsymbol{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.



- *Exercise*: Can the linear combinations of the above two vectors (i.e., the new vertex of the parallelogram) define all points in the 2-dimensional plane?

Let's now proceed to Section 1.2!

1.2 Length of a vector and angle between two vectors

- How to find the length of a vector $\boldsymbol{v}$?

Answer: By Pythagoras formula, $\|\boldsymbol{v}\| = \sqrt{v_1^2 + v_2^2}$ if $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$.

We can repeat using the Pythagoras formula to prove that $\|\boldsymbol{v}\| = v_1^2 + v_2^2 + v_3^2$ for a 3-dimensional vector by noting that $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ v_3 \end{bmatrix}$.

The length of a vector is sometimes referred to as its **norm**.

- How to find the angle between two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$?

Answer: Using **dot product** or **inner product** as

$$\boldsymbol{v} \cdot \boldsymbol{w} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = v_1 w_1 + v_2 w_2.$$

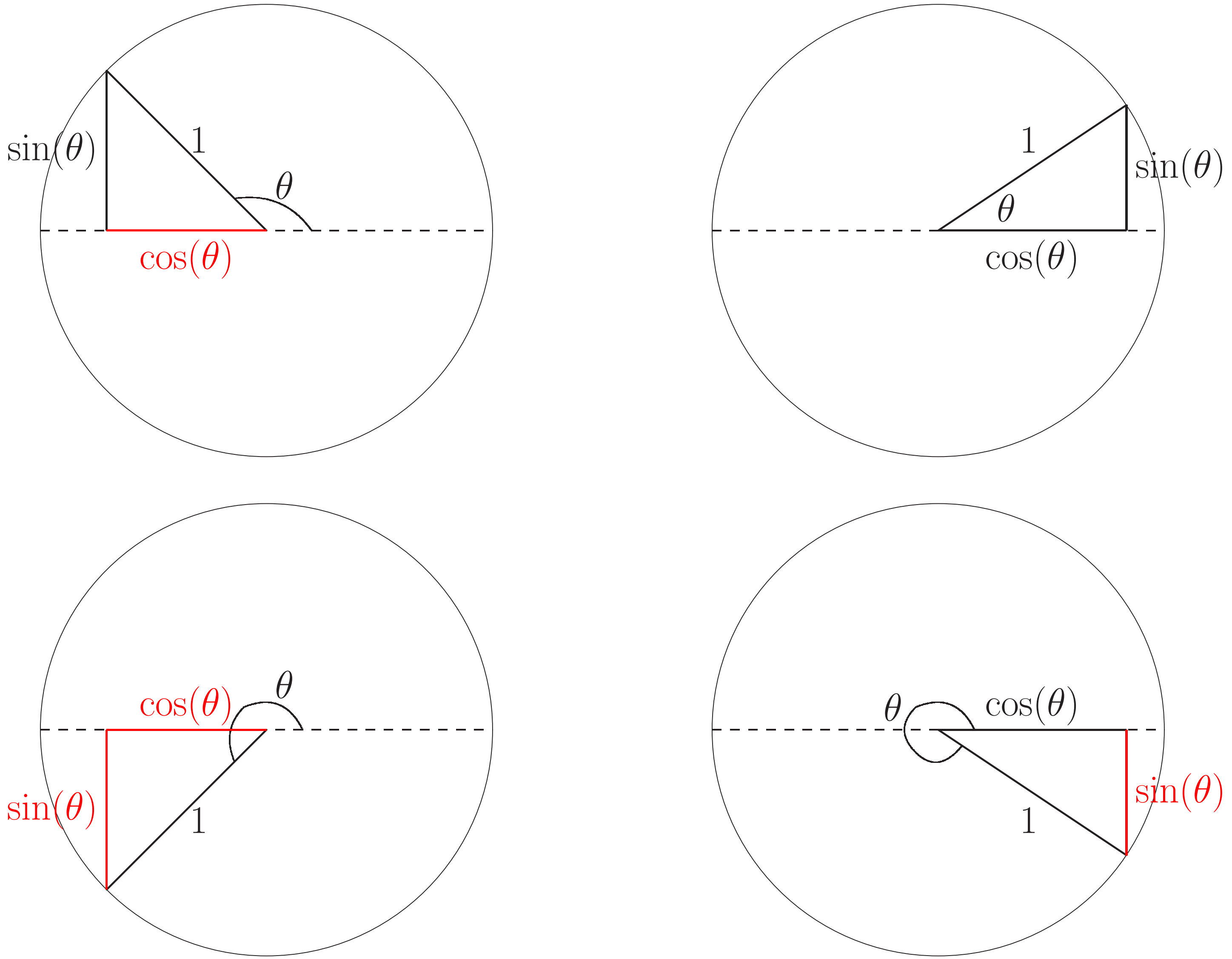
1.2 Sine and cosine versus inner product

- Since it is named **dot product**, the operator is *reasonably* denoted by a **dot** as $\boldsymbol{v} \cdot \boldsymbol{w}$.
- Since it is the sum of termwise products of *inner* components, it is also named **inner product**.

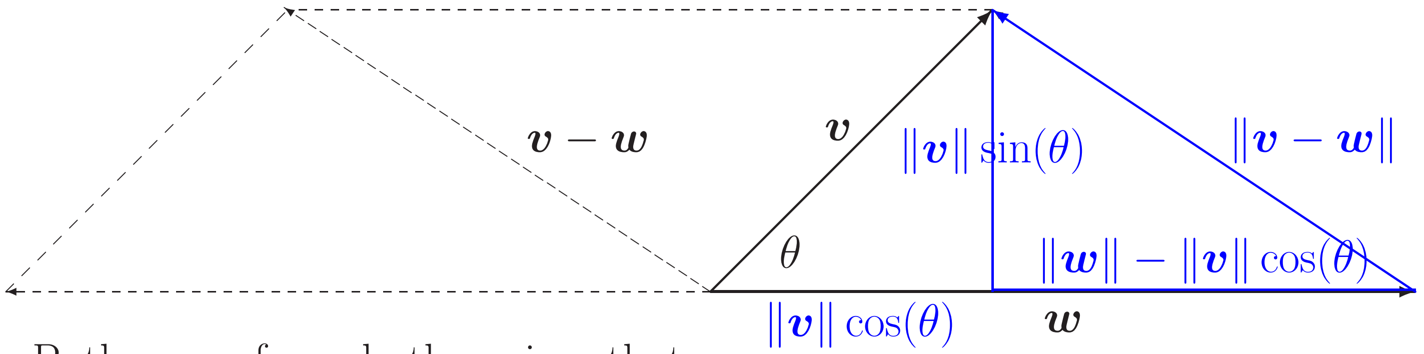
- Why can dot product or inner product be used to measure angle of two vectors?

Before we answer the question, let us recall the usual trigonometric functions, sin() and cos(), for measurement of angles.

1.2 Sine and cosine versus inner product



1.2 Sine and cosine versus inner product



- Pythagoras formula then gives that

$$(\|\boldsymbol{v}\|\sin(\theta))^2+(\|\boldsymbol{w}\|-\|\boldsymbol{v}\|\cos(\theta))^2=\|\boldsymbol{v}-\boldsymbol{w}\|^2$$

$$\Rightarrow\ \|\boldsymbol{v}\|^2+\|\boldsymbol{w}\|^2-2\|\boldsymbol{v}\|\|\boldsymbol{w}\|\cos(\theta)=\|\boldsymbol{v}-\boldsymbol{w}\|^2$$

$$\begin{aligned}\Rightarrow\ \|\boldsymbol{v}\|\|\boldsymbol{w}\|\cos(\theta)&=\frac{\|\boldsymbol{v}\|^2+\|\boldsymbol{w}\|^2-\|\boldsymbol{v}-\boldsymbol{w}\|^2}{2}\\&=\frac{(v_1^2+v_2^2)+(w_1^2+w_2^2)-[(v_1-w_1)^2+(v_2-w_2)^2]}{2}\\&=v_1w_1+v_2w_2=\boldsymbol{v}\cdot\boldsymbol{w}\end{aligned}$$

$$\Rightarrow\ \cos(\theta)=\frac{\boldsymbol{v}\cdot\boldsymbol{w}}{\|\boldsymbol{v}\|\|\boldsymbol{w}\|}$$

1.2 Examples of inner product

- *Example.* By

$$\cos(\theta) = \frac{\boldsymbol{v} \cdot \boldsymbol{w}}{\|\boldsymbol{v}\| \|\boldsymbol{w}\|},$$

if two vectors are perpendicular to each other, what will their inner product be? □

- *Example.* Is $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{w} \cdot \boldsymbol{v}$? This is called *commutative.* □

- *Example of using the inner product.* For a see-saw, put a weight of 4 on the left at a distance of 1, and put a weight of 2 on the right at a distance of 2. Will a balance be achieved?

  Answer: We obtain a weight vector $\boldsymbol{v} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ and a distance vector $\boldsymbol{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.
  Then, $\boldsymbol{v} \cdot \boldsymbol{w} = 4 \times (-1) + 2 \times 2 = 0$. Hence, a balance is achieved. □

1.2 Length and inner product

- The length of a vector $\boldsymbol{v}$ is the square root of its self inner product.

$$\boldsymbol{v} \cdot \boldsymbol{v} = v_1^2 + v_2^2 = \|\boldsymbol{v}\|^2$$

- Hence, the *inner product operation* can be used to determine both *length* and *angle* of vectors.

- *Example.* Two vectors, $\boldsymbol{v}$ and $\boldsymbol{w}$, are perpendicular if, and only if, $\boldsymbol{v} \cdot \boldsymbol{w} = 0$. Furthermore, the positive or negative of $\boldsymbol{v} \cdot \boldsymbol{w}$ tells whether the angle between the two vectors is below or above a right angle. $\square$

1.2 Length and inner product

After knowing that the length of a vector $\boldsymbol{v}$ can be represented by its inner product $\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$, the proof on Slide 1-12 can be generalized (to three or higher dimension) in the following way.

For vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ with dimension higher than 2,

$$(\|\boldsymbol{v}\| \sin(\theta))^2 + (\|\boldsymbol{w}\| - \|\boldsymbol{v}\| \cos(\theta))^2 = \|\boldsymbol{v} - \boldsymbol{w}\|^2$$

$$\Rightarrow \quad \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 - 2\|\boldsymbol{v}\|\|\boldsymbol{w}\| \cos(\theta) = \|\boldsymbol{v} - \boldsymbol{w}\|^2$$

$$\begin{aligned} \Rightarrow \quad \text{From } \|\boldsymbol{v} - \boldsymbol{w}\|^2 &= (\boldsymbol{v} - \boldsymbol{w}) \cdot (\boldsymbol{v} - \boldsymbol{w}) \\ &= \boldsymbol{v} \cdot \boldsymbol{v} - \boldsymbol{w} \cdot \boldsymbol{v} - \boldsymbol{v} \cdot \boldsymbol{w} + \boldsymbol{w} \cdot \boldsymbol{w} \\ &= \|\boldsymbol{v}\|^2 - 2\boldsymbol{v} \cdot \boldsymbol{w} + \|\boldsymbol{w}\|^2 \end{aligned}$$

we know that $\cos(\theta) = \dfrac{\boldsymbol{v} \cdot \boldsymbol{w}}{\|\boldsymbol{v}\|\|\boldsymbol{w}\|}$

1.2 Unit vector and standard unit vector

Definition (Unit vector): A **unit vector** is a vector with length equal to one.

- "unit": a quantity chosen as a standard, in terms of which other quantities may be expressed, e.g., a unit of measurement. For example, unit circle, unit cube, unit vector, ..., etc.
- *Example.* $\boldsymbol{v}/\|\boldsymbol{v}\|$ is a unit vector in the same direction as $\boldsymbol{v}$. $\square$

Definition (Standard unit vector): A **standard unit vector** is a unit vector with only one non-zero (positive) component.

- *Example.*

$$\boldsymbol{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \boldsymbol{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$\square$

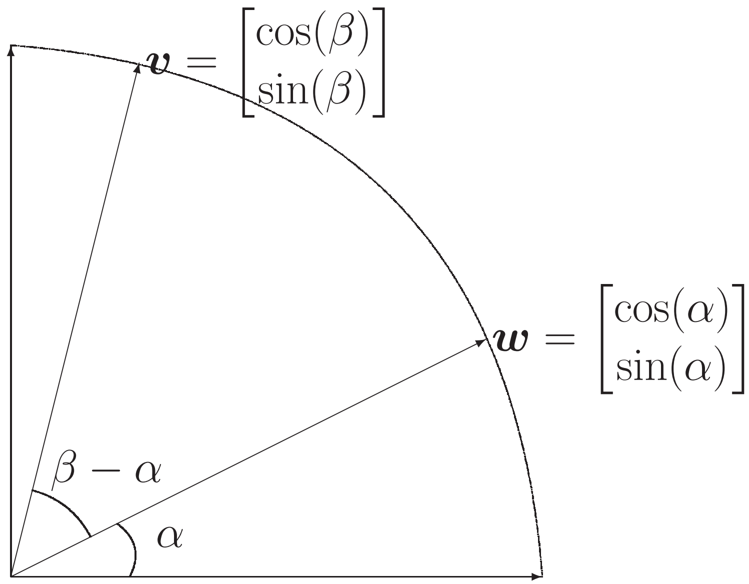
1.2 Theories regarding inner product

- *Angle between unit vectors.* The inner product of two **unit vectors** is equal to the cosine of the angle between these two vectors.

$$\cos(\theta) = \frac{\boldsymbol{v} \cdot \boldsymbol{w}}{\|\boldsymbol{v}\|\|\boldsymbol{w}\|} = \boldsymbol{v} \cdot \boldsymbol{w}$$

- *Trigonometric function.* Prove that $\cos(\beta-\alpha) = \cos(\beta)\cos(\alpha)+\sin(\beta)\sin(\alpha)$ by means of inner product.

  Hint: $\boldsymbol{v} \cdot \boldsymbol{w} = \cos(\beta - \alpha)$.



1.2 Theories regarding inner product

- $\boxed{\textit{Triangle inequality.}}$ Prove that $\|\boldsymbol{v}+\boldsymbol{w}\| \leq \|\boldsymbol{v}\|+\|\boldsymbol{w}\|$.

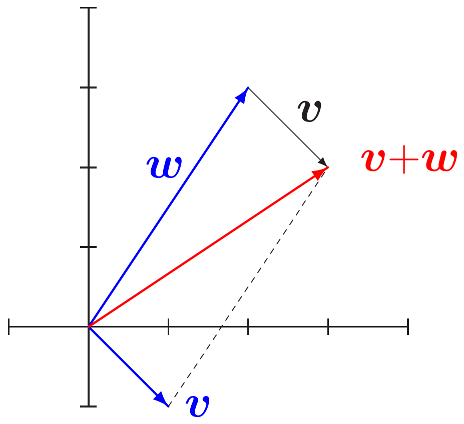
  Proof: From Slide 1-12, we obtain that when θ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$,

  $$\|\boldsymbol{v}-\boldsymbol{w}\|^2 = \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 - 2\|\boldsymbol{v}\|\|\boldsymbol{w}\|\cos(\theta).$$

  Similarly, since the angle between $\boldsymbol{v}$ and $-\boldsymbol{w}$ is $\pi-\theta$,

  $$\begin{aligned}\|\boldsymbol{v}+\boldsymbol{w}\|^2 &= \|\boldsymbol{v}-(-\boldsymbol{w})\|^2\\ &= \|\boldsymbol{v}\|^2 + \|-\boldsymbol{w}\|^2 - 2\|\boldsymbol{v}\|\|-\boldsymbol{w}\|\cos(\pi-\theta)\\ &= \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2\|\boldsymbol{v}\|\|\boldsymbol{w}\|\cos(\theta)\\ &\leq \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2\|\boldsymbol{v}\|\|\boldsymbol{w}\|\\ &= (\|\boldsymbol{v}\|+\|\boldsymbol{w}\|)^2.\end{aligned}$$

  $\square$



1.2 Theories regarding inner product

- $\boxed{\textit{Schwarz inequality.}}$ Prove that $|\boldsymbol{v}\cdot\boldsymbol{w}| \le \|\boldsymbol{v}\|\|\boldsymbol{w}\|$.

  Proof:

$$\left|\frac{\boldsymbol{v}\cdot\boldsymbol{w}}{\|\boldsymbol{v}\|\|\boldsymbol{w}\|}\right| = |\cos(\theta)| \le 1$$

$\square$

After the introduction of inner product, is there an outer product?

Hint: *Inner product* and *outer product* are respectively

$$\boldsymbol{v}\cdot\boldsymbol{w} = \boldsymbol{v}'\boldsymbol{w} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = v_1w_1 + v_2w_2$$

and

$$\boldsymbol{v}\otimes\boldsymbol{w} = \boldsymbol{v}\boldsymbol{w}' = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\begin{bmatrix} w_1 & w_2 \end{bmatrix} = \begin{bmatrix} v_1w_1 & v_1w_2 \\ v_2w_1 & v_2w_2 \end{bmatrix},$$

where superscript "$'$" denotes the **vector (matrix) transpose** operation. (Notably, the product of two matrices will be introduced in Section 1.3.)

In short, an *inner product* produces the **sum** of termwise products, while an *outer product* lists **all** termwise products in a matrix form.

Definition (Trace): The trace of a square matrix is the sum of the diagonals.

- By this definition, we observe that

$$\texttt{trace}(\boldsymbol{v} \otimes \boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{w}.$$

Definition (Vector transpose): The transpose of a vector is simply to re-express a column vector into a row vector (and vice versa).

In **notations**, people may use

$$\boldsymbol{v}' \quad \text{or} \quad \boldsymbol{v}^{\mathtt{T}}$$

to denote the transpose. For **complex** vectors, we may also wish to **complement** the vector components in addition to transpose. In such case,

$$\boldsymbol{v}^{\dagger} \quad \text{or} \quad \boldsymbol{v}^{\mathtt{H}}$$

will be used instead. For example, $\begin{bmatrix} 1+\imath \\ 2+3\imath \end{bmatrix}^{\dagger} = \begin{bmatrix} 1-\imath & 2-3\imath \end{bmatrix}$. Based on this, the inner product of complex vectors will be defined as

$$\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{v}^{\dagger}\boldsymbol{w}.$$

In Chinese, we would say 橫列(row)直行(column). In Excel, we would use 橫列(row)直欄(column).

1.3 Matrices

Definition (Matrix): A **matrix** is a rectangular array of numbers placed in rows and columns, and is treated as a single entity, which is manipulated according to particular rules.

- A matrix can be regarded as a collection of multiple (column) vectors, listed in a row. For example,

$$A = \begin{bmatrix} \boldsymbol{u} & \boldsymbol{v} & \boldsymbol{w} \end{bmatrix} = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}_{3\times 3} \quad \text{or} \quad B = \begin{bmatrix} \boldsymbol{v} & \boldsymbol{w} \end{bmatrix} = \begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \end{bmatrix}_{3\times 2}$$

- A matrix can be regarded as a collection of multiple row vectors (i.e., transpose of column vectors), listed in a column. For example,

$$A = \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \boldsymbol{x}_3' \end{bmatrix} = \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}_{3\times 3} \quad \text{or} \quad B = \begin{bmatrix} \boldsymbol{y}_1' \\ \boldsymbol{y}_2' \\ \boldsymbol{y}_3' \end{bmatrix} = \begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \end{bmatrix}_{3\times 2}$$

where $\boldsymbol{x}_1 = \begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix}$, $\boldsymbol{x}_2 = \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix}$, $\boldsymbol{x}_3 = \begin{bmatrix} u_3 \\ v_3 \\ w_3 \end{bmatrix}$,

and $\boldsymbol{y}_1 = \begin{bmatrix} v_1 \\ w_1 \end{bmatrix}$ $\boldsymbol{y}_2 = \begin{bmatrix} v_2 \\ w_2 \end{bmatrix}$, $\boldsymbol{y}_3 = \begin{bmatrix} v_3 \\ w_3 \end{bmatrix}$.

Definition (Product of matrices): The product of two **matrices** is defined as the inner products of row vectors and column vectors respectively from the first matrix and the second matrix, counted from the left. Specifically as an example,

$$A_{2\times 3}B_{3\times 2} = \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 \end{bmatrix} \triangleq \begin{bmatrix} \boldsymbol{a}_1'\boldsymbol{b}_1 & \boldsymbol{a}_1'\boldsymbol{b}_2 \\ \boldsymbol{a}_2'\boldsymbol{b}_1 & \boldsymbol{a}_2'\boldsymbol{b}_2 \end{bmatrix}_{2\times 2} = \begin{bmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{b}_1 & \boldsymbol{a}_1 \cdot \boldsymbol{b}_2 \\ \boldsymbol{a}_2 \cdot \boldsymbol{b}_1 & \boldsymbol{a}_2 \cdot \boldsymbol{b}_2 \end{bmatrix}_{2\times 2},$$

where for the two matrices A and B on the extreme left of the above equation, "2" is the number of vectors, and "3" is the dimension of the vectors.

Note that the dimension of the vectors shall be the same, otherwise the inner product operation cannot be performed.

- In order to be consistent, a vector is always by convention a **column** vector. I will always add the transpose mark "**′**" for a **row** vector.
- A vector can be treated as a matrix consisting of only one column. For example,

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}.$$

  Based on this, the product of two matrices can certainly be applicable to compute the product of a matrix and a vector.

Definition (Product of a matrix and a vector): The product of a **matrix** and a **vector** is defined as the inner products of the row vectors from the matrix in the front and the vector. Specifically as an example,

$$A_{2\times 3}\boldsymbol{x}_{3\times 1} = \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \end{bmatrix} \boldsymbol{x} \triangleq \begin{bmatrix} \boldsymbol{a}_1'\boldsymbol{x} \\ \boldsymbol{a}_2'\boldsymbol{x} \end{bmatrix}_{2\times 1} = \begin{bmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{x} \\ \boldsymbol{a}_2 \cdot \boldsymbol{x} \end{bmatrix}_{2\times 1},$$

where for the matrix A and the vector $\boldsymbol{x}$ on the extreme left of the above equation, "2" and "1" are the numbers of vectors, and "3" is the dimension of the vectors.

Note that the dimension of the vectors shall be the same, otherwise the inner product operation cannot be performed.

- In its description from the textbook, $A\boldsymbol{x}$ can be regarded as
  - a dot product (i.e., inner product) with row (vectors of A),
  - or a linear combination of column vectors of A with coefficients from $\boldsymbol{x}$.

  For the second description, it can be understood as

$$A\boldsymbol{x} = \begin{bmatrix} \boldsymbol{u} & \boldsymbol{v} & \boldsymbol{w} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\boldsymbol{u} + x_2\boldsymbol{v} + x_3\boldsymbol{w}.$$

1.3 Matrices and linear equations

- In what way the matrices can be useful?

Answer: Solving linear equations.

Example. Difference matrix A and sum matrix S.

Definition (Difference matrix): A **difference matrix** A produces the difference of adjacent components in $\boldsymbol{x}$ through $A\boldsymbol{x}$. In mathematics, a three-by-three difference matrix is given by

$$A = \begin{array}{c} \text{-1} \\ \\ \\ \end{array}\begin{bmatrix} 1 & 0 & 0 \\ \text{-1} & 1 & 0 \\ 0 & \text{-1} & 1 \end{bmatrix} \Rightarrow A\boldsymbol{x} = \begin{bmatrix} 1 & 0 & 0 \\ \text{-1} & 1 & 0 \\ 0 & \text{-1} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix}$$

where there is an invisible "-1" beyond the first row of A.

Based on the above definition, we are wondering how to solve the linear equation

$$A\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \triangleq \boldsymbol{b}?$$

Answer: $\begin{cases} x_1 & = b_1 \\ x_2 \ (= b_2 + x_1) & = b_2 + b_1 \\ x_3 \ (= b_3 + x_2) & = b_3 + b_2 + b_1 \end{cases}$ or equivalently, $\boldsymbol{x} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

Definition (Sum matrix): A **sum matrix** S produces the sum of components in $\boldsymbol{b}$ through $S\boldsymbol{b}$. In mathematics, a three-by-three sum matrix is given by

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \Rightarrow S\boldsymbol{b} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_1 + b_2 \\ b_1 + b_2 + b_3 \end{bmatrix}$$

In summary,

$$\boxed{A\boldsymbol{x} = \boldsymbol{b}} \quad \Rightarrow \quad \boxed{SA\boldsymbol{x} = \boldsymbol{x} = S\boldsymbol{b}}.$$

In terminology, we say that the sum matrix S is the **left inverse matrix** of the difference matrix A, since

$$SA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I \triangleq \text{Identity matrix.}$$

As self-revealed by its name, an **identity** element i for some operation "$\circ$" is one that $x \circ i = x$ for every x.

Here, the matrix I is a *multiplicative identity element* for matrix multiplication since $AI = A$ for every matrix A.

End of example of difference matrix and sum matrix □

Now we learn that the **sum matrix** S is the left inverse matrix of the **difference matrix** A because $SA = I$.

Analogously, the **difference matrix** A is the right inverse matrix of the **sum matrix** S because $SA = I$.

Now, let us ask ourselves a few questions, and try to answer them.

- Is the left inverse matrix of a matrix always equal to its right inverse matrix?

  Answer: Not necessarily.

- Why isn't the left inverse matrix always equal to the right inverse matrix?

  Answer: Because matrix multiplication is not *commutative*.
  In other words, $AB = BA$ is not necessarily true! For example, the 2×3 left inverse matrix of $A_{3\times 2}$ is usually not the right inverse of $A_{3\times 2}$.

1.3 Matrix and its inverse

Definition (Inverse matrix). If the *left inverse matrix* of a matrix A is equal to its *right inverse matrix*, we say that A is invertible, and its *unique* inverse matrix is denoted by A^{-1}.

- This notation is in parallel to the *multiplicative inverse* of a real number a is denoted by a^{-1} (if $a \neq 0$).
- From the definition, we see $AA^{-1} = A^{-1}A = I$.
- Does the inverse matrix always exist?

Example. A cyclic difference matrix $C = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ has no inverse. $\square$

1.3 Inverse and linear equations

- Why is the inverse matrix useful in solving linear equations?

Answer:

- $\boxed{A\boldsymbol{x} = \boldsymbol{b}} \implies \boxed{\boldsymbol{x} = A^{-1}\boldsymbol{b}}$.
- If the left inverse matrix of A does not exist, then $A\boldsymbol{x} = \boldsymbol{b}$ either has no solution or has no unique solution.

1.3 Matrices and linear independence of vectors

Definition (Linear independence of vectors) A set of vectors $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, ..., $\boldsymbol{u}_k$ is said to be *linearly independent* or simply *independent* if the matrix

$$A = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_k \end{bmatrix}$$

has left inverse.

- If A has left inverse S, then $A\boldsymbol{x} = \boldsymbol{0}$ has unique solution $\boldsymbol{x} = S(A\boldsymbol{x}) = S\boldsymbol{0} = \boldsymbol{0}$.
- If $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are *linearly* **independent**, we cannot represent $\boldsymbol{u}_k$ as a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k-1}$. (In a sense, "linear independence" = "linear-combination independence.")
- If $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ are *linearly* **dependent**, then $A\boldsymbol{x} = \boldsymbol{0}$ has solutions other than $\boldsymbol{x} = \boldsymbol{0}$.
- *Exercise.* Show that $\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \boldsymbol{x} = \boldsymbol{0}$ has a non-zero solution.

1.3 Matrices and linear independence of vectors

In summary, there are four **products** (i.e., multiplications) in linear algebra. They are respectively denoted by

- a dot $\cdot$ for dot product or inner product, e.g., $\boldsymbol{v} \cdot \boldsymbol{w}$;
- a circled cross $\otimes$ for outer product, e.g., $\boldsymbol{v} \otimes \boldsymbol{w}$;
- nothing for matrix product or scalar product, e.g., AB and vw;
- a skewed cross $\times$ for scalar product, e.g., $v \times w$.