Chapter 3 Vector Spaces

這章的主角由矩陣換成向量，而且將向量空間的定義一般化，只要滿足少數的條件，就可以有很多好的性質，例如線性獨立、向量基底等，這是抽象代數的性質。第一、二節定義向量空間相關名詞，很重要，建議後面卡住時，回來看看定義。第三、四節介紹向量空間的相關定理，會感覺繞來繞去，不知所云，但是要把住握住重要的結果。第五節是基底變換，可以看成是第四章的特例，學的好第四章會如魚得水。第六節回到矩陣，把矩陣看成一堆向量擺在一起，利用向量空間的性質，矩陣又有更多新的性質。

3.1 Definition

定義一般性的向量空間。

- Let V be a set on which the operations of <u>scalar multiplication</u> and <u>addition</u> are defined.
  （1） for each vector $\mathbf{x}$ in V and a scalar α, one can associate a unique element $\alpha\mathbf{x}$ in V,
  （2） for each pair of elements $\mathbf{x}$ and $\mathbf{y}$ in V, one can associate a unique element $\mathbf{x} + \mathbf{y}$ that is also in V.

- The <u>**closure properties**</u> of addition and scalar multiplication operations:
  C1. If $\mathbf{x} \in V$ and α is a scalar, then $\alpha\mathbf{x} \in V$
  C2. If $\mathbf{x}, \mathbf{y} \in V$, then $\mathbf{x} + \mathbf{y} \in V$

Example: Let $W = \{(a, 1) \mid a \text{ real}\}$
- By C1. $(a, 1) \in W$, $\alpha(a, 1) = (\alpha a, \alpha) \notin W$
- By C2. $(a, 1) \in W$ and $(b, 1) \in W$, $(a, 1) + (b, 1) = (a+b, 2) \notin W$

The operations of addition and scalar multiplication are ***not*** defined for *W.*

- The <u>set V</u> together with the <u>operations</u> of addition and scalar multiplication is said to form a **vector space** if the following axioms are satisfied.
  A1. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ for any $\mathbf{x}$ and $\mathbf{y}$ in V
  A2. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ for any $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ in V
  A3. There exist an element $\mathbf{0}$ in V such that $\mathbf{x} + \mathbf{0} = \mathbf{x}$ for each $\mathbf{x} \in V$
  A4. For each $\mathbf{x} \in V$, there exist an element $-\mathbf{x}$ in V such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$
  A5. $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$ for each scalar α and any $\mathbf{x}$ and $\mathbf{y}$ in V
  A6. $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$ for any scalars α and β and any $\mathbf{x} \in V$
  A7. $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$ for any scalars α and β and any $\mathbf{x} \in V$
  A8. $1 \cdot \mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in V$

Example: Let $C[a, b]$ denote the set of all real-valued functions that are defined and continuous on the closed intervals $[a, b]$. In this case, our universal set is a set of functions. Thus, our vectors are the functions in $C[a, b]$.

- If f and g are functions in $C[a, b]$ and a is a real number: For all x in $[a, b]$

$$(f+g)(x) = f(x) + g(x)$$
$$(\alpha f)(x) = \alpha f(x)$$

- Clearly, αf is in $C[a, b]$, since a constant times a continuous function is always continuous. For any x in $[a, b]$ the function

$$(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x)$$

$z(x) = 0 \quad$ for all x in $[a, b]$ acts as the zero vector: $f + z = f \quad$ for all f in $C[a, b]$

Example: Let P_n denote the set of all polynomials of degree less than n. Define $p+q$ and αp by $(p+q)(x) = p(x) + q(x)$ and $(\alpha p)(x) = \alpha p(x)$ for all real numbers x. P_n is a vector space.

- **Theorem 3.1.1** If V is a vector space and $\mathbf{x}$ is any element of V, then
  (i) $0\mathbf{x} = \mathbf{0}$
  (ii) $\mathbf{x} + \mathbf{y} = \mathbf{0}$ implies that $\mathbf{y} = -\mathbf{x}$ (i.e., the additive inverse of $\mathbf{x}$ is unique)
  (iii) $(-1)\mathbf{x} = -\mathbf{x}$

3.2 Subspaces

定義一般性的子空間。注意 null space 和 span 的定義和意義，之後會一直出現。

- If S is a nonempty subset of a vector space V, and S satisfies the following conditions:
  (i) $\alpha\mathbf{x} \in S$ whenever $\mathbf{x} \in S$ for any scalar α (closed under scalar multiplication)
  (ii) $\mathbf{x} + \mathbf{y} \in S$ whenever $\mathbf{x} \in S$ and $\mathbf{y} \in S$ (closed under addition)
  then S is said to be a **subspace** of V
- Remarks
  - A subapace of V is a subset S that is <u>closed under the operations</u> of V.
  - Every subspace of a vector space is a <u>vector space in its own right</u>.
  - If V is a vector space, then $\{\mathbf{0}\}$ and V are subspaces of V. All other subspaces are referred to as *proper subspaces*, $\{\mathbf{0}\}$ is referred to as the *zero subspace*.

Example: which of the followings are subspaces

- $S = \{(x_1, x_2, x_3)^T \mid x_1 = x_2\}$. is S a subspace of R^3?
- Let $S = \{ \begin{bmatrix} x \\ 1 \end{bmatrix} \mid x$ is a real number$\}$, is S a subspace of R^2?

- Let $S = \{A \in R^{2\times 2} \mid a_{12} = -a_{21}\}$, is S a subspace of $R^{2\times 2}$?
- Let S be the set of all polynomials of degree less than n with the property that $p(0) = 0$. The set S is nonempty since it contains the zero polynomial.
- Let $C^n[a, b]$ be the set of all functions f that have a continuous nth derivative on $[a, b]$, then $C^n[a, b]$ is a subspace of $C^n[a, b]$.

- Let A be an $m \times n$ matrix. Let $N(A)$ denote the set of all solutions to the homogeneous system $A\mathbf{x} = \mathbf{0}$. Thus $N(A) = \{\mathbf{x} \in R^n \mid A\mathbf{x} = 0\}$. $N(A)$ is called the **nullspace** of A.
- Is N(A) a subspace of R^n?
  - By C1:
  - By C2:
  - $\therefore$ $N(A)$ is a subspace of R^n

Example: Find the nullapce of

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

- Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be vectors in a vector space V. A sum of the form $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n$, where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are scalars, is called a **linear combination** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.
- The set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is called the **span** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, denoted by Span($\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$).

- **Theorem 3.2.1**: If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are elements of a vector space V, then Span($\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$) is a subspace of V.

Proof: Let $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n$ be an arbitrary element of Span($v_1, v_2, \ldots, v_n$)

By C1:

By C2:

- Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be vectors in a vector space V. Span($\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$) is referred to as the subspace *spanned* by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.
- If Span($\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$) = V, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is said to *span V* or that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a *spanning set* for V.
- The set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a **spanning set** for V iff every vector in V can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

Example: Which of the following are spanning sets for R^3?

(a) $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, (1, 2, 3)^T\}$

(b) $\{(1, 1, 1)^T, (1, 1, 0)^T, (1, 0, 0)^T\}$

(c) $\{(1, 0, 1)^T, (0, 1, 0)^T\}$

(d) $\{(1, 2, 4)^T, (2, 1, 3)^T, (4, -1, 1)^T\}$

- **Theorem 3.2.2**: If the linear system $A\mathbf{x} = \mathbf{b}$ is consistent and $\mathbf{x}_0$ is a particular solution, then a vector y will also be a solution if and only if $\mathbf{y} = \mathbf{x}_0 + \mathbf{z}$ where $\mathbf{z} \in N(A)$.

Proof:

3.3 Linear Independence

定義甚麼是線性獨立和線性相依，以及它和線性系統的可除性有何關聯。

- **Theorem**: If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span a vector space V and one of these vectors can be written as a linear combination of the other $n - 1$ vectors, then these $n - 1$ vectors span V.

Proof: Suppose $\mathbf{v}_n$ can be written as a linear combination of the other $n - 1$ vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$:

$\mathbf{v}_n$ = ____

- Let $\mathbf{v}$ be any vectors in V, since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span V, we can write

  $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_{n-1}\mathbf{v}_{n-1} + \alpha_n\mathbf{v}_n$

  = ____

  = ____
- Thus, any vectors $\mathbf{v}$ in V can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$

- **Theorem**: Given n vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, it is possible to write one of the vectors as a linear combination of the other $n - 1$ vectors iff there exist scalars $c_1, c_2, \ldots, c_n$ <u>not all zero</u> such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$$

Proof: (⇨) Suppose $\mathbf{v}_n$ can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$:

$$\mathbf{v}_n = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_{n-1}\mathbf{v}_{n-1}$$

Subtracting $\mathbf{v}_n$ from both sides of the equation, $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_{n-1}\mathbf{v}_{n-1} - \mathbf{v}_n = \mathbf{0}$

If we set $c_i = \alpha_i$ for $i = 1, 2, \ldots, n\text{-}1$ and $c_n = -1$, then $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$

(⇦) Conversely, if $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$ and at least one of the c_i's, say c_n, is nonzero, then

- The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in a vector space V are said to be **linearly independent** if $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$ implies that $c_1 = c_2 = \ldots = c_n = 0$.
- The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in a vector space V are said to be **linearly dependent** if there exist scalars $c_1, c_2, \ldots, c_n$ <u>not all zero</u> such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$
- If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a *minimum spanning set*, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are <u>linearly independent</u>.
  Conversely, if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent and span V, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a *minimum*

spanning set for V.

- A *minimum spanning set* is called a **basis**.

Example: Which of the following collections of vectors are linearly independent in R^3?

(a) $(1, 1, 1)^T$, $(1, 1, 0)^T$, $(1, 0, 0)^T$

(b) $(1, 0, 1)^T$, $(0, 1, 0)^T$

(c) $(1, 2, 4)^T$, $(2, 1, 3)^T$, $(4, -1, 1)^T$

- **Theorem 3.3.1**: Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be n vectors in R^n and let $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ then the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ will be <u>*linearly dependent*</u> iff <u>*X is singular*</u> (i.e., $\det(X) = 0$)

Proof: $c_1x_1 + c_2x_2 + \ldots + c_nx_n = 0$

- To test whether k vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are linearly independent in R^n, we can rewrite the equation

  $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_k\mathbf{x}_k = \mathbf{0}$

  as a linear system $X\mathbf{c} = \mathbf{0}$, where $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$.
  - If k=n, Form a matrix X whose <u>columns</u> are the vectors being tested. det(X) ____ iff the vectors are linearly dependent
  - If $k \neq n$, the matrix X is not square, the system is <u>homonegeous</u>. A <u>trivial solution</u> $\mathbf{c} = \mathbf{0}$.
  - It will have <u>nontrivial solutions</u> iff *the row echelon form of X involve <u>free variables</u>*.

  If there are <u>nontrivial solutions</u>, then the vectors are <u>linearly dependent</u>.

  If there are <u>no free variables</u>, then $\mathbf{c} = \mathbf{0}$ is the only solution ⇨ linearly independent

Example: Are x1, x2 and x3 linearly independent?

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} -2 \\ 3 \\ 1 \\ -2 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 7 \\ 7 \end{bmatrix}$$

- **Theorem 3.3.2**: Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be vectors in a vector space V. A vector $\mathbf{v} \in \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ can be written <u>uniquely</u> as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ iff $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are <u>linearly independent</u>.

Proof: If $\mathbf{v} \in \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, then $\mathbf{v}$ can be written as a linear combination

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n \quad (1)$$

- Suppose that $\mathbf{v}$ can also be expressed as a linear combination

$$\mathbf{v} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \ldots + b_n\mathbf{v}_n \quad (2)$$

  We will show that , if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent, that $b_i = a_i$, $i = 1, 2, \ldots, n$ and if $\mathbf{v}_1$,

$\mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly dependent, then it is possible to choose the b_i's different from the a_i's.

- From (1)-(2), we have

- On the other hand, if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly dependent, then there exist $c_1, c_2, \ldots, c_n$, not all 0, such that $\mathbf{0} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n$

3.4 Basis and Dimension

介紹線性空間特有的性質，向量基底和向量維度，當作基底的向量必須是線性獨立而且要能組合出空間中所有的向量，由向量基底數量的唯一性可以定義向量維度。

- The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ form a **basis** for a vector space V if and only if
  (1) $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent (minimal spanning set)
  (2) $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span V (spanning set)

- **Theorem 3.4.1**: If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a spanning set for a vector space V, then any collection of m vectors in V, where $m > n$, is linearly dependent.

Proof: Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ be m vectors in V, when $m > n$.

Since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span V, $\mathbf{u}_i$ = ____

A linearly combination $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m$ can be written in the form

- **Corollary 3.4.2**: If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ are both bases for a vector space V, then $n = m$.

Proof: Let both $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ be abses for V. Since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span V and $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ are linearly independent, it follows from Theorem 3.4.1 that $m \le n$. By same reasoning, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ span V and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent, so $n \le m$.

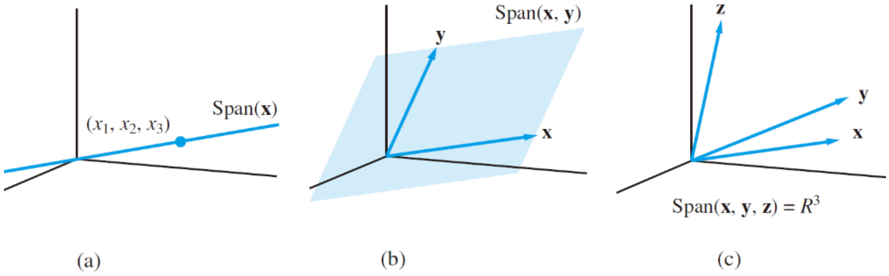
- Let V be vector space. If V has a basis consisting of n vectors, we say that V has **dimension** n.
- The subspace $\{\mathbf{0}\}$ of V is said to have dimension 0.
- V is said to be **finite-dimensional** if there is a finite set of vectors that spans V; otherwise we say that V is **infinite-dimensional**.

Example:

(a) Span(**x**) = {α**x** | α is a scalar}: line

(b) Span(**x, y**) = {α**x** + β**y** | α, β are scalars}: plane

(c) Span(**x, y, z**) = R^3



- **Theorem 3.4.3**: If V is a vector space of dimension $n > 0$

  (1) Any set of n <u>linearly independent</u> vectors spans V;

  (2) Any n vectors that span V are <u>linearly independent</u>.

Proof:

- **Theorem 3.4.4**: If V is a vector space of dimension $n > 0$, then

(I) No set of less then n vectors can span V

(II) Any subset of less then n <u>linearly independent</u> vectors can be extended to form a basis for V

(III) Any <u>spanning set</u> containing more than n vectors can be pared down to form a basis for V.

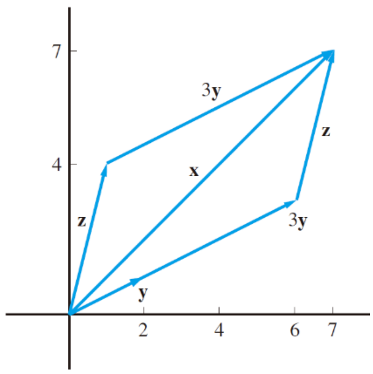
Proof:

3.5 Change of Basis

介紹如何變換向量基底，這是一種線性變換（第四章介紹）。

Example: Let {y, z} be a basis of R^2. Find x's coordinate to the basis [y, z]

$$\mathbf{y} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} 7 \\ 7 \end{bmatrix}$$



Example: Let $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ be the new basis

Q1: Given a vector $\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$, find its coordinates with respect to $\mathbf{e}_1$ and $\mathbf{e}_2$

$\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 = c_1(3\mathbf{e}_1 + 2\mathbf{e}_2) + c_2(\mathbf{e}_1 + \mathbf{e}_2) = (3c_1 + c_2)\mathbf{e}_1 + (2c_1 + c_2)\mathbf{e}_2$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3c_1+c_2 \\ 2c_1+c_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \quad U = (\mathbf{u}_1, \mathbf{u}_2) = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \qquad \mathbf{x} = U\,\mathbf{c}$$

- U is called the **transition matrix** from the ordered basis $[\mathbf{u}_1, \mathbf{u}_2]$ to the standard basis $[\mathbf{e}_1, \mathbf{e}_2]$

Q2: Given a vector $\mathbf{x} = (x_1, x_2)^T$, find its coordinates with respect to $\mathbf{u}_1$ and $\mathbf{u}_2$

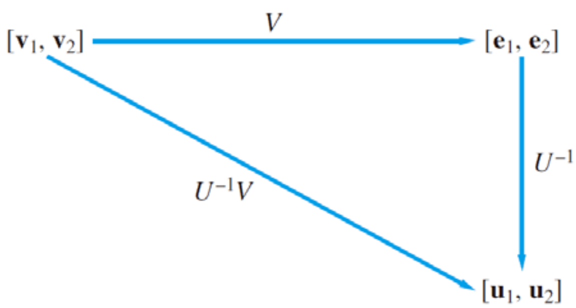
- Since U is nonsingular (why?) $\Rightarrow$ $\mathbf{c} = U^{-1}\,\mathbf{x}$
- U^{-1} is the **transition matrix** from $[\mathbf{e}_1, \mathbf{e}_2]$ to $[\mathbf{u}_1, \mathbf{u}_2]$

Example: Let $\mathbf{u}_1 = (3, 2)^T$, $\mathbf{u}_2 = (1, 1)^T$, $\mathbf{x} = (7, 4)^T$, find the coordinates of $\mathbf{x}$ with respect to $\mathbf{u}_1$ and $\mathbf{u}_2$

Example: Let $\mathbf{b}_1 = (1, -1)^T$, $\mathbf{b}_2 = (-2, 3)^T$. Find the transition matrix from $[e_1, e_2]$ to $[b_1, b_2]$, and the coordinates of $\mathbf{x} = (1, 2)^T$ with respect to $[\mathbf{b}_1, \mathbf{b}_2]$

- Assume a given vector $\mathbf{x}$, its coordinates with respect to $\{\mathbf{v}_1, \mathbf{v}_2\}$ are known: $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Find scalars d_1 and d_2 so that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = d_1\mathbf{u}_1 + d_2\mathbf{u}_2$. If we set $V = (\mathbf{v}_1, \mathbf{v}_2)$ and $U = (\mathbf{u}_1, \mathbf{u}_2)$, then

  $V\mathbf{c} = U\mathbf{d}$ and $\mathbf{d} = U^{-1}V\mathbf{c}$
- $U^{-1}V$ is the transition matrix from $[\mathbf{v}_1, \mathbf{v}_2]$ to $[\mathbf{u}_1, \mathbf{u}_2]$



- Let V be a vector space and let $E = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ be an ordered basis for V. For any $\mathbf{v} \in V$, then $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n$ for scalar $c_1, c_2, \ldots, c_n$. The vector $\mathbf{c}$ defined in this way is called the **coordinate vector** of $\mathbf{v}$ with respect to the ordered basis E and is denoted $[\mathbf{v}]_E$. The c_i's are called the **coordinates** of $\mathbf{v}$ relative to E.

Example: Let $E=[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]=[(1, 1, 1)^T, (2, 3, 2)^T, (1, 5, 4)^T]$, $F=[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3]=[(1, 1, 0)^T, (1, 2, 0)^T, (1, 2, 1)^T]$,

(1) Find the transition matrix from E to F.

(2) If $\mathbf{x} = 3\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3$ and $\mathbf{y} = \mathbf{v}_1 - 3\mathbf{v}_2 + 2\mathbf{v}_3$

Find the coordinates of $\mathbf{x}$ and $\mathbf{y}$ with respect to the ordered basis F.

3.6 Row Space and Column Space

這一章回到矩陣，把矩陣 A 當成是向量的集合，限定義矩陣的稚（就是行向量空間的維度），接著由行向量空間的性質看線性系統是否有解，最後，導出「線性代數基本定理」：稚加上 null space 的維度等於整體空間的維度。

- If A is an $m \times n$ matrix, the m vectors in $R^{1\times n}$ corresponding to the rows of A is referred to as the **row vectors of** A and the n vectors in R^m corresponding to the columns of A is referred to as the **column vectors of** A**.**
- The **rank** of a matrix A is the dimension of the row space of *A.*

- **Theorem 3.6.1**: Two row equivalent matrices have the same row space.

Proof:

- To determine the rank of a matrix, we can reduce the matrix to row echelon form. The nonzero rows of the row echelon matrix will form a basis for the row space.

Example:

$$A=\begin{bmatrix}1 & -2 & 3\\ 2 & -5 & 1\\ 1 & -4 & -7\end{bmatrix}\Rightarrow\begin{bmatrix}1 & -2 & 3\\ 0 & 1 & 5\\ 0 & 0 & 0\end{bmatrix}=U$$

$\Rightarrow$ (1, –2, 3) and (0, 1, 5) will form a basis for the row space of $U \Rightarrow$ rank(A) = rank(U) = 2

- **Theorem 3.6.2** A linear system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A.

Proof:

- **Theorem 3.6.3**: Let A be an $m \times n$ matrix. The linear system $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b} \in R^m$ iff the column vectors of A span R^m. The system $A\mathbf{x} = \mathbf{b}$ has at most one solution for every $\mathbf{b} \in R^m$ iff the column vector of A are linearly independent.

Proof:

- **Corollary 3.6.4**: An $n \times n$ matrix A is nonsingular if and only if the column vectors of A form a basis for R^n.

Proof:

- The dimension of the nullspace of a matrix is called the **nullity** of the matrix (dim $N(A)$).

- **Theorem 3.6.5 (fundamental theorem of linear algebra):** If A is an $m \times n$ matrix, then the rank of A plus the nullity of A equals n.

Proof: Let U be the row echelon form of A. Rank(A) = r = the number of nonzero rows in U (r lead variables) Nullity of A = the number of free variables = $n - r$

Example: Find a basis for the row space of A and a basis for N(A). Verify that dim N(A) = n-r.

$$\begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 4 & -3 & 0 \\ 1 & 2 & 1 & 5 \end{bmatrix}$$

- The reduced row echelon form of A is U =
- Thus, {______________________________} is a basis for the row space of A and rank(A) = __
- $x_1 + 2x_2 + \quad 3x_4 = 0$

  $x_3 + 2x_4 = 0$

  lead variable: x_1, $x_3 \Rightarrow$ rank = 2 and free variable: x_2, $x_4 \Rightarrow$ dim $N(A)$ = 2
- Let $x_2 = \alpha$, $x_4 = \beta$, then $\Rightarrow (-2, 1, 0, 0)^T$ and $(-3, 0, -2, 1)^T$ form a basis for $N(A) \Rightarrow$ dim $N(A) = 2 = n - r = 4 - 2$.

- If U is the row echelon form of A, then A and U have the same row space. But A and U have the different column space, since $A\mathbf{x} = \mathbf{0}$ if and only if $U\mathbf{x} = \mathbf{0}$, their column vectors satisfy the same dependency relations.

- **Theorem 3.6.6**: If A is an $m \times n$ matrix, the dimension of the row space of A equals the dimension of the column space of A.

Proof:

- To find the column space of A, we can use the row echelon form U of A by determining the columns of U that corresponds to the lead 1's. These same columns of A will be linearly independent and form a basis for the column space of A.

Example: Let $A = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5)$

The row echelon form of A is

$$\begin{bmatrix} 1 & -2 & 1 & 1 & 2 \\ -1 & 3 & 0 & 2 & -2 \\ 0 & 1 & 1 & 3 & 4 \\ 1 & 2 & 5 & 13 & 5 \end{bmatrix}$$

Chapter 4. Linear Transformations

矩陣在 3.6 被視為向量的集合，但是在這一章變成了另外一個腳色，運算元 operator，可以將向量或是基底做線性轉換，線性轉換有很多應用，對於 Computer Graphics 有興趣的同學要多注意。

4.1 Definition and Examples

定義何謂一般性的線性轉換。

- A mapping L from a vector space V into a vector Space W is said to be a **linear transformation** if

$$L(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2) \qquad (1)$$

  for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and for all scalars α and β.

- If L is a linear transformation mapping a vector space V into W, from (1) we get

$$\boldsymbol{L}(\mathbf{v}_1 + \mathbf{v}_2) = \boldsymbol{L}(\mathbf{v}_1) + \boldsymbol{L}(\mathbf{v}_2) \qquad (\alpha = \beta = 1) \qquad (2)$$

  and

$$\boldsymbol{L}(\alpha\mathbf{v}) = \alpha\boldsymbol{L}(\mathbf{v}) \qquad (\mathbf{v} = \mathbf{v}_1, \beta = 0) \qquad (3)$$

  Conversely, if L satisfies (2) and (3), then $L(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = L(\alpha\mathbf{v}_1) + L(\beta\mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2)$.

  L is a linear transformation if and only if L satisfies (2) and (3).

Example: Let L be the operator defined by $L(\mathbf{x}) = (x_1, -x_2)^T$ for each $\mathbf{x} = (x_1, x_2)^T$ in R^2. For each $\mathbf{x} = (x_1, x_2)^T$ and $\mathbf{y} = (y_1, y_2)^T$

$$L(\alpha\mathbf{x} + \beta\mathbf{y}) = L\left(\alpha\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \beta\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \boldsymbol{L}\left(\begin{bmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \end{bmatrix}\right) = \begin{bmatrix} \alpha x_1 + \beta y_1 \\ -(\alpha x_2 + \beta y_2) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha x_1 \\ -\alpha x_2 \end{bmatrix} + \begin{bmatrix} \beta y_1 \\ -\beta y_2 \end{bmatrix} = \alpha\begin{bmatrix} x_1 \\ -x_2 \end{bmatrix} + \beta\begin{bmatrix} y_1 \\ -y_2 \end{bmatrix} = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$$

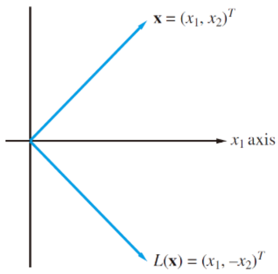


Figure 4.1.3.

$\Rightarrow$ L is a linear operator.

Example: Consider the mapping M defined by $M(\mathbf{x}) = (x_1^2 + x_2^2)^{1/2}$.

Example: L: $R^2 \rightarrow R^3$ defined by $L(\mathbf{x}) = (x_2, x_1, x_1+x_2)^T$.

- In general, if A is any $m \times n$ matrix, we can define a linear transformation L_A from R^n to R^m by

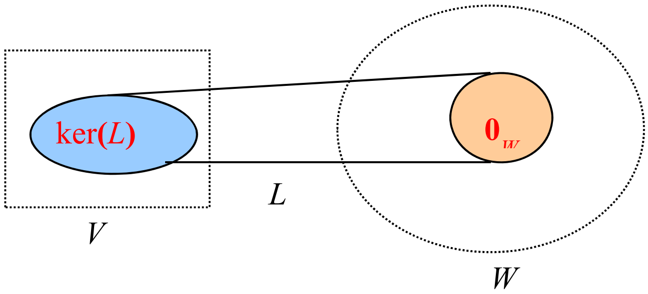
$$L_A(\mathbf{x}) = A\mathbf{x} \text{ for each } \mathbf{x} \in R^n$$

  The transformation L_A is linear since $L_A(\alpha\mathbf{x} + \beta\mathbf{y}) = A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A\mathbf{x} + \beta A\mathbf{y} = \alpha L_A(\mathbf{x}) + \beta L_A(\mathbf{y})$

  We can think of each $m \times n$ matrix as defining a linear transformation from R^n to R^m.

- If L is a linear transformation mapping a vector space V into a vector space W, then
  (1) $L(\mathbf{0}_V) = \mathbf{0}_W$ (where $\mathbf{0}_V$ and $\mathbf{0}_W$ are zero vectors in V and W)
  (2) $L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n) = \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \ldots + \alpha_n L(\mathbf{v}_n)$
  (3) $L(-\mathbf{v}) = -L(\mathbf{v})$ for all $\mathbf{v} \in V$.
- Let $L: V \to W$ be a linear transformation. The **kernel** of L, denoted ker(L), is define by

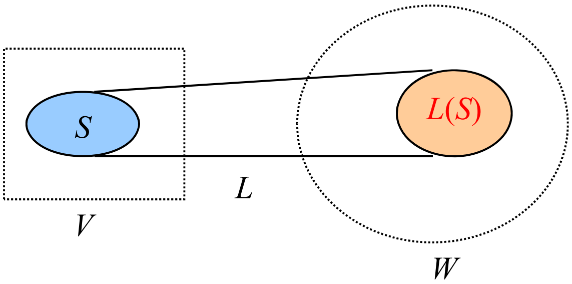
  $$\ker(L) = \{\mathbf{v} \in V \mid L(\mathbf{v}) = \mathbf{0}_W\}$$

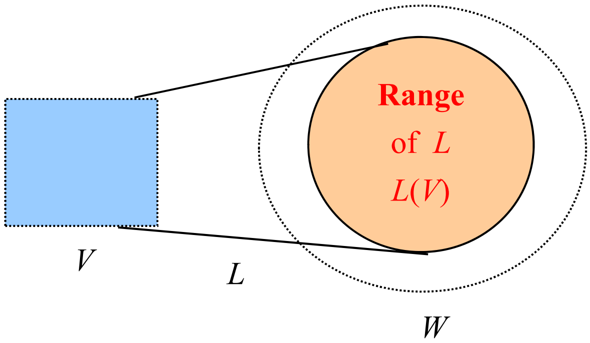


- Let $L: V \to W$ be a linear transformation and let S be a subspace of V.
  The **image** of S, denoted $L(S)$, is defined by

  $$L(S) = \{\mathbf{w} \in W \mid \mathbf{w} = L(\mathbf{v}) \text{ for some } \mathbf{v} \in S\}$$

  The **image** of the entire vector space, $L(V)$, is called the **range** of L.





- **Theorem 4.1.1**: If $L: V \to W$ is a linear transformation and S is a subspace of V, then
  (1) ker(L) is a subspace of V
  (2) $L(S)$ is a subspace of W

Proof: (1) by C_1: If $\mathbf{v} \in \ker(L)$ and α is a scalar $\Rightarrow L(\mathbf{v}) = \mathbf{0}_W$

$L(\alpha\mathbf{v}) = \alpha L(\mathbf{v}) = \alpha\mathbf{0}_W = \mathbf{0}$ $\Rightarrow \alpha\mathbf{v} \in \ker(L)$

by C_2: If $\mathbf{v}_1$ and $\mathbf{v}_2 \in \ker(L) \Rightarrow L(\mathbf{v}_1) = L(\mathbf{v}_2) = \mathbf{0}_W$, $L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W$

$\Rightarrow \mathbf{v}_1 + \mathbf{v}_2 \in \ker(L)$.

ker(L) is a subspace of V.

(2) by C_1: If $\mathbf{w} \in L(S)$, then $\mathbf{w} = L(\mathbf{v})$ for some $\mathbf{v} \in S$, $\alpha\mathbf{w} = \alpha L(\mathbf{v}) = L(\alpha\mathbf{v})$

Since S is a subspace $\Rightarrow \alpha\mathbf{v} \in S \Rightarrow \alpha\mathbf{w} \in L(S)$

by C_2: If $\mathbf{w}_1$ and $\mathbf{w}_2 \in L(S)$, then there exist $\mathbf{v}_1$ and $\mathbf{v}_2 \in S$ such that $L(\mathbf{v}_1) = \mathbf{w}_1$ and $L(\mathbf{v}_2) = \mathbf{w}_2$

$\Rightarrow \mathbf{w}_1+\mathbf{w}_2 = L(\mathbf{v}_1) + L(\mathbf{v}_2) = L(\mathbf{v}_1+\mathbf{v}_2)$

Since S is a subspace $\Rightarrow \mathbf{v}_1+\mathbf{v}_2 \in S \Rightarrow \mathbf{w}_1+\mathbf{w}_2 \in L(S)$

$L(S)$ is a subspace of L.

4.2 Matrix Representations of Linear Transformations

所有的線性變換居然都可以用一個矩陣來表示，反之，一個矩陣就是一個線性變換，像是旋轉、投影等。對於一般性的線性空間也適用。基底變換那邊有點混亂，可以參考 3.5 的內容，基本上是一樣的。

- Each $m \times n$ matrix A defines a linear transformation L_A from R^n to R^m: $L_A(\mathbf{x}) = A\mathbf{x}$ for each $\mathbf{x} \in R^n$.
- For each linear transformation L mapping R^n into R^m there is an $m \times n$ matrix A such that $L(\mathbf{x}) = A\mathbf{x}$

- **Theorem 4.2.1**: If L is a linear transformation mapping R^n into R^m, there is an $m \times n$ matrix A such that $L(\mathbf{x}) = A\mathbf{x}$ for each $\mathbf{x} \in R^n$. In fact, the jth column vector of A is given by $\mathbf{a}_j = L(\mathbf{e}_j)$, $j = 1, 2, \ldots, n$

Proof: If $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$ is any element of R^n:

$L(\mathbf{x}) = L(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n = x_1L(\mathbf{e}_1) + x_2L(\mathbf{e}_2) + \ldots + x_nL(\mathbf{e}_n) = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a} = A\mathbf{x}$

Example: L: $R^3 \to R^2$ defined by $L(\mathbf{x}) = (x_1 + x_2, x_2 + x_3)^T$ for each $\mathbf{x} = (x_1, x_2, x_3)^T$ in R^3.

Let $L(\mathbf{x}) = A\mathbf{x}$

$$\mathbf{a}_1 = L(\mathbf{e}_1) = L(\begin{bmatrix}1\\0\\0\end{bmatrix}) = \begin{bmatrix}1+0\\0+0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} \quad \mathbf{a}_2 = L(\mathbf{e}_2) = L(\begin{bmatrix}0\\1\\0\end{bmatrix}) = \begin{bmatrix}0+1\\1+0\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix} \quad \mathbf{a}_3 = L(\mathbf{e}_3) = L(\begin{bmatrix}0\\0\\1\end{bmatrix}) = \begin{bmatrix}0+0\\0+1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}$$

$$A = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = \begin{bmatrix}1 & 1 & 0\\0 & 1 & 1\end{bmatrix}, \quad A\mathbf{x} = \begin{bmatrix}1 & 1 & 0\\0 & 1 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}x_1 + x_2\\x_2 + x_3\end{bmatrix}$$

Example: Let L: $R^2 \to R^2$ which rotates each vector by an angle θ in the counterclockwise direction.

Let $L(\mathbf{x}) = A\mathbf{x}$. Since

$\mathbf{a}_1 = L(\mathbf{e}_1) = (\cos\theta, \sin\theta)^T$, and

$\mathbf{a}_2 = L(\mathbf{e}_2) = (-\sin\theta, \cos\theta)^T$

$$A = (\mathbf{a}_1, \mathbf{a}_2) = \begin{bmatrix}\cos\ \theta & -\sin\ \theta\\ \sin\ \theta & \cos\ \theta\end{bmatrix}.$$

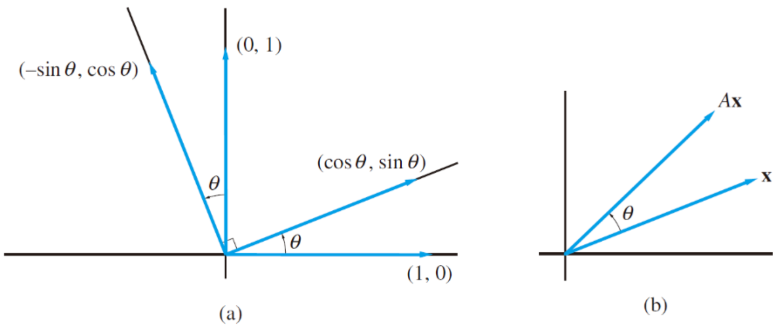


Figure 4.2.1.

- Question: How to find a similar representation for linear transformations from an n-dimensional vector space V into an m-dimensional vector space W?
- Let $E = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $F = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ be ordered bases for vector spaces V and W, and L be a linear transformation mapping V into W. If $\mathbf{v} \in V$, then $\mathbf{v}$ can be expressed in terms of the basis E:

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n$$

- There exists an $m \times n$ matrix A representing the linear transformation L:

$$A\mathbf{x} = \mathbf{y} \text{ iff } L(\mathbf{v}) = y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \ldots + y_m\mathbf{w}_m$$

- If $\mathbf{x}$ is the coordinate vector of $\mathbf{v}$ w. r. t. (with respect to) E, then the coordinate vector of $L(\mathbf{v}) = \mathbf{y}$ w. r. t. F is given by:

$$\mathbf{y} = [L(\mathbf{v})]_F = A\mathbf{x}$$

- **Question**: How to determine the matrix representation A?
- Let $\mathbf{a}_j = (a_{1j}, a_{2j}, \ldots, a_{mj})^T$ be the <u>coordinate vector</u> of $L(\mathbf{v}_j)$ w. r. t. $F = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$, $j = 1, 2, \ldots, n$

  $L(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \ldots + a_{mj}\mathbf{w}_m \quad 1 \le j \le n$
- Let $A = (a_{ij}) = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$. If $\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n$ then

$$L(\mathbf{v}) = L(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n) = x_1L(\mathbf{v}_1) + x_2L(\mathbf{v}_2) + \ldots + x_nL(\mathbf{v}_n)$$

$$= L\left(\sum_{j=1}^{n} x_j\mathbf{v}_j\right) = \sum_{j=1}^{n}\left(x_j\, L(\mathbf{v}_j)\right) = \sum_{j=1}^{n}\left(x_j\left(\sum_{i=1}^{m} a_{ij}\mathbf{w}_i\right)\right) = \sum_{i=1}^{m}\left(\left(\sum_{j=1}^{n} a_{ij}x_j\right)\mathbf{w}_i\right)$$

Let $y_i = \sum_{j=1}^{n} a_{ij}x_j = \mathrm{a}(i,:)^T\mathbf{x}$, for $i = 1, 2, \ldots, m$

$\mathbf{y} = (y_1, y_2, \ldots, y_m)^{\mathrm{T}} = A\mathbf{x}$ is the coordinate vector of $L(\mathbf{v})$ w. r. t. $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$.

- **Theorem 4.2.2** If $E = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $F = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ are ordered bases for vector spaces V and W, respectively, then corresponding to each linear transformation $L: V \to W$ there is an $m \times n$ matrix A such that $[L(\mathbf{v})]_F = A\,[\mathbf{v}]_E$ for each $\mathbf{v} \in V$. A is the matrix representing L relative to the ordered bases E and F. In fact, $\mathbf{a}_j = [L(\mathbf{v}_j)]_F$ $j = 1, 2, \ldots, n$

- $\mathbf{x} = [\mathbf{v}]_E$: the coordinate vector of $\mathbf{v}$ with respect to E

  $\mathbf{y} = [\mathbf{w}]_F$: the coordinate vector of $\mathbf{w}$ with respect to F

  $\Rightarrow$ L maps $\mathbf{v}$ into $\mathbf{w}$ iff A maps $\mathbf{x}$ into $\mathbf{y}$

$$\begin{tikzcd}
\mathbf{v} \in V \arrow[r, "L = L_A"] & \mathbf{w} = L(\mathbf{v}) \in W \\
\mathbf{x} = [\mathbf{v}]_E \in R^n \arrow[u] \arrow[r, "A"] & A\mathbf{x} = [\mathbf{w}]_F \in R^m \arrow[u]
\end{tikzcd}$$

Figure 4.2.2.

Example: $L: R^3 \to R^2$ defined by $L(\mathbf{x}) = x_1\mathbf{b}_1 + (x_2 + x_3)\mathbf{b}_2$ for each $\mathbf{x} \in R^3$, where

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Find the matrix A representing L w. r. t. the ordered bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{b}_1, \mathbf{b}_2\}$.

- **Theorem 4.2.3**: Let $E = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $F = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m\}$ be ordered bases for R^n and R^m, respectively. If $L: R^n \to R^m$ is a linear transformation and A is the matrix representing L with respect to E and F, then $\mathbf{a}_j = B^{-1}L(\mathbf{u}_j)$ for $j = 1, 2, \ldots, n$ where $B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m)$.

Proof: If A is representing L with respect to E and F, then for $j = 1, 2, \ldots, n$

$$L(\mathbf{u}_j) = a_{1j}\mathbf{b}_1 + a_{2j}\mathbf{b}_2 + \ldots + a_{-mj}\mathbf{b}_m \; (\underline{\textbf{Note}}: \mathbf{a}_j = [L(\mathbf{u}_j)]_F) = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m)\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = B\mathbf{a}_j$$

$\Rightarrow \mathbf{a}_j = B^{-1}L(\mathbf{u}_j)$ for $j = 1, 2, \ldots, n$

4.3 Similarity

在不同的基底（座標系統）下，同一個線性變換有不同的變換矩陣，我們稱這些變換矩陣彼此「相似」。相似的概念在第六章會發揮很大的作用。在這邊先了解一下他們之間如何轉換。

Example: Let L be the linear transformation mapping R^2 into itself defined by$L(\mathbf{x}) = (2x_1, x_1 + x_2)^T$

Since $L(\mathbf{e}_1) = L\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}2\\1\end{bmatrix}$ and $L(\mathbf{e}_2) = L\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\1\end{bmatrix}$. The matrix representing L w. r. t. $\{\mathbf{e}_1, \mathbf{e}_2\}$ is $A = \begin{bmatrix}2 & 0\\1 & 1\end{bmatrix}$

- If we use $\mathbf{u}_1 = \begin{bmatrix}1\\1\end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix}-1\\1\end{bmatrix}$ as the basis for R2, then
  $L(\mathbf{u}_1) = A\mathbf{u}_1 = \begin{bmatrix}2 & 0\\1 & 1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}2\\2\end{bmatrix}$ $L(\mathbf{u}_2) = A\mathbf{u}_2 = \begin{bmatrix}2 & 0\\1 & 1\end{bmatrix}\begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}-2\\0\end{bmatrix}$
- Since the transition matrix from $\{\mathbf{u}_1, \mathbf{u}_2\}$ to $\{\mathbf{e}_1, \mathbf{e}_2\}$ is $U = (\mathbf{u}_1, \mathbf{u}_2) = \begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}$
  the transition matrix from $\{\mathbf{e}_1, \mathbf{e}_2\}$ to $\{\mathbf{u}_1, \mathbf{u}_2\}$ is $U^{-1} = \begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\-\frac{1}{2} & \frac{1}{2}\end{bmatrix}$
- Then, the coordinates of $L(\mathbf{u}_1)$ and $L(\mathbf{u}_2)$ w. r. t. $\{\mathbf{u}_1, \mathbf{u}_2\}$ is

  $$U^{-1}L(\mathbf{u}_1) = \begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\-\frac{1}{2} & \frac{1}{2}\end{bmatrix}\begin{bmatrix}2\\2\end{bmatrix} = \begin{bmatrix}2\\0\end{bmatrix} \quad U^{-1}L(\mathbf{u}_2) = \begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\-\frac{1}{2} & \frac{1}{2}\end{bmatrix}\begin{bmatrix}-2\\0\end{bmatrix} = \begin{bmatrix}-1\\1\end{bmatrix}$$

  $\Rightarrow$ $L(\mathbf{u}_1) = 2\mathbf{u}_1 + 0\mathbf{u}_2$, $L(\mathbf{u}_2) = -1\mathbf{u}_1 + 1\mathbf{u}_2$

  $\Rightarrow$ The matrix representing L w. r. t. $\{\mathbf{u}_1, \mathbf{u}_2\}$ is $B = \begin{bmatrix}2 & -1\\0 & 1\end{bmatrix}$
- Since $(2, 0)^T = U^{-1}L(\mathbf{u}_1) = U^{-1}A\mathbf{u}_1$ and $(-1, 1)^T = U^{-1}L(\mathbf{u}_2) = U^{-1}A\mathbf{u}_2$.
  Hence, $B = (U^{-1}A\mathbf{u}_1, U^{-1}A\mathbf{u}_2) = U^{-1}A(\mathbf{u}_1, \mathbf{u}_2) = U^{-1}AU$

- Conclusion: If (i) B is the matrix representing L w. r. t. $\{\mathbf{u}_1, \mathbf{u}_2\}$
  (ii) A is the matrix representing L w. r. t. $\{\mathbf{e}_1, \mathbf{e}_2\}$
  (iii) U is the transition matrix from $\{\mathbf{u}_1, \mathbf{u}_2\}$ to $\{\mathbf{e}_1, \mathbf{e}_2\}$
  then $B = U^{-1}AU$.

- **Theorem 4.3.1**: Let $E = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $F = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ be two ordered bases for a vector space V and let L be a linear operator on V. Let S be the transition matrix representing the change from F to E. If A is the matrix representing L w. r. t. E and B is the matrix representing L w. r. t. F, then $B = S^{-1}AS$.

Basis E: $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$

$$\begin{tikzcd}
\mathbf{y} \arrow[r, "A"] & \mathbf{t} \arrow[d, "S^{-1}"] \\
\mathbf{x} \arrow[u, "S"] \arrow[r, "B"'] & \mathbf{z}
\end{tikzcd}$$

Basis F: $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$

Proof: Let $\mathbf{x}$ be any vector in R^n and let $\mathbf{x} = [\mathbf{v}]_F$

$\mathbf{v} = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \ldots + x_n\mathbf{w}_n$

Since S is the transition matrix representing the change from F to E. Let

$\mathbf{y} = S\mathbf{x}$, $\mathbf{t} = A\mathbf{y}$, $\mathbf{z} = B\mathbf{x}$

Let $\mathbf{x}$ be any vector in R^n and let $\mathbf{y} = [\mathbf{v}]_E \Rightarrow \mathbf{v} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \ldots + y_n\mathbf{v}_n$

Since A represents L w. r. t. E and B represents L w. r. t. F, we have $\mathbf{t} = [L(\mathbf{v})]_E$ and $\mathbf{z} = [L(\mathbf{v})]_F$

Since the transition from E to F is S^{-1}, $S^{-1}\mathbf{t} = \mathbf{z}$ $\Rightarrow$ $S^{-1}\mathbf{t} = S^{-1}A\mathbf{y} = S^{-1}AS\mathbf{x} = \mathbf{z} = B\mathbf{x}$ $\Rightarrow S^{-1}AS = B$

- Let A and B be two $n \times n$ matrices. B is said to be **similar** to A if there exists a nonsingular matrix S such that $B = S^{-1}AS$.

Example: L: $R^3 \to R^3$ defined by $L(\mathbf{x}) = A\mathbf{x}$, where $A = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix}$

Thus the matrix A represents L with respect to $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, find the matrix representing L with respect to $\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3\}$ where $\mathbf{y}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{y}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \mathbf{y}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.