Chapter 3

Week3

3.1. Monday for MAT3040

Reviewing.

1. Complementation. Suppose $\dim(V) = n < \infty$, then $W \leq V$ implies that there exists W' such that
$$W \oplus W' = V.$$
2. Given the linear transformation $T : V \to W$, define the set $\ker(T)$ and $\mathrm{Im}(T)$.
3. Isomorphism of vector spaces: $T : V \cong W$
4. Rank-Nullity Theorem

3.1.1. Remarks on Isomorphism

Proposition 3.1 If $T : V \to W$ is an isomorphism, then

1. the set $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ is linearly independent in V if and only if $\{T\boldsymbol{v}_1, \dots, T\boldsymbol{v}_k\}$ is linearly independent.
2. The same goes if we replace the linearly independence by spans.
3. If $\dim(V) = n$, then $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ forms a basis of V if and only if $\{T\boldsymbol{v}_1, \dots, T\boldsymbol{v}_n\}$ forms a basis of W. In particular, $\dim(V) = \dim(W)$.
4. Two vector spaces with finite dimensions are isomorphic if and only if they have the same dimension:

Proof. It suffices to show the reverse direction. Let $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ be two basis of V, W, respectively. Define the linear transformation $T : V \to W$ by

$$T(a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n) = a_1\boldsymbol{w}_1 + \cdots + a_n\boldsymbol{w}_n$$

Then T is surjective since $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ spans W; T is injective since $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ is linearly independent. ■

3.1.2. Change of Basis and Matrix Representation

Definition 3.1 [Coordinate Vector] Let V be a finite dimensional vector space and $B = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ an **ordered** basis of V. Any vector $\boldsymbol{v} \in V$ can be uniquely written as

$$\boldsymbol{v} = \alpha_1\boldsymbol{v}_1 + \cdots + \alpha_n\boldsymbol{v}_n,$$

Therefore we define the map $[\cdot]_{\mathcal{B}} : V \to \mathbb{F}^n$, which maps any vector in $\boldsymbol{v}$ into its **coordinate vector**:

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

R Note that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and $\{\boldsymbol{v}_2, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ are distinct ordered basis.

■ **Example 3.1** Given $V = M_{2\times 2}(\mathbb{F})$ and the ordered basis

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

Any matrix has the coordinate vector w.r.t. $\mathcal{B}$, i.e.,

$$\left[\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 4 \\ 2 \\ 3 \end{pmatrix}$$

However, if given another ordered basis

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \right\},$$

the matrix may have the different coordinate vector w.r.t. $\mathcal{B}_1$:

$$\left[\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\right]_{\mathcal{B}_1} = \begin{pmatrix} 4 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

Theorem 3.1 The mapping $[\cdot]_{\mathcal{B}} : V \to \mathbb{F}^n$ is an isomorphism.

Proof. 1. First show the operator $[\cdot]_{\mathcal{B}}$ is well-defined, i.e., the same input gives the same output. Suppose that

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \quad [\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1' \\ \vdots \\ \alpha_n' \end{pmatrix},$$

then we imply

$$\begin{aligned} \boldsymbol{v} &= \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n \\ &= \alpha_1' \boldsymbol{v}_1 + \cdots + \alpha_n' \boldsymbol{v}_n. \end{aligned}$$

By the uniqueness of coordinates, we imply $\alpha_i = \alpha_i'$ for $i = 1, \dots, n$.

2. It's clear that the operator $[\cdot]_{\mathcal{B}}$ is a linear transformation, i.e.,

$$[p\boldsymbol{v} + q\boldsymbol{w}]_{\mathcal{B}} = p[\boldsymbol{v}]_{\mathcal{B}} + q[\boldsymbol{w}]_{\mathcal{B}} \quad \forall p, q \in \mathbb{F}$$

3. The operator $[\cdot]_{\mathcal{B}}$ is surjective:

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \implies \boldsymbol{v} = 0\boldsymbol{v}_1 + \cdots + 0\boldsymbol{v}_n = \mathbf{0}.$$

4. The injective is clear, i.e., $[\boldsymbol{v}]_{\mathcal{B}} = [\boldsymbol{w}]_{\mathcal{B}}$ implies $\boldsymbol{v} = \boldsymbol{w}$.

Therefore, $[\cdot]_{\mathcal{B}}$ is an isomorphism. ■

We can use the Theorem (3.1) to simplify computations in vector spaces:

■ **Example 3.2** Given a vector sapce $V = P_3[x]$ and its basis $B = \{1, x, x^2, x^3\}$.

To check if the set $\{1 + x^2, 3 - x^3, x - x^3\}$ is linearly independent, by part (1) in Proposition (3.1) and Theorem (3.1), it suffices to check whether the corresponding coordinate vectors

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \right\}$$

is linearly independent, i.e., do Gaussian Elimination and check the number of pivots. ■

Here gives rise to the question: if $\mathcal{B}_1, \mathcal{B}_2$ form two basis of V, then how are $[\boldsymbol{v}]_{\mathcal{B}_1}, [\boldsymbol{v}]_{\mathcal{B}_2}$ related to each other?

Here we consider an easy example first:

■ **Example 3.3** Consider $V = \mathbb{R}^n$ and its basis $\mathcal{B}_1 = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$. For any $\boldsymbol{v} \in V$,

$$\boldsymbol{v} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_n \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n \implies [\boldsymbol{v}]_{\mathcal{B}_1} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Also, we can construct a different basis of V:

$$\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\},$$

which gives a different coordinate vector of $\boldsymbol{v}$:

$$[\boldsymbol{v}]_{\mathcal{B}_2} = \begin{pmatrix} \alpha_1 - \alpha_2 \\ \alpha_2 - \alpha_3 \\ \vdots \\ \alpha_{n-1} - \alpha_n \\ \alpha_n \end{pmatrix}$$

■

Proposition 3.2 — Change of Basis. Let $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\mathcal{A}' = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ be two ordered basis of a vector space V. Define the **change of basis** matrix from $\mathcal{A}$ to $\mathcal{A}'$, say $\mathcal{C}_{\mathcal{A}',\mathcal{A}} := [\alpha_{ij}]$, where

$$\boldsymbol{v}_j = \sum_{i=1}^{m} \alpha_{ij} \boldsymbol{w}_i$$

Then for any vector $\boldsymbol{v} \in V$, the *change of basis amounts to left-multiplying the change of basis matrix*:

$$\mathcal{C}_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}]_A = [\boldsymbol{v}]_{A'} \tag{3.1}$$

Define matrix $\mathcal{C}_{\mathcal{A},\mathcal{A}'} := [\beta_{ij}]$, where

$$\boldsymbol{w}_j = \sum_{i=1}^{n} \beta_{ij}\boldsymbol{v}_i$$

Then we imply that

$$(\mathcal{C}_{\mathcal{A},\mathcal{A}'})^{-1} = \mathcal{C}_{\mathcal{A}',\mathcal{A}}$$

Proof. 1. First show (3.1) holds for $\boldsymbol{v} = \boldsymbol{v}_j$, $j = 1, \dots, n$:

$$\text{LHS of (3.1)} = [\alpha_{ij}]\boldsymbol{e}_j = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

$$\text{RHS of (3.1)} = [\boldsymbol{v}_j]_{\mathcal{A}'} = \left[\sum_{i=1}^{n} \alpha_i \boldsymbol{w}_i\right]_{\mathcal{A}'} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

Therefore,

$$\mathcal{C}_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}_j]_{\mathcal{A}} = [\boldsymbol{v}_j]_{\mathcal{A}'}, \quad \forall j = 1, \dots, n. \tag{3.2}$$

2. Then for any $\boldsymbol{v} \in V$, we imply $\boldsymbol{v} = r_1\boldsymbol{v}_1 + \cdots + r_n\boldsymbol{v}_n$, which implies that

$$\begin{aligned}
\mathcal{C}_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}]_{\mathcal{A}} &= \mathcal{C}_{\mathcal{A}',\mathcal{A}}[r_1\boldsymbol{v}_1 + \cdots + r_n\boldsymbol{v}_n]_{\mathcal{A}} && \text{(3.3a)} \\
&= \mathcal{C}_{\mathcal{A}',\mathcal{A}}\left(r_1[\boldsymbol{v}_1]_{\mathcal{A}} + \cdots + r_n[\boldsymbol{v}_n]_{\mathcal{A}}\right) && \text{(3.3b)} \\
&= \sum_{j=1}^{n} r_j \mathcal{C}_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}_j]_{\mathcal{A}} && \text{(3.3c)} \\
&= \sum_{j=1}^{n} r_j [\boldsymbol{v}_j]_{\mathcal{A}'} && \text{(3.3d)} \\
&= \left[\sum_{j=1}^{n} r_j \boldsymbol{v}_j\right]_{\mathcal{A}'} && \text{(3.3e)} \\
&= [\boldsymbol{v}]_{\mathcal{A}'} && \text{(3.3f)}
\end{aligned}$$

where (3.3a) and (3.3e) is by applying the lineaity of $[\cdot]_{\mathcal{A}}$ and $[\cdot]_{\mathcal{A}'}$; (3.3d) is by applying the result (3.12). Therefore (3.1) is shown for $\forall \boldsymbol{v} \in V$.

3. Now we show that $(\mathcal{C}_{\mathcal{A}\mathcal{A}'}\mathcal{C}_{\mathcal{A}'\mathcal{A}}) = \boldsymbol{I}_n$. Note that

$$\begin{aligned}\boldsymbol{v}_j &= \sum_{i=1}^{n} \alpha_{ij}\boldsymbol{w}_i \\ &= \sum_{i=1}^{n} \alpha_{ij} \sum_{k=1}^{n} \beta_{ki}\boldsymbol{v}_k \\ &= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) \boldsymbol{v}_i\end{aligned}$$

By the uniqueness of coordinates, we imply

$$\left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) = \delta_{jk} := \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

By the matrix multiplication, the (k, j)-th entry for $\mathcal{C}_{\mathcal{A}\mathcal{A}'}\mathcal{C}_{\mathcal{A}'\mathcal{A}}$ is

$$[\mathcal{C}_{\mathcal{A}\mathcal{A}'}\mathcal{C}_{\mathcal{A}'\mathcal{A}}]_{kj} = \left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) = \delta_{jk} \implies (\mathcal{C}_{\mathcal{A}\mathcal{A}'}\mathcal{C}_{\mathcal{A}'\mathcal{A}}) = \boldsymbol{I}_n$$

Noew, suppose

$$\begin{aligned}\boldsymbol{v}_j &= \sum_{i=1}^{n} \alpha_{ij}\boldsymbol{w}_i \\ &= \sum_{i=1}^{n} \alpha_{ij} \sum_{k=1}^{n} \beta_{ki}\boldsymbol{v}_k \\ &= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) \boldsymbol{v}_i\end{aligned}$$

By the uniqueness of coordinates, we imply

$$\left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

where

$$\left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) = (\mathcal{C}_{AA'}\mathcal{C}_{A'A}).$$

Therefore, $(\mathcal{C}_{AA'}\mathcal{C}_{A'A}) = \boldsymbol{I}_n$. ■

■ **Example 3.4** Back to Example (3.3), write $\mathcal{B}_1, \mathcal{B}_2$ as

$$\mathcal{B}_1 = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}, \quad \mathcal{B}_2 = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$$

and therefore $\boldsymbol{w}_i = \boldsymbol{e}_1 + \cdots + \boldsymbol{e}_i$. The change of basis matrix is given by

$$C_{\mathcal{B}_1,\mathcal{B}_2} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

which implies that for $\boldsymbol{v}$ in the example,

$$C_{\mathcal{B}_1,\mathcal{B}_2}[\boldsymbol{v}]_{\mathcal{B}_2} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 - \alpha_2 \\ \vdots \\ \alpha_{n-1} - \alpha_n \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = [\boldsymbol{v}]_{\mathcal{B}_1}$$

Definition 3.2 Let $T : V \to W$ be a linear transformation, and

$$\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}, \quad \mathcal{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$$

be basis of V and W, respectively. The **matrix representation** of T with respect to (w.r.t.) $\mathcal{A}$ and $\mathcal{B}$ is defined as $(T)_{\mathcal{B}\mathcal{A}} \in M_{m\times m}(\mathbb{F})$, where

$$T(\boldsymbol{v}_j) = \sum_{i=1}^{m} \alpha_{ij}\boldsymbol{w}_j$$

3.2. Monday for MAT3006

Reviewing.

1. Compactness/Sequential Compactness:
   - Equivalence for metric space
   - Stronger than closed and bounded
2. Completeness:
   - The metric space (E,d) is complete if every Cauchy sequence on E is convergent.
   - $\mathbb{P}[a,b] \subseteq \mathcal{C}[a,b]$ is not complete:

$$f_N(x) = \sum_{n=0}^{N} (-1)^n \frac{x^{2n}}{(2n)!} \to \cos x,$$

   while $\cos x \notin \mathcal{P}[a,b]$.

3.2.1. Remarks on Completeness

Proposition 3.3 Let (X,d) be a metric space.

1. If X is complete and $E \subseteq X$ is closed, then E is complete.
2. If $E \subseteq X$ is complete, then E is closed in X.
3. If $E \subseteq X$ is compact, then E is complete.

Proof.

1. Every Cauchy sequence $\{e_n\}$ in $E \subseteq X$ is also a Cauchy sequence in E. Therefore we imply $\{e_n\} \to x \in X$, due to the completeness of X. Due to the closedness of E, the limit $x \in E$, i.e., E is complete.
2. Consider any convergent sequence $\{e_n\}$ in E, with some limit $x \in X$. We imply $\{e_n\}$ is Cauchy and thus $\{e_n\} \to e \in E$, due to the completeness of E. By the uniqueness of limits, we must have $x = z \in E$, i.e., E is closed.
3. Consider a Cauchy sequence $\{e_n\}$ in E. There exists a subsequence $\{e_{n_j}\} \to e \in E$, due to the sequential compactness of E.

It follows that for large n and j,

$$d(e_n,e) \overset{\text{(a)}}{\leq} d(e_n,e_{n_j}) + d(e_{n_j},e) \overset{\text{(b)}}{<} \varepsilon$$

where (a) is due to triangle inequality and (b) is due to the Cauchy property of $\{e_n\}$ and the convergence of $\{e_{n_j}\}$.

Therefore, we imply $\{e_n\} \to e \in E$, i.e., E is complete.

■

R Given any metric space that may not be necessarily complete, we can make the union of it with another space to make it complete, e.g., just like the completion from $\mathbb{Q}$ to $\mathbb{R}$.

3.2.2. Contraction Mapping Theorem

The motivation of the contraction mapping theorem comes from solving an equation $f(x)$. More precisely, such a problem can be turned into a problem for fixed points, i.e., it suffices to find the fixed points for $g(x)$, with $g(x) = f(x) + x$.

Definition 3.3 Let (X,d) be a metric space. A map $T : (X,d) \to (X,d)$ is a **contraction** if there exists a constant $\tau \in (0,1)$ such that

$$d(T(x),T(y)) < \tau \cdot d(x,y), \quad \forall x,y \in X$$

A point x is called a fixed point of T if $T(x) = x$.

R All contractions are continuous: Given any convergence sequence $\{x_n\} \to x$, for $\varepsilon > 0$, take N such that $d(x_n,x) < \frac{\varepsilon}{\tau}$ for $n \geq N$. It suffices to show the convergence of $\{T(x_n)\}$:

$$d(T(x_n),T(x)) < \tau \cdot T(x_n,x) < \tau \cdot \frac{\varepsilon}{\tau} = \varepsilon.$$

Therefore, the contraction is Lipschitz continuous with Lipschitz constant τ.

Theorem 3.2 — **Contraction Mapping Theorem / Banach Fixed Point Theorem.** Every contraction T in a **complete** metric space X has a unique fixed point.

■ **Example 3.5**
1. The mapping $f(x) = x + 1$ is not a contraction in $X = \mathbb{R}$, and it has no fixed point.
2. Consider an in-complete space $X = (0,1)$ and a contraction $f(x) = \frac{x+1}{2}$. It doesn't admit a fixed point on X as well.

■

Proof. Pick any $x_0 \in X$, and define a sequence recursively by setting $x_{n+1} = T(x_n)$ for $n \geq 0$.

1. Firstly show that the sequence $\{x_n\}$ is Cauchy.

   We can upper bound the term $d(T^n(x_0), T^{n-1}(x_0))$:

$$d(T^n(x_0), T^{n-1}(x_0)) \leq \tau d(T^{n-1}(x_0), T^{n-2}(x_0)) \leq \cdots \leq \tau^{n-1} d(T(x_0), x_0) \tag{3.4}$$

   Therefore for any $n \geq m$, where m is going to be specified later,

$$\begin{aligned}
d(x_n, x_m) &= d(T^n(x_0), T^m(x_0)) && \text{(3.5a)}\\
&\leq \tau d(T^{n-1}(x_0), T^{m-1}(x_0)) \leq \cdots \leq \tau^m d(T^{n-m}(x_0), x_0) && \text{(3.5b)}\\
&\leq \tau^m \sum_{j=1}^{n-m} \tau^{n-m-j} d(T(x_0), x_0) && \text{(3.5c)}\\
&< \frac{\tau^m}{1-\tau} d(T(x_0), x_0) && \text{(3.5d)}\\
&\leq \varepsilon && \text{(3.5e)}
\end{aligned}$$

   where (3.5b) is by repeatedly applying contraction property of d; (3.5c) is by applying the triangle inequality and (3.4); (3.5e) is by choosing sufficiently large m such that $\frac{\tau^m}{1-\tau} d(T(x_0), x_0) < \varepsilon$.

   Therefore, $\{x_n\}$ is Cauchy. By the completeness of X, we imply $\{x_n\} \to x \in X$.

2. Therefore, we imply

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} T(x_n) = T(\lim_{n\to\infty} x_n) = T(x),$$

i.e., x is a fixed point of T.

Now we show the uniqueness of the fixed point. Suppose that there is another fixed point $y \in X$, then

$$d(x,y) = d(T(x),T(y)) < \tau \cdot d(x,y) \implies d(x,y) < \tau d(x,y), \quad \tau \in (0,1),$$

and we conclude that $d(x,y) = 0$, i.e., $x = y$.

■

■ **Example 3.6** [Convergence of Newton's Method] The Newton's method aims to find the root of $f(x)$ by applying the iteration

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x)}$$

Suppose r is a root for f, the pre-assumption for the convergence of Newton's method is:

1. $f'(r) \neq 0$
2. $f \in \mathcal{C}^2$ on some neighborhood of r

Proof. 1. We first show that there exists $[r-\varepsilon, r+\varepsilon]$ such that the mapping

$$T : \mathcal{C}[r-\varepsilon, r+\varepsilon] \to \mathbb{R}, \quad f(x) \mapsto x - \tfrac{f(x)}{f'(x)}$$

satisfies $|T'(x)| < 1$ for $\forall x \in [r-\varepsilon, r+\varepsilon]$:

Note that $T'(x) = \frac{f(x)}{[f'(x)]^2} f''(x)$, and we define $h(x) = |T'(x)|$.

It's clear that $h(r) = 0$ and $h(x)$ is continuous, which implies

$$r \in h^{-1}((-1,1)) \implies B_\rho(r) \subseteq h^{-1}((-1,1)) \text{ for some } \rho > 0.$$

Or equivalently, $h((r-\rho,r+\rho)) \subseteq (-1,1)$. Take $\varepsilon = \frac{\rho}{2}$, and the result is obvious.

2. Therefore, any $x,y \in [r-\varepsilon, r+\varepsilon]$,

$$
\begin{aligned}
d(T(x),T(y)) &:= |T(x)-T(y)| && (3.6a)\\
&= |T'(\xi)||x-y| && (3.6b)\\
&\le \max_{\xi\in[r-\varepsilon,r+\varepsilon]} |T'(\xi)||x-y| && (3.6c)\\
&< m \cdot |x-y| && (3.6d)
\end{aligned}
$$

where (3.6b) is by applying MVT, and ξ is some point in $[r-\varepsilon, r+\varepsilon]$; we assume that $\max_{\xi\in[r-\varepsilon,r+\varepsilon]} |T'(\xi)| < m$ for some $m < 1$ in (3.6d).

Therefore, $T \in \mathcal{C}[r-\varepsilon, r+\varepsilon]$ is a contraction. By applying the contraction mapping theorem, there exists a unique fixed point near $[r-\varepsilon, r+\varepsilon]$:

$$x - \frac{f(x)}{f'(x)} = x \implies \frac{f(x)}{f'(x)} = 0 \implies f(x) = 0,$$

i.e., we obtain a root $x = r$. ■

Summary: if we use Newton's method on any point between $[r-\varepsilon, r+\varepsilon]$ where $f(r) = 0$ and ε is sufficiently small, then we will eventually get close to r.

3.2.3. Picard Lindelof Theorem

We will use Banach fixed point theorem to show the existence and uniqueness of the solution of ODE

$$
\begin{cases}
\dfrac{\mathrm{d}y}{\mathrm{d}x} = f(x,y(x)) \\
y(x_0) = y_0
\end{cases}
\quad \textbf{Initial Value Problem, IVP} \qquad (3.7)
$$

■ **Example 3.7** Consider the IVP

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = x^2 y^{1/5} \\ y(x_0) = c > 0 \end{cases} \Longrightarrow y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}$$

which can be solved by the separation of variables:

$$c > 0 \Longrightarrow y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}.$$

However, when $c = 0$, the ODE does not have a unique solution. One can verify that y_1, y_2 given below are both solutions of this ODE:

$$y_1 = (\frac{4x^3}{15})^{5/4}, \qquad y_2 = 0$$

This example shows that even when f is very nice, the IVP may not have unique solution. The Picard-Lindelof theorem will give a clean condition on f ensuring the unique solvability of the IVP (3.7). ■

3.3. Monday for MAT4002

3.3.1. Remarks on Basis and Homeomorphism

Reviewing.

1. $A \subseteq A_S \subseteq \overline{A}$, where A_S is sequential closure and $\overline{A}$ denotes closure.
2. Subspace topology.
3. Homeomorphism. Consider the mapping $f : X \to Y$ with the topogical space X, Y shown below, with the standard topology, the question is whether f is continuous?

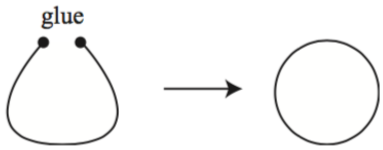

Figure 3.1: Diagram for mapping f

The answer is no, since the left in (3.1) can be isomorphically mapped into $(0,1)$; the right can be isomorphically mapped into $[0,1]$, and the mapping $(0,1) \to [0,1]$ cannot be isomorphism:

Proof. Assume otherwise the mapping $g : (0,1) \to [0,1]$ is isomorphism, and therefore $f^{-1}(U)$ is open for any open set U in the space $[0,1]$.
Construct $U = (1-\delta, 1]$ for $\delta \leq 1$, and therfore $f^{-1}((1-\delta, 1])$ is open, and therfore for the point $x = f^{-1}(1)$, there exists $\varepsilon > 0$ such that

$$B_\varepsilon(x) \subseteq f^{-1}((1-\delta,1]) \implies [x-\varepsilon, x) \subseteq f^{-1}((1-\delta,1)), \text{ and } (x, x+\varepsilon] \subseteq f^{-1}((1-\delta,1)).$$

which implies that there exists a, b such that $[x-\varepsilon, x) = f^{-1}((a,1))$ and $(x, x+\varepsilon] = f^{-1}((b,1))$, i.e., $f^{-1}((a,b) \cap (b,1))$ admits into two values in $[x-\varepsilon, x)$ and $(x, x+\varepsilon]$, which is a contradiction. ■

4. Basis of a topology $\mathcal{B} \subseteq (X, \mathcal{T})$ is a collection of open sets in the space such that the whole space can be recovered, or equivalently

(a) $\mathcal{B} \subseteq \mathcal{T}$

(b) Every set in $\mathcal{T}$ can expressed as a union of sets in $\mathcal{B}$

Example: Let $\mathbb{R}^n$ be equipped with usual topology, then

$$B = \{B_q(x) \mid x \in \mathbb{Q}^n, q \in \mathbb{Q}^+\} \text{ is a basis of } \mathbb{R}^n.$$

It suffices to show $U \subseteq \mathbb{R}^n$ can be written as

$$U = U_{x \in \mathbb{Q}} B_{q_x}(x)$$

Proposition 3.4 Let X, Y be topological spaces, and $\mathcal{B}$ a basis for topology on Y. Then

$$f : X \to Y \text{ is continuous} \iff f^{-1}(B) \text{ is open in X, } \forall B \in \mathcal{B}$$

Therefore checking $f^{-1}(U)$ is open for all $U \in \mathcal{T}_Y$ suffices to checking $f^{-1}(N)$ is open for all $B \in \mathcal{B}$.

Proof. The forward direction follows from the fact $B \subseteq \mathcal{T}_Y$.

To show the reverse direction, let $U \in \mathcal{T}_Y$, then $U = \bigcup_{i \in I} B_i$, where $B_i \in \mathcal{B}$, which implies

$$f^{-1}(U) = f^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f^{-1}(B_i)$$

which is open in X by our hypothesis. ■

Corollary 3.1 Let $f : X \to Y$ be a bijection. Suppose there is a basis $\mathcal{B}_X$ of $\mathcal{T}_X$ such that $\{f(B) \mid B \in \mathcal{B}_X\}$ forms a basis of $\mathcal{T}_Y$. Then $X \cong Y$.

Proof. Suppose $W \in \mathcal{T}_Y$, then by our hypothesis,

$$W = \bigcup_{i \in I} f(B_i),\ B_i \in \mathcal{B}_X \implies f^{-1}(W) = \bigcup_{i \in I} B_i \in \mathcal{T}_X,$$

which implies f is continuous.

Suppose $U \in \mathcal{T}_X$, then

$$U = \bigcup_{i \in I} B_i \implies f(U) = \bigcup_{i \in I} f(B_i) \in \mathcal{T}_Y \implies [f^{-1}]^{-1}(U) \in \mathcal{T}_Y,$$

i.e., f is continuous. ■

Question: *how to recognise whether a family of subsets is a basis for some given topology?*

Proposition 3.5 Let X be a set, $\mathcal{B}$ is a collection of subsets satisfying

1. X is a union of sets in $\mathcal{B}$, i.e., every $x \in X$ lies in some $B_x \in \mathcal{B}$
2. The intersection $B_1 \cap B_2$ for $\forall B_1, B_2 \in \mathcal{B}$ is a union of sets in $\mathcal{B}$, i.e., for each $B_1, B_2 \in \mathcal{B}$, and $x \in B_1 \cap B_2$, then there exists $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

Then the collection of subsets $\mathcal{T}_{\mathcal{B}}$, formed by taking any union of sets in $\mathcal{B}$, is a topology, and $\mathcal{B}$ is a basis for $\mathcal{T}_{\mathcal{B}}$.

Proof. 1. $\emptyset \in \mathcal{T}_{\mathcal{B}}$ (taking nothing from $\mathcal{B}$); for $x \in X, B_x \in \mathcal{B}$, by hypothesis (1),

$$X = \bigcup_{x \in X} B_x \in \mathcal{T}_{\mathcal{B}}$$

2. Suppose $T_1, T_2 \in \mathcal{T}_{\mathcal{B}}$. Let $x \in T_1 \cap T_2$, where T_i is a union of subsets in $\mathcal{B}$. Therefore,

$$\begin{cases} x \in B_1 \subseteq T_1, & B_1 \in \mathcal{B} \\ x \in B_2 \subseteq T_2, & B_2 \in \mathcal{B} \end{cases}$$

which implies $x \in B_1 \cap B_2$, i.e., $x \in B_x \subseteq B_1 \cap B_2$ for some $B_x \in \mathcal{B}$. Therefore,

$$\bigcup_{x \in B_1 \cap B_2} \{x\} \subseteq \bigcup_{x \in B_1 \cap B_2} B_x \subseteq B_1 \cap B_2,$$

i.e., $B_1 \cap B_2 = \bigcup_{x \in B_1 \cap B_2} B_x$, i.e., $B_1 \cap B_2 \in \mathcal{T}_{\mathcal{B}}$.

3. The property that $\mathcal{T}_{\mathcal{B}}$ is closed under union operations can be checked directly.

The proof is complete. ■

3.3.2. Product Space

Now we discuss how to construct new topological spaces out of given ones is by taking Cartesian products:

Definition 3.4 Let $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ be topological spaces. Consider the family of subsets in $X \times Y$:

$$\mathcal{B}_{X\times Y} = \{U \times V \mid U \in \mathcal{T}_X, V \in \mathcal{T}_y\}$$

This $\mathcal{B}_{X\times Y}$ forms a basis of a topology on $X \times Y$. The induced topology from $\mathcal{B}_{X\times Y}$ is called **product topology**.

For example, for $X = \mathbb{R}, Y = \mathbb{R}$, the elements in $\mathcal{B}_{X\times Y}$ are rectangles.

Proof for well-definedness in definition (3.4). We apply proposition (3.5) to check whether $B_{X\times Y}$ forms a basis:

1. For any $(x, y) \in X \times Y$, we imply $x \in X, y \in Y$. Note that $X \in \mathcal{T}_X, Y \in \mathcal{T}_Y$, we imply $(x, y) \in X \times Y \in \mathcal{B}_{X\times Y}$.
2. Suppose $U_1 \times V_1, U_2 \times V_2 \in \mathcal{B}_{X\times Y}$, then

$$(U_1 \times V_1) \cap (U_2 \times V_2) = (U_1 \cap U_2) \times (V_1 \cap V_2),$$

where $U_1 \cap U_2 \in \mathcal{T}_X, V_1 \cap V_2 \in \mathcal{T}_Y$. Therefore, $(U_1 \times V_1) \cap (U_2 \times V_2) \in \mathcal{B}_{X\times Y}$.

■

R However, the product topology may not necessarily become the largest topology in the space $X \times Y$. Consider $X = \mathbb{R}, Y = \mathbb{R}$, the open set in the space $X \times Y$ may not necessarily be rectangles. However, all elements in $\mathcal{B}_{X\times Y}$ are rectangles.

Example 3.8 The space $\mathbb{R} \times \mathbb{R}$ is isomorphic to $\mathbb{R}^2$, where the product topology is defined on $\mathbb{R} \times \mathbb{R}$ and the standard topology is defined on $\mathbb{R}^2$:

Construct the function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ with $(a, b) \to (a, b)$.

Obviously, $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ is a bijection.

Take the basis of the topology on $\mathbb{R}$ as open intervals,

$$B_X = \{(a,b) \mid a < b \text{ in } \mathbb{R}\}$$

Therefore, one can verify that the set $\mathcal{B} := \{(a,b) \times (c,d) \mid a < b, c < d\}$ forms a basis for the product topology, and

$$\{f(B) \mid B \in \mathcal{B}\} = \{(a,b) \times (c,d) \mid a < b, c < d\}$$

forms a basis of the usual topology in $\mathbb{R}^2$.

By Corollary (3.1), we imply $\mathbb{R} \times \mathbb{R} \cong \mathbb{R}^2$. ■

We also raise an example on the homeomorphism related to product spaces:

■ **Example 3.9** Let $S^1 = \{(\cos x, \sin x \mid x \in [0, 2\pi])\}$ be a unit circle on $\mathbb{R}^2$.

Consider $f : S^1 \times (0, \infty) \to \mathbb{R}^2 \setminus \{\mathbf{0}\}$ defined as

$$f(\cos x, \sin x, r) \mapsto (r \cos x, r \sin x)$$

It's clear that f is a bijection, and f is continuous. Moreover, the inverse $g := f^{-1}$ is defined as

$$g(a,b) = \left(\frac{a}{\sqrt{a^2+b^2}}, \frac{b}{\sqrt{a^2+b^2}}, \sqrt{a^2+b^2}\right)$$

which is continuous as well. Therefore, the $f : \mathcal{S}^1 \times (0, \infty) \to \mathbb{R}^2 \setminus \{\mathbf{0}\}$ is a homeomorphism.

■

3.4. Wednesday for MAT3040

3.4.1. Remarks for the Change of Basis

Reviewing.

- $[\cdot]_{\mathcal{A}} : V \to \mathbb{F}^n$ denotes coordinate vector mapping
- Change of Basis matrix: $\mathcal{C}_{\mathcal{A}',\mathcal{A}}$
- $T : V \to W$, $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\boldsymbol{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$.

  $\mathrm{Hom}_{\mathbb{F}}(V, W) \to M_{m \times n}(\mathbb{F})$

■ **Example 3.10** Let $V = \mathbb{P}_3[x]$ and $\mathcal{A} = \{1, x, x^2, x^3\}$.

Let $T : V \to V$ defined as $p(x) \mapsto p'(x)$:

$$\begin{cases} T(1) = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\ T(x) = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\ T(x^2) = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\ T(x^3) = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2 + 0 \cdot x^3 \end{cases}$$

We can define the change of basis matrix for a linear transformation T as well, w.r.t. $\mathcal{A}$ and $\mathcal{A}$:

$$\mathcal{C}_{\mathcal{A},\mathcal{A}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Also, we can define a different basis $\mathcal{A}' = \{x^3, x^2, x, 1\}$ for the output space for T, say $T : V_{\mathcal{A}} \to V_{\mathcal{A}'}$:

$$(T)_{\mathcal{A},\mathcal{A}'} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Our observation is that the corresponding coordinate vectors before and after linear transformation admits a matrix multiplication:

$$(2x^2 + 4x^3) \xrightarrow{T} ((4x + 12x^2))$$

$$(2x^2 + 4x^3)_{\mathcal{A}} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 4 \end{pmatrix} \qquad (4x + 12x^2)_{\mathcal{A}} = \begin{pmatrix} 0 \\ 4 \\ 12 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 12 \\ 0 \end{pmatrix}$$

$$\mathcal{C}_{\mathcal{A}\mathcal{A}} \cdot (2x^2 + 4x^3)_{\mathcal{A}} = (4x + 12x^2)_{\mathcal{A}}$$

■

Theorem 3.3 — **Matrix Representation.** Let $T : V \to W$ be a linear transformation of finite dimensional vector sapces. Let $\mathcal{A}, \mathcal{B}$ the ordered basis of V, W, respectively. Then the following diagram holds:

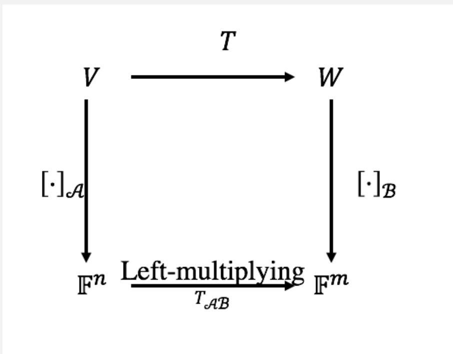

Figure 3.2: Diagram for the matrix reprentation, where $n := \dim(V)$ and $m := \dim(W)$

namely, for any $\boldsymbol{v} \in V$,

$$(T)_{\mathcal{B},\mathcal{A}}(\boldsymbol{v})_{\mathcal{A}} = (T\boldsymbol{v})_{\mathcal{B}}$$

Therefore, we can compute $T\boldsymbol{v}$ by matrix multiplication.

R Linear transformation corresponds to coordinate matrix multiplication.

Proof. Suppose $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\mathcal{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$. The proof of this theorem follows the same procedure of that in Theorem (3.1)

1. We show this result for $\boldsymbol{v} = \boldsymbol{v}_j$ first:

$$\text{LHS} = [\alpha_{ij}]\boldsymbol{e}_j = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

$$\text{RHS} = (T\boldsymbol{v}_j)_{\mathcal{B}} = \left(\sum_{i=1}^{m} \alpha_{ij}\boldsymbol{w}_i\right)_{\mathcal{B}} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

2. Then we show the theorem holds for any $\boldsymbol{v} := \sum_{j=1}^{n} r_j \boldsymbol{v}_j$ in V:

$$\begin{aligned}
(T)_{\mathcal{B}\mathcal{A}}(\boldsymbol{v})_{\mathcal{A}} &= (T)_{\mathcal{B}\mathcal{A}} \left(\sum_{j=1}^{n} r_j \boldsymbol{v}_j\right)_{\mathcal{A}} && (3.8a) \\
&= (T)_{\mathcal{B}\mathcal{A}} \left(\sum_{j=1}^{n} r_j (\boldsymbol{v}_j)_{\mathcal{A}}\right) && (3.8b) \\
&= \sum_{j=1}^{n} r_j (T)_{\mathcal{B}\mathcal{A}} (\boldsymbol{v}_j)_{\mathcal{A}} && (3.8c) \\
&= \sum_{j=1}^{n} r_j (T\boldsymbol{v}_j)_{\mathcal{B}} && (3.8d) \\
&= \left(\sum_{j=1}^{n} r_j (T\boldsymbol{v}_j)\right)_{\mathcal{B}} && (3.8e) \\
&= \left[T\left(\sum_{j=1}^{n} r_j \boldsymbol{v}_j\right)\right]_{\mathcal{B}} && (3.8f) \\
&= (T\boldsymbol{v})_{\mathcal{B}} && (3.8g)
\end{aligned}$$

The justification for (3.8a) is similar to that shown in Theorem (3.1). The proof is complete.

■

R Consider a special case for Theorem (3.3), i.e., $T = \text{id}$ and $\mathcal{A}, \mathcal{A}'$ are two ordered basis for the input and output space, respectively. Then the result in Theorem (3.3) implies

$$C_{\mathcal{A}',\mathcal{A}}(\boldsymbol{v})_{\mathcal{A}} = (\boldsymbol{v})_{\mathcal{A}'}$$

i.e., the matrix representation theorem (3.3) is a general case for the change of basis theorem (3.1)

Proposition 3.6 — Functionality. Suppose V, W, U are finite dimensional vector spaces, and let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be the ordered basis for V, W, U, respectively. Suppose that

$$T : V \to W, \quad S : W \to U$$

are given two linear transformations, then

$$(S \circ T)_{\mathcal{C},\mathcal{A}} = (S)_{\mathcal{C},\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}$$

Composition of linear transformation corresponds to the multiplication of change of basis matrices.

Proof. Suppose the ordered basis $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$, $\mathcal{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$, $\mathcal{C} = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_p\}$. By defintion of change of basis matrices,

$$T(\boldsymbol{v}_j) = \sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \boldsymbol{w}_i$$

$$S(\boldsymbol{w}_i) = \sum_k (S_{\mathcal{C},\mathcal{B}})_{ki} \boldsymbol{u}_k$$

We start from the j-th column of $(S \circ T)_{\mathcal{C},\mathcal{A}}$ for $j = 1, \dots, n$, namely

$$
\begin{aligned}
(S \circ T)_{\mathcal{C},\mathcal{A}}(\boldsymbol{v}_j)_{\mathcal{A}} &= (S \circ T(\boldsymbol{v}_j))_{\mathcal{C}} && (3.9a)\\
&= \left[S \circ \left(\sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \boldsymbol{w}_i \right) \right]_{\mathcal{C}} && (3.9b)\\
&= \sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \, (S(\boldsymbol{w}_i))_{\mathcal{C}} && (3.9c)\\
&= \sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \left(\sum_k (S_{\mathcal{C},\mathcal{B}})_{ki} \boldsymbol{u}_k \right)_{\mathcal{C}} && (3.9d)\\
&= \sum_k \sum_i (S_{\mathcal{C},\mathcal{B}})_{ki} (T_{\mathcal{B},\mathcal{A}})_{ij} (\boldsymbol{u}_k)_{\mathcal{C}} && (3.9e)\\
&= \sum_k (S_{\mathcal{C},\mathcal{B}} T_{\mathcal{B},\mathcal{A}})_{kj} (\boldsymbol{u}_k)_{\mathcal{C}} && (3.9f)\\
&= \sum_k (S_{\mathcal{C},\mathcal{B}} T_{\mathcal{B},\mathcal{A}})_{kj} \boldsymbol{e}_k && (3.9g)\\
&= j\text{-th column of } [S_{\mathcal{C}\mathcal{B}} T_{\mathcal{B},\mathcal{A}}] && (3.9h)
\end{aligned}
$$

where (3.9a) is by the result in theorem (3.3); (3.9b) and (3.9d) follows from definitions of $T(\boldsymbol{v}_j)$ and $S(\boldsymbol{w}_i)$; (3.9c) and (3.9e) follows from the linearity of $\mathcal{C}$; (3.9f) follows from the matrix multiplication definition; (3.9g) is because $(\boldsymbol{u}_k)_{\mathcal{C}} = \boldsymbol{e}_k$.

Therefore, $(S \circ T)_{\mathcal{C}\mathcal{A}}$ and $(S_{\mathcal{C},\mathcal{B}})(T_{\mathcal{B},\mathcal{A}})$ share the same j-th column, and thus equal to each other. ■

Corollary 3.2 Suppose that S and T are two identity mappings $V \to V$, and consider $(S)_{\mathcal{A}'\mathcal{A}}$ and $(T)_{\mathcal{A},\mathcal{A}'}$ in proposition (3.6), then

$$(S \circ T)_{\mathcal{A}',\mathcal{A}'} = (S)_{\mathcal{A}'\mathcal{A}} (T)_{\mathcal{A},\mathcal{A}'}$$

Therefore,

$$\text{Identity matrix} = \mathcal{C}_{\mathcal{A}',\mathcal{A}} \mathcal{C}_{\mathcal{A},\mathcal{A}'}$$

Proposition 3.7 Let $T : V \to W$ with $\dim(V) = n, \dim(W) = m$, and let

- $\mathcal{A}, \mathcal{A}'$ be ordered basis of V
- $\mathcal{B}, \mathcal{B}'$ be ordered basis of W

then the change of basis matrices admit the relation

$$(T)_{\mathcal{B}',\mathcal{A}'} = \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}\mathcal{C}_{\mathcal{A}\mathcal{A}'} \tag{3.10}$$

Here note that $(T)_{\mathcal{B}',\mathcal{A}'}, (T)_{\mathcal{B},\mathcal{A}} \in \mathbb{F}^{m\times n}$; $\mathcal{C}_{\mathcal{B}',\mathcal{B}} \in \mathbb{F}^{m\times m}$; and $\mathcal{C}_{\mathcal{A}\mathcal{A}'} \in \mathbb{F}^{n\times n}$.

Proof. Let $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}, \mathcal{A}' = \{\boldsymbol{v}_1', \dots, \boldsymbol{v}_n'\}$. Consider simplifying the j-th column for the LHS and RHS of (3.10) and showing they are equal:

$$\begin{aligned}
\text{LHS} &= (T)_{\mathcal{B}',\mathcal{A}'}\boldsymbol{e}_j \\
&= (T)_{\mathcal{B}',\mathcal{A}'}(\boldsymbol{v}_j')_{\mathcal{A}'} \\
&= (T\boldsymbol{v}_j')_{\mathcal{B}'}
\end{aligned}$$

$$\begin{aligned}
\text{RHS} &= \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}\mathcal{C}_{\mathcal{A}\mathcal{A}'}\boldsymbol{e}_j \\
&= \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}\mathcal{C}_{\mathcal{A}\mathcal{A}'}(\boldsymbol{v}_j')_{\mathcal{A}'} \\
&= \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}(\boldsymbol{v}_j')_{\mathcal{A}} \\
&= \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T\boldsymbol{v}_j')_{\mathcal{B}} \\
&= (T\boldsymbol{v}_j')_{\mathcal{B}'}
\end{aligned}$$

■

R Let $T : V \to V$ be a linear operator with $\mathcal{A}, \mathcal{A}'$ being two ordered basisof V, then

$$(T)_{\mathcal{A}'\mathcal{A}'} = \mathcal{C}_{\mathcal{A}',\mathcal{A}}(T)_{\mathcal{A}\mathcal{A}}\mathcal{C}_{\mathcal{A},\mathcal{A}'} = (\mathcal{C}_{\mathcal{A},\mathcal{A}'})^{-1}(T)_{\mathcal{A}\mathcal{A}}\mathcal{C}_{\mathcal{A},\mathcal{A}'}$$

Therefore, the change of basis matrices $(T)_{\mathcal{A}'\mathcal{A}'}$ and $(T)_{\mathcal{A}\mathcal{A}}$ are similar to each other, which means they share the same eigenvalues, determinant, trace.

Therefore, two similar matrices cooresponds to same linear transformation using different basis.

3.5. Wednesday for MAT3006

3.5.1. Remarks on Contraction

Reviewing.

- Suppose $E \subseteq X$ with X being complete, then E is closed in X iff E is complete
- Suppose $E \subseteq X$, then E is closed in X if E is complete.
- Contraction Mapping Theorem
- Classification for the Convergence of Newton's method: the Newton's method aims to find the fixed point of T.

  $$T : \mathbb{R} \to \mathbb{R}, \quad T(x) = x - \frac{f(x)}{f''(x)}$$

  In the last lecture we claim that there exists $[r - \varepsilon, r + \varepsilon]$ such that $\sup_{[r-\varepsilon, r+\varepsilon]} |T'(x)| < 1$.

  Note that we doesn't make our statement rigorous enough. we need to furthermore show that $T(X) \subseteq X$:

  - $T : [r - \varepsilon, r + \varepsilon] \to [r - \varepsilon, r + \varepsilon]$, since

    $$|T(x) - r| = |T(x) - T(r)| = |T'(s)||x - r| \le \sup_{[r-\varepsilon, r+\varepsilon]} |T'(s)||x - r| < |x - r|$$

    Therefore, if $x \in [r - \varepsilon, r + \varepsilon]$, then $T(x) \in [r - \varepsilon, r + \varepsilon]$.
  - T is a contraction:

    $$|T(x) - T(y)| < \tau \cdot |x - y|$$

  Therefore, applying contraction mapping theorem gives the desired result.

3.5.2. Picard-Lindelof Theorem

Consider solving the the initival value problem given below

$$\begin{cases} \frac{\mathrm{d}y}{\mathrm{d}x} = f(x,y) \\ y(\alpha) = \beta \end{cases} \implies y(x) = \beta + \int_{\alpha}^{x} f(t,y(t))\,\mathrm{d}t \tag{3.11}$$

Definition 3.5 Let $R = [\alpha - a, \alpha + a] \times [\beta - b, \beta + b]$. Then the function $f(x,y)$ satisfies the **Lipschitz condition** on R if there exists $L > 0$ such that

$$|f(x,y_1) - f(x,y_2)| < L \cdot |y_1 - y_2|, \quad \forall (x,y_i) \in R \tag{3.12}$$

The smallest number $L^* = \inf\{L \mid \text{The relation (3.12) holds for } L\}$ is called the **Lipschitz constant** for f.

■ **Example 3.11** A $\mathcal{C}^1$-function f(x,y) in a rectangle automatically satisfies the Lipschitz condition:

$$f(x,y_1) - f(x,y_2) \overset{\text{Appling MVT}}{=} \frac{\partial f}{\partial y}(x,\tilde{y})(y_1 - y_2)$$

Since $\frac{\partial f}{\partial y}$ is continuous on R and thus bounded, we imply

$$|f(x,y_1) - f(x,y_2)| < L \cdot |y_1 - y_2|, \quad \forall (x,y_i) \in R$$

where

$$L = \max\left\{ \text{abs}\left(\frac{\partial f}{\partial y}\right) \middle| (x,y) \in R \right\}$$

Theorem 3.4 — Picard-Lindelof Theorem (existence part). Suppose $f \in \mathcal{C}(R)$ be such that f satisfies the Lipschitz condition, then there exists $a'' \in (0,a]$ such that (**??**) is solvable with $y(x) \in \mathcal{C}([\alpha - a'', \alpha + a''])$.

Proof. Consider the complete metric space

$$X = \{y(x) \in \mathcal{C}([\alpha - a, \alpha + a]) \mid \beta - b \le y(x) \le \beta + b\},$$

with the mapping $T : X \to X$ defined as

$$(Ty)(x) = \beta + \int_{\alpha}^{x} f(t, y(t))\,\mathrm{d}t$$

It suffices to show that T is a contraction, but here we need to estrict a a smaller number as follows:

1. Well-definedness of T: Take $M := \sup\{f(x,y) \mid (x,y) \in R\}$ and construct $a' = \min\{b/M, a\}$. Consider the complete matric space

   $$X = \{y(x) \in \mathcal{C}([\alpha - a', \alpha + a']) \mid \beta - b \le y(x) \le \beta + b\}$$

   which implies that

   $$|(Ty)(x) - \beta| \le \left| \int_{\alpha}^{x} f(t, y(t))\,\mathrm{d}t \right| \le M|x - \alpha| \le Ma' \le b,$$

   i.e., $T(X) \subseteq X$, and therefore $T : X \to X$ is well-defined.
2. Contraction for T: Construct $a'' \in \min\{a', \frac{1}{2L^*}\}$, where L^* is the Lipschitz constant for f. and consider the complete metric space

   $$X = \{y(x) \in \mathcal{C}([\alpha - a'', \alpha + a'']) \mid \beta - b \le y(x) \le \beta + b\}$$

   Therefore for $\forall x \in [\alpha - a'', \alpha + a'']$ and the mapping $T : X \to X$,

   $$\begin{aligned} |[T(y_1) - T(y_2)](x)| &\le \left| \int_{\alpha}^{x} [f(t, y_2(t)) - f(t, y_1(t))]\,\mathrm{d}t \right| \\ &\le \int_{\alpha}^{x} |f(t, y_2) - f(t, y_1)|\,\mathrm{d}t \le \int_{\alpha}^{x} L^*|y_2(t) - y_1(t)|\,\mathrm{d}t \\ &\le L^*|x - \alpha| \sup|y_2(t) - y_1(t)| \le L^* a'' d_\infty(y_2, y_1) \le \frac{1}{2} d_\infty(y_2, y_1) \end{aligned}$$

Therefore, we imply $d_\infty(Ty_2, Ty_1) \le \frac{1}{2} d_\infty(y_2, y_1)$, i.e., T is a contraction.

Applying contraction mapping theorem, there exists $y(x) \in X$ such that $Ty = y$, i.e.,

$$y = \beta + \int_\alpha^x f(t, y(t))\,\mathrm{d}t$$

Thus y is a solution for the IVP (3.11). ■

R On $[\alpha - a'', \alpha + a'']$, we can solve the IVP (3.11) by recursively applying T:

$$\begin{aligned} y_0(x) &= \beta, \qquad \forall x \in [\alpha - a'', \alpha + a''] \\ y_1 &= T(y_0) = \beta + \int_\alpha^x f(t, \beta)\,\mathrm{d}t \\ y_2 &= T(y_1) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

By studying (3.11) on different rectangles, we are able to show the uniqueness of our solution:

Proposition 3.8 Suppose f satisfies the Lipschitz conditon, and y_1, y_2 are two solutions of (3.11), where y_1 is defined on $x \in I_1$, and y_2 is defined on $x \in I_2$. Suppose $I_1 \cap I_2 \ne \emptyset$ and y_1, y_2 share the same initial value condition $y(\alpha) = \beta$. Then $y_1(x) = y_2(x)$ on $I_1 \cap I_2$.

Proof. Suppose $I_1 \cap I_2 = [p, q]$ and let $z := \sup\{x \mid y_1 \equiv y_2 \text{ on } [\alpha, x]\}$. It suffices to show $z = q$. Now suppose on the contrary that $z < q$, and consider the subtraction $|y_1 - y_2|$:

$$y_i = \beta + \int_\alpha^x f(t, y_i)\,\mathrm{d}t \implies |y_1 - y_2| = \left| \int_z^x f(t, y_1) - f(t, y_2)\,\mathrm{d}t \right|.$$

Construct an interval $I^* = [z - \frac{1}{2L^*}, z + \frac{1}{2L^*}] \cap [p, q]$, and let $x_m = \arg\max_{x \in I^*} |y_1(x) -$

$y_2(x)|$, which implies for $\forall x \in I^*$,

$$\begin{aligned}
|y_1(x) - y_2(x)| &= \left| \int_z^x f(t,y_1) - f(t,y_2) \, \mathrm{d}t \right| \\
&\leq \int_z^x |f(t,y_1(t)) - f(t,y_2(t))| \, \mathrm{d}t \\
&\leq L^* \int_z^x |y_1(x) - y_2(x)| \, \mathrm{d}t \\
&\leq L^* |x - z| |y_1(x_m) - y_2(x_m)| \\
&\leq \frac{1}{2} |y_1(x_m) - y_2(x_m)|.
\end{aligned}$$

Taking $x = x_m$, we imply $y_1 \equiv y_2$ for $\forall x \in I^*$, which contradicts the maximality of z. ■

Combining Theorem (3.4) and proposition (3.8), we imply the existence of a unique "maximal" solution for the IVP (3.11), i.e., the unique solution is defined on a maximal interval.

Corollary 3.3 Let $U \subseteq \mathbb{R}^2$ be an open set such that $f(x,y)$ satisfies the Lipschitz condition for any $[a,b] \times [c,d] \subseteq U$, then there exists x_m and x_M in $\overline{\mathbb{R}}$ such that

- The IVP (3.11) admits a solution $y(x)$ for $x \in (x_m, x_M)$, and if y^* is another solution of (3.11) on some interval $I \subseteq (x_m, x_M)$, then $y \equiv y^*$ on I.
- Therefore $y(x)$ is maximally defined; and $y(x)$ is unique.

■ **Example 3.12** Consider the IVP

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = x^2 y^{1/5} \\ y(0) = C \end{cases} \implies \frac{\partial f}{\partial y} = \frac{x^2}{5y^{4/5}}.$$

- Taking $U = \mathbb{R} \times (0, \infty)$ implies $y = \left(\frac{4x^3}{15} + c^{4/5} \right)^{5/4}$, defined on $(\sqrt[3]{-15/4c^{4/5}}, \infty)$.
- When $c = 0$, $f(x,y)$ does not satisfy the Lipschitz condition. The uniqueness of solution does not hold.

■

3.6. Wednesday for MAT4002

3.6.1. Remarks on product space

Reviewing.

- Product Topology: For topological space $(X, \mathcal{T}_X)$ and $(Y, \mathcal{Y})$, define the basis

$$\mathcal{B}_{X\times Y} = \{U \times V \mid U \in \mathcal{T}_X, V \in \mathcal{T}_Y\}$$

  and the family of union of subsets in $\mathcal{B}_{X\times Y}$ forms a product topology.

Proposition 3.9 a ring torus is homeomorphic to the Cartesian product of two circles, say $S^1 \times S^1 \cong T$.

Proof. Define a mapping $f : [0,2\pi] \times [0,2\pi] \to T$ as

$$f(\theta,\phi) = \Big((R + r\cos\theta)\cos\phi, \quad (R + r\cos\theta)\sin\phi, \quad r\sin\theta\Big)$$

Define $i : T \to \mathbb{R}^3$, we imply

$$i \circ f : [0,2\pi] \times [0,2\pi] \to \mathbb{R}^3 \text{ is continuous}$$

Therefore we imply $f : [0,2\pi] \times [0,2\pi] \to T$ is continuous. Together with the condition that

$$\begin{cases} f(0,y) = f(2\pi,y) \\ f(x,0) = f(x,2\pi) \end{cases}$$

we imply the function $f : S^1 \to S^1 \to T$ is continuous. We can also show it is bijective. We can also show f^{-1} is continuous. ■

Proposition 3.10 1. Let $X \times Y$ be endowed with product topology. The projection mappings defined as

$$p_X : X \times Y \to X, \text{ with } p_X(x,y) = x$$
$$p_Y : X \times Y \to Y, \text{ with } p_Y(x,y) = y$$

are continuous.

2. (an equivalent definition for product topology) The product topology is the **coarest topology** on $X \times Y$ such that p_X and p_Y are both continuous.
3. (an equivalent definition for product topology) Let Z be a topological space, then the product topology is the unique topology that the red and the blue line in the diagram commutes:

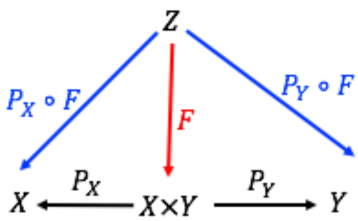

Figure 3.3: Diagram summarizing the statement (*)

namely,

the mapping $F : Z \to X \times Y$ is continuous iff both $P_X \circ F : Z \to X$ and $P_Y \circ F : Z \to Y$ are continuous. (*)

Proof. 1. For any open U, we imply $p_X^{-1}(U) = U \times Y \in \mathcal{B}_{X\times Y} \subseteq \mathcal{T}_{X\times Y}$, i.e., $p_X^{-1}(U)$ is open. The same goes for p_Y.

2. It suffices to show any topology $\mathcal{T}$ that meets the condition in (2) must contain $\mathcal{T}_{\text{product}}$. We imply that for $\forall U \in \mathcal{T}_X, V \in \mathcal{T}_Y$,

$$\begin{cases} p_X^{-1}(U) = U \times X \in \mathcal{T} \\ p_Y^{-1}(V) = X \times V \in \mathcal{T} \end{cases} \implies (U \times Y) \cap (X \times V) = (U \cap X) \times (Y \cap V) = U \times V \in \mathcal{T},$$

which implies $\mathcal{B}_{X\times Y} \subseteq \mathcal{T}$. Since $\mathcal{T}$ is closed for union operation on subsets, we imply $\mathcal{T}_{\text{product topology}} \subseteq \mathcal{T}$.

3. (a) Firstly show that $\mathcal{T}_{\text{product}}$ satisfies (*).

    - For the forward direction, by (1) we imply both $p_X \circ F$ and $p_Y \circ F$ are continuous, since the composition of continuous functions are continuous as well.
    - For the reverse direction, for $\forall U \times \mathcal{T}_X, V \in \mathcal{T}_Y$,

    $$F^{-1}(U \times V) = (p_X \circ F)^{-1}(X) \cap (p_Y \circ F)^{-1}(Y),$$

    which is open due to the continuity of $p_X \circ F$ and $p_Y \circ F$.

   (b) Then we show the uniqueness of $\mathcal{T}_{\text{product}}$. Let $\mathcal{T}$ be another topology $X \times Y$ satisfying (*).

    - Take $Z = (X \times Y, \mathcal{T})$, and consider the identity mapping $F = \text{id} : Z \to Z$, which is continuous. Therefore $p_X \circ \text{id}$ and $p_Y \circ \text{id}$ are continuous, i.e., p_X and p_Y are continuous. By (2) we imply $\mathcal{T}_{\text{product}} \subseteq \mathcal{T}$.
    - Take $Z = (X \times Y, \mathcal{T}_{\text{product}})$, and consider the identity mapping $F = \text{id} : Z \to Z$. Note that $p_X \circ F = p_X$ and $p_Y \circ F = p_Y$, which is continuous by (1). Therefore, the identity mapping $F : (X \times Y, \mathcal{T}_{\text{product}}) \to (X \times Y, \mathcal{T})$ is continuous, which implies

    $$U = \text{id}^{-1}(U) \subseteq \mathcal{T}_{\text{product}} \text{ for } \forall U \in \mathcal{T},$$

    i.e., $\mathcal{T} \subseteq \mathcal{T}_{\text{product}}$.

   The proof is complete.

■

Definition 3.6 [Disjoint Union] Let $X \times Y$ be two topological spaces, then the **disjoint union** of X and Y is

$$X \sqcup Y := (X \times \{0\}) \cup (Y \times \{1\})$$

R

1. We define that U is open in $X \sqcup Y$ if

   (a) $U \cap (X \times \{0\})$ is open in $X \times \{0\}$; and

   (b) $U \cap (Y \times \{1\})$ is open in $Y \times \{1\}$.

   We also need to show the weill-definedness for this definition.

2. S is open in $X \perp Y$ iff S can be expressed as

$$S = (U \times \{0\}) \cup (V \times \{1\})$$

   where $U \subseteq X$ is open and $V \subseteq Y$ is open.

3.6.2. Properties of Topological Spaces

3.6.2.1. Hausdorff Property

Definition 3.7 [First Separation Axiom] A topological space X satisfies the **first separation axiom** if for any two distinct points $x \neq y \in X$, there exists open $U \ni x$ but not including y.

Proposition 3.11 A topological space X has first separation property if and only if for $\forall x \in X$, $\{x\}$ is closed in X.

Proof. Sufficiency. Suppose that $x \neq y$, then construct $U := X \setminus \{y\}$, which is a open set that contains x but not includes y.

Necessity. Take any $x \in X$, then for $\forall y \neq x$, there exists $y \in U_y$ that is open and $x \notin U_y$. Thus

$$\{y\} \subseteq U_y \subseteq X \setminus \{x\}$$

which implies

$$\bigcup_{y \in X \setminus \{x\}} \{y\} \subseteq \bigcup_{y \in X \setminus \{x\}} U_y \subseteq X \setminus \{x\},$$

i.e., $X \setminus \{x\} = \bigcup_{y \in X \setminus \{x\}} U_y$ is open in X, i.e., $\{x\}$ is closed in X. ■

Definition 3.8 [Second separation Axiom] A topological space satisfies the **second separation axiom** (or X is Hausdorff) if for all $x \neq y$ in X, there exists open sets U, V such that

$$x \in U, \quad y \in V, \quad U \cap V = \varnothing$$

■

■ **Example 3.13** All metrizable topological spaces are Hausdorff.

Suppose $d(x,y) = r > 0$, then take $B_{r/2}(x)$ and $B_{r/2}(y)$ ■

■ **Example 3.14** Note that a topological space that is **first separable** may not necessarily be **second separable**:

Consider $\mathcal{T}_{\text{co-finite}}$, then X is first separable but not Hausdorff:

Suppose on the contrary that for given $x \neq y$, there exists open sets U, V such that $x \in U, y \in V$, and

$$U \cap V = \varnothing \implies X = X \setminus (U \cap V) = (X \setminus U) \cup (X \setminus V),$$

implying that the union of two finite sets equals X, which is infinite, which is a contradiction.

■