Tutorial Notes Information

- Tutor: Nelson Lam (nelson.lam.lcy@gmail.com)
- Prerequisite Knowledge:
  - Linear Independence, Span, Basis
  - Matrix Operations, Computation of Determinant
  - Matrix Representation of Rotation, Reflection
- Table of Content & Outline
  - HKDSE Vector, Inner Product , Cross Product in $\mathbb{R}^3$
  - Orthogonal & Orthnormal Basis & Orthogonalization
  - Composition of Linear Transformations
  - Orthogonal Matrix & Isometry
  - Vector Valued Functions
- References:
  - **Lecture Notes** of Dr. LAU & Dr. CHENG
- All the suggestions and feedback are welcome. Any report of typos is appreciated.

1 Quick Review of HKDSE Vector

In HKDSE M2, we use $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ to denote the standard ordered basis of $\mathbb{R}^3$. In general, we also use coordinates to represent the position of a vector, e.g. $\mathbf{v} = 3\,\mathbf{i} + 4\,\mathbf{j} - 8\,\mathbf{k} = (3, 4, -8)$. Most operations are naturally coordinate-wise applied.

Suppose $\mathbf{v} = v_1\,\mathbf{i} + v_2\,\mathbf{j} + v_3\,\mathbf{k} = (v_1, v_2, v_3)$ and $\mathbf{w} = w_1\,\mathbf{i} + w_2\,\mathbf{j} + w_3\,\mathbf{k} = (w_1, w_2, w_3)$, let $\alpha \in \mathbb{R}$, then:

1. $\mathbf{v} \pm \mathbf{w} = (v_1 \pm w_1)\,\mathbf{i} + (v_2 \pm w_2)\,\mathbf{j} + (v_3 \pm w_3)\,\mathbf{k} = (v_1 \pm w_1, v_2 \pm w_2, v_3 \pm w_3)$
2. $\alpha\,\mathbf{v} = (\alpha v_1)\,\mathbf{i} + (\alpha v_2)\,\mathbf{j} + (\alpha v_3)\,\mathbf{k} = (\alpha v_1, \alpha v_2, \alpha v_3)$
3. $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}$

Definition 1.1. We say a vector $\mathbf{v}$ is a unit vector if $\|\mathbf{v}\| = 1$

Key Question. What if we are given a non unit vector $\mathbf{v}$? $\Rightarrow$ Construct $\hat{\mathbf{v}} = \dfrac{\mathbf{v}}{\|\mathbf{v}\|}$

Definition 1.2. We say two vectors $\mathbf{v}, \mathbf{w}$ are parallel if $\mathbf{v} = \alpha\,\mathbf{w}$ for some $\alpha \in \mathbb{R}$

Key Question. On top of parallel, how to formulate perpendicular in vector sense ?

We need to introduce the concept of Inner Product & Cross Product, before that:

1. Inner Product $\Rightarrow$ Input two vectors, Output a real number.

   Example: $\langle 3\,\mathbf{i} + 2\,\mathbf{j} + 5\,\mathbf{k}, -4\,\mathbf{i} + 2\,\mathbf{j} + \mathbf{k} \rangle = -3$
2. Cross Product $\Rightarrow$ Input two vectors, Output a vector.

   Example: $(2\,\mathbf{i} - 2\,\mathbf{j} - 2\,\mathbf{k}) \times (22\,\mathbf{i} + 11\,\mathbf{j} - 11\,\mathbf{k}) = 44\,\mathbf{i} - 22\,\mathbf{j} + 66\,\mathbf{k}$

Properties and calculation of these vector products will be discussed in the next section

2 Inner Product

2.1 Basic Definitions

Definition 2.1. Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$, We say $\langle \bullet, \bullet \rangle \to \mathbb{R}$ is an inner product if:

1. $\langle \alpha\,\mathbf{u} + \beta\,\mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$
2. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
3. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, with equality hold if and only if $\mathbf{v} = 0$

Definition 2.2. We say two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are orthogonal (perpendicular) if:

$$\langle \mathbf{v}, \mathbf{w} \rangle = 0$$

Key Observation. It seems there are freedom in defining inner product. Yes ! There are many inner products in $\mathbb{R}^n$. In our course, we only consider the most geometrically meaningful one !

Definition 2.3. Let $\mathbf{v} = v_1\,\mathbf{i} + v_2\,\mathbf{j} + v_3\,\mathbf{k}$ and $\mathbf{w} = w_1\,\mathbf{i} + w_2\,\mathbf{j} + w_3\,\mathbf{k}$, we define the inner product of $\mathbf{v}$ and $\mathbf{w}$ as:

$$\langle \mathbf{v}, \mathbf{w} \rangle \overset{\text{def}}{=} v_1 w_1 + v_2 w_2 + v_3 w_3$$

If two vectors $\mathbf{v}, \mathbf{w}$ are column vectors, then orthogonality requires $\mathbf{w}^T\,\mathbf{v} = 0$

Example 2.4. Let $\mathbf{u} = (3, 4, 0)$ and $\mathbf{v} = (4, -3, 0)$, then $\langle \mathbf{u}, \mathbf{v} \rangle = (3)(4) + (4)(-3) = 0$

2.2 Geometric Meaning of Inner Product

It is tempting to understand the geometric meaning of Definition 2.3. Here is an outline:

Compute $\|\mathbf{v}\|$ $\to$ Cauchy Schwarz Inequality $\to$ Cosine Law in $\mathbb{R}^n$ and Triangle Inequality

Theorem 2.1 (Compute Norm)**.** For any $\mathbf{u} \in \mathbb{R}^n$, we have $\langle \mathbf{u}, \mathbf{u} \rangle = \|\mathbf{u}\|^2$

Theorem 2.2 (C-S Inequality)**.** For any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we have $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$

Theorem 2.3 (Cosine Law)**.** For any triangle $\Delta ABC \subset \mathbb{R}^n$, we have: $c^2 = a^2 + b^2 - 2ab\cos\theta$

Theorem 2.4 (Cosine Law)**.** For any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we have $\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \|\mathbf{v}\| \cos\theta$

Theorem 2.5 ($\Delta-$Inequality)**.** Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we have: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

Proof. Will be demonstrated in tutorial □

2.3 Vector Projection

Theorem 2.6. Projection of $\mathbf{u} = \overrightarrow{PQ}$ onto $\mathbf{v} = \overrightarrow{PS}$ is a vector $\overrightarrow{PR}$, where R is the foot of the perpendicular from Q to the line PS. Then we have:

$$\overrightarrow{PR} = \text{Proj}_{\mathbf{v}}\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2}\,\mathbf{v}$$

Proof. Direction of $\overrightarrow{PR}$ is $\mathbf{v}$. Signed Magnitude of $\overrightarrow{PR}$ is $\|\overrightarrow{PQ}\| \times \cos(\angle QPS)$

Therefore $\overrightarrow{PR} = \left(\|\overrightarrow{PQ}\| \times \cos(\angle QPS)\right) \dfrac{\mathbf{v}}{\|\mathbf{v}\|} = \|\mathbf{u}\| \dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \dfrac{\mathbf{v}}{\|\mathbf{v}\|} = \dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2}\,\mathbf{v}$ □

The vector projection brings us to the fact that every vector $\mathbf{u}$ can be expressed as a sum of a vector parallel to $\mathbf{v}$ and another vector orthogonal to $\mathbf{v}$.

2.4 Orthogonal Basis & Linear Combination

Recall Tutorial 1: Let $\{\mathbf{u},\mathbf{v},\mathbf{w}\}$ be a basis of $\mathbb{R}^3$, then any vector $\mathbf{r}=\alpha\,\mathbf{u}+\beta\,\mathbf{v}+\gamma\,\mathbf{w}$, for some α,β,γ. You were taught to use RREF to solve for α,β,γ. In this tutorial, if the basis vectors are mutually orthogonal, then we can compute α,β,γ easily

Theorem 2.7. Suppose the basis vectors $\{\mathbf{u},\mathbf{v},\mathbf{w}\}\subset\mathbb{R}^3$ are mutually orthogonal, that is:

$$\langle\mathbf{u},\mathbf{v}\rangle=\langle\mathbf{v},\mathbf{w}\rangle=\langle\mathbf{w},\mathbf{u}\rangle=0$$

Then $\forall\,\mathbf{r}\in\mathbb{R}^3,\mathbf{r}=\dfrac{\langle\mathbf{r},\mathbf{u}\rangle}{\langle\mathbf{u},\mathbf{u}\rangle}\,\mathbf{u}+\dfrac{\langle\mathbf{r},\mathbf{v}\rangle}{\langle\mathbf{v},\mathbf{v}\rangle}\,\mathbf{v}+\dfrac{\langle\mathbf{r},\mathbf{w}\rangle}{\langle\mathbf{w},\mathbf{w}\rangle}\,\mathbf{w}$, that is: $\alpha=\dfrac{\langle\mathbf{r},\mathbf{u}\rangle}{\langle\mathbf{u},\mathbf{u}\rangle},\beta=\dfrac{\langle\mathbf{r},\mathbf{v}\rangle}{\langle\mathbf{v},\mathbf{v}\rangle},\gamma=\dfrac{\langle\mathbf{r},\mathbf{w}\rangle}{\langle\mathbf{w},\mathbf{w}\rangle}$

Corollary 2.8. Further suppose $\mathbf{u},\mathbf{v},\mathbf{w}$ are unit vectors, then $\mathbf{r}=\langle\mathbf{r},\mathbf{u}\rangle\,\mathbf{u}+\langle\mathbf{r},\mathbf{v}\rangle\,\mathbf{v}+\langle\mathbf{r},\mathbf{w}\rangle\,\mathbf{w}$

Warm Reminder. Here are some general tips to read mathematical texts

- When facing abstract theorems, consider simple examples to get a feeling of it.

  Try to paraphrase it using humane language
- Make sure you understand EVERY words. Cross-referencing if necessary.

  Example: "unit vector" / "mutually orthogonal"

Example 2.5. Let $\mathbf{u}=(3,-4,0),\mathbf{v}=(4,3,0),\mathbf{w}=(0,0,1)$. Suppose $\mathbf{r}=(1,2,3)$. We want to find α,β,γ such that $\mathbf{r}=\alpha\,\mathbf{u}+\beta\,\mathbf{v}+\gamma\,\mathbf{w}$. Tutorial 1 techniques suggests us:

$$\begin{pmatrix}3&4&0\\-4&3&0\\0&0&1\end{pmatrix}\begin{pmatrix}\alpha\\\beta\\\gamma\end{pmatrix}=\begin{pmatrix}1\\2\\3\end{pmatrix}$$

By RREF or Matrix Inverse or Cramer's Rule (HKDSE M2), we have: $\alpha=-\dfrac{1}{5},\beta=\dfrac{2}{5},\gamma=3$

Now consider method 2: Let $(1,2,3)=\alpha(3,-4,0)+\beta(4,3,0)+\gamma(0,0,1)$

Step 1 Taking inner product of LHS and RHS with $(3,-4,0)$

$$\underbrace{\langle(1,2,3),(3,-4,0)\rangle}_{-5}=\underbrace{\alpha\langle(3,-4,0),(3,-4,0)\rangle}_{25\alpha}+\underbrace{\beta\langle(4,3,0),(3,-4,0)\rangle}_{0}+\underbrace{\gamma\langle(0,0,1),(3,-4,0)\rangle}_{0}$$

So we have $-5=25\alpha$ and $\alpha=-\dfrac{1}{5}$

Step 2 Taking inner product of LHS and RHS with $(4,3,0)$

$$\underbrace{\langle(1,2,3),(4,3,0)\rangle}_{10}=\underbrace{\alpha\langle(3,-4,0),(4,3,0)\rangle}_{0}+\underbrace{\beta\langle(4,3,0),(4,3,0)\rangle}_{25\beta}+\underbrace{\gamma\langle(0,0,1),(4,3,0)\rangle}_{0}$$

So we have $10=25\beta$ and $\beta=\dfrac{2}{5}$

Step 3 Taking inner product of LHS and RHS with $(0,0,1)$

$$\underbrace{\langle(1,2,3),(0,0,1)\rangle}_{3}=\underbrace{\alpha\langle(3,-4,0),(0,0,1)\rangle}_{0}+\underbrace{\beta\langle(4,3,0),(0,0,1)\rangle}_{0}+\underbrace{\gamma\langle(0,0,1),(0,0,1)\rangle}_{\gamma}$$

So we have $3=\gamma$ and $\gamma=3$

Exercise 2.1. Try to prove Theorem 2.7. Hints: $\underbrace{\langle\mathbf{r},\mathbf{u}\rangle}_{?}=\underbrace{\alpha\langle\mathbf{u},\mathbf{u}\rangle}_{?}+\underbrace{\beta\langle\mathbf{v},\mathbf{u}\rangle}_{?}+\underbrace{\gamma\langle\mathbf{w},\mathbf{u}\rangle}_{?}$

Remark. Theorem 2.7 was examined in HKALE Pure Math Long Questions (before 2005)

Warm Reminder. A little bit examination techniques:

- In advanced mathematics, given $LHS = RHS$, we often do the same operation on LHS and RHS to get a new identity. For example:

$$(1). \quad (1+x)^n = \sum_{k=0}^{n} C_k^n x^k \xrightarrow{d/dx} n(1+x)^{n-1} = \sum_{k=1}^{n-1} kC_k^n x^{k-1}$$

$$(2). \quad \frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1} \xrightarrow{\Sigma} \sum_{k=1}^{n} \frac{1}{k(k+1)} = 1 - \frac{1}{n+1}$$

- In principle, a good exam question should allow multiple approaches. Brute Force Explosion is one of approach while computationally time-consuming. However, if you have clever insight, you might finish the question much faster. This is what you should expect in TDG assessments – Good questions with strong discriminators

- In undergraduate mathematics, you often study problem sets and 'recite' them. 'Recite' means summarise the key steps / tricks / techniques of the problem. And then try to apply it. Esp. Mathematical Analysis. Of course, such principle WILL BE TESTED IN TDG ASSESSMENTS

2.5 Orthonormal Basis

Definition 2.6. Let $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ be a basis for $\mathbb{R}^3$. We say $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a orthonormal basis if:

1. (orthogonal basis) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{u} \rangle = 0$
2. (unit length) $\|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{w}\| = 1$

Theorem 2.9. Suppose $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a orthonormal basis, then for any $\mathbf{r} \in \mathbb{R}^3$, we have:

$$\mathbf{r} = \langle \mathbf{r}, \mathbf{u} \rangle \mathbf{u} + \langle \mathbf{r}, \mathbf{v} \rangle \mathbf{v} + \langle \mathbf{r}, \mathbf{w} \rangle \mathbf{w}$$

Key Takeaway. Motivation of investigating "orthonormal basis": Easier Computation

2.6 Gram-Schmidt Orthogonalization

Key Question. Orthonormal basis is "good & nice". How to generate orthonormal basis ? Given any basis $B = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$, how to construct a orthonormal basis $\widetilde{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ from arbitary basis B ?

Key Observation. Once we get orthogonal basis, divide each vector by its norm will get orthonormal basis. The problem reduces to: how to construct orthogonal basis from arbitary basis

$$\underbrace{\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}}_{\text{arbitary}} \xrightarrow{???} \underbrace{\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}}_{\text{orthogonal}} \xrightarrow{\text{easy}} \underbrace{\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}}_{\text{orthonormal}}$$

Warm Reminder. Find a "bridge", "middlemen" as intermediate step to simplify harder problems. Try to work backward to find "pre-condition" [Common Technique in HKALE Inequality]

Theorem 2.10 (Gram-Schmidt Orthogonalization Process)**.**
Suppose $B = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is a linearly independent subset of $\mathbb{R}^n$. Let $\mathbf{v}_1 = \mathbf{w}_1$ and define:

$$\mathbf{v}_l \stackrel{\text{def}}{=} \mathbf{w}_l - \sum_{i=1}^{l-1} \frac{\langle \mathbf{w}_l, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} \mathbf{v}_i, l = 2, 3, \ldots, n$$

Then $\{\mathbf{v}_1, \ldots \mathbf{v}_n\}$ is a orthogonal set of vectors with the same span as B

Proof. Step by step construction, will be demonstrated in tutorial ☐

3 Vector Cross Product

Key Question. Given 2 non-parallel vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, by common sense, there is a "direction" $\mathbf{w}$ which is orthogonal to BOTH $\mathbf{u}, \mathbf{v}$. How to determine this "direction" ?

Theorem 3.1 (Computation of Cross roduct in $\mathbb{R}^3$)**.** Suppose that $\mathbf{u} = u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k}$ and $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$, then:

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

Moreover, $(\mathbf{u} \times \mathbf{v})$ is a direction orthogonal to BOTH $\mathbf{u}, \mathbf{v}$ (two rows the same, then det $= 0$)

Example 3.1. Find a vector perpendicular to the plane containing $P(1, -1, 0)$, $Q(2, 1, -1)$,$R(-1, 1, 2)$. Hence, find a unit normal vector perpendicular to the plane

Solution. The required vector is: $\overrightarrow{PQ} \times \overrightarrow{PR}$

By computation: $\overrightarrow{PQ} = \mathbf{i} + \mathbf{j} - \mathbf{k}, \overrightarrow{PR} = -2\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$, hence $\overrightarrow{PQ} \times \overrightarrow{PR} = 6\mathbf{i} + 6\mathbf{k}$

The unit normal vector is: $\mathbf{n} = \dfrac{\overrightarrow{PQ} \times \overrightarrow{PR}}{|\overrightarrow{PQ} \times \overrightarrow{PR}|} = \dfrac{1}{\sqrt{2}}\mathbf{i} + \dfrac{1}{\sqrt{2}}\mathbf{k}$

Example 3.2. Find $\mathbf{u} \times \mathbf{v}$ and $\mathbf{v} \times \mathbf{u}$ if $\mathbf{u} = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$ and $\mathbf{v} = -4\mathbf{i} + 3\mathbf{j} + \mathbf{k}$

Solution.

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & 1 \\ -4 & 3 & 1 \end{vmatrix} = -2\mathbf{i} - 6\mathbf{j} + 10\mathbf{k}, \mathbf{v} \times \mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -4 & 3 & 1 \\ 2 & 1 & 1 \end{vmatrix} = 2\mathbf{i} + 6\mathbf{j} - 10\mathbf{k}$$

Theorem 3.2 (Properties of Cross Product)**.**
Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ and $\alpha, \beta \in \mathbb{R}$, then

1. $(r\mathbf{u}) \times (s\mathbf{v}) = (rs)(\mathbf{u} \times \mathbf{v})$
2. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ and $(\mathbf{v} + \mathbf{w}) \times \mathbf{u} = \mathbf{v} \times \mathbf{u} + \mathbf{w} \times \mathbf{u}$
3. $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$
4. $\mathbf{0} \times \mathbf{u} = \mathbf{0}$
5. $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\| \sin\theta$ (area of parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$)
6. $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are parallel ($\mathbf{0}$ is parallel to every vector)

Example 3.3 (2019 HKDSE M2 Q12)**.** Let $A(1, -4, 2), B(-5, -4, 8), C(-5, -12, 2)$, Π is a plane containing A, B, C. Given that P, Q, R are points lying on Π such that $\overrightarrow{OP} = p\,\mathbf{i}, \overrightarrow{OQ} = q\,\mathbf{j}, \overrightarrow{OR} = r\,\mathbf{k}$. Find p, q, r.

Solution. By computation, $\overrightarrow{AB} = -6\,\mathbf{i} + 6\,\mathbf{k}, \overrightarrow{AC} = -6\,\mathbf{i} + 8\,\mathbf{j}$, hence $\overrightarrow{AB} \times \overrightarrow{AC} = 48\,\mathbf{i} - 36\,\mathbf{j} + 48\,\mathbf{k}$
Suppose $G(x, y, z)$ lies on Π, a equation of plane is just a relation (constraint) on (x, y, z)
Notice that $\overrightarrow{AG} = (x - 1)\,\mathbf{i} + (y + 4)\,\mathbf{j} + (z - 2)\,\mathbf{k}$ is a vector lies on Π, thus $\overrightarrow{AG} \perp (\overrightarrow{AB} \times \overrightarrow{AC})$
Therefore $\langle \overrightarrow{AG}, \overrightarrow{AB} \times \overrightarrow{AC} \rangle = 0$
That is: $48(x - 1) - 36(y + 4) + 48(z - 2) = 0$
Simplifying gives: $\Pi : 4x - 3y + 4z = 24$
Since $P(p, 0, 0), Q(0, q, 0), R(0, 0, r)$, they are $x-$intercept, $y-$intercept, $z-$intercept of Π
Substituting into Π gives $p = 6, q = -8, r = 6$

4 Linear Transformations Review

Recall from Tutorial 1: Linear Transformations are characterised by "Basis vector goes where"

- Rotation by θ anticlockwise: $\begin{pmatrix}1\\0\end{pmatrix} \overset{rot}{\to} \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$ & $\begin{pmatrix}0\\1\end{pmatrix} \overset{rot}{\to} \begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}$

  We "collect" those column vectors and form a matrix: $\text{Rot}(\theta) = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$

- Reflection along $y = (\tan\theta)x$: $\begin{pmatrix}1\\0\end{pmatrix} \overset{ref}{\to} \begin{pmatrix}\cos 2\theta\\ \sin 2\theta\end{pmatrix}$ & $\begin{pmatrix}0\\1\end{pmatrix} \overset{ref}{\to} \begin{pmatrix}\sin 2\theta\\ -\cos 2\theta\end{pmatrix}$

  We "collect" those column vectors and form a matrix: $\text{Ref}(\theta) = \begin{pmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{pmatrix}$

- Enlargement by k times: $\begin{pmatrix}1\\0\end{pmatrix} \overset{\times k}{\to} \begin{pmatrix}k\\0\end{pmatrix}$ & $\begin{pmatrix}0\\1\end{pmatrix} \overset{\times k}{\to} \begin{pmatrix}0\\k\end{pmatrix}$

  We "collect" those column vectors and form a matrix: $kI = \begin{pmatrix}k & 0\\ 0 & k\end{pmatrix}$

Key Question. Why we need matrix representation for linear transformation ?

Theorem 4.1. Let $T : V \to V$ be a linear transformation with matrix representation A, then

$$T(\mathbf{x}) = A\,\mathbf{x}, \forall\, \mathbf{x} \in V$$

Key Takeaway. Apply the transformation T on vector $\mathbf{x} \equiv$ Multiply $\mathbf{x}$ by A

Exercise 4.1 (HKDSE 2021 Mathematics Compulsory Part Paper 2 Question 23)**.**
The coordinates of the point P are $(7, -5)$. P is reflected with respect to the $y-$axis to the point Q. Q is the rotated clockwise about the origin through $90°$ to the point R. Find the coordinates of R

Solution. We first calculate two matrix of transformations

- $\text{Rot}(-90°) = \begin{pmatrix}\cos(-90°) & -\sin(-90°)\\ \sin(-90°) & \cos(-90°)\end{pmatrix} = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$

- $\text{Ref}(90°) = \begin{pmatrix}\cos 2(90°) & \sin 2(90°)\\ \sin 2(90°) & -\cos 2(90°)\end{pmatrix} = \begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}$

Then $\overrightarrow{R} = \text{Rot}(-90°)\left[\text{Ref}(90°)\begin{pmatrix}7\\-5\end{pmatrix}\right] = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\left[\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}7\\-5\end{pmatrix}\right]$

Method 1: We directly compute the matrix multiplications

$$\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\left[\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}7\\-5\end{pmatrix}\right] = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\left[\begin{pmatrix}-7\\-5\end{pmatrix}\right] = \begin{pmatrix}-5\\7\end{pmatrix}$$

Method 2: We apply associative law of matrix multiplication here

$$\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\left[\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}7\\-5\end{pmatrix}\right] = \left[\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\right]\begin{pmatrix}7\\-5\end{pmatrix} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}7\\-5\end{pmatrix} = \begin{pmatrix}-5\\7\end{pmatrix}$$

Key Takeaway. Multiplication of Matrices $\equiv$ Composition of Linear Transformation

Warm Reminder. Mathematicians make definitions / develop theories with motivations. It must somehow "make life easier". There must be deep theories behind stupid computation

5 Orthogonal Matrix

Definition 5.1. Let Q be an $n \times n$ matrix, we say Q is an orthogonal matrix if $QQ^T = Q^TQ = I$

Key Question. Why orthogonal matrix Q is worth to investigate ? How to translate $QQ^T = Q^TQ = I$ into humane language ?

5.1 Geometric Meaning of Orthogonal Matrix

Warm Reminder. To get a feeling of the definition, let's consider small size examples: $n = 2$ and ask: how does 2×2 orthogonal matrix looks like ?

Theorem 5.1. Any 2×2 orthogonal matrix Q is either a reflection or a rotation

Proof. Let $Q = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $Q^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ and $\det(Q) = ad - bc$

$QQ^T = I$ implies $\begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Comparing entries, we have: $\begin{cases} a^2+b^2 = 1 & \ldots(1) \\ c^2+d^2 = 1 & \ldots(2) \\ ac+bd = 0 & \ldots(3) \Rightarrow a^2c^2 = b^2d^2 \end{cases}$

From (1) and (2), we have $b^2 = 1 - a^2$ and $d^2 = 1 - c^2$ respectively

Substituting into (3): $a^2c^2 = (1-a^2)(1-c^2)$

Simplifying, we have $a^2 + c^2 = 1 \qquad \ldots(4)$

Hence the system of equations reduce to:

$$\begin{cases} a^2+b^2 = 1 & \ldots(1) \\ c^2+d^2 = 1 & \ldots(2) \\ a^2+c^2 = 1 & \ldots(4) \end{cases}$$

This is actually a system of non-linear equations with infinitely many solutions

By HKDSE M2 methodology, let $\theta \in \mathbb{R}$ be a parameter, we express a, b, c, d in terms of θ

Let $a = \cos\theta$ and substitute into $(1), (2), (4)$, we have:

$$a = \cos\theta \qquad b = \pm\sin\theta \qquad c = \pm\sin\theta \qquad d = \pm\cos\theta$$

Therefore

$$TaQ = \begin{pmatrix} \cos\theta & \pm\sin\theta \\ \pm\sin\theta & \pm\cos\theta \end{pmatrix}$$

which is a rotation matrix or a reflection matrix, depending on $\pm$ signs. □

Corollary 5.2. If $\det(Q) = 1$, then Q represents rotation. If $\det(Q) = -1$, then Q represents reflection.

Key Takeaway. 2×2 orthogonal matrix is a generalisation of rotation & reflection. That is: Distance & Angle Preserving Linear Mappings.

Key Question. We observe that " 2×2 orthogonal matrix represents "Distance & Angle Preserving Linear Mappings" in $\mathbb{R}^2$. How about 3×3 orthogonal matrix, or even $n \times n$ orthogonal matrix ? Can we generalise our observation to higher dimension ?

Warm Reminder. In higher dimensions, it is hard to write out the explicit form of rotation / reflection matrices. So "checking by looking at the matrix form" is not feasible in $\mathbb{R}^n, n \geq 3$. There should be some clever methods to verify.

Let's recall your memory: What is distance & angle ?

- Length of a vector $\mathbf{v}$ is calculated by $\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$
- Angle between two vectors $\mathbf{u}, \mathbf{v}$ are calculated by $\cos\theta = \dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}$

Key Observation. Distance & Angle are determined by Inner Product. We claim that "$n \times n$ orthogonal matrix Q represents distance & angle preserving map". It suffices to prove: "$n \times n$ orthogonal matrix Q preserves inner product". We will prove it below:

Theorem 5.3 (Lecture Notes Chapter 1 Exercise 9)**.**
For any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ (write them as column vectors), let Q be an $n \times n$ matrix, then:

1. If Q is an orthogonal matrix, then $\langle Q\,\mathbf{u}, Q\,\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$
2. If $\langle Q\,\mathbf{u}, Q\,\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$, then Q is an orthogonal matrix

Proof. Will be demonstrated in tutorial (need prep, the skills may be hard for some students) □

Warm Reminder. The art of problem formulating, reducing and solving, via observation and guessing. The whole theory starts at a simple definition: $QQ^T = Q^TQ = I$

5.2 Isometry & Orthonormal Basis

Key Question. The column vectors of an orthogonal matrix constitute a orthonormal basis, mathematically why ? Geometrically why ?

Warm Reminder. Observing some interesting patterns is good. After that, formulate your observation into a mathematical statement

Theorem 5.4. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subset \mathbb{R}^n$. Construct $n \times n$ matrix $A = [\,\mathbf{v}_1|\mathbf{v}_2|\ldots|\mathbf{v}_{n-1}|\mathbf{v}_n\,]$ using $\mathbf{v}_i$ as the i^{th} column vector. Then we have the following:

1. If A is an orthogonal matrix, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a orthonormal basis
2. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a orthonormal basis, then A is an orthogonal matrix

Proof. By block multiplication, $[A^TA]_{ij} = \mathbf{v}_j^T\,\mathbf{v}_i = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$, hence:

$$A^TA = I \iff [A^TA]_{ij} = \delta_{ij} \iff \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij} \iff \|\mathbf{v}_i\| = 1 \text{ and } \underbrace{\langle \mathbf{v}_i, \mathbf{v}_j \rangle}_{\text{if } i \neq j} = 0$$ □

Remark. The above proof is deliberately written in a very shitty notation. Humane proof will be demonstrated in tutorial.

Warm Reminder. Most proofs in university mathematical texts are like this. Confusion and struggle is common. The right way to read them is to get your hands dirty: Write by yourself, then compare with the proof in the text, asking yourself: "Anything you overlooked ?" (Sometimes your proof may be much shorter than the textbook proof. It does NOT necessarily imply you are wrong. Just because the textbook is SHIT)

Finally, one question remains to answer: What is the geometric picture, which Theorem 5.4 want to tell us ? Why is it naturally true ?

Since isometry preserves length and angle, hence it maps orthonormal basis to orthonormal basis. Consider the standard ordered basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \subset \mathbb{R}^n$, it is am orthonormal basis. On the other hand, $\{Q\,\mathbf{e}_1, \ldots, Q\,\mathbf{e}_n\}$ is also an orthonormal basis. Moreover, the i^{th} column of Q is exactly $Q\,\mathbf{e}_i$. Therefore column vectors of an orthogonal matrix constitute a orthonormal basis

6 Vector-valued Function

6.1 Basic Definitions

Definition 6.1. A vector valued function $\mathbf{r}$ is a function with input $t \in \mathbb{R}$ and output $\mathbf{r} \in \mathbb{R}^n$

- t is known as "parameter"
- $\mathbf{r}(t) = (r_1(t), r_2(t), \ldots, r_{n-1}(t), r_n(t))$, where $r_i(t)$ is the i^{th} coordinate function

Example 6.2 (Unknown Curves)**.**

1. $\mathbf{r}(t) = (\underbrace{5 + 13\cos t}_{r_1(t)}, \underbrace{12 + 5\sin t}_{r_2(t)}, \underbrace{13 - 12\sin t}_{r_3(t)})$ (Circle) (2015 TDG Quiz 2)
2. $\mathbf{r}(t) = \left(\dfrac{3t^2 + 4t + 1}{t^2 - 3t + 2}, \dfrac{t^2 + t + 1}{t^2 - 3t + 2}\right)$ (Hyperbola) (Conic Section Old Textbook)

Example 6.3 (Parameterization of Famous Curves)**.**

1. Quadratic: $y = x^2 \Longrightarrow \mathbf{r}(t) = (t, t^2), t \in \mathbb{R}$
2. Circle: $x^2 + y^2 = a^2 \Longrightarrow \mathbf{r}(t) = (a\cos t, a\sin t), t \in \mathbb{R}$
3. Circle: $x^2 + y^2 = a^2 \Longrightarrow \mathbf{r}(t) = (a\cos(123t), a\sin(123t)), t \in \mathbb{R}$
4. Circle: $x^2 + y^2 = a^2 \Longrightarrow \mathbf{r}(t) = \left(\dfrac{2t}{1+t^2}, \dfrac{1-t^2}{1+t^2}\right), t \in \mathbb{R}$
5. Ellipse: $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1 \Longrightarrow \mathbf{r}(t) = (a\cos t, b\sin t), t \in \mathbb{R}$

Few Important Notes.

- Vector valued function with parameter t traces a curve in $\mathbb{R}^3$ as t varies
- Parameterization is NOT unique. E.g. 3 different parameterization of circle
- There are 3 questions types: $\begin{cases} \text{Given an equation} \leftrightarrow \text{Find a parameterization} \\ \text{Given geometrical constraint, derive parameterization (HKDSE Locus)} \end{cases}$

Key Observation. From Example 6.3, the trace of (2), (3) are circle. But the "speed" of (2), (3) are fundamentally different. To talk about "speed" "rate of change of XXX w.r.t time", we introduce differentiation of vector-valued functions. (Note: (4) is known as steographical projection)

Definition 6.4. Let $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ be a vector-valued function. Write $\mathbf{r}(t) = (r_1(t), r_2(t), \ldots, r_{n-1}(t), r_n(t))$ as its coordinate function expansion, then

$$\mathbf{r}'(t) \overset{\text{def}}{=} (r_1'(t), r_2'(t), \ldots, r_{n-1}'(t), r_n'(t))$$

In other words: Coordinate-wise differentiation

Example 6.5. For $\mathbf{r}(t) = (\underbrace{5 + 13\cos t}_{r_1(t)}, \underbrace{12 + 5\sin t}_{r_2(t)}, \underbrace{13 - 12\sin t}_{r_3(t)})$, $\mathbf{r}'(t) = (\underbrace{-13\sin t}_{r_1'(t)}, \underbrace{5\cos t}_{r_2'(t)}, \underbrace{-12\cos t}_{r_3'(t)})$

Example 6.6 (Speed of Circular Motion)**.**
For $\mathbf{r}(t) = (a\cos t, a\sin t)$, $\mathbf{r}'(t) = (-a\sin t, a\cos t)$, $\|\mathbf{r}'(t)\| = |a|$

For $\mathbf{r}(t) = (a\cos(123t), a\sin(123t))$, $\mathbf{r}'(t) = (-123a\sin t, 123a\cos t)$, $\|\mathbf{r}'(t)\| = 123|a|$

Definition 6.7. A regular parameterized curve is a differentiable function $\mathbf{r} : (a, b) \subset \mathbb{R} \to \mathbb{R}^n$ such that $\mathbf{r}'(t) \neq \mathbf{0}, \forall t \in (a, b)$

Example 6.8. For $\mathbf{r}(t) = (a\cos t, a\sin t)$, $\mathbf{r}'(t) = (-a\sin t, a\cos t)$

Denote $x(t) = -a\sin t$ and $y(t) = a\cos t$. Notice that $x(t), y(t)$ not simultaneously zero. Hence regular !

Example 6.9. For distinct $A(a_1, a_2), B(b_1, b_2) \in \mathbb{R}^2$, the line segment AB can be parameterized by:

$$\mathbf{r}(t) = ((1-t)a_1 + tb_1, (1-t)a_2 + b_2)$$

By computation, $\mathbf{r}'(t) = (b_1 - a_1, b_2 - a_2) \neq \mathbf{0}$, hence regular

Example 6.10 (Example 2.1.3 in Lecture Notes)**.** Let $\mathbf{r}(t) = (t^2, t^3)$, then $\mathbf{r}'(t) = (2t, 3t^2) = \mathbf{0}$ if $t = 0$. Therefore **not** regular

Key Question. Differentiability is obviously essential. But what is the significance of $\|\mathbf{r}'(t)\| \neq \mathbf{0}$? If the curve is not regular, what will happen ? [The answer will be revealed in next subsection]

6.2 Arc-Length Calculation (may not taught by Dr. Cheng at Lecture 2) (for preview pupose)

Definition 6.11. Let $\mathbf{r} : (a, b) \subset \mathbb{R} \to \mathbb{R}^n$ be a regular parameterized curve, then the arc-length of $\mathbf{r}$ is defined as:

$$L[a, b] \stackrel{\text{def}}{=} \int_a^b \|\mathbf{r}'(t)\| dt$$

Warm Reminder. You can ask "what is the motivation / underlying picture of the definition". But never ask "how to prove the definition".

Important "Counter Example" (2020 TDG Quiz 1: Astroid)

Let $C : x^{2/3} + y^{2/3} = 1$, parameterize C by: $\mathbf{r}(t) = (\cos^3 t, \sin^3 t)$ for $t \in [0, 2\pi)$, then:

$$\mathbf{r}'(t) = (-3\cos^2 t \sin t, 3\sin^2 t \cos t)$$

$$\|\mathbf{r}'(t)\|^2 = (-3\cos^2 t \sin t)^2 + (3\sin^2 t \cos t)^2 = 9\cos^2 t \sin^2 t(\cos^2 t + \sin^2 t) = 9\cos^2 t \sin^2 t$$

$$\|\mathbf{r}'(t)\|^2 = 9\cos^2 t \sin^2 t$$

Note that $\mathbf{r}(t)$ is **not** regular at $t = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{2}, 2\pi$

$$L[0, 2\pi] = \int_0^{2\pi} \|\mathbf{r}'(t)\| dt = 3\int_0^{2\pi} \sin t \cos t dt = \frac{3}{2}\int_0^{2\pi} \sin 2t dt = 0$$

Remark. This counter example is deliberately written in an inaccurate way. Explanation will be given during tutorial.

Key Observation. If the curve is **not** regular, then there will be "bug". (zero arc-length ???)

Exercise 6.1. Try to generalise the definition of arc-length so that it works for astroid. Can we relax the requirement of "$\mathbf{r}$ is regular" ?

Warm Reminder. This is what mathematicians do: Construct counter examples, generalise results, remove unnecessary conditions. Especially Mathematical Analysis !

6.3 Differentiation By Part

Theorem 6.1 (Differentiation by Part)**.** Let $\mathbf{u}, \mathbf{v} : \mathbb{R} \to \mathbb{R}^n$ are vector-valued functions, then:

$$\frac{d}{dt}\langle \mathbf{u}(t), \mathbf{v}(t) \rangle = \langle \mathbf{u}'(t), \mathbf{v}(t) \rangle + \langle \mathbf{u}(t), \mathbf{v}'(t) \rangle$$

Proof. Expand $\mathbf{u}, \mathbf{v}$ in coordinate functions: $\mathbf{u} = (u_1, \ldots, u_n), \mathbf{v} = (v_1, \ldots v_n)$
Then $\mathbf{u}'(t) = (u_1'(t), \ldots, u_n'(t)), \mathbf{v}'(t) = (v_1'(t), \ldots v_n'(t))$

$$LHS = \frac{d}{dt}\langle \mathbf{u}(t), \mathbf{v}(t) \rangle = \frac{d}{dt}\sum_{i=1}^{n} u_i(t)v_i(t) = \sum_{i=1}^{n} [u_i'(t)v_i(t) + u_i(t)v_i'(t)]$$

$$RHS = \langle \mathbf{u}'(t), \mathbf{v}(t) \rangle + \langle \mathbf{u}(t), \mathbf{v}'(t) \rangle = \sum_{i=1}^{n} u_i'(t)v_i(t) + \sum_{i=1}^{n} u_i(t)v_i'(t)$$

Hence $LHS = RHS$ □

Remark. Differentiation By Part has numerous applications in moving frame theory (Frenet Serret Equation) and Surface Theory (Chapter 3 of Lecture Notes). Let's have a very brief investigation.

Theorem 6.2. Let $\mathbf{r} : \mathbb{R} \to \mathbb{R}^2$ be a regular parameterized plane curve, with $\|\mathbf{r}'(s)\| = 1$, then:

1. $\langle \mathbf{r}'(s), \mathbf{r}''(s) \rangle = 0$
2. $\mathbf{r}'(s) \perp \mathbf{r}''(s)$

Proof. Will be demonstrated in tutorial □

Key Observation. Suppose Γ is the trace of $\mathbf{r}(s)$. Let s_0 be fixed, then $\mathbf{r}(s_0)$ is a point on Γ

- $\mathbf{r}'(s_0)$ is the vector pointing at tangential direction of Γ at $\mathbf{r}(s_0)$
- $\mathbf{r}''(s_0)$ is the vector pointing at the normal direction of Γ at $\mathbf{r}(s_0)$

Hence $\mathbf{r}'(s_0)$ and $\dfrac{\mathbf{r}''(s_0)}{\|\mathbf{r}''(s_0)\|}$ is a frame (unit square) indicating the curvature of Γ