§ 6.1. Matrices as transformations.

Def. A <u>transformation</u> is a function whose domain and codomain are the set of vectors, and it is denoted by capital letters such as F, T, or L.

If T is a transformation that maps $\mathbf{x}$ into $\mathbf{w}$, then $\mathbf{w} = T(\mathbf{x})$ can be written as $\mathbf{x} \xrightarrow{T} \mathbf{w}$.

If T is a transformation whose domain is $\mathbb{R}^n$ and codomain is $\mathbb{R}^m$, then we will write $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

If $T: \mathbb{R}^n \to \mathbb{R}^n$, then T is called an <u>operator</u> on $\mathbb{R}^n$.

A matrix transformation is a transformation induced by a matrix.

For an $m \times n$ matrix A, the matrix transformation T_A is

$$T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m \quad \& \quad \mathbf{x} \xrightarrow{T_A} A\mathbf{x}.$$

If A is $n \times n$, then T_A is called a matrix operator on $\mathbb{R}^n$.

E.g.) ① 0 : $m \times n$ zero matrix

$\Rightarrow T_0$ is the zero transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$

② I : $n \times n$ identity matrix

$\Rightarrow T_I$ is the identity operator on $\mathbb{R}^n$.

Note For an $m \times n$ matrix,

$\boxed{\text{Solve a linear system } A\mathbf{x} = \mathbf{b}}$

$= \boxed{\text{Find } \mathbf{x} \in \mathbb{R}^n \text{ whose image under } T_A \text{ is } \mathbf{b} \in \mathbb{R}^n}$

Recall that an $m \times n$ matrix A defines a matrix transformation

$$T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m \quad \& \quad \mathbf{x} \xrightarrow{T_A} A\mathbf{x}.$$

If A is an $n \times n$ matrix, then T_A is called a matrix operator.

Since $A(c\mathbf{u}+\mathbf{v}) = cA\mathbf{u} + A\mathbf{v}$ for $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ & $c \in \mathbb{R}$,

$$T_A(c\mathbf{u}+\mathbf{v}) = cT_A(\mathbf{u}) + T_A(\mathbf{v}).$$

i.e., T_A satisfies "linearity property".

<u>Def.</u> A transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a <u>linear transformation</u> if $\forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ & $\forall\, c \in \mathbb{R}$, it satisfies

$$\begin{cases} \text{(i)} \quad T(c\mathbf{u}) = cT(\mathbf{u}) & \text{[homogeniety]} \\ \text{(ii)} \quad T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) & \text{[additivity]} \end{cases}$$

If $n=m$, then T is a <u>linear operator.</u>

$\therefore$ Every matrix operator is a linear transformation.

Conversely, every linear transformation is a matrix transformation.

<u>Thm 6.1.4</u> If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear, then there exists an $m \times n$ matrix A such that $T(\mathbf{x}) = A\mathbf{x}$.

proof) $\forall\, \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$, $\quad \mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$.

$$\therefore\ T(\mathbf{x}) = x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) + \cdots + x_n T(\mathbf{e}_n)$$

$$= \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Putting $A = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix}$, we obtain $T(\mathbf{x}) = A\mathbf{x}$. $\square$

Denote $[T] := \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix}$.

$\uparrow$ the <u>standard matrix</u> for T.

If $[T] = A$, then we say that T is the transformation corr. to A, or T is the transformation represented by A.

$\therefore\ T$: linear $\Rightarrow T(\mathbf{x}) = [T]\mathbf{x}$.

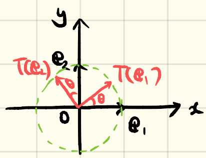
Let V be a subspace of $\mathbb{R}^n$. Then $T: V \to \mathbb{R}^m$ is <u>linear</u> if T satisfies homogeneity & additivity properties.

※ We will define the notion of basis for a subspace V of $\mathbb{R}^n$, and then find a matrix which represents a linear transformation $T: V \to \mathbb{R}^m$.

< Some special linear operators on $\mathbb{R}^2$ >

① The reflection about the origin through an angle θ is the transformation represented by the matrix

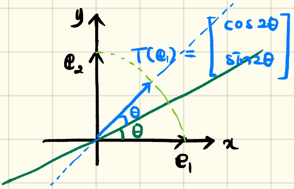
$$R_\theta = \begin{bmatrix} | & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) \\ | & | \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

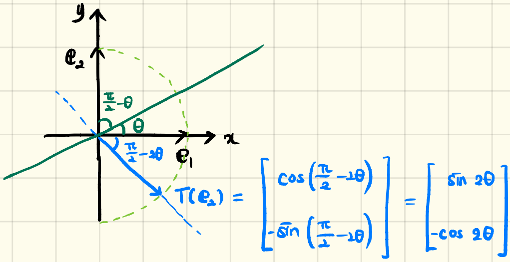


② The reflection about a line through the origin is the transformation represented by the matrix

$$H_\theta = \begin{bmatrix} T(e_1) & T(e_2) \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix},$$

where the line makes an angle θ with the positive x-axis.

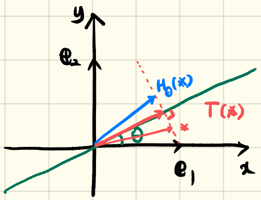




③ The orthogonal projection onto a line through the origin is the transformation represented by the matrix

$$P_\theta = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix},$$

where the line makes an angle θ with the positive x-axis.



Note that $T(x) - x = \frac{1}{2} H_\theta(x) - x$

$$\therefore T(x) = \frac{1}{2}(H_\theta + I)\,x$$

$$\therefore P_\theta = \frac{1}{2}\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos\theta \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$$

$$\left(\begin{aligned} \because\ \cos 2\theta &= \cos^2\theta - \sin^2\theta \\ \sin 2\theta &= 2\sin\theta\cos\theta \\ \cos^2\theta + \sin^2\theta &= 1 \end{aligned} \right)$$

Reading assignment: P. 276 ~ 277.

§6.2. Geometry of linear operators

Def. A linear operator $T: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ with the length-preserving property (i.e., $\|T(\mathbf{x})\| = \|\mathbf{x}\| \ \forall \mathbf{x} \in \mathbb{R}^n$) is called an orthogonal operator or a linear isometry.

e.g.) R_θ, H_θ, P_θ are linear operators on $\mathbb{R}^2$.

Thm 6.2.1 Let $T: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be linear. TFAE.

(a) $\|T(\mathbf{x})\| = \|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^n$

(b) $T(\mathbf{x}) \cdot T(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$.

proof) (b) $\Rightarrow$ (a)

$$\|T(\mathbf{x})\| = \sqrt{T(\mathbf{x}) \cdot T(\mathbf{x})} = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^n$$

(a) $\Rightarrow$ (b)

Note that $\mathbf{x} \cdot \mathbf{y} = \frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right) \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

$$\begin{aligned} \therefore \ T(\mathbf{x}) \cdot T(\mathbf{y}) &= \frac{1}{4}\left(\|T(\mathbf{x})+T(\mathbf{y})\|^2 - \|T(\mathbf{x})-T(\mathbf{y})\|^2\right) \\ &= \frac{1}{4}\left(\|T(\mathbf{x}+\mathbf{y})\|^2 - \|T(\mathbf{x}-\mathbf{y})\|^2\right) \\ &= \frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right) \\ &= \mathbf{x} \cdot \mathbf{y} \end{aligned}$$

$\blacksquare$

$\therefore$ A linear operator is orthogonal iff it preserves dot product.

$\therefore$ Every orthogonal operator preserves angles, but the converse is not true.

Thm 6.2.4 Let A be an $m\times n$ matrix. TFAE.

(a) $A^TA = I$ (b) $\|A\mathbf{x}\| = \|\mathbf{x}\|$ $\forall \mathbf{x}\in\mathbb{R}^n$ (c) $A\mathbf{x}\cdot A\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$ $\forall \mathbf{x},\mathbf{y}\in\mathbb{R}^n$

(d) The column vectors of A are orthonormal.

proof) (a) $\Rightarrow$ (b)

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})\cdot(A\mathbf{x}) = (A\mathbf{x})^TA\mathbf{x} = \mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2$$

$\therefore \|A\mathbf{x}\| = \|\mathbf{x}\|$

(b) $\Rightarrow$ (c) (Similar to the proof of Thm 6.2.1)

$$(A\mathbf{x})\cdot(A\mathbf{y}) = \frac{1}{4}\left(\|A\mathbf{x}+A\mathbf{y}\|^2 - \|A\mathbf{x}-A\mathbf{y}\|^2\right) = \frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right) = \mathbf{x}\cdot\mathbf{y}$$

(c) $\Rightarrow$ (d) Note that $\mathbf{c}_i(A) = A\mathbf{e}_i$.

Hence $\mathbf{c}_i(A)\cdot\mathbf{c}_j(A) = (A\mathbf{e}_i)\cdot(A\mathbf{e}_j) = \mathbf{e}_i\cdot\mathbf{e}_j = \begin{cases} 1 & (i=j) \\ 0 & (i\neq j)\end{cases}$

(d) $\Rightarrow$ (a) $(A^TA)_{ij} = \mathbf{r}_i(A^T)\,\mathbf{c}_j(A) = \mathbf{c}_i(A)\cdot\mathbf{c}_j(A) = \begin{cases} 1 & (i=j) \\ 0 & (i\neq j)\end{cases}$

$\blacksquare$

Def. A square matrix A is orthogonal if $A^{-1} = A^T$.

e.g.) Set E_{ij} be the elementary matrix obtained from I_n by interchanging row i and row j. Then $E_{ij}^{-1} = E_{ij} = E_{ij}^T$

$\therefore E_{ij}$ is orthogonal.

Thm 6.2.3 Let A and B be orthogonal matrices of the same size.

(a) A^T is orthogonal.

(b) A^{-1} is orthogonal.

(c) AB is orthogonal.

(d) $\det(A) = \pm 1$

$\therefore$ Every permutation matrix is orthogonal.

Thm 6.2.5 Let A be an $n\times n$ matrix. TFAE.

(a) A is orthogonal

(b) $\|A\mathbf{x}\| = \|\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^n$

(c) $A\mathbf{x}\cdot A\mathbf{y} = \mathbf{x}\cdot\mathbf{y} \quad \forall \mathbf{x},\mathbf{y} \in \mathbb{R}^n$

(d) The column vectors of A are orthonormal.

(e) The row vectors of A are orthonormal.

pf) Note that A is orthogonal iff $AA^T = A^TA = I$.

The first four statements follow from Thm 6.2.4.

(e) $\Leftrightarrow$ $AA^T = I$. ◩

Thm 6.2.6 A linear operator $T: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is orthogonal iff $[T]$ is orthogonal.

$$\left(\because \quad T: \text{orthogonal} \Leftrightarrow T(\mathbf{x})\cdot T(\mathbf{y}) = \mathbf{x}\cdot\mathbf{y} \Leftrightarrow ([T]\mathbf{x})\cdot[T]\mathbf{y} = \mathbf{x}\cdot\mathbf{y} \Leftrightarrow [T] \text{ is orthogonal.} \right)$$

e.g.) $R_\theta^T R_\theta = I$ & $H_\theta^T H_\theta = I$.

Thm 6.2.7 An orthogonal operator $T:\mathbb{R}^2\to\mathbb{R}^2$ is either the rotation about the origin through an angle θ or the reflection about a line through the origin.

pf) Set $[T] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Since $\begin{bmatrix} a \\ c \end{bmatrix}$ and $\begin{bmatrix} b \\ d \end{bmatrix}$ are orthonormal,

$$a^2+c^2 = 1, \quad b^2+d^2=1, \text{ and } \quad ab+cd=0.$$

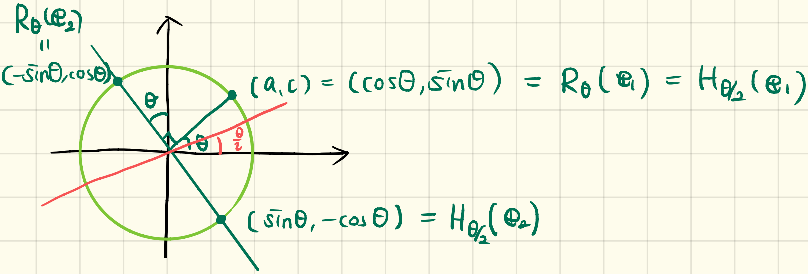
Since $P(a,c)$ and $Q(b,d)$ are on the unit circle, we can set

$$\begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \text{ and } \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} \cos(\theta\pm\frac{\pi}{2}) \\ \sin(\theta\pm\frac{\pi}{2}) \end{bmatrix}.$$

If $\begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} \cos(\theta + \frac{\pi}{2}) \\ \sin(\theta + \frac{\pi}{2}) \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$, then $[T] = R_\theta$.

If $\begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} \cos(\theta - \frac{\pi}{2}) \\ \sin(\theta - \frac{\pi}{2}) \end{bmatrix} = \begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix}$, then $[T] = H_{\theta/2}$. $\boxed{}$

※ $\det(R_\theta) = 1$ & $\det(H_{\theta/2}) = -1$.



Reading assignment: page 286 ~