Orthogonal Transformations

Prerequisites:

- Linear Transformations
- Dot products

A particularly important class of linear transformations on $\mathbb{R}^N$ are the **orthogonal** transformations. Orthogonal transformations transform $\mathbb{R}^N$ as though it was a *rigid* object. Thus, the *lengths* of lines and the *angles* between them are preserved. For example, *rotations* and *reflections* are orthogonal transformations.

Theorem 1: *Orthogonal Transformations*
Let $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be a linear transformation, equivalent to multiplication by the matrix $\boxed{\mathsf{F}}$. The following are equivalent:

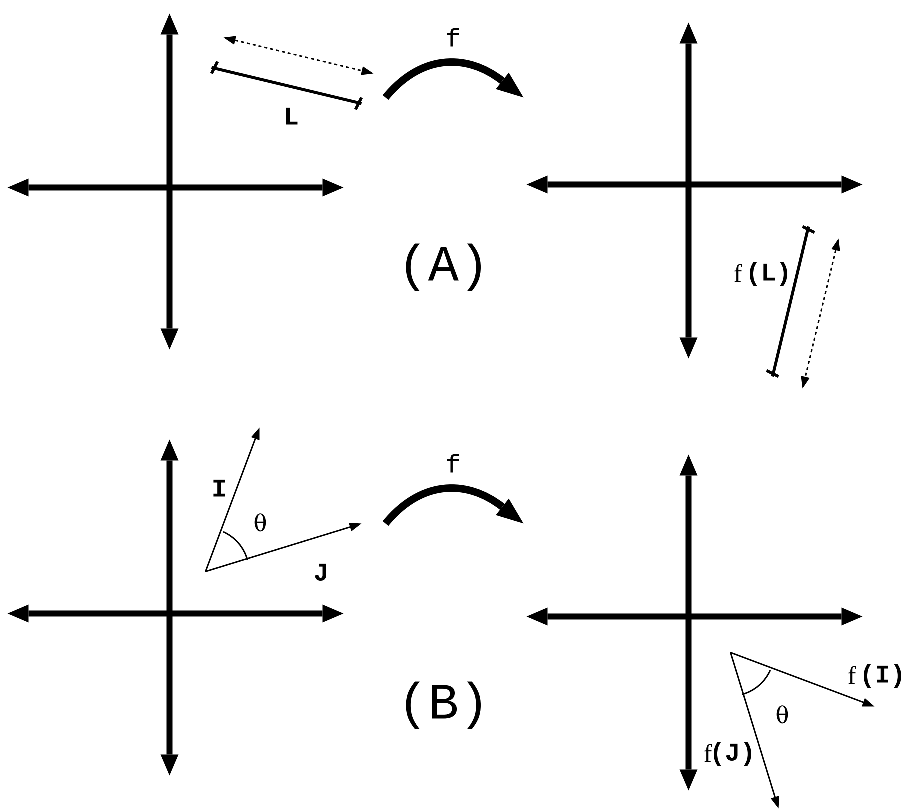

Figure 1: Orthogonal transformations preserves lengths and angles

1. f preserves *lengths* and *angles*. In other words, if $\mathbb{L}$ is a line-segment in $\mathbb{R}^N$, then $\ell_{ength}[f(\mathbb{L})] = \ell_{ength}[\mathbb{L}]$ (see **Part (A)** of Figure 1). If $\mathbb{I}$ and $\mathbb{J}$ are two

line-segments which intersect with an angle of θ between them, then the line segments $f(\mathbb{I})$ and $f(\mathbb{J})$ also intersect with an angle of θ (see **Part (B)** of Figure 1 on the page before).

2. f preserves the *dot products* of vectors. In other words, for any $\mathbf{v}, \mathbf{w} \in \mathbb{R}^N$,

$$(f(\mathbf{v})) \bullet (f(\mathbf{w})) \quad = \quad \mathbf{v} \bullet \mathbf{w}$$

3. If $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_N\}$ is any **orthonormal basis** for $\mathbb{R}^N$, then $\{f(\mathbf{b}_1), \ \ldots, \ f(\mathbf{b}_N)\}$ is *also* an orthonormal basis for $\mathbb{R}^N$.

4. If $\boxed{\mathsf{F}} = \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}$, then the column vectors $\{\mathbf{c}_1, \ldots, \mathbf{c}_N\}$ form an orthonormal basis for $\mathbb{R}^N$.

5. $\boxed{\mathsf{F}}^{-1} = \boxed{\mathsf{F}}^{t}$

6. If $\boxed{\mathsf{F}} = \begin{bmatrix} \longleftarrow & \mathbf{r}_1 & \longrightarrow \\ \longleftarrow & \mathbf{r}_2 & \longrightarrow \\ \vdots & \vdots & \vdots \\ \longleftarrow & \mathbf{r}_N & \longrightarrow \end{bmatrix}$, then the row vectors $\{\mathbf{r}_1, \ldots, \mathbf{r}_N\}$ form an orthonormal basis for $\mathbb{R}^N$.

Proof:

Proof of "(1) $\iff$ (2)": Recall that the *length* of a vector $\mathbf{v}$ is given by: $\|\mathbf{v}\| = \sqrt{\mathbf{v} \bullet \mathbf{v}}$, and that the *angle* between two vectors is given:

$$angle(\mathbf{v}, \mathbf{w}) \quad = \quad \arccos\left(\frac{\mathbf{v} \bullet \mathbf{w}}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|}\right)$$

Thus, a linear transformation preserves *all angles* and *all lengths* if and only if it preserves *all dot products.*

Proof of "(2)$\Longrightarrow$(3)": If $\{\mathbf{b}_1, \ldots, \mathbf{b}_N\}$ is an orthonormal set, and f preserves all inner products, then $\{f(\mathbf{b}_1), \ \ldots, \ f(\mathbf{b}_N)\}$ must *also* be an orthonormal set, because for any i, j,

$$f(\mathbf{b}_i) \bullet f(\mathbf{b}_j) \quad = \quad \mathbf{i} \bullet \mathbf{j}$$

Proof of "(3)$\Longrightarrow$(4)": Let $\mathcal{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ be the standard basis. Then $\mathcal{E}$ is an orthonormal basis; thus $\{f(\mathbf{e}_1), \ldots, f(\mathbf{e}_N)\}$ must also be an orthonormal basis. But $f(\mathbf{e}_1) = \mathbf{c}_1, \ \ldots, \ f(\mathbf{e}_N) = \mathbf{c}_N$.

Proof of "(4)$\Longrightarrow$(2)": Let $\mathbf{v} = (v_1, \ldots, v_N)$ and $\mathbf{w} = (w_1, \ldots, w_N)$ be vectors in $\mathbb{R}^N$. Then

$$
\begin{aligned}
(f(\mathbf{v})) \bullet (f(\mathbf{w})) &= \left(f\left(\sum_{n=1}^{N} v_n \mathbf{e}_n\right)\right) \bullet \left(f\left(\sum_{m=1}^{N} w_m \mathbf{e}_m\right)\right) \\
&=_{(1)} \left(\sum_{n=1}^{N} v_n f(\mathbf{e}_n)\right) \bullet \left(\sum_{m=1}^{N} w_m f(\mathbf{e}_m)\right) \\
&=_{(2)} \left(\sum_{n=1}^{N} v_n \mathbf{c}_n\right) \bullet \left(\sum_{m=1}^{N} w_m \mathbf{c}_m\right) \\
&=_{(3)} \sum_{n=1}^{N} \sum_{m=1}^{N} v_n w_m \left(\mathbf{c}_n \bullet \mathbf{c}_m\right) \\
&=_{(4)} \sum_{n=1}^{N} v_n w_n \\
&= \mathbf{v} \bullet \mathbf{w}.
\end{aligned}
$$

(1) Because f is linear.

(2) Because $f(\mathbf{e}_n) = \mathbf{c}_n$.

(3) Because $\bullet$ distributes through addition.

(4) Because $\{\mathbf{c}_1, \ldots, \mathbf{c}_N\}$ is an orthonormal set.

Proof of "(4) $\Longleftrightarrow$ (5)": Consider the matrix $\boxed{\mathrm{A}} = \boxed{\mathrm{F}}^t \cdot \boxed{\mathrm{F}}$. Now,

$$
\boxed{\mathrm{F}} = \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}, \qquad \text{while } \boxed{\mathrm{F}}^t = \begin{bmatrix} \longleftarrow & \mathbf{c}_1 & \longrightarrow \\ \longleftarrow & \mathbf{c}_2 & \longrightarrow \\ \vdots & \vdots & \vdots \\ \longleftarrow & \mathbf{c}_N & \longrightarrow \end{bmatrix}
$$

Thus, if a_{ij} is the (i, j)th entry of $\boxed{\mathrm{A}}$, then

$$
a_{ij} = \mathbf{c}_i \bullet \mathbf{c}_j.
$$

Hence,

$$
\begin{aligned}
\left(\boxed{\mathrm{F}}^{-1} = \boxed{\mathrm{F}}^t\right) &\Longleftrightarrow \left(\boxed{\mathrm{A}} = \boxed{\mathrm{F}}^t \cdot \boxed{\mathrm{F}} = \boxed{\mathbf{Id}}\right) \\
&\Longleftrightarrow \left(a_{ij} = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq j \end{cases}\right)
\end{aligned}
$$

$$\iff \left(\mathbf{c}_i \bullet \mathbf{c}_j = \begin{cases} 1 & if \quad i = j \\ 0 & if \quad i \neq j \end{cases} \right)$$

$$\iff \left(\{\mathbf{c}_1, \ldots, \mathbf{c}_N\} \text{ are an orthonormal basis} \right).$$

Proof of "(5) $\iff$ (6)": The proof is identical to "(4) $\iff$ (5)"; now consider the matrix $\boxed{\mathrm{A}} = \boxed{\mathrm{F}} \cdot \boxed{\mathrm{F}}^t$, and show that this matrix is the identity if and only if the *row vectors* of $\boxed{\mathrm{F}}$ are orthonormal.

□ [Theorem 1]

Definition 2: *Orthogonal Transformation*

Let $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be a linear transformation. f is called an **orthogonal transformation** if it satisfies any (and thus, *all*) of the conditions of the previous theorem.

If $\boxed{\mathrm{F}}$ is the matrix of an orthogonal transformation, then we say $\boxed{\mathrm{F}}$ is an **orthogonal matrix**.

Examples 3:

1. The **identity map** is an orthogonal transformation.

2. If $\theta \in [0, 2\pi]$, and $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is **rotation by** θ about the origin, then f is an orthogonal transformation. The matrix of f is

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Thus, the vectors $[\cos(\theta),\ \sin(\theta)]$ and $[-\sin(\theta),\ \cos(\theta)]$ form an orthonormal basis for $\mathbb{R}^2$.

3. If $\theta \in [0, 2\pi]$, and $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is **rotation by** θ about the second ("y") axis, then f is an orthogonal transformation. The matrix of f is

$$\begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix}$$

Thus, the vectors $[\cos(\theta),\ 0,\ \sin(\theta)]$, $[0, 1, 0]$, and $[-\sin(\theta),\ 0,\ \cos(\theta)]$ form an orthonormal basis for $\mathbb{R}^3$.

4. If $\mathbb{V} \subset \mathbb{R}^N$ is a linear subspace, and $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is the **reflection across** $\mathbb{V}$:

$$f(\mathbf{v}) \;=\; 2\mathbf{pr}_{\mathbb{V}}(\mathbf{v}) - \mathbf{v},$$

then f is an orthogonal transformation.

Remark 4:

1. If $f, g : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ are orthogonal transformations, then so are f^{-1} and $f \circ g$.

2. If $\boxed{\mathrm{A}}$, $\boxed{\mathrm{B}}$ are orthogonal $N \times N$ matrices then so are $\boxed{\mathrm{A}}^{-1}$ and $\boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}}$.

3. If $\boxed{\mathrm{A}}$ is the matrix of an orthogonal transformation f, then $\boxed{\mathrm{A}}^{t}$ is *also* the matrix of an orthogonal transformation: f^{-1}.

4. Suppose $g : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is some linear transformation, whose matrix representation (relative to the standard basis) is $\boxed{\mathrm{G}}$. Suppose $\mathcal{B}$ is an orthonormal basis for $\mathbb{R}^N$, and we want to compute the matrix representation of g with respect to $\mathcal{B}$. Let $\boxed{\mathrm{B}}$ be the matrix whose **column vectors** are the elements of $\mathcal{B}$. Then we know that

$$\widetilde{\boxed{\mathrm{G}}} = \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{G}} \cdot \boxed{\mathrm{B}}$$

is the matrix representation of g with respect to $\mathcal{B}$. But $\boxed{\mathrm{B}}$ is an orthogonal matrix, hence $\boxed{\mathrm{B}}^{-1} = \boxed{\mathrm{B}}^{t}$. Hence, the matrix representation of g relative to $\mathcal{B}$ can also be written:

$$\widetilde{\boxed{\mathrm{G}}} = \boxed{\mathrm{B}}^{t} \cdot \boxed{\mathrm{G}} \cdot \boxed{\mathrm{B}}$$